Chromatographic Analysis of Compounds of Pharmaceutical Interest

Ming-Shang Kuo, F. Barclay Shilliday, and **David A. Yurek**
Discovery Technologies, Pharmacia, Kalamazoo, MI 49007

Abstract

A set of 17,600 samples belonging to our compound collection is selectively examined by liquid chromatography with UV, evaporative light scattering, and mass spectrometric detection methods. At least 70% of this set consists of pure samples with the expected structures. Subsequent studies by flow injection mass spectrometry show that this value is a conservative estimate and that the actual percentage of pure and correct compounds is close to 80%. Because this is the first time that sample quality information becomes available on such a large scale for the compound collection, it offers an opportunity to perform chemi-informatic studies for which structural integrity is essential. Results of these studies can be used to improve the selection of compounds for screening and evaluate the quality of compounds from particular sources.

Introduction

Since the early 1970s, numerous high-performance separation methods have been developed allowing for the routine examination of the purity of chemicals with pharmaceutical properties. Especially noteworthy is the rapid development of highly pure, silica-based reverse-phase high-performance liquid chromatographic (HPLC) resins with smaller pore size, smaller particle diameters, and a more complete covering of uncapped silica sites. These resins allow enhanced resolution and faster throughput with minimum assay development time. Most recently, several companies have introduced columns packed with resins containing a polar moiety in the derivatization of the silica, allowing for the use of 100% aqueous mobile phases and the retention of more polar compounds than traditional reversed-phase columns (1). The past decade also witnessed the explosive development of many online detection methods. One of the most important among them is the atmosphere pressure ionization mass spectrometer (MS) interface, a technique ideally suited for the introduction of polar, thermally labile compounds into an MS (2–5). The sensitivity and stability of performance were improved over other kinds of MS interfaces for highly aqueous eluent at high flow rates (> 200 µL/min). Lately, the synergistic use of MS detection with other quantitative detectors (UV, light scattering, and chemiluminescent nitrogen detector) has provided quantitative (structure identification, sample purity, and concentration) analyses within one run (6,7). Lastly, the rapid development of information technology has made the data acquisition, coordination, processing, reporting, and archiving for such a large set of samples technologically feasible at a reasonable cost. In this report, we have applied these recent advances to the study of sample quality in compound collections used for high-throughput screening.

A critical component of any high-throughput screening organization is to maintain and expand a group of chemically diverse high-quality organic compounds. A significant challenge for curators is to assess and maintain the quality of these libraries, because the uncertainty surrounding the structural integrity and purity of the compounds found as hits from biological screens has caused great concern. If one were to analyze the entire library (normally numbering in the range of hundreds of thousands of compounds), it would take years to complete, even with current-day technologies. However, with the aid of computer software and statistical design, one can select a group of compounds representing the entire library and have them completely analyzed within months. We therefore implemented an analytical procedure for a group of 17,600 compounds assembled by their structure dissimilarities (8). In order to define quality for the purposes of this study, the most important criteria was the presence of the indicated structure and its relative purity percentage in each sample. Thus, all samples in question were analyzed for these criteria. These parameters were measured by HPLC with detection by UV absorbance, evaporative light scattering (ELSD) (9–13), and MS. The mass spectra were used to confirm the expected molecular weight (MW) and identify the peak containing the target compound. The two analog channels were used to compute the relative purity of this peak.

Experimental

Liquid chromatographic–UV–ELSD–MS methods

Arrangement of hardware employed

A chart showing the equipment used for the liquid chromatographic (LC)–UV–ELSD–MS method and the flow of samples is shown in Figure 1. Contact closures, used to activate valve switches for the purging of the void volume and column selec-

tion, were programmed in the HPLC pump method via the MS software. Although this software controlled most of the components, the autosampler was controlled via a separate computer external to the MS system.

Sample preparation for HPLC

Samples were received as 5-µL, 2mM DMSO solutions in 88 wells of a 96-well microtiter plate. The wells in column 12 of each plate were empty. Four 5-µL samples of PNU-100766 (Linezolid, 2mM DMSO) were added to each plate at positions B12, D12, F12, and G12, respectively, as external positive controls. The remainder of the empty wells served as negative control samples. Samples were stored frozen at –20°C. Immediately prior to analysis, the samples were thawed and diluted with 45 µL of 50% aqueous acetonitrile and 0.05% formic acid. Samples were diluted and injected using a Gilson (Middleton, WI) 235 autosampler controlled by Unipoint Version 1.8 software (Gilson).

HPLC conditions

Gradient analyses were performed on an HP 1100 HPLC system controlled by the software of the MS employing two columns (Aquasep C-8 5µ, 5 cm × 2.0 mm, ES Industries, West Berlin, NJ). Solvent 1 consisted of water and 0.05% formic acid by volume, and solvent 2 was acetonitrile with 0.05% formic acid by volume. The injection volume was 10 µL. Column-switching valves were used to re-equilibrate one column while the other was performing a separation. Re-equilibration solvent was delivered by a Gilson 305 HPLC pump in a stand-alone program mode set to pump solvent 1 at 1 mL/min for 4.2 min. The solvent program is presented in Table I.

Column-switching system

A Model CS3010 Column Swapper Valve system from Chiralizer Services (Plainsboro, NJ) was used to select the analysis column. The switch was controlled by timed contact closures from the HP 1100 HPLC pump.

Flow splitter

The effluent from the column was split downstream from the UV detector by an Acurate Post-Column Splitter from LC Packings (San Francisco, CA). This device splits the flow at a constant ratio regardless of the composition of the solvent.

MS

A PE Sciex 150EX single quadrupole MS equipped with an APCI Heated Nebulizer interface set to 400°C was used to identify the components of the HPLC effluent. The MS was programmed to scan from 130 to 800 mass units in both the positive and negative ionization modes in alternating fashion. In the positive MS mode, the expected ions were molecular ions with adducts of a proton (M+1), sodium ion (M+23), or ammonium ion (M+18). In the negative mode, the expected ions were the molecular ions with the loss of a proton (M–1) and those as a formate adduct minus a proton (M+46–1).

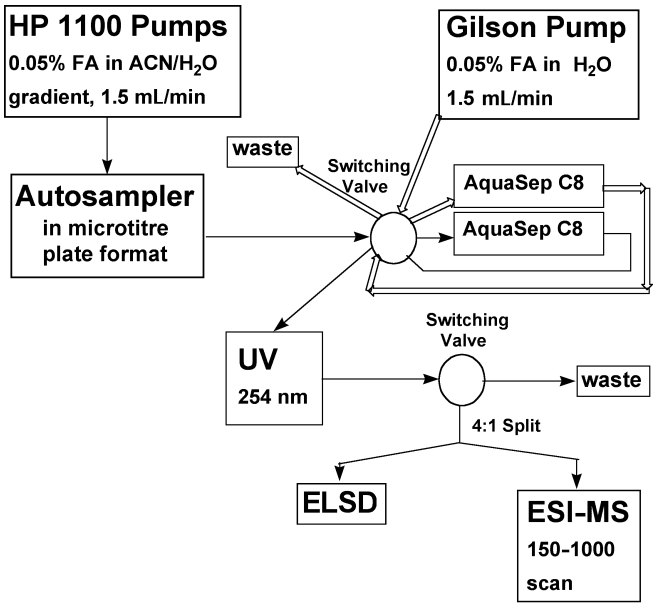

Figure 1. Diagram of an LC–UV–ELSD–MS analytical system.

Table I. Solvent Program

Step	Time (min)	Flow (µL/min)	Solvent 1	Solvent 2	External events
1	0.00	1000	100%	0%	flow to waste
2	1.00	1000	100%	0%	flow to MS–ELSD
3	4.00	1000	0%	100%	flow to waste, column switched

UV detector

An HP 1100 Series Variable wavelength detector was used with the wavelength set at 254 nm.

ELSD

A Sedex 55C ELSD was used with the following settings: the temperature of the drift tube was 70°C, the gain was 11, and the N_2 pressure was 4 bar.

Data analysis

LC–MS data with analog channels from UV and ELSD were acquired using the PE-Sciex software package MassChrom Version 1.1 on the Macintosh platform. Each plate generated 88 sample files containing the chromatographic data, plus the quality control samples. Data files, along with the well positions and expected MWs of the samples, were loaded into the PE-Sciex LC Combiview Version 1.0 software for analysis. LC Combiview searched the chromatograms for compounds with the expected MW and provided purity calculations based on the UV and ELSD analog channels. The individual information in each well was initially processed and then assimilated together in an output format resembling a microtiter plate that was used to review the results. In the format, a green rectangle represented a well in which the expected MWs were detected, and a red rectangle indicated that the expected ion was not observed

in the chromatogram. The results for each plate were exported to an Excel table (as partially shown in Table II). On this table, the purity values were reviewed and edited, if necessary.

Flow injection analysis methods

Sample preparation for flow injection analysis

Sample plates were dried on a Speed Vac Plus for 8 h to remove the DMSO. Immediately prior to analysis, samples in each column of each microtiter plate were reconstituted in 50% aqueous methanol containing 0.05% by volume of formic acid using a Gilson 215 Liquid Handler with an 8-probe attachment. The samples were simultaneously loaded into the sample loops of eight injection valves on the 889 Multiple Injector System (MIS). This system allowed the reconstitution and loading of eight samples simultaneously to eight individually controlled injection valves corresponding to the eight wells in each column of a microtiter plate.

8-Probe injector

The 8-probe Gilson 215, 889 MIS, and Gilson 306 HPLC pump were controlled by Unipoint software. In order to place each of the eight sample injection valves sequentially into the flow to the detector, the pump, autosampler, and MS were interfaced with an electronically activated 10-port switching valve (Model #EMTMA-CE, Valco Instruments, Houston, TX). This valve contained a common input port, ten pairs of loading ports, and a common exit port, enabling the routing of solvent flow from the pump sequentially through the eight valves to one detector. Thus, pump flow (0.4 mL/min of 50% aqueous methanol containing 0.05% by volume of formic acid) entered the common input port and was routed to a selected injector. Flow from the selected injector proceeded back through the 10-port valve to the common exit port connected to the MS. The valve was then switched to route the flow to the next injector on the 889 MIS. The same APCI Heated Nebulizer interface and MS conditions were used for flow injection analysis (FIA) as well as for the HPLC–UV–ELSD–MS analysis.

Data analysis

Data were stored as electronic files containing data from one pass of the 8-probe, or eight injections. There were twelve such files for each plate of samples. The files were split into eight separate packets and loaded in the data program Combiview FIA Version 1.0 provided by PE-Sciex (Toronto, Canada). The splitting and loading of the files were accomplished by means of a modified Applescript macro originally provided by PE-Sciex. The Combiview FIA program read the mass spectra from each sample and computed the relative intensity of the expected molecular ion. The output was stored in text files, each of which contained data from an entire plate. The data for the entire 17,600-member library was contained in 200 files, which were loaded into Microsoft Excel and concatenated into a single file. This file was then merged with the HPLC–UV–ELSD–MS data in order to create a master file containing all the purity and identification results from both protocols.

Compound selection procedures and statistics

Compounds were selected from the collection based on their dissimilarity using a molecular modeling program developed in-house. A set of 17,600 compounds was selected from a total collection containing > 250,000 compounds. This set was considered to be a randomized selection because the collection covered a widely diverse chemistry space (8).

Results

Analytical method development

Two complementary methods were employed to analyze the samples. The LC–UV–ELSD–MS method was used initially to obtain primary identity and purity information. Then, the entire sample set was subjected to FIA in order to distinguish compounds that may have been overlooked in the initial analysis because of their chromatographic properties.

HPLC–UV–ELSD–MS

The analysis system is shown in Figure 1. The conditions for analysis were developed with the goal of designing a high-throughput system for determining the purity of the samples in the compound collection. Conditions were therefore chosen that would serve to detect a wide range of compounds. The chromatography was conducted on a reverse-phase column (Aquasep C-8 5µ, 2.1 × 50 mm). This column contained a polar grouping in the derivatization of the silica particle and enabled a greater retention of polar compounds and no phase collapse at the initial 100% aqueous mobile phase conditions. Because of the presence of DMSO in the samples, a switching valve was used to divert the void volume from the column to waste in order to avoid the interference in the MS. After the void volume eluted, the mobile phase went to 100% acetonitrile over 4 min in order to elute nonpolar compounds. The mobile phase contained a constant concentration of 0.05% formic acid as a pH buffer to assist the ionization of compounds for facile detection of the analytes by MS. In order to increase throughput, two identical columns were used in conjunction with a column-switching valve so that one column will be re-equilibrated while an analysis is performed on the other. The total time between injections was 5 min.

The optimal detector for a universal system would have a response proportional to a compound's concentration over a wide range. The response would also be equal for all chemical types. Such a detector does not yet exist for HPLC. The alternative is to use several detectors that rely on different properties of the analyte. In these studies, the entire LC eluent passed through a UV detector whose response depended on the electronic properties of each analyte. Because the UV detector was not destructive, the eluent was then split 25:75 to the MS and ELSD, respectively. The ELSD worked by converting the effluent to a mist, allowing the solvent to evaporate, and passing the nonvolatile components through a light beam. The response for a particular compound depended on its volatility and the light scattering caused by the aggregated particles of the compound (9,10). The ELSD thus provided a tool to quantitate those compounds that did not possess a chromophore and assess the relative purity of samples that had multiple components (11–13).

The main role of the MS was to provide data to support the

expected structures of the samples. The response in the MS depended on the ionization properties of the analytes. There are several methods available for the ionization of compounds in an HPLC effluent. Atmospheric pressure chemical ionization was chosen for this method because it has been shown to be the most general for compounds of pharmaceutical interest (2). Because the resultant ions may carry a positive or a negative charge, the spectrometer was set to scan the selected mass range in both positive and negative modes. Although this decreased the resolution of the total ion chromatogram, it increased the chance of identifying a particular peak.

The use of multiple detectors increased the chance of detecting compounds, both the expected analytes, and impurities. It is conceivable that by using many different detectors we would be able to collect data on most if not all of the compounds in the sample collection. There is a tradeoff, however, in adding more detectors, because each mode adds complexity and thus time requirements for data analysis. Using this system, we can process 300 samples per day. However, even when using automated analysis tools (that will be described), many of the data sets need to be interpreted manually with the result that the rate-limiting step is data analysis. Therefore, improvements in sample handling and chromatography will not increase throughput until better data interpretation tools are developed.

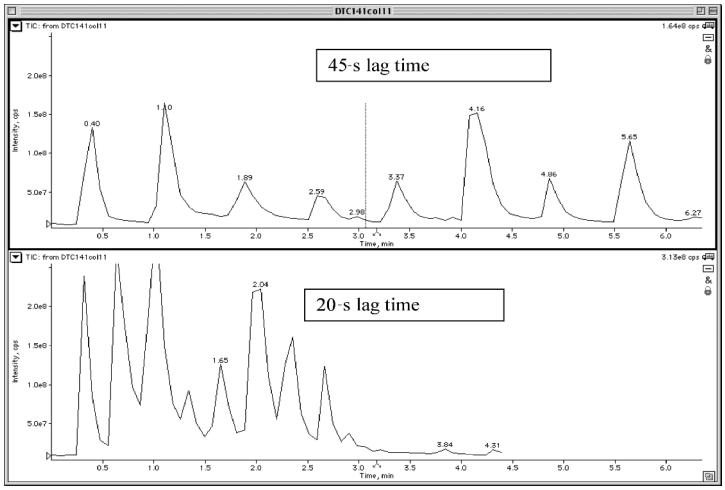

Figure 2. Comparison of 20- and 45-s time lags between injections in 8-arm flow injection systems.

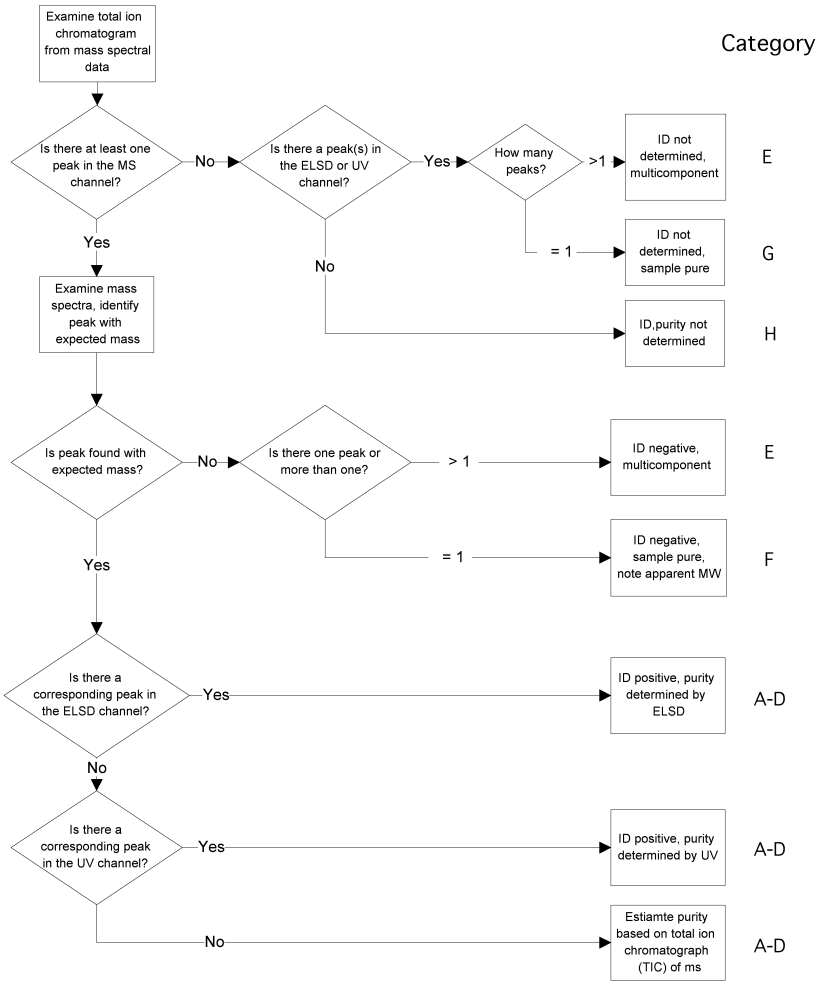

Figure 3. Decision tree for ID and purity analysis.

8-Probe flow injection system

The FIA system was set up for high-throughput analysis and could handle up to two plates per hour. The higher throughput was attained by the capability of the instrument to process eight samples simultaneously. Because there were eight probes moving together, the time required for sampling and washing was cut by 87.5%. A major concern of operating such a fast injection system is the carryover effect. In order to find an optimal injection speed, we experimented with different lag times between injections and found that, ideally, the injections should be separated by at least 20 s to minimize carryover between samples. Two such traces are shown in Figure 2. The use of the 10-port VICI valve, controlled by contact closures set in the Unipoint software, ensured that equal flow rate was experienced by each injector resulting in uniform elution times for each sample. Previous attempts using standard serial arrangements proved to give injection valve dependent peak widths and elution times.

Data analysis

HPLC–UV–ELSD–MS

Depending on the outcome for each detection channel, each sample was assigned to one of eight categories (Table III) indicating whether the expected MW was detected and if it had relative purity. The logic tree for assigning the categories is shown in Figure 3. If the LC–MS chromatogram contained ions consistent with the expected structure, the sample fell into one of four categories dependent on the relative ELSD or UV purity of the peak containing the expected ions (ranging from

"purity 75–100%", category A, to "purity 5–25%", category D). However, if the MS results were not consistent with the assigned structure, the sample fell into categories E (impure) or F (pure). The compounds that did not lend themselves to MS detection but are perceived as being pure by UV and ELSD fell into category G. Samples with no MS response that were impure by UV or ELSD were grouped in category E. Lastly, category H included compounds not observed by any of the detection channels, but were most likely caused by either a lack of retention of the compounds on the HPLC column or a lack of solubility in the HPLC solvent.

As an example, Figure 4 shows the results of the analysis of a real sample. The chromatograms showed two major peaks in all data channels. The mass spectrum identified the later eluting peak as a compound with the expected MW. The earlier (major) peak had an MW corresponding to a process impurity. This sample was assigned to be 45% pure based on the ELSD and the UV outputs.

The overall results for all samples are shown in Table III. Out of the 17,600 samples analyzed, we were able to use LC–UV–ELSD–MS to obtain quality information for 14,954 samples. The remaining 2646 samples gave no MS response, and the presence of the expected compound could not be confirmed or denied. The former group of MS responsive samples is summarized in the first six rows of Table III representing categories A–F. Samples containing the expected ions are summarized in categories A–D. Samples without the expected ions were listed as impure (category E) or pure (category F) based on the relative purity in the UV or ELSD detectors. As readily seen from the first row, the largest group (category A, having a purity of ≥ 75%) represented 70% of the samples (10,474 out of 14,954). The next three rows (50–75%, 25–50%, and 5–25% relative purity) in the table indicated those samples with intermediate purity in descending order. Samples in these three categories showed more than one peak, one of which

Table II. Partial Excel Table for Review of a Plate

PNU #	Formula	Weight	Well	XIC purity	UV purity	ELSD purity	Comments	No response	Pure, MW not seen	Pure, MW contrary	Multi, MW not seen	5–25%	25–50%	50–75%	75–100%
0030031	C11 H13 N3 O4	251.0906	A2	100	100	100									1
0126993	C7 H9 N5	163.0858	A1					1							
0126996	C18 H18 N4 O3 S	370.11	A3	0	99.9	98.6									1
0126997	C13 H12 N2 S2	260.0442	A4	100	96.7	94.1									1
0126998	C17 H16 N4 O3	324.1222	A5	100	100	100									1
0126999	C17 H11 Cl F3 N3 O2	381.0492	A6	100	99.9	96									1
0127000	C18 H17 N3 O3 S	355.0991	A7	100	95.1	95.2									1
0127002	C18 H13 Cl2 N5 S	401.0269	A8	0	99.1	85									1
0127016	C11 H7 Cl3 N2	271.9675	A9	100	93.6	93.6									1
0127017	C14 H12 Cl2 N2 O3	326.0225	A10	100	79.4	92.6									1
0127023	C12 H18 N2 O3	238.1317	A11				MW=290			1					
0127024	C18 H16 N O4 P	341.0817	B1	100	100	100									1
0127027	C12 H14 N4 O2	246.1117	B2	95.6	99.5	96.6									1
0127031	C9 H11 N3 O3 S	241.0521	B3	100	94.3	71.5									1
0127033	C16 H16 F3 N3 O	323.1245	B4	100	98.6	40.7									1
0127037	C17 H11 F3 N2 O5	380.062	B5	100	99.7	97.9									1
0127038	C16 H13 Cl2 N3 O2 S	381.0106	B6		22	30	remainder dehydrogenated						1		
0127045	C16 H11 F3 N4 O2	348.0834	B7	0	99.5	98.2									1
0127054	C19 H15 F3 N2 O2 S	392.0806	B8	100	100	100									1
0127055	C8 H19 N3 O3 S	237.1147	B9					1							
0127056	C16 H7 F6 N3 O3	403.0392	B10	98.4	99.8	95.1									1
0127058	C17 H16 Cl3 N O4	403.0145	B11	100	98.6	84.2									1
0127062	C13 H10 Cl F3 N4 O3	362.0394	C1	100	100	80.4									1
0127064	C18 H18 F3 N3 O3	381.13	C2	0	95.7	96.8									1
0127065	C12 H11 F3 N4 O3 S	348.0504	C3	0							1				
0127066	C14 H7 Cl2 N3 O5	366.9763	C4						1						
0127068	C15 H10 F3 N3 O2 S	353.0446	C5	100	100	100									1
0127069	C19 H13 Cl2 F3 N2 O3 S	475.9976	C6	100	98	98.7									1
0127074	C21 H17 F3 N4 O2 S	446.1024	C7	100	100	100									1
0127075	C16 H12 Cl2 N4 O	346.0388	C8	0	99.9	98.2									1
0127077	C15 H11 F3 N2 O5	356.062	C9	100	27.9	97.6									1
0127079	C13 H13 F3 N4	282.1092	C10	100	100	100									1
0127081	C19 H16 Cl2 N4 O2	402.065	C11								1				
0127083	C18 H19 Cl F3 N3 O2	401.1118	D1	100	99.5	92.5									1
0127084	C15 H12 N2 O7	332.0645	D2	100	100	100									1
0127085	C12 H9 F N2 S	232.047	D3	80.8	55.9	67.6								1	

had the expected MW. These three categories together constituted 1377 samples, or 9% of 14,954. As stated previously, samples that did not contain the expected ions were placed in categories E and F. Category E, totaling 1133 samples (7.6%), included samples with at least two peaks in either the UV or ELSD channels and had mass spectra contrary to that expected or had no MS response. These were considered as the lowest quality samples. Category F, totaling 1970 samples or 13.2%, represented those samples for which the main chromatographic peak was ≥ 75% pure, but the mass spectrum contradicted the expected structure. The structural data for these samples may have been entered incorrectly or the compound may have completely fragmented in the MS. After a follow-up to identify the compound present, these samples may be placed in category A.

A summary of the 14,954 samples that displayed a response in the MS is shown in the pie chart in Figure 5.

Samples giving no response in the MS were contained in categories G and H and represented a total of 2646 samples, or 15% of the total 17,600 assayed. Out of this group, 594 samples (category G) showed a single peak in the UV or ELSD channels (or both). This result indicated that the general ionization conditions employed for MS were not appropriate for these compounds. Nonetheless, these compounds were presumed to be pure as judged by the UV and ELSD outputs. The remaining 2052 samples (category H) showed no response in any data channel. Possible explanations for no response in this category include compound insolubility in the sample solution after the DMSO was diluted with 50% aqueous acetonitrile and that the sample was not retained on the HPLC column. Unless further analysis is performed, no conclusion as to the identity of the samples contained in these categories can be made. FIA was used as a rapid follow-up method to verify the results of all the samples, with a special focus on category E, G, and H samples.

FIA

Because only MS data were obtained in this experiment, the data were only analyzed for the presence or absence of the predicted molecular ions for the expected structures. The results of the FIA were collected as a text file that listed each sample with the intensity of the expected molecular ions. In order to set a cutoff for assigning the samples as positive or negative for the expected compounds, the data for several blank injections was analyzed. The cutoff of 10^5 counts per second was set to avoid noise in the mass spectra from being identified as a real ion. This proved adequate in avoiding false positives as well as detecting true and weakly ionized samples. The results are shown in Table IV. Of the 2052 samples from category H that showed no response in any detector originally, 927 samples displayed the expected MS by FIA. The obvious explanation is that these samples contained the expected structures but the compound eluted with the solvent front during the LC run. This argument also applied to 430 of the 1133 samples from category E (listed as impure and MW inconsistent with the expected structure) that had a positive ID by FIA. It may be that the expected structures eluted at the solvent front but the impurities eluted later. In both cases, the sample would be wrongly categorized (H or E instead of A through D as appropriate).

Table III. Categorization of Samples in DTC Dissimilarity Set by Purity and Identification

Label	No. of samples	Explanation
A, purity > 75%	10474	Single peak in ELSD, UV, or both; mass spectrum matched formula
B, purity 50–75%	684	Peak area in ELSD, UV, or both with expected mass spectrum integrated between 75% and 50% total
C, purity 25–50%	420	Peak area in ELSD, UV, or both with expected mass spectrum integrated between 5% and 25% total
D, purity 5–25%	273	Peak area in ELSD, UV, or both with expected mass integrates
E, impure, MW inconsistent	1133	> 1 peak, no mass spectrum matched formula
F, pure, MW inconsistent	1970	Single peak in UV, ELSD, or both; mass spectrum did not agree with formula
G, pure, no MS signal	594	Single peak in UV, ELSD, or both, but no mass spectrum
H, no response	2052	No peaks in MS, ELSD, or UV later than void volume

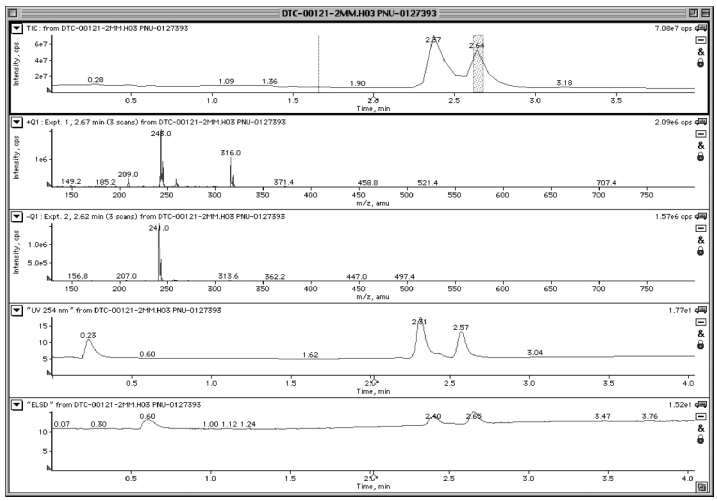

Figure 4. Analysis of PNU-127393 by HPLC–UV–ELSD–MS.

Discussion

When analyzing results, it is important to recognize that both LC and FIA methods are subject to certain biases, some of which overlap. First and foremost, if the analyte does not ionize well in the MS, it does not matter which method is used for analysis because neither one will correctly identify it. Some of the samples in category G and H that failed the FIA test may fall into this class. For the FIA method, suppressed ionization resulting from coinjected impurities can also cause false nega-

tives. Ion suppressions are eliminated if the interfering components are separated by LC. Another cause for FIA bias is difference in sample solubility. In this study, FIA samples were dried to remove DMSO because it suppresses ionization of other sample components. Apparently, on reconstitution in 50% aqueous methanol containing 0.05% by volume of formic acid, some samples failed to redissolve. This was evidenced by the fact that almost 25% of category A samples failed to be correctly identified by the FIA method. Differences in solubility and possibly compound stability in the sample solvent between the two protocols may lead to different conclusions about a particular sample. For those cases in which the result from HPLC–UV–ELSD–MS was negative and the FIA was positive, it was suspected that most of these samples were not seen in the former protocol because they were not retained on the LC column and thus shunted to waste. Another possible cause for false negative identification was that the missed compounds were thermally labile. The samples were exposed briefly to a 400°C zone in the interface, and thermally labile compounds may not survive such treatment. For the samples in this class, other techniques such as electrospray ionization (ESI)–MS, nuclear magnetic resonance, or both were needed to determine the structures and correct the data if necessary.

When considering all of the possible reasons for discrepancies, there is a much greater potential for false negative identifications than exists for false positive ones. Unfortunately, no single set of conditions would be optimal for all compounds. If the FIA results are taken into consideration, the percentage of "pure and correct" compounds in category A is very likely higher by another 10%. Of the remaining categories, B–D (lesser purity but containing the expected structures) constitute 9.2% and E (impure, MS inconsistent) constitutes 7.6% (Figure 5). As discussed previously, the estimation of the number of compounds contained in this latter class of compounds may be on the high side because it contains target compounds eluted at the solvent front or may be thermally labile.

Table IV. Flow Injection MS Correlated with LC–UV–ELSD–MS Results

	No. of samples			
LC–UV–ELSD–MS category	FIA ID*	FIA ID†	FIA ID‡	Totals
A, purity > 75%	8300	2125	45	10470
B, purity 50–75%	530	152	4	686
C, purity 25–50%	299	108	8	415
D, purity 5–25%	172	98	2	272
E, impure, MW inconsistent	430	697	9	1136
F, pure, MW inconsistent	659	1323	6	1988
G, pure, no MS signal	237	347	3	587
H, no response	927	1108	11	2046
Total	**11554**	**5958**	**88**	**17600**

* Yes.
† No.
‡ Not detected.

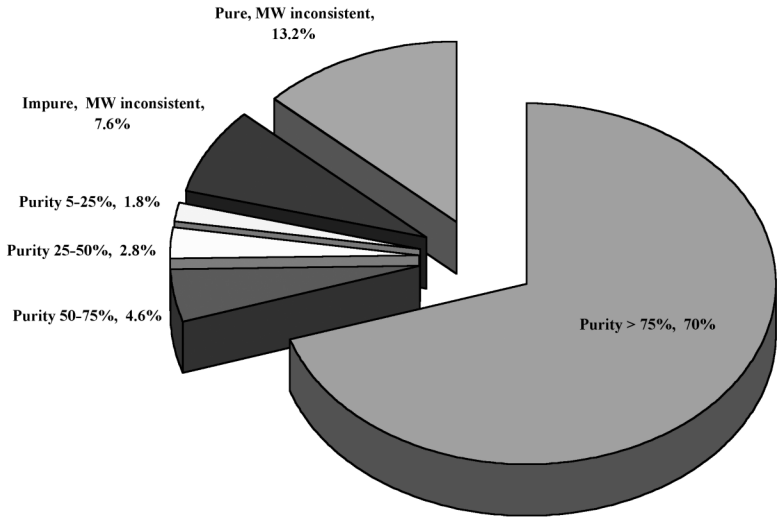

Figure 5. Summary of observed identification and purity of samples from DTC 1–200 (*n* = 14954).

Conclusion

Moderate and high-throughput analytical methods were developed to analyze the quality of a large set of samples selected from the compound collection based on chemical dissimilarity. The results from the moderate throughput HPLC–UV–ELSD–MS method indicated that more than 70% of the samples were pure and correct. The next largest category (13%) contained compounds deemed as pure but gave mass spectra inconsistent with the expected structure. Our FIA results suggested that this may be an overestimation because 33% of them displayed the expected mass spectra using FIA. The overall percentage of compounds that may be considered "pure and correct" was very likely to be near 80% after considering the FIA results. Of the remaining categories, those less pure but containing the expected structures constituted 9.2% and the "impure and incorrect" group constituted 7.6%. The availability of such data has provided the users and curator of the compound collection with useful information and boosted their overall confidence in continuing the screening of these compounds for biological activities. Moreover, the ready availability of individual compound analytical data has allowed screeners to immediately assess the quality of their leads. Finally, this set of data has allowed us to retroactively assess the quality of commercially purchased compounds. Such information can be a useful guide for future procurements.

Acknowledgments

Some of the results discussed in this study have been presented in an oral session at the Society of

Biological Screenings Annual Conference in September 2000 (Vancouver, BC, Canada).

References

1. Mark J. Hayward, Leonard O. Hargiss, and James L. Munson III. "Coordinated Selection of Stationary and Mobile Phases for Fast LC/MS Used in Parallel Synthesis and Open Access Applications". Presented at the 48th ASMS Conference on Mass Spectrometry and Allied Topics, Los Angeles, CA, June 11–15, 2000. Paper WPH 301.
2. *The Atmospheric Pressure Ionization (API) Technique*, 5th ed. PESciex, Central Productions, Mississauga, ON, Canada, 1990.
3. T.R. Covey, E.D. Lee, A.P. Bruins, and J.D. Henion. Liquid chromatography/mass spectrometry. *Anal. Chem.* **58(14):** 1451A–61A (1986).
4. B.A. Thomson and J.V. Irbarne. Field induced ion evaporation from liquid surfaces at atmospheric pressure. *Chem. Phys.* **71(11):** 4451–63 (1979).
5. A.P. Bruins, T.R. Covey, and J.D. Henion. Determination of sulfonated azo dyes by liquid chromatography/atmospheric pressure ionization mass spectrometry. *Anal. Chem.* **59:** 2642 (1987).
6. E.W. Taylor, M.G. Qian, and G.D. Dollinger. Simultaneous on-line characterization of small organic molecules derived from combinatorial libraries for identity, quantity, and purity by reversed-phase HPLC with chemiluminescent nitrogen, UV and mass spectrometric detection. *Anal. Chem.* **70(16):** 3339–47 (1998).
7. D.A. Yurek and Ming-Shang Kuo. "Development and Application of a System to Evaluate Compound Identity, Purity, and Concentration in One Experiment Using HPLC-PDA-ELSD-CLND-TOFMS". Presented at the 48th ASMS Conference on Mass Spectrometry and Allied Topics, June 11–15, 2000. Paper ThPF 208.
8. M.S. Lajiness. Dissimilarity-based compound selection techniques. *Perspect. Drug Discovery Des.* **7/8:** 65–84 (1997).
9. T.H. Mourey and L.E. Oppenheimer. Principles of operation of an evaporative light scattering detector for liquid chromatography. *Anal. Chem.* **56:** 2427–34 (1984).
10. C. Henry. In the Spray with ELSDs. *Anal. Chem. News Features* 561–3A (1997).
11. A.I. Hopia and V.-M. Ollilainen. Comparison of the evaporative light scattering detector (ELSD) and refractive index detector (RID) in lipid analysis. *J. Liq. Chromatogr.* **16(12):** 2469–82 (1993).
12. S. Kermasha, S. Kubow, and M. Goetghebeur. Comparative high-performance liquid chromatographic analyses of cholesterol and its oxidation products using diode-array ultraviolet and laser light- scattering detection. *J. Chromatogr. A* **685(2):** 229 (1995).
13. H.J. Chaves das Neves and Z.B. Morais. HPLC assay of underivatized free amino acids with column switching and evaporative light-scattering detection. *J. High Resolut. Chromatogr.* **20(2):** 115 (1997).

Manuscript accepted September 28, 2001.