Selectivity of Substance Identification by HPLC–DAD in Toxicological Analysis using a UV Spectra Library of 2682 Compounds

Matthias Herzler, Sieglinde Herre, and **Fritz Pragst***
Institute of Legal Medicine, Humboldt-University, Hannoversche Straße 6, D-10115 Berlin, Germany

Abstract

The UV spectra and relative retention times (RRT) of 2682 toxicologically relevant substances were measured by high-performance liquid chromatography with diode array detection (HPLC–DAD) in an acetonitrile phospate buffer (pH 2.3) mixture on an RP8 column and were arranged in a database. A complete survey of the molecular structures of all database entries showed the presence of 1650 different chromophores or chromophore combinations. The specificity of the UV spectrum for substance identification was determined by calculation of the similarity indices (SI) of all possible substance pairs within the database with an SI > 0.9990, which indicated spectral identity. In a similar way, the RRT was evaluated for all possible pairs: two compounds were declared indistinguishable because the RRT of at least one of them fell into the RRT error window of the other. Although the use of the RRT alone produced unsatisfactory identification results, 1619 substances (60.4%) were unambiguously identified by their UV spectrum only. This rate was increased to 84.2% by the combination of spectrum and RRT. The selectivity parameters discrimination power (DP) and mean list length (MLL) were calculated (DP = 0.9999, MLL = 1.253) and compared with literature data, which proved HPLC–DAD to be one of the most reliable methods for substance identification in toxicological analysis. The practical relevance of the results for systematic toxicological analysis is demonstrated by the example of a multidrug intoxication and in the context of sample preparation methods routinely used.

Introduction

Systematic toxicological analysis (STA, general unknown analysis) is one of the most challenging tasks of chemical analysis (1–3). Besides gas chromatography (GC)–mass spectrometry (MS) (4–6) and more recently liquid chromatography (LC)–MS (7–10), high-performance liquid chromatography with photodiode array detection (HPLC–DAD) is one of the most common methods used for substance identification in this context. The identification is based on the comparison of the UV absorption spectrum and a retention parameter [absolute or relative retention time (RT, RRT, respectively) or retention index] with corresponding data stored in a spectra library of toxic compounds.

UV spectra libraries used for this purpose were described by Bogusz and Wu (11) (225 substances), Bogusz and Erkens (12) (383 substances), Elliott and Hale (13) (250 substances), Gaillard and Pépin (14) (600 substances), Koves (15) (272 substances), Lai et al. (16) (300 drugs or metabolites), Lambert et al. (17) (130 substances), Logan et al. (18) (100 substances), Maier and Bogusz (19) (370 substances), Tracqui et al. (20) (311 substances), and Turcant et al. (21) (350 substances).

In most cases, RTs or RTs related to a standard compound and corrected for the dead time were used as retention parameters. Bogusz et al. (11), Hill and Kind (22), and Elliott and Hale (13) used a system of retention indices for HPLC gradient systems on the basis of nitroalkanes or acidic and basic standard substances to increase reproducibility.

A UV spectra library of 2682 compounds was produced by the authors of this paper (23). It was found that of a total of 2888 compounds, only 206 displayed no significant UV absorption in the spectral range examined. The spectra and RTs were measured under isocratic conditions in an acetonitrile phosphate buffer (pH 2.3) mixture. This mobile phase is used widely in pharmacological and toxicological analysis.

Over the years, HPLC–DAD has been proven to be a reliable and reproducible method for substance identification in clinical as well as forensic toxicology by many laboratories, where it is appreciated for its ruggedness and the relatively simple sample preparation required. Nevertheless, it is held in low esteem by many analysts because UV spectra display only a limited number of maxima and minima with no simple correlation to chemical structure. This controversy between practical experience and general analytical feeling should be resolved in a more exact way. Therefore, in this paper the question of the selectivity of this method for substance identification was systematically investigated on the basis of 2682 sets of UV spectra,

* Author to whom correspondence should be addressed. E-mail: fritz.pragstr@charite.de.

RRT, and structural formulae contained in the library mentioned previously (23).

Material and methods

Toxicologically relevant compounds

The 2682 substances included in the database and 206 further compounds, which were found to have no UV absorption, were selected from all fields of toxicological and pharmacological interest. All groups of therapeutic drugs were included with a particular consideration for those pharmaceuticals most frequently applied. In most cases, the drugs were generously donated by a large number of manufacturers, the enumeration of which is impossible in the frame of this section. Opiates, cannabinoids, cocaine and its metabolites, amphetamines, and many designer drugs, as well as a number of hallucinogens were purchased from the companies Sigma (Taufkirchen, Germany), Radian (Austin, Texas), or Lipomed (Arlesheim, Switzerland). All types of pesticides were included. These compounds were mostly purchased from the firms Promochem (Wesel, Germany), Riedel de Haën (Seelze, Germany), or Labor Dr. Ehrenstorfer (Augsburg, Germany). From these manufacturers, a number of ecotoxic compounds (polycyclic aromatic hydrocarbons, polychlorinated biphenyls, and chlorinated phenols) were also obtained. Finally, a series of alkaloids including strychnine, brucine, nicotine or toxic glycosides, some aromatic or unsaturated solvents, and some frequently applied chemical reagents were also included. Almost all substances were organic compounds, and only a very small number of inorganic anions such as NO_2^- and CrO_4^- were analyzed. A complete list of all compounds is given by Pragst et al. (23) and can also be obtained directly from the authors. All substances were given a code that consisted of the first letter of the substance name and a three-digit number assigned in the order of registration in the laboratory (e.g., B145 for bufotenine).

Instruments

Each compound was measured with the following three HPLC–DAD devices: (1) A Shimadzu HPLC device (Shimadzu Europe GmbH, Duisburg, Germany) that consisted of a pump LC-6A, an autosampler SIL-9A, a DAD SPD-M10Avp, and a computer (standard IBM PC). The SPD-M10Avp was operated by the software Class-VP 5.032; (2) A Bio-Tek HPLC device (Bio-Tek Instruments GmbH, Neufahrn, Germany) that consisted of a pump 525, sampler 560, and DAD 540 operated with the software Kroma 2000; and (3) a Hewlett-Packard (Agilent) HPLC device HP 1090 Series II (Agilent Technologies, Waldbronn, Germany) operated with the software HPLC-Chemstation Rev. A.02.05.

Wavelength calibration and accuracy checks of the DAD were performed regularly. The UV spectra obtained from the three devices were in excellent agreement with each other with respect to wavelength resolution, as well as to wavelength accuracy (24). For the evaluation presented in this paper, the spectra recorded with the Shimadzu SPD-M10Avp detector were used.

The mobile phase was degassed by vacuum degassers DG-1210 (Optilab Chromatographie Technik, Berlin, Germany).

HPLC–DAD conditions

A reversed-phase HPLC column Lichrospher-RP8ec (5 µm, 250 × 4.0 mm) (Merck, Darmstadt, Germany) and two isocratic mobile phases A and B that consisted of acetonitrile (UV grade, Merck), and 0.1M phosphate buffer pH 2.3 in different ratios (A 37:63, v/v; B 63:37, v/v) were used. The measurements were carried out at room temperature (20–25°C) at a flow rate of 1 mL/min with a run time of 30 min.

For the control of accurate performance, 10 µL of a test solution that consisted of the following substances dissolved in the mobile phase was measured at the beginning of each series: 0.1 mg/mL histamine for the determination and control of t_0 (time of an unretained peak); 0.1 mg/mL caffeine for control of the autosampler (check of peak area); 0.1 mg/mL of 5-(4-methylphenyl)-5-phenylhydantoin (MPPH, reference substance for the calculation of the RRT in mobile phase A) or 0.1 mg/mL 4-phenylbenzophenone (PBP, reference substance for the calculation of the RRT in mobile phase B), and 1 mg/mL benzene for control of the resolution, accuracy, and reproducibility of the UV spectrum (check of vibration bands). The test solution was stored in a brown glass vessel at 0–4°C and was stable under these conditions for at least three months.

Measurement of the UV spectra and RRT

For the measurement of the UV spectra, 1 µg of each substance (10 µL of a 0.1-mg/mL solution in mobile phases A and B or in acetonitrile) was injected. The UV spectrum between 195 and 380 nm at the maximum of the chromatographic peak was corrected with respect to the absorbance of the solvent and stored. This wavelength range was chosen as a reasonable compromise between analytical demands and the aim of limiting data file size. Only 173 (6.4%) of the 2682 compounds displayed considerable absorption above 380 nm, however the corresponding spectra were also sufficiently specific below this limit.

The RRT of a compound i (RRT_i) was calculated from the absolute retention time of i (RT_i), the time of an unretained peak (t_0), and the RT of the reference substance (RT_{MPPH} or RT_{PBP}) using:

$$RRT_i = (RT_i - t_0)/(RT_{MPPH} - t_0) \qquad \text{Eq. 1}$$

The spectra and RRTs of all 2682 compounds were published and are available as a book, as well as on CD-ROM (23) for use with customary HPLC software.

Software operations

The structural formulae of all compounds were drawn using ISIS Draw 2.3 and stored in a substance database using ISIS Base 2.1.1 (both MDL Molecular Design Ltd., San Leandro, CA). All chromophore substructures were determined and registered into an additional chromophore database.

UV spectra were recorded using the DAD software CLASS-VP 5.032 (Shimadzu Corp., Kyoto, Japan) and were afterwards converted into the ASCII format. All 2682 spectra were then im-

ported into Excel 2000 (Microsoft Corp., Redmond, WA), which resulted in a data matrix of 2682 substance rows and 181 absorbance columns that corresponded to the wavelength region from 200 to 380 nm. Absorbance values from 195 to199 nm were omitted because of potentially increased disturbances by the solvent. The calculation of the similarity indices (SI) between all pairs of spectra was performed from this matrix using MatLab Version 5.3.0.10183 (R11) (The Math Works Inc., Natick, MA).

Analysis of real samples

The procedures used for sample preparation in the routine application of HPLC–DAD for STA were described in detail in previous papers (3,25).

Results and Discussion

The absorption spectrum in the UV wavelength range investigated in this paper originates mainly from the interaction of UV radiation with a conjugated system of π-electrons and free electron pairs of hetero atoms in the form of $\pi \rightarrow \pi^*$ or $n \rightarrow \pi^*$ transitions. The involvement of σ-electrons ($\sigma \rightarrow \sigma^*$ or $n \rightarrow \sigma^*$ transitions) is rather an exception and is limited to some compounds containing sulfur, bromine, or iodine. Therefore, not the complete molecule, but only the conjugated system of π-electrons and free electron pairs of hetero atoms (the chromophore) is responsible for UV absorption. In general, unsaturated, but nonconjugated parts of the molecule, as well as isolated substituents do not or only indirectly contribute to the absorption. Nevertheless, hyperconjugation effects also play a certain role. Thus, in the authors' definition of a chromophore, substituents without *n*- or π-electrons conjugated to the chromophore were included up to the first saturated carbon atom (see examples that follow).

A high selectivity of UV spectra for substance identification would mean that the spectrum of each compound could be unambiguously identified in the presence of other compounds randomly picked from the investigated substance pool. A highly specific relationship between spectrum and chemical structure would be the prerequisite. In order to find out to what extent this demand was met for the library spectra, two strategies were chosen: (i) statistical assessment of the structural diversity of the registered chromophores and (ii) determination of the distinctness of each spectrum by similarity calculations between all spectra.

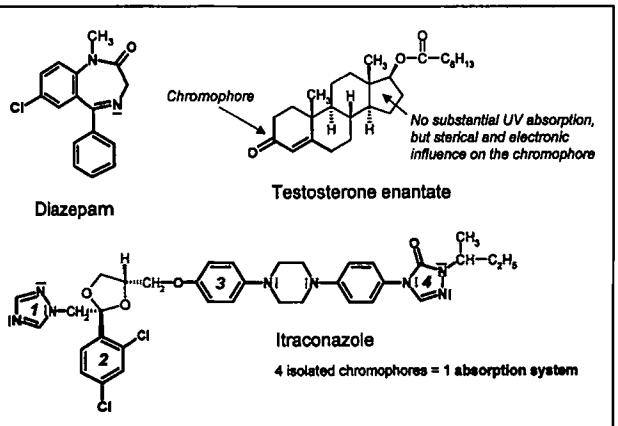

Figure 1. Structural formulae and chromophores of diazepam, testosterone enantate, and itraconazole. Chromophores are drawn bold.

Diversity of chromophores and absorption systems

In Figure 1, the structural formulae of three compounds are shown with their chromophores drawn bold. For diazepam, the chromophore is almost identical with the whole molecule. However, in the case of testosterone enantate the absorption is limited to the α,β-unsaturated carbonyl group. The predominant part of the molecule has no absorption in the investigated wavelength range, but it may change the absorption of the chromophore by electronic or sterical effects. Itraconazole possesses four different chromophores separated from each other by saturated carbon atoms. In this case, the UV spectrum of the molecule is roughly the sum of the spectra of the four chromophores. In the following, those structural parts of a molecule involved in UV absorption (one or more isolated chromophores) will be referred to as the absorption system.

For 2560 of the 2682 compounds, such an absorption system could be defined. In the remaining 122 cases (saturated substances or inorganic anions), the explanation for the UV absorption was more complicated. By counting out, it was found that in the 2560 compounds with a defined absorption system 1486 different chromophores and a total of 1650 different absorption systems were present. Some of the chromophores never appeared in combination with others, whereas some combinations would contain chromophores not found alone. The frequency distribution is shown in Figure 2. Only once did 1370 absorption system occur, 170 twice, and 47 three times. Nevertheless, some chromophores were found with a high frequency such as the toluene system (125×) or the 4-methylpent-3-en-2-one substructure of the steroids (64×).

Practical experience, as well as the findings from the spectral similarity calculations of this work (see following) lead to the conclusion that structurally different absorption systems will almost always be distinguishable from each other by their UV

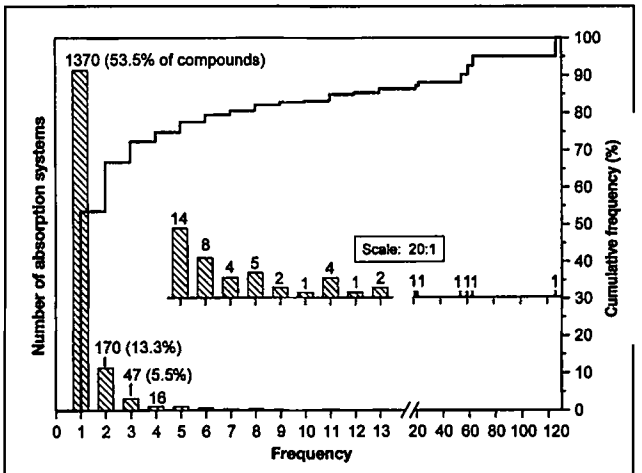

Figure 2. Frequency of different absorption systems in 2560 toxicologically relevant compounds.

spectra. At least 1370 compounds (53.5%) possessing a unique absorption system should, therefore, exhibit a unique UV spectrum, as well. On the other hand, different compounds with the same chromophore or absorption system are likely to produce similar spectra. However, because of sterical or electronic effects of the residual molecule, even in theses cases, spectral differences were often found. As an example, in Figure 3 the spectra of 32 steroids with the 3-methylcyclohexadienone chromophore are shown. The isolated substituents R_1 and R_2 at ring B are either hydrogen or fluorine. Depending on the number and position of the fluorine atoms, the spectra can roughly be subdivided into four groups. The more distant substituents R_3–R_5 lead to a further differentiation within these groups, probably owed to sterical effects.

Selectivity parameters

For the characterization of the selectivity of an analytical method, several parameters have been introduced in the past, of which two were chosen here. The discrimination power (DP) (26) is a measure for the probability that two compounds selected by chance from the substance pool under consideration can be distinguished by the method. It is calculated by:

$$DP = 1 - [(u - n)/n(n - 1)] \quad \text{Eq. 2}$$

where u is the number of indistinguishable substance pairs (all fields with a 1 in the identification matrix, see following), n is the total number of substances under consideration, DP_{max} is 1 (all compounds can be distinguished from each other), and DP_{min} is 0 (all compounds are indistinguishable).

The mean list length (MLL) (27) is given by the average number of indistinguishable substance pairs for each compound and is calculated by:

$$MLL = u/n \quad \text{Eq. 3}$$

where MLL_{min} is 1 (each substance is only indistinguishable from itself) and MLL_{max} is n (all compounds are indistinguishable).

The advantages or drawbacks of these parameters are discussed by Herzler (28). In short, it can be said that DP is a rather insensitive parameter with respect to small differences between high-resolution analytical methods, whereas MLL does not consider the number of substances examined (certainly it makes a big difference if a MLL of 1.5, for example, was found for a total of 100 or 1000 compounds). A truly fair comparison between different analytical methods only seems possible if an identical substance collective was examined or if this collective reached a size and diversity in chemical structure representative for the given task.

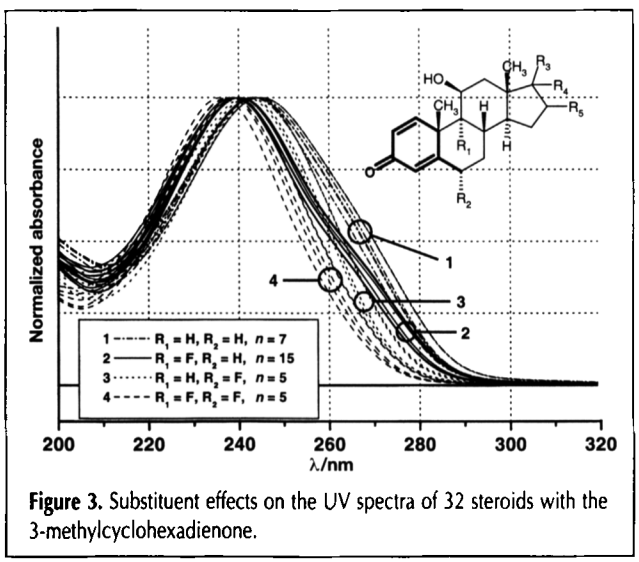

Figure 3. Substituent effects on the UV spectra of 32 steroids with the 3-methylcyclohexadienone.

Specificity of the UV spectra evaluated from SI

Mathematical models for the assessment of spectral similarity use the description of the spectrum as a vector in n-dimensional space, where n is the number of absorbance-wavelength pairs measured. Details are described by Huber and George (29). For the wavelength range from 200 to 380 nm recorded with a spectral resolution of 1 nm, n equals 181. In the case of complete identity of two spectra, both vectors point in exactly the same direction (i.e., the angle between them is θ = 0°). Different concentrations only have an effect on vector length, but not on its direction in space. The SI is defined as cos θ and is calculated by Equation 4. For two identical spectra, cos θ equals 1.0000. Because near 1.0000 the cosine function is relatively insensitive for changes of Θ, already small deviations from 1.0000 can indicate significant differences between the spectra.

$$SI = \cos\theta_{\bar{s}_1\bar{s}_2} = \frac{\bar{s}_1 \bullet \bar{s}_2}{|\bar{s}_1| \times |\bar{s}_2|} \quad \text{Eq. 4}$$

where s_i equals the vectorized spectrum of compound i.

By using Equation 4, the SI of all possible pairs among the 2682 substances of the database (approximately 7.2×10^6 indices) were calculated and entered into a so-called similarity matrix. A small excerpt of this huge matrix is shown in Figure 4a. Because the spectra of all compounds are identical with themselves, the diagonal fields of the matrix always show an SI of 1.0000. All other substance pairs occur twice in the matrix.

With up-to-date DADs, UV spectra can be measured with high sensitivity and reproducibility. Nevertheless, in practical analysis, spectra may be changed to a small extent by noise or matrix effects, and an index of 1.0000 for complete identity is not always met. Therefore, it was necessary to define a practical threshold value of SI above which two spectra were regarded as identical. For this purpose, a large number of library search operations were considered. As an example, in Figure 5 the spectrum of temazepam is compared with the spectra of three other benzodiazepines. It is seen that all three spectra differ from that of temazepam with an increasing SI from lormetazepam (SI = 0.9908) through flurazepam (SI = 0.9965) to camazepam (SI = 0.9993). As a result of the evaluation of such examples, a threshold SI of 0.9990 was chosen. Thus, in Figure 5, only the spectrum of camazepam would be considered identical with that of temazepam.

This threshold value is also in agreement with the authors' experience from the use of the spectra library in daily toxicolog-

ical routine analysis. From undisturbed chromatographic peaks originating from an injected amount > 10 ng, substances were always identified by their library spectrum with SI values above 0.9990. Using customary sample preparation techniques, this substance amount corresponds to serum concentrations in the order of 50–100 ng/mL, or even less.

By use of this criterion, the similarity matrix was transformed into a so-called identification matrix (Figure 4b), in which all substance pairs with SI > 0.9990 were assigned a value of 1 (identical) and all substance pairs with SI ≤ 0.9990 received a value of 0 (different). In this matrix, for each compound the number of identical spectra was counted. The results are shown in Figure 6. Of the 2682 substances, 1619 (60.4%) showed unique spectra. For 399 compounds (14.9%) there was one other and for 175 (6.5%) there were two further substances, each with an indistinguishable spectrum. However, there were also spectra less specific, by far. For example, 28 substances were found to have 15 further candidates each with an identical spectrum. Using Equations 2 and 3, a DP of 0.9994 and MLL of 2.701 were calculated.

Additional use of the RRT

Besides the UV spectrum for all compounds, the RRT in mobile phases A (1993 compounds) or B (689 compounds) was available as an additional criterion for identification. The usefulness of the RRT for identification depends on the accuracy with which it can be reproduced.

Over a period of more than two years, a standard test solution (see HPLC–DAD conditions) was measured with every series of measurements for the spectrum library. The variation coefficients of the RT ($VK_{RT} = (s_{RT}/RT) \times 100\%$) of histamine (dead time), as well as of the reference standards MPPH (mobile phase A) and PBP (mobile phase B) were found to be 3% for mobile phase A (n = 167 for t_0 and 166 for RT_{MPPH}) and 5% for eluent B (n = 49 for t_0 and 71 for RT_{PBP}). Although, in this way a rather intermediate precision (same laboratory, different analysts, different equipment) of the method was obtained; interlaboratory method reproducibility should be comparable because of the long time period covered by these experiments. The intraday VK_{RT} (repeatability) determined by 15 consecutive analyses, each separated by a blank sample in eluent A for his-

(a) Similarity matrix (spectrum only)

CODE	P306	P307	P308	P310	P311	P312
P306	1.0000	0.8589	0.9997	0.4609	0.5938	0.6380
P307	0.8589	1.0000	0.8533	0.3365	0.4417	0.5355
P308	0.9997	0.8533	1.0000	0.4598	0.5929	0.6363
P310	0.4609	0.3365	0.4598	1.0000	0.7979	0.7647
P311	0.5938	0.4417	0.5929	0.7979	1.0000	0.9480
P312	0.6380	0.5355	0.6363	0.7647	0.9480	1.0000
P313	0.7737	0.7921	0.7746	0.5774	0.6599	0.7263
P314	0.6102	0.4642	0.6093	0.7689	0.9939	0.9710
P316	0.6527	0.5590	0.6508	0.8482	0.9482	0.9813
P317	0.6974	0.6575	0.7040	0.3677	0.4471	0.5059
P319	0.7153	0.6780	0.7123	0.6272	0.8922	0.9678
P320	0.7684	0.8121	0.7630	0.6643	0.8138	0.8790
P321	0.8300	0.7866	0.8298	0.5602	0.7552	0.8674
P322	0.6520	0.5443	0.6504	0.7756	0.9635	0.9977
P323	0.8889	0.6485	0.8896	0.3579	0.3724	0.3457
P324	0.7316	0.6935	0.7287	0.7251	0.8981	0.9675
P325	0.6220	0.4878	0.6207	0.8159	0.9936	0.9726
P326	0.5892	0.4355	0.5883	0.8032	0.9998	0.9425
P327	0.6348	0.5060	0.6335	0.8693	0.9759	0.9666

(c) Identification matrix (RRT only)

CODE	P306	P307	P308	P310	P311	P312
P306	1	0	0	0	0	0
P307	0	1	0	0	0	0
P308	0	0	1	0	0	0
P310	0	0	0	1	0	1
P311	0	0	0	0	1	0
P312	0	0	0	1	0	1
P313	0	0	0	0	0	0
P314	0	0	1	0	0	0
P316	0	0	0	0	0	0
P317	0	0	0	1	0	1
P319	0	0	0	0	0	0
P320	0	0	0	0	0	0
P321	0	0	1	0	0	0
P322	0	0	0	0	0	0
P323	0	0	0	0	0	0
P324	0	1	0	0	0	0
P325	0	0	0	1	0	0
P326	0	0	0	0	1	0
P327	0	0	1	0	0	0

(b) Identification matrix (spectrum only)

CODE	P306	P307	P308	P310	P311	P312
P306	1	0	1	0	0	0
P307	0	1	0	0	0	0
P308	1	0	1	0	0	0
P310	0	0	0	1	0	0
P311	0	0	0	0	1	0
P312	0	0	0	0	0	1
P313	0	0	0	0	0	0
P314	0	0	0	0	0	0
P316	0	0	0	0	0	0
P317	0	0	0	0	0	0
P319	0	0	0	0	0	0
P320	0	0	0	0	0	0
P321	0	0	0	0	0	0
P322	0	0	0	0	0	0
P323	0	0	0	0	0	0
P324	0	0	0	0	0	0
P325	0	0	0	0	0	0
P326	0	0	0	0	1	0
P327	0	0	0	0	0	0

(d) Combined identification matrix

CODE	P306	P307	P308	P310	P311	P312
P306	1	0	0	0	0	0
P307	0	1	0	0	0	0
P308	0	0	1	0	0	0
P310	0	0	0	1	0	0
P311	0	0	0	0	1	0
P312	0	0	0	0	0	1
P313	0	0	0	0	0	0
P314	0	0	0	0	0	0
P316	0	0	0	0	0	0
P317	0	0	0	0	0	0
P319	0	0	0	0	0	0
P320	0	0	0	0	0	0
P321	0	0	0	0	0	0
P322	0	0	0	0	0	0
P323	0	0	0	0	0	0
P324	0	0	0	0	0	0
P325	0	0	0	0	0	0
P326	0	0	0	0	1	0
P327	0	0	0	0	0	0

Figure 4. Excerpt from the matrices used for the evaluation of the selectivity of HPLC–DAD for substance identification with a database of 2682 compounds. Fields with indistinguishable substance pairs were printed inversely. Similarity matrix (spectrum only) (a), SI of all possible substance pairs were calculated by Equation 4; identification matrix (spectrum only) (b), substance pairs with SI > 0.9990 were assigned the value 1 (identical), and substance pairs with SI ≤ 0.9990 received the value 0 (different); identification matrix (RRT only) (c), substance pairs whose RRTs fell into their mutual RRT error windows (see text) were set to 1 (identical), otherwise they were assigned 0 (different); and combined identification matrix (d), obtained by multiplication of the spectrum and RRT identification matrices. In this example, only P311 and P326 (phenelzine and phenethylamine) remain indistinguishable, even by the combination of spectrum and RRT.

tamine, MPPH, and a third standard compound (caffeine) ranged from 0.2 to 0.9% (n = 15). It was concluded from these measurements that the variation coefficients found were applicable over the entire RT range. By setting the RT error window to ± $2s_{RT}$ (corresponding to a 95% confidence interval), an individual error window for RRT was calculated for every compound from the long-term VK_{RT} (3 and 5%, respectively) by application of the rules of error propagation to equation 1.

Two compounds were then rated identical if the RRT of at least one of them would fall into the RRT error window of the other; otherwise they were considered distinguishable. Using this criterion, all compounds measured in mobile phase A were compared with each other in an RRT identification matrix (Figure 4c). In this matrix, again, compound pairs regarded as identical received a 1, and distinguishable pairs were set to 0. The same procedure was performed for all compounds measured in mobile phase B.

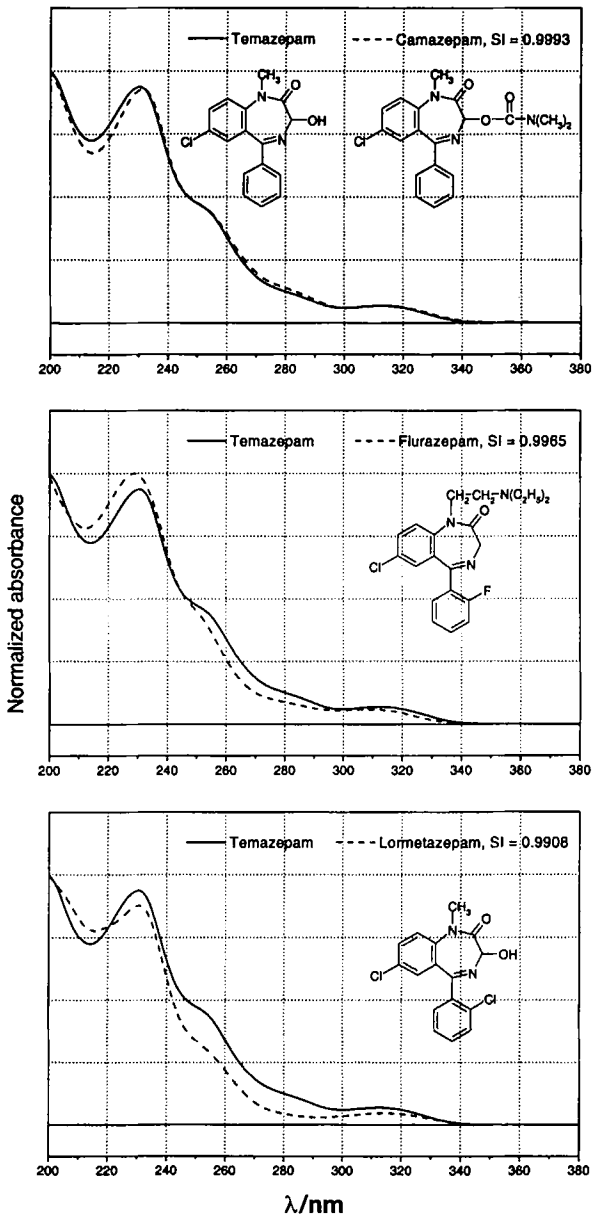

Figure 5. UV spectra with a different degree of similarity to that of temazepam. With a threshold SI = 0.9990, the spectrum of camazepam is assessed to be identical, those of flurazepam and lormetazepam are considered distinguishable.

From the evaluation of these matrices, the RRT alone turned out to be a rather uncharacteristic parameter for substance identification (under the separation conditions used here). In mobile phase A (1993 compounds), no compound was unambiguously identified, and on average each compound had 124 others from which it could not be distinguished by the RRT alone (MLL = 125). The situation in mobile phase B (689 compounds) was only slightly improved (MLL = 94). This unfavorable result was at least partly caused by the isocratic HPLC conditions, which were chosen for practical reasons (recycling of the mobile phase, no conditioning of the column between runs, and higher reproducibility) and lead to an uneven distribution of RRT with an accumulation of substances in the early minutes of the chromatogram.

Nevertheless, in combination with the UV spectrum, the RRT should significantly decrease the number of false-positive search results. Therefore, for both mobile phases the spectrum and RRT identification matrices were combined by multiplication of the contents of the corresponding fields (Figure 4d). In this combined identification matrix, two substances were considered identical only if SI > 0.9990 and their RRTs did not fall into their mutual RRT error windows. The results for the 1993 compounds measured in mobile phase A are shown in Figure 7. It can be seen that the results achieved on the basis of the UV spectrum alone are very similar to those shown in Figure 6, which lead to the conclusion that a representative substance collective was investigated. By inclusion of the RRT, the number of unambiguously identified substances increased from 60.6% to 84.2%, the number of substances with two hits (positive search results) decreased from 14.2% to 10.6%, and of those with three positive hits decreased from 7.0% to 2.6%. There were no compounds with more than six hits. The MLL was decreased to

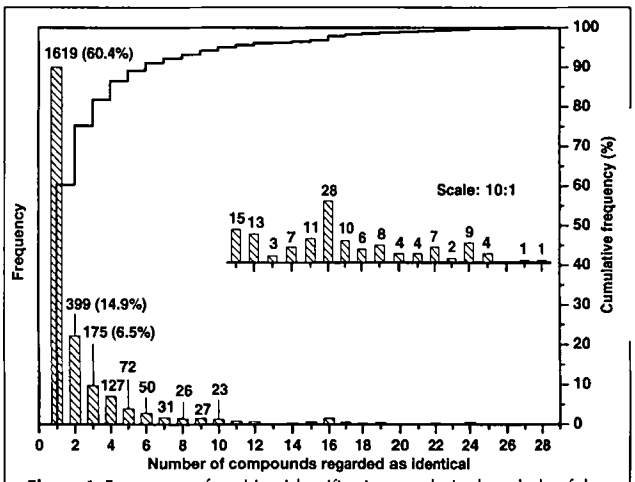

Figure 6. Frequency of positive identification results in the whole of the spectra library (2682 compounds) on the basis of the UV spectrum alone. The comparison of every substance with itself is included. Thus, 1 compound regarded as identical indicates that the corresponding substances can be successfully distinguished from every other compound in the spectra library, 2 means there is one other compound indistinguishable, and so on.

1.25, which signified that, on average, every substance was indistinguishable only from 0.25 others (put differently, out of 1993 substances only every fourth could have been mistaken for one other compound).

Example of indistinguishable compounds

An example of a compound found to be indistinguishable from five other substances, even by the combination of UV spectrum and RRT (LL = 6), is shown in Figure 8. All six compounds are esters of benzilic acid with a β- or γ-aminoalcohol or a β- or γ-ammoniumalcohol (please note that at the pH of the mobile phase the nitrogen in benactyzine and tropine benzilate will also be protonated, which explained the chromatographic behavior similar to that of the other compounds).

In this case, the ambiguity of identification arises from high structural similarity extending far into the residual part of the molecule, remote from the actual chromophores (2 toluene chromophores). This need not always be a disadvantage; when working with a large spectra library, an essential part of the structure of an unknown substance may be deducible from its UV spectrum. On the other hand, it should be pointed out, again, that there are numerous cases where small structural changes at some distance from the chromophore lead to discernible spectra (Figure 3).

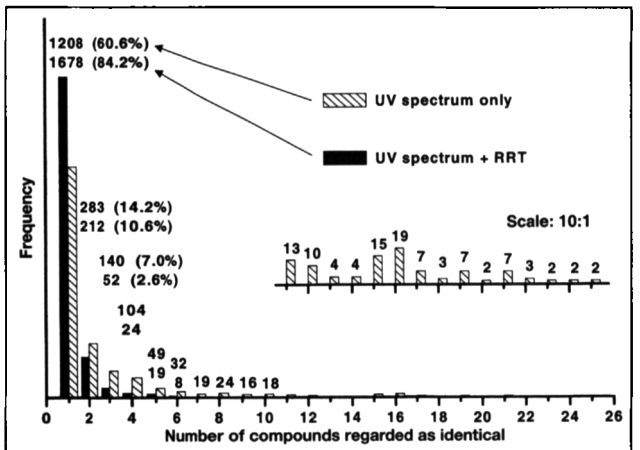

Figure 7. Frequency of positive search results for the 1993 compounds measured in mobile phase A using UV spectra alone, as well as in combination with RRT (for explanation cf. Figure 6).

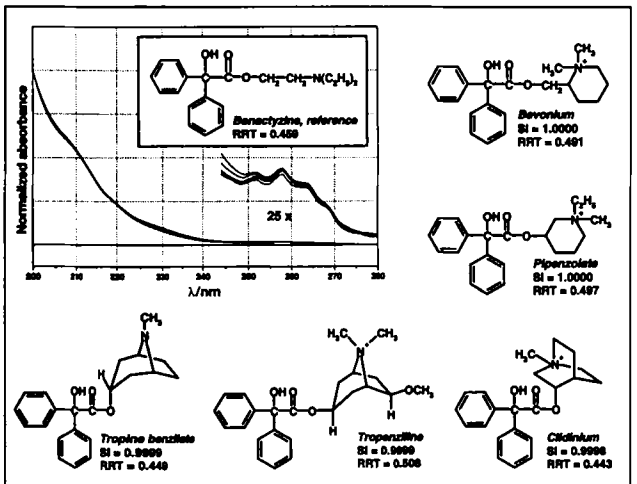

Figure 8. Example of five compounds that cannot be distinguished from the reference compound benactyzine by the combination of UV spectrum and RRT.

It also should not be overlooked that multiple identification results constitute—in numbers—only a minor problem; only 5% of the substances can be confused with more than one other compound (Figure 7). Cases where there are exactly two candidates for identification (10.6%) should then be resolvable by an independent confirmation method (or sometimes even by changing the HPLC separation system). Also, a repeated analysis of the sample spiked with a sufficient amount of the compound in question should reduce the number of identification candidates significantly because the intraday variance of RT (below 1%, see previous) is much lower than on the two-year basis taken for this work.

Comparison with literature data

In Table I, DP and MLL found in the present work are compared with the results of other authors. Unfortunately, the number of substances evaluated for DP or MLL in the literature was approximately one order of magnitude smaller than in the HPLC–DAD library described in this paper.

Taking this into account, HPLC–DAD was confirmed to be a highly selective method for substance identification with results similar to those found for GC–MS by Demme et al. (30) or de Zeeuw et al. (31) for much smaller substance groups. The rather poor results for HPLC–DAD in combination with RT or a retention index scale published by Maier et al. (19) and de Zeeuw et al. (31) could be explained by the fact that in these papers only UV maxima seem to have been used for substance identification. To the knowledge of the authors, the present work is the first to use the whole UV spectrum as a basis for the evaluation of the

Table I. DP and MLL Results of the Present Investigation and Comparison with Literature Data

Method	Number of substances	DP	MLL	Remarks	Authors, reference
HPLC–DAD, UV spectra	2682	0.9994	2.701	Complete library	This paper
HPLC–DAD, UV spectra + RRT	1993	0.9999	1.253	Mobile phase A	This paper
HPLC–DAD, UV spectra + RRT	689	0.9998	1.118	Mobile phase B	This paper
GC–MS (ion trap)	244	0.9997	1.073		Demme et al. (30)
HPLC–DAD, UV maxima + retention index	372	0.9930	5.390		Maier et al. (19)
HPLC–DAD, UV maxima + RT	99	0.9743	3.520		de Zeeuw et al. (31)
GC–MS	99	0.9990	1.100		de Zeeuw et al. (31)

selectivity of HPLC–DAD. It is also the first to use a substance collective as large as 2682 compounds for the evaluation of method selectivity.

Application to real cases

The results presented previously are based on the analysis of pure substance solutions. However, the spectra measured in real cases could be altered by noise at low concentrations, as well as by overlapping with insufficiently separated peaks or with peaks of matrix constituents. Furthermore, the HPLC–DAD measurement is only one step of the STA, and it is another indispensable prerequisite that the substance can be obtained in a measurable state by an adequate sample preparation.

The authors have been using HPLC–DAD in combination with the steadily growing spectra library as the main method for general unknown analysis in about 350 forensic and 500 clinical cases per year since 1992. Immunoassays for drugs of abuse and headspace GC for alcohol and volatile substances were also regularly used, and GC–MS was available for supplementary and confirmatory analyses. The sample preparation routinely used for HPLC–DAD and data about the extractability and detection limits under these conditions for drugs frequently occurring in intoxications, as well as examples of application were previously published in detail (3,25). Briefly, in a routine case, from each 0.5 mL whole blood or serum, an acidic and basic extract (pH 2 and 9, respectively) in dichloromethane were prepared, and from a further aliquot of the whole blood or serum sample the protein was precipitated with acetonitrile (1:1). The basic extract and the supernatant of the protein precipitation were analyzed in mobile phase A and the acidic extract in mobile phase B parallel on three HPLC–DAD instruments. In special cases, other liquid–liquid or solid-phase extraction conditions were applied or the HPLC conditions were changed.

With respect to the quality of the spectra obtained from the extracts, there was no difference to those from the pure substances down to sample concentrations of 0.1 µg/mL, and the problem of interference by badly resolved peaks or matrix constituents was, in most cases, not severe. As a more complicated example, in Figure 9 the chromatogram of the basic CH_2Cl_2 extract of an emergency case blood sample is shown. The 45-year-old man had taken in a drug mixture, presumably 24 h before admission to the hospital. The immunoassays of urine for the usual drugs of abuse were only positive for benzodiazepines; blood alcohol was 0.0 mg/g. Of the approximately 20 peaks of the chromatogram, 9 were identified by library search (UV spectrum + RRT) as first hits with SI < 0.999 for 6 peaks and between 0.991 and 0.998 for the other 3 peaks. Deviations of the actual RRT from the library data were found in the range of 0–1.8%. A multidrug intoxication by diazepam, flurazepam, nitrazepam, zolpidem, and zopiclone resulted. Approximate concentrations were calculated from the peak areas in the chromatogram and independently measured specific peak areas and extraction yields of the drugs or metabolites. Flurazepam itself (RT marked by an arrow in Figure 9, half-life 1–3 h) was not unambiguously detected after the long survival time, but four of its active metabolites were clearly identified. In order to illustrate different degrees of spectral identity at different concentrations, the spectra of zolpidem (peak 4, 1.38 mg/mL) and 2-hydroxyethylflurazepam (peak 7, ≈ 0.05 µg/mL) were overlaid with the corresponding library spectra in the right part of Figure 9. It is seen that there is a complete agreement for zolpidem, and that despite the noise at this low concentration, the spectrum of the flurazepam metabolite still agrees very well with the library spectrum. For the remaining peaks, no satisfactory search results were obtained. Two of these peaks (RT = 5.3 and 6.0 min) had typical benzodiazepine spectra and could originate from further flurazepam metabolites. For the other peaks no suitable spectra could be extracted because of overlapping or the absorbance being too small. In the present case, this was not pursued any further because these peaks represented only small concentrations and the symptoms of the patient were sufficiently explained by the results already obtained. In other cases the extraction of a higher sample amount (2 instead of 0.5 mL whole blood, if available), the use of more selective extraction conditions (e.g., 1-chlorobutane instead of dichloromethane for benzodiazepines), and the improvement of the HPLC separation by switching to a weaker mobile phase (with a greater buffer content) would have been applied in order to improve peak identification.

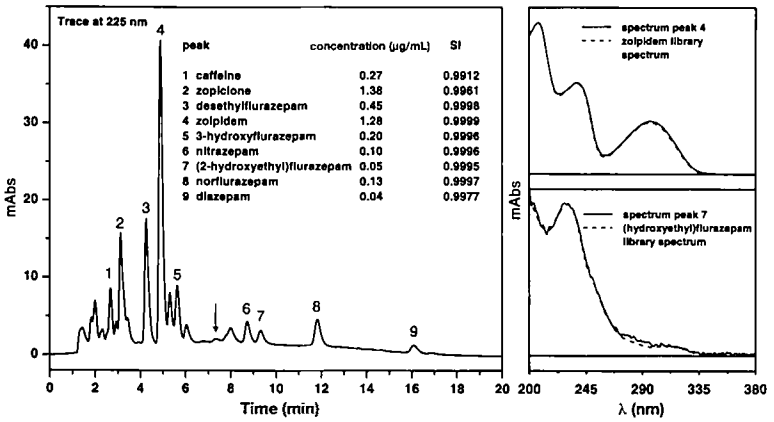

Figure 9. HPLC–DAD analysis of a multidrug intoxication. The sample was collected approximately 24 h after intake. The chromatogram was obtained from the CH_2Cl_2 extract of 0.5 mL whole blood at pH 9. By library search, zopiclone, zolpidem, nitrazepam, diazepam, and four flurazepam metabolites were identified. Flurazepam (RT marked by the arrow, $t_{1/2}$ = 1–3 h) could not be identified unambiguously from the very small peak (< 0.01 µg/mL).

From the acidic CH_2Cl_2 extract, no additional information was gained. Also, by investigation of the supernatant of the protein precipitation, many hydrophilic substances with low extraction yields in CH_2Cl_2 (e.g., paracetamol or salicylic acid) could be excluded.

Generally, a complicated analytical task such as the toxicological general unknown analysis cannot always be solved by schematic application of a fixed procedure, but needs the availability of alternative methods and a flexible choice or change of the conditions for sample preparation, chromatographic separation, and spectrometric detection by an experienced analyst.

Conclusions

In this evaluation of an extensive library of UV spectra and RRTs, the role of HPLC–DAD as one of the methods best suited for substance identification in the frame of systematic toxicological analysis was confirmed. More than 90% of the toxicologically relevant substances under examination showed a significant UV absorption in the investigated wavelength range. The great structural diversity of a substance collective that contained compounds from all important fields of toxicological relevance in combination with the excellent performance of modern DADs (in terms of resolution, accuracy, repeatability, and reproducibility) were found to provide a high degree of method selectivity. Although literature data for a substance pool of similar magnitude was not available, DP and MLL values found for HPLC–DAD were comparable to those of competing GC–MS methods.

The results of this investigation suggest that HPLC–DAD in combination with a suitable spectra library will continue to make indispensable contributions to the field of toxicological screening and systematic toxicological analysis. The advantages of an HPLC method such as a comparatively simple sample preparation (e.g., without the need for a derivatization required), cost-effectiveness, and ruggedness are combined with a highly selective detector. The online coupling of a DAD with a modern LC–MS system, which enhances selectivity by the use of two independent detection modes, is seen as a particularly promising prospect for the future.

Acknowledgments

The authors thank Mrs. S. Menzel, Mrs. K. Malzacher, and Mr. M. Witteweg for their thorough assistance in the measurement of the HPLC–DAD library. Furthermore, they thank Dr. B.-T. Erxleben (Shimadzu Europe GmbH, Duisburg, Germany), Mrs. A. Herrndobler (Bio-Tek Instruments, Neufahrn, Germany), and Mrs. A. Kohn (Agilent Technologies, Waldbronn, Germany) for helpful discussions and technical support. They are also grateful to Dr. M. Rothe (Lung and Thorax Diagnostics GmbH, Berlin, Germany) for measuring all UV spectra with the DAD HP 1090 and Dr. W. Wlodarczyk (Department of Radiotherapy, University Hospital Charité, Berlin, Germany) for providing a MatLab workplace.

References

1. R.A. de Zeeuw. Drug screening in biological fluids. The need for a systematic approach. *J. Chromatogr. B: Biomed. Sci. Appl.* **689:** 71–79 (1997).
2. A. Polettini. Systematic toxicological analysis of drugs and poisons in biosamples by hyphenated chromatographic and spectroscopic techniques. *J. Chromatogr. B: Biomed. Sci. Appl.* **733:** 4763 (1999).
3. F. Pragst, H.H. Maurer, J. Hallbach, U. Staerk, W.R. Külpmann, F. Degel, and H.J. Gibitz. Suchverfahren (General unknown). In *Klinisch-toxikologische Analytik, Verfahren und Interpretation,* W.R. Külpmann, Ed. Wiley-VCH, Weinheim, Germany, 2002, pp 49–95.
4. K. Pfleger, H.H. Maurer, and A. Weber. *Mass Spectral and GC Data of Drugs, Poisons, Pesticides, Pollutants and Their Metabolites,* 2nd ed., 1992–2000. Wiley-VCH, Weinheim, Germany, 2000.
5. B. Aebi and W. Bernhard. Advances in the use of mass spectral libraries for forensic toxicology. *J. Anal. Toxicol.* **26:** 149–156 (2002).
6. A. Polettini, A. Groppi, C. Vignali, and M. Montagna. Fully-automated systematic toxicological analysis of drugs, poisons, and metabolites in whole blood, urine, and plasma by gas chromatography-full scan mass spectrometry. *J. Chromatogr. B: Biomed. Sci. Appl.* **713:** 265–279 (1998).
7. P. Marquet. Is LC-MS suitable for a comprehensive screening of drugs and poisons in clinical toxicology? *Ther. Drug Monit.* **24:** 125–133 (2002).
8. W. Weinmann, A. Wiedemann, B. Eppinger, M. Renz, and M. Svoboda. Screening for drugs in serum by electrospray ionization/collision-induced dissociation and library searching. *J. Am. Soc. Mass Spec.* **10:** 1028–1037 (1999).
9. H.H. Maurer. Liquid chromatography-mass spectrometry in forensic and clinical toxicology. *J. Chromatogr. B: Biomed. Sci. Appl.* **713:** 3–25 (1998).
10. M. Rittner, F. Pragst, W.R. Bork, and J. Neumann. Screening method for seventy psychoactive drugs or drug metabolites in serum based on high-performance liquid chromatography-electrospray ionization mass spectrometry. *J. Anal. Toxicol.* **25:** 115–124 (2001).
11. M. Bogusz and M. Wu. Standardized HPLC/DAD system, based on retention indices and spectral library, applicable for systematic toxicological screening. *J. Anal. Toxicol.* **15:** 188–197 (1991).
12. M. Bogusz and M. Erkens. Reversed-phase high-performance liquid chromatographic database of retention indices and UV spectra of toxicologically relevant substances and its interlaboratory use. *J. Chromatogr. A* **674:** 97–126 (1994).
13. S.P. Elliott and K.A. Hale. Applications of an HPLC-DAD drug-screening system based on retention indices and UV spectra. *J. Anal. Toxicol.* **22:** 279–289 (1998).
14. Y. Gaillard and G. Pépin. Use of high-performance liquid chromatography with photodiode-array UV detection for the creation of a 600-compound library. Application to forensic toxicology. *J. Chromatogr. A* **763:** 149–163 (1997).
15. E.M. Koves. Use of high-performance liquid chromatography-diode array detection in forensic toxicology. *J. Chromatogr. A* **692:** 103–119 (1995).
16. C.-K. Lai, T. Lee, K.-M. Au, and A.Y.W. Chan. Uniform solid-phase extraction procedure for toxicological drug screening in serum and urine by HPLC with photodiode-array detection. *Clin. Chem.* **43:** 312–325 (1997).
17. W.E. Lambert, E. Meyer, Y. Xue-Ping, and A.P. De Leenheer. Screening, identification, and quantitation of benzodiazepines in postmortem samples by HPLC with photodiode array detection. *J. Anal. Toxicol.* **19:** 35–40 (1995).
18. B.K. Logan, D.T. Stafford, I.R. Tebbett, and C.M. Moore. Rapid screening for 100 basic drugs and metabolites in urine using cation exchange solid-phase extraction and high-performance liquid chromatography with diode array detection. *J. Anal. Toxicol.* **14:** 154–159 (1990).
19. R.D. Maier and M. Bogusz. Identification power of a standardized HPLC-DAD system for systematic toxicological analysis. *J. Anal. Toxicol.* **19:** 79–83 (1995).
20. A. Tracqui, P. Kintz, and P. Mangin. Systematic toxicological analysis using HPLC/DAD. *J. Forensic Sci.* **40:** 254–262 (1995).
21. A. Turcant, A. Premel-Cabic, A. Cailleux, and P. Allain. Toxicological screening of drugs by microbore high-performance liquid chromatography with photodiode-array detection and ultraviolet spectral library searches. *Clin. Chem.* **37:** 1210–1215 (1991).
22. D.W. Hill and A.J. Kind. Reversed-phase solvent-gradient HPLC retention indexes of drugs. *J. Anal. Toxicol.* **18:** 233–242 (1994).
23. F. Pragst, M. Herzler, S. Herre, B.-T. Erxleben, and M. Rothe. *UV Spectra of Toxic Compounds.* Verlag Dr. Dieter Helm, Heppenheim, Germany, 2001.
24. M. Herzler, F. Pragst, S. Herre, and M. Rothe. Selectivity of pho-

todiode array UV spectra for substance identification in systematic toxicological analysis. In *Proceedings of the 37th TIAFT Triennial Meeting*, T. Lech, Ed. Institute of Forensic Research Publishers, Kraków, Poland, 2000, pp 122–129.
25. F. Pragst, M. Herzler, and S. Herre. Use of high performance liquid chromatography with photodiode array detection (HPLC-DAD) in clinical toxicology. *Klin. Biochem. Metab.* **1:**13–16 (2000).
26. A.C. Moffat, K.W. Smalldon, and C. Brown. Optimum use of paper, thin-layer and gas-liquid chromatography for the identification of basic drugs I. Determination of effectiveness for a series of chromatographic systems. *J. Chromatogr.* **90:** 1–7 (1974).
27. P.G. Schepers, J.P. Franke, and R.A. de Zeeuw. System evaluation and substance identification in systematic toxicological analysis by the mean list length approach. *J. Anal. Toxicol.* **7:** 272–278 (1983).
28. M. Herzler. *Untersuchungen zur Selektivität der HPLC-DAD in der Systematischen Toxikologischen Analyse.* Dissertation for the Post-graduate Study Course in Toxicology. University of Leipzig, Leipzig, Germany, 2001.
29. L. Huber and S.A. George, Eds. *Diode Array Detection in HPLC.* Marcel Dekker, Inc., New York, NY, 1993.
30. U. Demme, B. Ahrens, A. Klein, and R. Werner. Discriminating power nearly one—ion-trap GC-MS as an ideal tool for STA. In: *Proceedings of the 38th Meeting of TIAFT*, I. Rasanen, Ed. Tallina Raamatutrükikoja OÜ, Helsinki, Finland, 2001, 19–26.
31. R.A. de Zeeuw, J. Hartstra, and J.P. Franke. Potential and pitfalls of chromatographic techniques and detection modes in substance identification for systematic toxicological analysis. *J. Chromatogr. A* **674:** 3–13 (1994).

Manuscript received March 18, 2002;
revision received August 7, 2002.