Journal of Chromatography A, 1217 (2010) 1616–1625

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

# A high throughput metabolic stability screening workflow with automated assessment of data quality in pharmaceutical industry

Rongda Xu*, Melinda Manuel, Joshua Cramlett, Daniel B. Kassel

*Takeda San Diego, Inc., 10410 Science Center Dr., San Diego, CA 92121, USA*

ARTICLE INFO

*Article history:*
Received 17 June 2009
Received in revised form
24 December 2009
Accepted 6 January 2010
Available online 14 January 2010

*Keywords:*
Intelligent
Workflow
High throughput
Metabolic stability assay
Microsomal stability assay
Sample pooling
Cassette analysis
UPLC
UHPLC
MS
Mass spectrometry
Automation

ABSTRACT

One of the most commonly performed *in vitro* ADME assays during the lead generation and lead optimization stage of drug discovery is metabolic stability evaluation. Metabolic stability is typically assessed in liver microsomes, which contain Phase I metabolizing enzymes, mainly cytochrome P450 enzymes (CYPs). The amount of parent drug metabolized by these CYPs is determined by LC/MS/MS. The metabolic stability data are typically used to rank order compounds for *in vivo* evaluation. We describe a streamlined and intelligent workflow for the metabolic stability assay that permits high throughput analyses to be carried out while maintaining the standard of high quality. This is accomplished in the following ways: a novel post-incubation pooling strategy based on *c* Log $D_{3.0}$ values, coupled with ultra-performance liquid chromatography/tandem mass spectrometry (UPLC/MS/MS), enables sample analysis times to be reduced significantly while ensuring adequate chromatographic separation of compounds within a group, so as to reduce the likelihood of compound interference. Assay quality and fast turnaround of data reports is ensured by performing automated real-time intelligent re-analysis of discrete samples for compounds that do not pass user-definable criteria during the pooling analysis. Intelligent, user-independent data acquisition and data evaluation are accomplished via a custom visual basic program that ties together every step in the workflow, including cassette compound selection, compound incubation, compound optimization, sample analysis and re-analysis (when appropriate), data processing, data quality evaluation, and database upload. The workflow greatly reduces labor and improves data turnaround time while maintaining high data quality.

© 2010 Elsevier B.V. All rights reserved.

## 1. Introduction

The pharmaceutical community has recognized that evaluation and optimization of pharmacokinetic (PK) properties of compounds during drug discovery is critical to their ultimate clinical development success. Over the past two decades, compound failures in the clinic due to poor PK properties have plummeted from 40% in 1990 to 10% in 2000 [1]. This significant drop in clinical development attrition due to poor PK may be directly attributable to increased attention to PK evaluation during drug discovery. In the last decade, advances in automated synthesis, high throughput screening, and high throughput analytical technologies have significantly increased the number of compounds that are synthesized during the discovery phase of a program. Prior to the advent of these high throughput technologies, medicinal chemistry project teams made relatively few compounds during the lead generation and lead optimization stages of drug discovery, and hence, it may have been possible (theoretically) to rely on an *in vivo* animal PK screening strategy as a sole means for optimizing and selecting the best candidate drugs for pre-clinical development. However, medicinal chemistry project teams now make far greater number of compounds during lead generation and lead optimization than they did previously (a conservative estimate is 10-fold to 100-fold more compounds since 2000 than 1990), and the shear number of compounds synthesized precludes the ability to be optimized solely based on an *in vivo* testing strategy. To address this, scientists have implemented *in vitro* absorption, distribution, metabolism, and elimination (ADME) assays as surrogates for *in vivo* evaluation as a way to streamline the evaluation and optimization of PK properties. These high throughput *in vitro* ADME assays allow for parallel optimization of structure–ADME property relationship along with structure–activity relationship [2–13]. Scaffolds identified with superior ADME properties are given higher priority for *in vivo* profiling, such as efficacy and toxicity evaluation, whereas those with inferior ADME properties are further optimized or dropped to minimize attrition at later stages.

Liquid chromatography/tandem mass spectrometry (LC/MS/MS) has been the technique of choice for *in vitro* ADME analysis

* Corresponding author at: Analytical Sciences and DMPK, Takeda San Diego, Inc., 10410 Science Center Dr., San Diego, CA 92121, USA. Tel.: +1 858 731 3712; fax: +1 858 550 0526.
*E-mail address:* ron.xu@takedasd.com (R. Xu).

0021-9673/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2010.01.009

due to its high sensitivity, high selectivity, versatility, and ease of automation [3–5,14–16]. However, analyses by LC/MS/MS are generally carried out one at a time (i.e., serial analysis mode). Consequently, the data acquisition step becomes a bottleneck when large numbers of compounds are being evaluated. Metabolic stability screening, one of the most commonly conducted ADME assays, is impacted most by such a data acquisition bottleneck. In the metabolic stability assay, test compounds are incubated in human and animal liver microsomes over multiple time points to assess the half life ($T_{1/2}$) and intrinsic clearance ($Cl'_{int}$) of the test compounds in the presence of the drug metabolizing Cytochrome P450 enzymes (CYPs). $Cl'_{int}$ can be used to predict the extent of *in vivo* clearance and/or to understand and propose potential *in vivo* clearance pathways when compared with *in vivo* PK results [7,8]. Taking an example of 48 compounds to be assayed each day with 4 time points in duplicate for two species, e.g., human and rat liver microsomes (HLM and RLM), there will be 768 samples to be analyzed. Using a typical LC/MS/MS cycle time of 2.5 min/sample, the total analysis time will exceed 32 h. Furthermore, additional instrument time (2–4 h) has to be spent to optimize MS/MS transitions and parameters individually for each test compound before the sample analysis. Consequently, one LC/MS/MS system is required for a relatively small number of samples. When demand grows further, more instrumentation will be needed if analysis efficiency is not improved.

Numerous groups have implemented methods that increase analysis throughput. Injection cycle times as short as <1 min may be achieved by incorporating short chromatographic columns operated at high flow rate [17–20]. Staggered elution from multiple columns makes use of the chromatographic dead-time, where peaks of interest are not eluted, which improves throughput while maintaining chromatographic separation. The staggered elution method works especially well with large number of samples of the same compound (e.g., clinical bioanalysis). Parallel LC/MS/MS (indexed and non-indexed) has been shown to greatly increase analysis throughput [5,21]. Recently, direct, non-chromatographic methods have been evaluated for high throughout ADME analysis [4,5,22]. A silicon chip-based nanospray infusion system was shown to reduce matrix suppression at nl/min flow rates [23]. A high repetition rate MALDI triple quadrupole mass spectrometer was shown recently to achieve sample analysis times approaching plate reader speed [24] in which failed samples could be transferred to an electrospray mass spectrometer for follow-up analysis [25]. An on-line SPE-MS/MS system was shown to permit analysis times of 6–10 s per sample [26].

Other methods for improving analysis efficiency include either pre-incubation or post-incubation pooling, sometimes referred to as cassette (or N-in-1) dosing and cassette analysis [4,5,27]. Numerous groups have implemented a post-incubation sample pooling strategy to accelerate analyses for the *in vitro* metabolic stability assay [3,4,8,11,16]. The pooling strategy generally considers molecular weight redundancy between compounds and sometimes metabolites [6,28] as well as potential signal suppression caused by coelution among compounds to be pooled. Experiments are normally conducted either to determine retention times for the selection of compounds in each cassette group before pooling or to develop LC methods to resolve the compounds in each cassette group.

However, fast data turnaround of the assay results is important for accelerating the iterative cycle of the lead optimization. Such requirement for large number of compounds in the metabolic stability assay leaves little time for LC method development. It is possible to simply avoid such pre-work by taking the calculated risk of potential compound interference and signal suppression based on the good correlation obtained from prior statistical data comparison between the cassette analysis and the discrete analysis of a set of validation compounds [11], but quality of data may be compromised for compounds with poor chromatographic separations or poor ionization efficiency. Although cassette analysis reduces analysis time, certain shortcomings come with it. For example, sample pooling reduces compound concentration, which affects ion intensity for compounds that do not ionize well. Also, post-incubation pooling may introduce additional liquid handling errors. Consequently, the analyst may have to spend significant effort and use more instrument time to troubleshoot.

We have established a strategy that maintains assay quality while increasing throughput. We accomplished this by introducing a pooling strategy that pre-selects compounds for analysis based on calculated Log $D$ ($c$ Log $D_{3.0}$) values coupled with ultra-performance liquid chromatography (UPLC) to ensure high resolution separation of complex mixtures and reduce potential peak coelution [29]. To address the potential shortcomings of a pooling strategy, we also implemented an automated real-time process to intelligently re-inject discrete samples for the compounds that do not pass user-definable criteria from the pooling analysis. We also created a method for automated QC evaluation of test compounds under various solution conditions, which proved useful for automated troubleshooting of compound solubility and/or degradation issues. These automated measures ensure quick data turnaround while at the same time ensuring quality of the reported assay data [30]. With such enhanced analysis capability, more test compounds can be analyzed without adding additional LC/MS/MS instruments or scientists to assess data quality. A custom Visual Basic program has been developed that interfaces with our corporate database and ties together every step in the workflow, including discrete compound incubations, cassette compound optimization, cassette compound analysis and discrete compound re-analysis (as appropriate), data processing, data quality evaluation, and database upload. This workflow greatly reduces labor and improves data turnaround time while maintaining high data quality.

## 2. Experimental

### 2.1. Chemicals and reagents

Optima grade acetonitrile (ACN) was obtained from Fisher Scientific (Fair Lawn, NJ, USA). Dimethylsulfoxide (DMSO), formic acid, trichloroacetic acid (TCA), potassium phosphate monobasic and dibasic, β-nicotinamide adenine dinucleotide phosphate reduced form (NADPH), buspirone hydrochloride, (±)-sulpiride, phenacetin, coumarin, bupropion hydrochloride, and amodiaquine dihydrochloride dihydrate were obtained from Sigma Chemical Co. (St. Louis, MO, USA). Proprietary drug discovery compounds tested routinely using the described workflow were synthesized by the Medicinal Chemistry Lab of Takeda San Diego. These compounds are non-peptide small molecules with diverse drug-like structures with molecular weights ranging from 200 to 600 and $c$ Log $D_{3.0}$ values ranging from −5 to 5 in general. Pooled human and rat liver microsomes (HLM and RLM) containing 20 mg/mL of total protein were obtained from BD Biosciences (San Jose, CA, USA). Potassium phosphate buffer was prepared as a 50 mM stock solution in deionized water by combining monobasic and dibasic phosphate to achieve pH 7.4. Stock solutions of test compounds were prepared to 10 mM concentration in DMSO. NADPH solution was prepared to 16 mM concentration in phosphate buffer prior to use. TCA solution was prepared to 0.38 M in deionized water. The microsomes were thawed and mixed with phosphate buffer to give a 2 mg/mL total protein concentration prior to use.

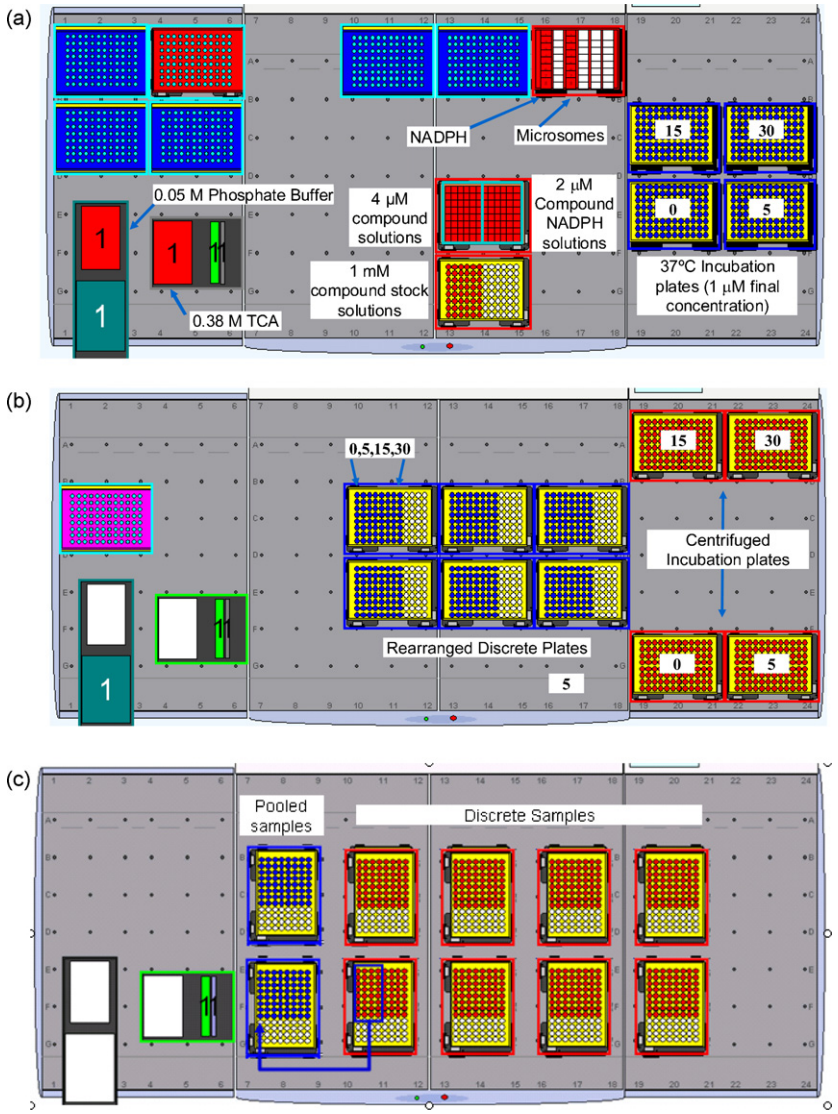


**Fig. 1.** MP II deck layout for the incubation of 48 compounds with one species (a) incubation, (b) post-incubation rearrangement to create discrete plates, (c) creating plates for pooled samples.

### 2.2. High throughput metabolic stability assay

A Multiprobe II HT EX (MP II) Liquid Handler (PerkinElmer, Waltham, MA) was used to carry out liver microsomal incubations, reaction quenching, and post-incubation sample pooling in 96-well microtiter plates. Plate-based incubations were conducted at 37 °C, using a Microplate Heater (PerkinElmer) fitted with custom aluminum blocks that provided close contacts with individual wells of the 500 μL conical incubation sample plates (Thomson Instrument Company, Oceanside, CA, USA). The incubations were performed in duplicate and quenched at 0, 5, 15, and 30 min. Each incubation set allowed for evaluation of either 48 compounds with single species, such as HLM or RLM, or 24 compounds with both species. This procedure has been applied to the metabolic stability studies of compounds in liver microsomes from various species, including human, rat, monkey, dog, and mouse, although only human and rat liver microsomal experiments are described here. Stock solutions of test compounds (10 mM in DMSO) along with reference compounds (buspirone as positive control and sulpiride as negative control) were prepared. The MP II liquid handler was programmed to rearrange the plated samples according to work list created for cassette groups by the custom program described later. The rearranged stock solutions were then manually diluted with ACN to 1 mM solutions and placed on MP II deck along with thawed and diluted 2 mg/mL microsomal solution in buffer, 16 mM freshly prepared NADPH solution in buffer, 0.38 M TCA solution, and empty incubation plates. Fig. 1a shows an MP II deck layout for the incubation of 48 compounds with HLM or RLM. The liquid handler started with one-step dilution of 1 mM solutions to 4 μM in buffer followed by 1:1 mixing with NADPH solution. For 5, 15, and 30 min incubations (*T*1, *T*2, and *T*3), the compound/NADPH solutions and microsomal solution were added into the heated plates to start the incubation. The final incubation solutions contained 1 μM test compound, 4 mM NADPH, 1 mg/mL (total protein) microsomes, and 50 mM phosphate (pH 7.4). Equal volume of TCA solution was added to quench the reactions at appropriate times. For 0 min samples (*T*0), microsomal solution and TCA solution were mixed before compound/NADPH solutions were added. After all incubations were quenched, the incubation plates were manually transferred into an Allergra 25R centrifuge (Beckman Coulter, Fullerton, CA, USA) and centrifuged at 6100 × *g* for 10 min. The liquid handler then transferred supernatants from the incubation plates and rearranged them into empty discrete analysis plates so that all samples for the same compound were in the same plate

(Fig. 1b). At the final step, samples (15 μL each) from four compounds in the discrete analysis plates were then pooled into empty pooling analysis plates (Fig. 1c). The pooling and discrete analysis plates were then transferred to the UPLC/MS/MS system for analysis. QC plates containing 0.5 μM neat compounds in 0.19 M TCA solution, phosphate buffer solution, and 1:1 ACN/buffer solution were also prepared.

### 2.3. Ultra-performance liquid chromatography/mass spectrometry (UPLC/MS/MS)

The samples were analyzed using a Quattro® Premier Tandem Quadrupole Mass Spectrometer coupled to an Acquity® Ultra-Performance Liquid Chromatograph (UPLC®) consisting of a Binary Solvent Manager, a Sample Manager with column compartment, and a Sample Organizer (Waters Corp., Milford, MA, USA). The mass spectrometer was fitted with an ESI probe and was controlled by MassLynx 4.1 software. The volume of the sample loop was 20 μL and the injection volume was 10 μL in "partial loop injection with needle overfill" mode. An Acquity UPLC BEH C18 column (1.0 mm × 50 mm, 1.7 μm, Waters Corp.) was used and maintained at 50 °C within the column compartment. The flow rate was 0.35 mL/min. The mobile phases used in this study were A: 0.04% formic acid in water and B: 0.04% formic acid in acetonitrile. The following mobile phase gradient was applied: 1% B hold for 0.3 min, 1–95% B gradient in 1 min, hold at 95% B for 0.3 min, and hold at 1% B for 0.4 min. The column was re-equilibrated at starting conditions while the next sample was injected. The mass spectrometer was configured to acquire data with multiple reaction monitoring (MRM) in the positive ionization mode. Q1 and Q3 ion transition, as well as cone voltage (CV) and collision energy (CE) were optimized using QuanOptimize module within MassLynx program. The following mass spectrometer parameters were used for sample analysis: MRM Dwell time: 0.02 s for pooled samples or 0.06 s for discrete samples, Interchannel delay: 0.01 s, interscan delay: 0.01 s, capillary voltage: 1.0 kV, extractor voltage: 5 V, RF lens voltage: 0.5 V, source temperature: 125 °C, desolvation temperature: 450 °C, desolvation gas flow: 850 L/h, cone gas flow: 50 L/h, ion energy 1 and 2: 1.0, entrance potential: −3, exit potential: 2, mutliplier: 650, collision gas flow: 0.23 mL/min. MRM data were acquired at near unit resolution: low mass resolution (LM) 1 and 2: 13, high mass resolution (HM) 1 and 2: 13.

### 2.4. Custom Visual Basic program for a streamlined workflow

A custom Visual Basic program, we coined Intelligent Metabolic Stability Analysis for Pooled Samples (IMSAPS), was developed for (a) sorting and grouping compounds into cassettes and creating MP II work list for the rearrangement of stock solutions; (b) creating all UPLC/MS/MS experimental files; (c) monitoring progress of analysis; (d) making intelligent decisions on compounds requiring discrete re-injection and QC analysis; (e) automatically submitting discrete re-injection samples and QC samples to MassLynx through AutoLynx module; and (f) merging data from pooling and discrete analyses for database upload. Fig. 2 shows the program graphical user interface, including (a) the main form, (b) the compound selection module, and (c) the settings window. IMSAPS enables a streamlined workflow that effectively ties together every step in the assay process, including compound incubation, compound optimization, sample analysis and re-analysis (when appropriate), data processing, data quality evaluation, and database upload (Fig. 3). The workflow is described in detail below.

Once receiving the stock plate(s), the ADME scientist uses the IMSAPS Compound Selection module (Fig. 2b) to retrieve the compound information from the corporate database, which includes the following information: plate barcodes, compound ID's, formula, monoisotopic mass, and lipophilicity values of the compound in unionized state ($c\log P$), at pH 7.4 ($c\log D_{7.4}$) and at pH 3.0 ($c\log D_{3.0}$) that are calculated using Marvin $c\log P$/$c\log D$ Calculators (ChemAxon, Budapest, Hungary). The IMSAPS program sorts the compounds according to $c\log D_{3.0}$ values and selects compounds for cassette analysis so that the $c\log D_{3.0}$ values of the compounds in the same group are separated as much as possible from one another. The program also ensures molecular weights of the compounds in a "pool" differ by more than 2 to avoid isobaric mass interferences. Once the compounds are grouped, the program registers new compound locations in a new assay plate into the corporate database (required for the data upload and reporting) and then generates a work list file for MP II to perform the rearrangement of stock solutions.

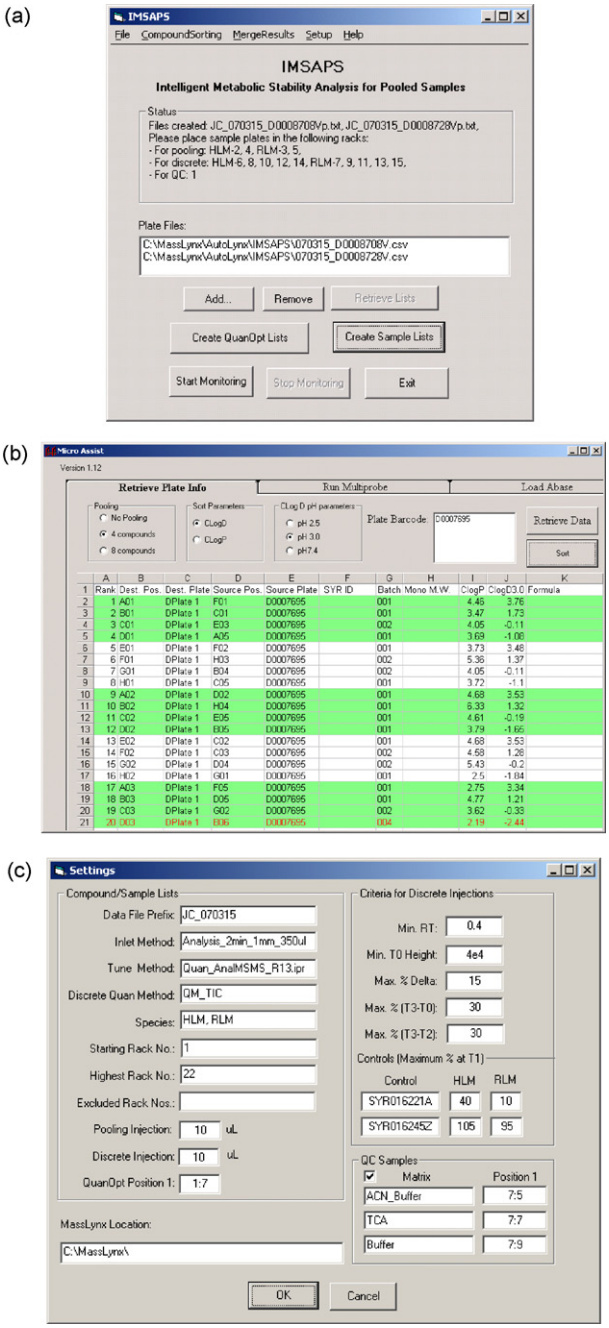


**Fig. 2.** The IMSAPS graphical user interfaces: (a) main form, (b) compound selection module (certain proprietary information removed), and (c) settings window.

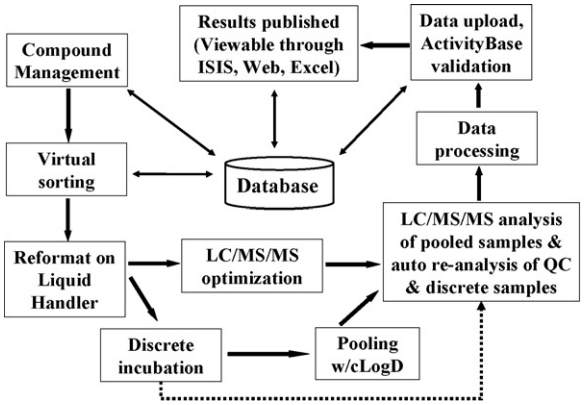


**Fig. 3.** A streamlined workflow using IMSAPS that effectively ties together every step in the assay process, including compound incubation, compound optimization, sample analysis, data processing, and database upload.

The scientist confirms or updates the Settings window (Fig. 2c) for all parameters that are required to generate UPLC/MS/MS compound list for optimization and sample lists for pooled sample analysis, including file naming prefix, LC and MS methods, microsomal species, available rack positions, injection volumes, etc. Intelligent data analysis is a key part of the IMSAPS program, allowing for automated evaluation of data quality and re-injecting discrete samples for compounds that fail to meet a set of user-definable quality metrics, such as minimum retention time, minimum signal intensity, acceptable variation between replicates, and data trend over the incubation time points. The IMSAPS program applies these user-definable criteria to intelligently evaluate the data from the pooled samples to determine what compounds would require re-analysis from the discrete samples and what compounds would require a troubleshooting for sensitivity with the QC samples.

IMSAPS (Fig. 2a) is used to import the new plate details to generate sample lists for UPLC/MS/MS optimization and pooled sample analysis after browsing in text files containing the compound information including compound formula and well positions. One command button generates compound list for QuanOptimize to optimize Q1, Q3, CV, and CE values for each test compound. Another command button generates sample lists to be imported into MassLynx for analysis of pooled samples. The MS/MS optimization of the test compounds are then performed using the MassLynx QuanOptimize with a plate of 5 μM solutions made from the reformatted stock solutions. While the MS setup for the compounds is optimized, the individual compound incubations are performed on MP II and, at the end of the incubation, a portion of the supernatants are pooled in groups of 4 at each incubation time point. The final analysis plates containing pooled samples as well as discrete samples are then placed in Quattro Premier Sample Manager for analysis according to sample rack positions prompted in IMSAPS status box. Once the Start button is pressed, IMSAPS monitors the progress of the UPLC/MS/MS analysis through the status.ini file generated by MassLynx. When the analysis of a sample list is completed, MassLynx processes each chromatogram with its QuanLynx module and generates a text file containing chromatographic data of each sample. IMSAPS detects the text file, retrieves peak areas and retention times of each sample, evaluates data quality of each compound, and intelligently submits failed compounds to be re-injected from discrete and/or QC plates. When the scientist arrives for work the next day, all data have been acquired for final processing and, if needed, troubleshooting. The scientist may inspect or adjust the peak integrations of the pooled and discrete samples in QuanLynx and export the final chromatographic data as text files, which are then merged by IMSAPS MergeResults tool, and uploaded into ActivityBase (IDBS, Guildford, Surrey, UK) for final curve fitting and validation. The validated results (half life, $T_{1/2}$, and intrinsic clearance, $Cl'_{int}$) are then automatically published in the corporate database and are available for the project teams to retrieve through various informatics tools, such as ISIS/Base, web, or Excel. This efficient process ensures routine data turnaround time of 1–2 days.

## 3. Results and discussion

### 3.1. Chromatographic retention times of test compounds correlated to their c Log $D_{3.0}$ values

Compound lipophilicity values (Log $D$) have been shown to correlate with retention time on C18 reversed phase columns under certain conditions [31]. Many software programs are available to calculate lipophilicity values (so-called $c$ Log $P$ and $c$ Log $D$). Since our mobile phase pH is ~3.0, the $c$ Log $D_{3.0}$ values were used to select compounds into cassette groups so that the compounds in each group would be more likely resolved chromatographically. The experimental results from 110 compounds indeed showed that $c$ Log $D_{3.0}$ values correlate well with retention times ($R^2 = 0.83$) (Fig. 4a). On the other hand, little correlation was observed when $c$ Log $P$ values were used (Fig. 4b), because $c$ Log $P$ values did not reflect the lipophilicity of the test compounds under conditions of the chromatographic mobile phase, i.e. pH 3.0.

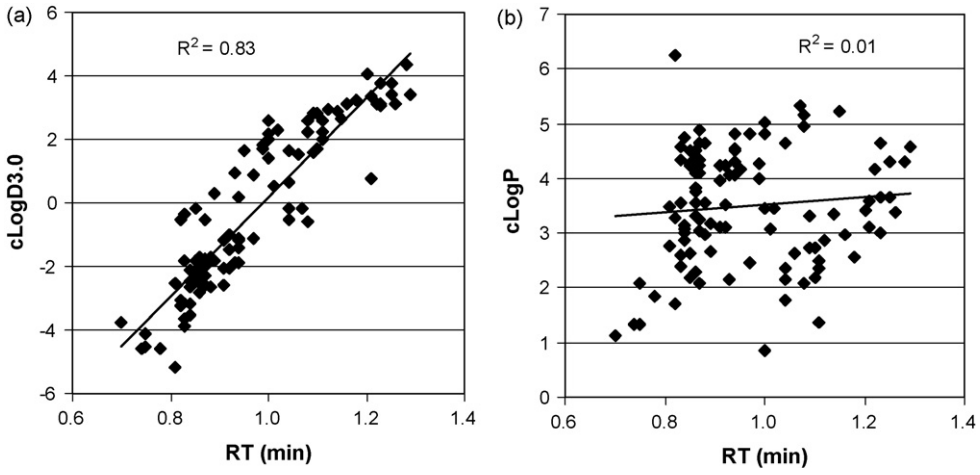


**Fig. 4.** Correlation of retention time with (a) $c$ Log $D_{3.0}$ value and (b) $c$ Log $P$ value from a test set of 110 compounds. Data show that retention time correlates much better with $c$ Log $D_{3.0}$ than with $c$ Log $P$.

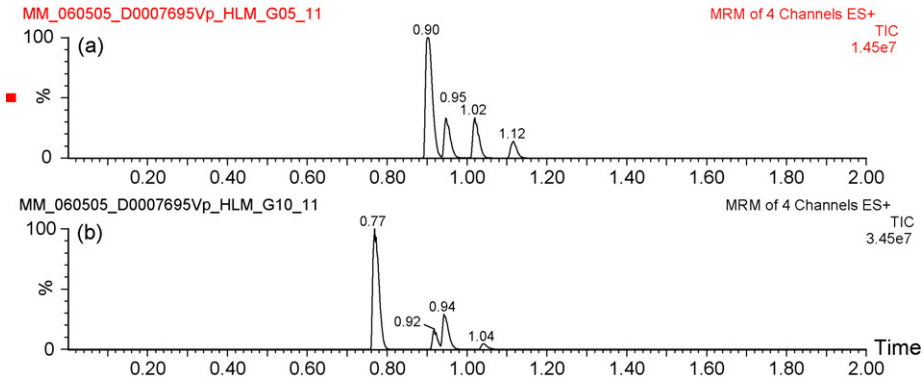


**Fig. 5.** Pooled microsomal samples: (a) baseline resolution achieved for two peaks separated by 3 s (retention time = 0.90 and 0.95 min, respectively); (b) nearly baseline-resolved for peaks separated by <2 s (retention time = 0.92 and 0.94 min, respectively).

### 3.2. Chromatographic resolution further enhanced by UPLC

In many cases the *c* Log $D_{3.0}$ values predict well the interaction of compound functional groups and the column stationary phase [31]. However, grouping based on *c* Log $D_{3.0}$ does not always ensure baseline resolution will be achieved in the chromatographic analysis. This is especially true when the compounds in a batch have very similar structures. In this instance, the compounds in a cassette group may have only moderately different lipophilicity values and may therefore present a challenge for conventional HPLC separation. UPLC, on the other hand, provides further chromatographic resolving power for compounds with similar lipophilicity values. Fig. 5a shows that baseline resolution is achieved for two peaks separated by as little as 3 s (retention time = 0.90 and 0.95 min, respectively). Even peaks separated by <2 s (retention time = 0.92 and 0.94 min, respectively) were nearly baseline-resolved (Fig. 5b).

### 3.3. Correlation of results between pooled samples and discrete samples

To evaluate whether or not the stability results would be affected by the sample pooling, a set of samples were analyzed in both pooling and discrete modes. Validation data demonstrated excellent correlation of results between pooled samples and discrete samples. Fig. 6 shows an example of a batch of 48 compounds assayed in both HLM and RLM. Two of the compounds were

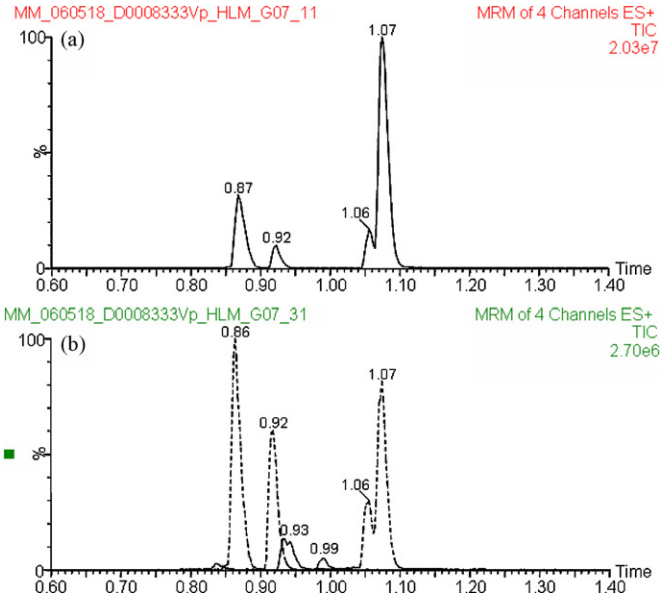


**Fig. 7.** Overlay of MRM chromatograms of a group of significantly metabolized compounds, showing good resolution from those of the oxidation metabolites. (a) MRM chromatograms of four test compounds at $t = 0$ min incubation; (b) overlay of MRM chromatograms of oxidation metabolites (solid line) overlaid with parent compounds (dashed line) in a 15-min incubation sample.

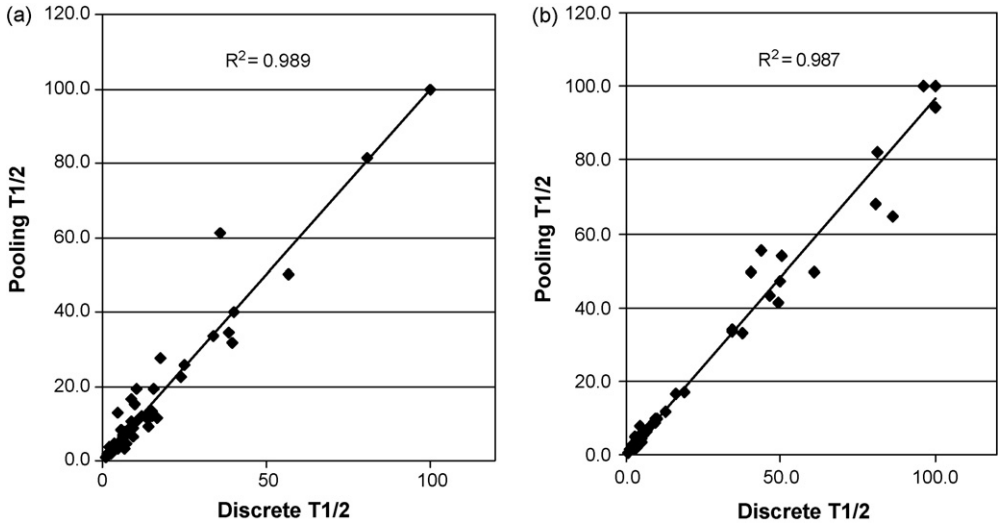


**Fig. 6.** Comparison of $T_{1/2}$ determined from pooled samples and discrete samples from a batch of 48 compounds assayed in both HLM and RLM (correlation coefficients 0.989 and 0.987, respectively). Two of the compounds were excluded from the plots due to low signal intensity.

excluded from the correlation plots due to low signal intensity. Metabolic half-life values ($T_{1/2}$) of pooled samples vs. discrete samples showed correlation coefficients of 0.989 and 0.987 for HLM and RLM, respectively.

### 3.4. Potential chromatographic and ionization interference due to metabolite formation

When grouping compounds based on molecular weight alone, some researchers [28] considered possible interference from potential common Phase I metabolites, such as +16, +32, −14, −28 for hydroxylated, dihydroxylated, and demethylated or deethylated metabolites, respectively. During the IMSAPS implementation, we chose not to consider the mass interference of potential metabolites since isobaric metabolites of a compound, if any, would likely have different fragment ion than other test compounds in the group, which would not cause interference in the MRM channel of the target compound. In the unlikely event both parent ion and fragment ion of a metabolite are the same as the target compound in the group, the UPLC system would very likely resolve the metabolite from the test compounds. Fig. 7 shows the chromatogram of a group of significantly metabolized test compounds overlaying with that of the hydroxylated metabolites, which showed no effect to the parent MRM chromatograms. We also intentionally constructed a set of commercially available compounds that could yield various combinations of peak coelution to further investigate such effect. We incubated phenacetin, coumarin, bupropion, and amodiaquin in both HLM and RLM and monitored additional MRM channels for their major metabolites of deethylation, hydroxylation, hydroxylation, and deethylation, respectively. Fig. 8 shows the coelution of (a) phenacetin and coumarin, (b) bupropion and hydroxylated metabolite of coumarin, and (c) amodiaquine and its own deethylated metabolite. Fig. 9 shows that such coelution did not affect $T_{1/2}$ determinations and results from pooling analysis correlated very well with those determined from discrete analysis (correlation coefficient = 0.982). This example demonstrated that even with intentionally forced coelution, there was no effect on the parent MRM channels that would cause incorrect $T_{1/2}$ measurement. In the actual sample analyses, such coelution will be much less likely to occur. It will be even

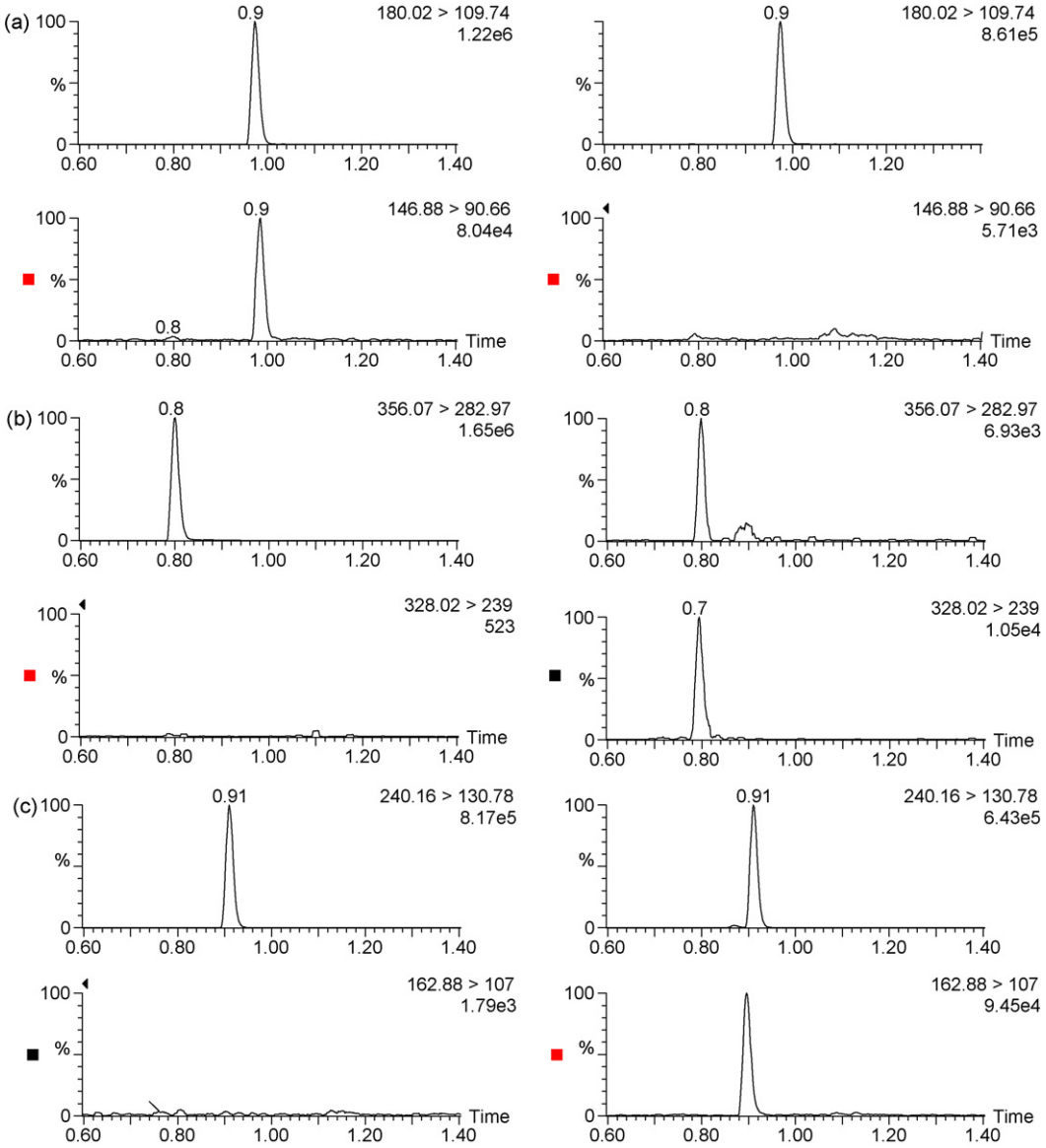


**Fig. 8.** A compound set intentionally constructed to yield various combinations of peak coelution in 0 min (left panels) and 15 min (right panels) HLM incubation samples: (a) phenacetin (upper chromatograms) and coumarin (lower chromatograms); (b) bupropion (upper chromatograms) and hydroxylated metabolite of coumarin (lower chromatograms); and (c) amodiaquine (upper chromatograms) and its own deethylated metabolite (lower chromatograms).

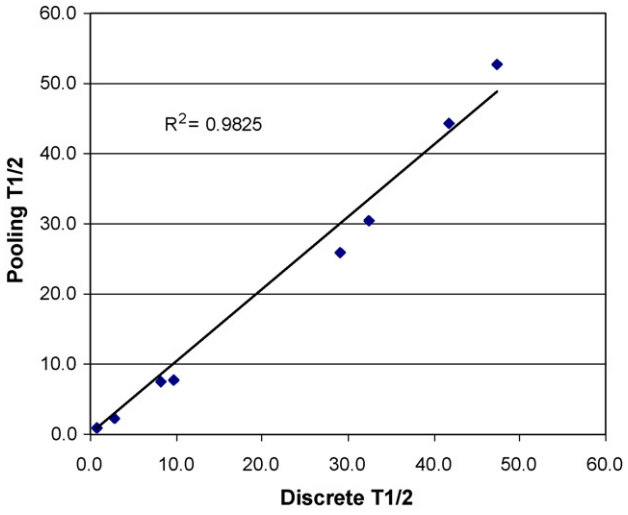


**Fig. 9.** Comparison of $T_{1/2}$ determined from pooled samples and discrete samples from in both HLM and RLM incubation for a compound set intentionally constructed to yield various combinations of peak coelution as shown in Fig. 8 (correlation coefficients 0.982).

more unlikely to cause interference in the MRM channels of parent compounds. Therefore, by focusing on grouping compounds with the most different $c\,\mathrm{Log}\,D_{3.0}$ values without considering unlikely interference from metabolite formation, IMSAPS will be more likely to achieve the best chromatographic peak resolution for the four pooled parent compounds.

### 3.5. IMSAPS program streamlines the microsomal assay process and enables fast data turnaround

Although sample pooling improves analysis throughput, the data is sometimes compromised for poorly ionizable compounds (due to dilution). For 4-compound pooling, each compound is effectively diluted 4-fold. Additionally, signal-to-noise (S/N) ratio is also reduced by having to use shorter MRM dwell time (0.02 s each channel over the 0.12 s cycle) to accommodate the narrow UPLC peak widths. As a result, compounds with poor ionization efficiency might not produce acceptable ion signals from the pooled sample analysis. If running pooled analyses overnight without the IMSAPS program, these low sensitivity compounds would have to be re-analyzed using discrete samples the next day after data review, which would significantly delay the data turnaround. In other instances, signal reduction of a test compound in the microsomal sample may occur as a result of matrix suppression, precipitation, or decomposition in the buffer or TCA solutions. Such situations would necessitate manual troubleshooting with compounds in various solutions, again slowing down the overall data turnaround. IMSAPS is able to perform re-analysis and troubleshooting during the overnight runs. It intelligently evaluates the data and submits discrete and/or QC samples for the compounds requiring re-analysis. All of the data would be available for final processing and reporting the next morning.

Multiple criteria were established for IMSAPS to evaluate the data quality of pooled samples and capture all information in a log file for future review. For example, *T*0 peak height is used for evaluating ion intensity. For those compounds that do not meet the pre-set ion intensity criterion, discrete samples and QC samples are injected. A set of IMSAPS criteria for assessing data quality is %Remaining variation of replicate samples at each time point or variation across time points. These parameters are used to identify potential robotic liquid handling errors that may occur during sample pooling. Reference compounds are included with all incubations to ensure that microsomal enzymes are active and experiments are performed properly. The entire dataset is rejected if the reference compounds give unexpected results. To ensure that the assay will not be disqualified erroneously due to analysis error of the reference compounds, IMSAPS re-injects discrete samples of the reference compounds if their turnover rates from the pooled samples are not as expected. All of these criteria are user-definable, as shown in Fig. 2c. Currently, three different solvent conditions are evaluated for the QC samples: (1) evaluation of compound in TCA/buffer (50:50) solution assesses compound instability in acid; (2) evaluation of compound in buffer solution assesses compound solubility under assay conditions; and (3) evaluation of compound in ACN/buffer (50:50) solution provides a reference signal intensity and determines if ACN may be used as an alternative quench solution if the compound is unstable in TCA.

Fig. 10 shows an example of the power of IMSAPS for automatically improving data quality. In this example large variations in the duplicate pooled samples were detected at both 5 and 30 min. IMSAPS automatically submitted the discrete samples from all time points for re-analysis that generated a much smaller variation. The automated re-analysis increased the confidence level of the analyst when reporting the results for this compound. Fig. 11 shows the data quality improvement achieved by IMSAPS for a poorly ionizable compound. In the pooled samples, IMSAPS automatically detected a low ion signal in the *T*0 sample and submitted the discrete samples for re-analysis. The discrete samples showed ~4-fold

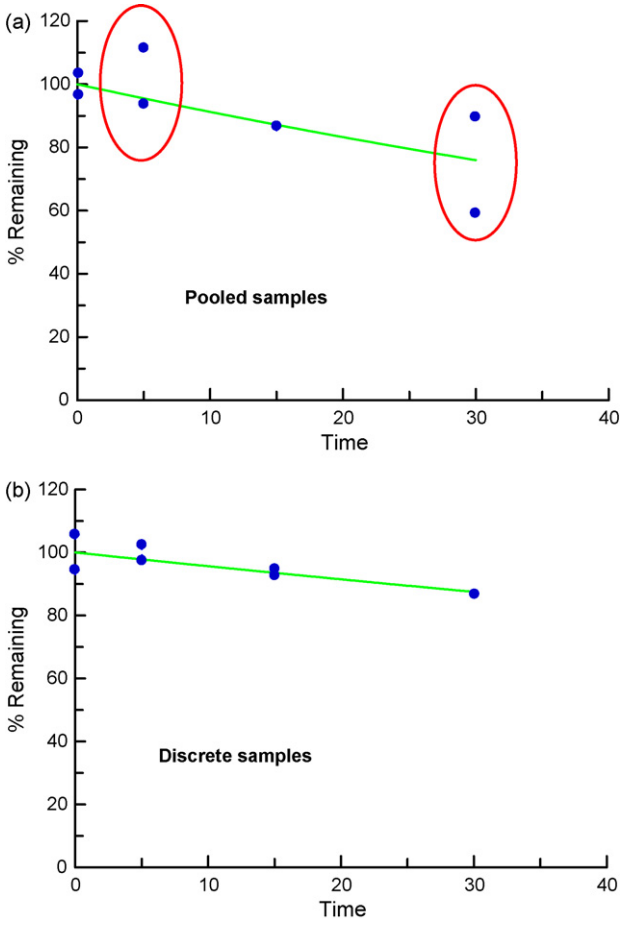


**Fig. 10.** (a) Large variations in the duplicate samples at both 5 and 30 min were detected by IMSAPS; (b) Much smaller variation resulted from the discrete samples re-injected automatically by IMSAPS.

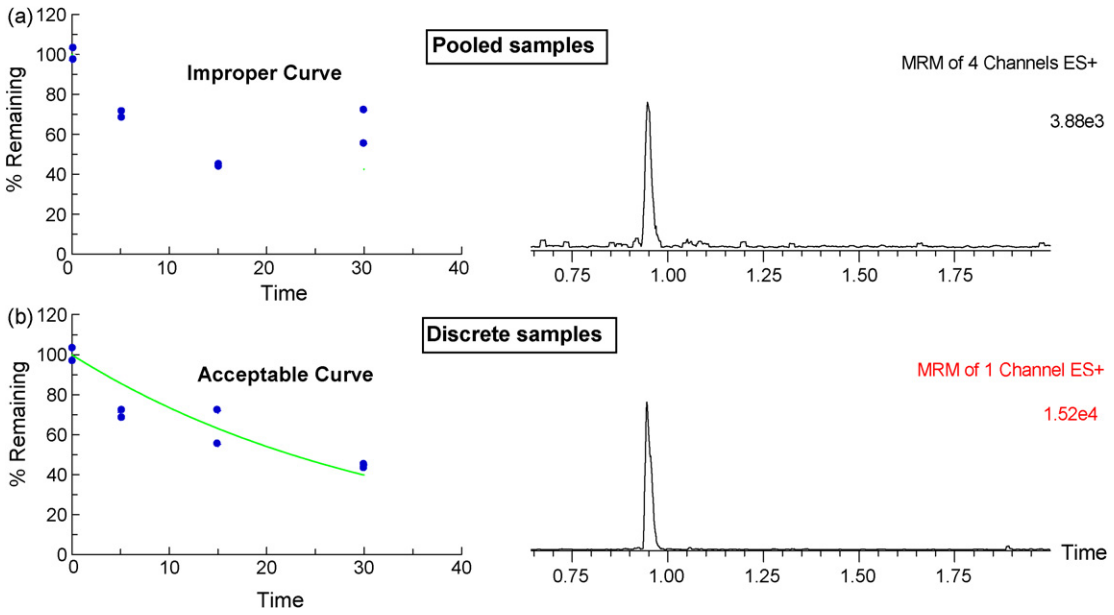


**Fig. 11.** (a) Low sensitivity in the pooled sample analysis resulted in poor curve fitting; (b) improved curve fitting and therefore, more accurate estimate of metabolic stability resulted from the discrete samples re-injected automatically by IMSAPS after detecting low *T*0 sensitivity.

increase of signal sensitivity, which resulted in better curve fitting and a better estimate of metabolic stability. Fig. 12 shows an example in which IMSAPS detected a very weak peak at 0.92 min in the *T*0 sample and submitted the discrete samples for re-analysis. The discrete samples revealed that the peak of interest was actually at 1.13 min and the 0.92 min peak was likely an impurity from another compound in the group. The discrete samples gave weak but acceptable peaks, which resulted in an acceptable curve. The data from automatically injected matrix QC samples provides a convenient and instant troubleshooting tool. In one occasion, strong signals were obtained for all QC solutions but no signal was detected in the microsomal samples for a test compound. This information led us to discover that the MP II had a liquid sensing problem with one of its tips when it was near its home position. This allowed us to quickly reject the data for this particular compound and include it for re-incubation in the next batch. In another occasion, the automated matrix QC showed that a compound generated strong ion signal in ACN/buffer solution, moderate ion signal in TCA/buffer solution, but very low ion signal in buffer solution (Fig. 13). The automated QC data revealed that the compound is poorly soluble in assay buffer.

After datasets of the pooling analysis and discrete analysis are acquired and processed, the final data from both pooling and discrete analysis have to be merged and arranged in the order that is consistent to the order of compounds in the incubation plates before they can be imported into ActivityBase for final curve fitting and validation. Manually merging these separate datasets into a final, complete dataset is time-consuming and prone to manual

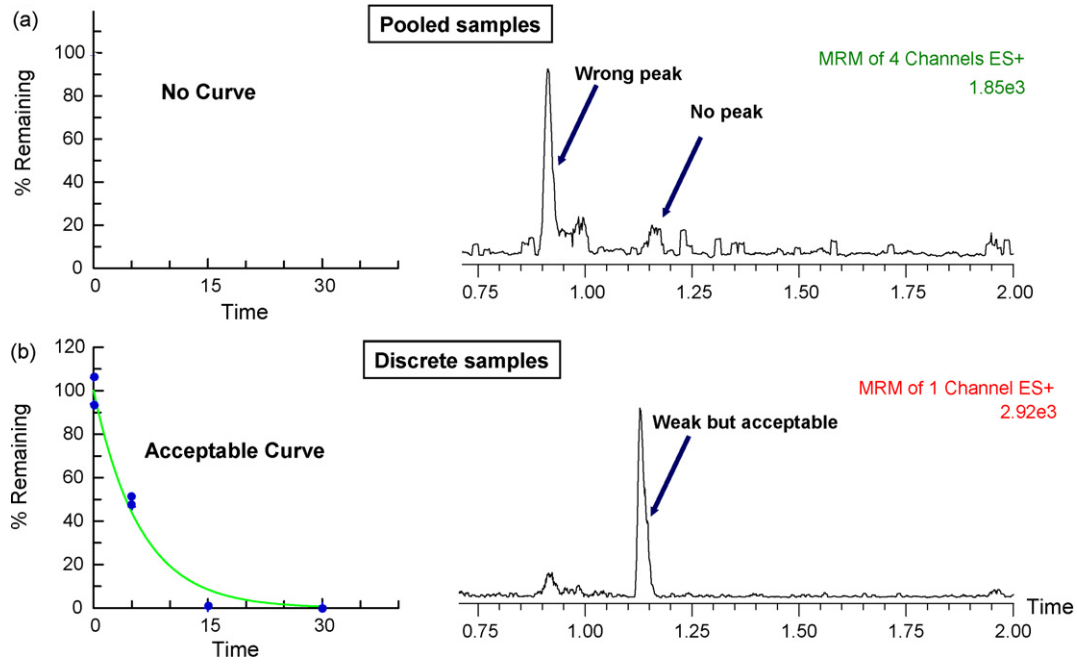


**Fig. 12.** (a) IMSAPS detected a very weak (and incorrect) peak in the *T*0 sample; (b) the discrete samples gave weak but detectable peaks, which resulted in improved curve fitting.

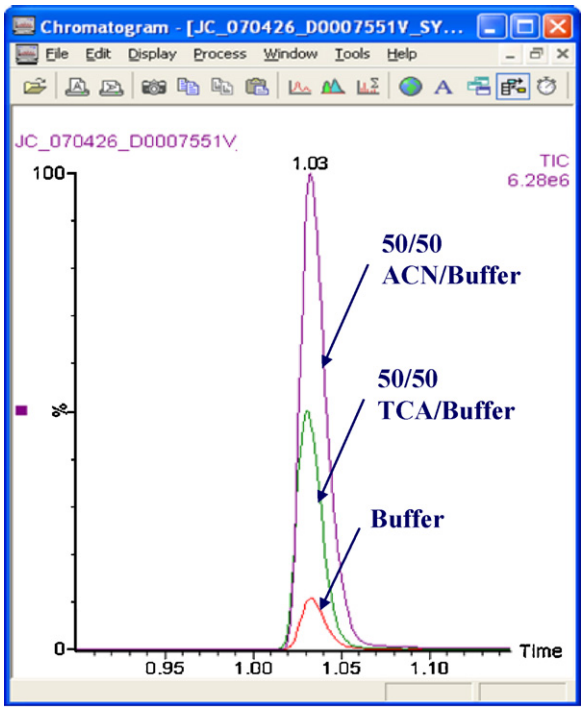


**Fig. 13.** Example of the automated matrix QC, which helps troubleshoot weak ion signals in microsomal samples and revealed a compound solubility issue under assay condition.

errors (taking the analyst hours). To address this, IMSAPS was modified to include a MergeResults module that automatically retrieves both datasets, finds the appropriate sections of the pooling dataset, and replaces that data with corresponding discrete data. The entire data merging process now takes 1 min or less.

A review of our drug discovery project samples assayed in the past few months showed that about 20% failed the acceptable criteria during the pooling analysis and would be automatically submitted for discrete analysis. The direct savings of the instrument time were about 55%. More importantly, the streamlined intelligent workflow improved efficiency and turnaround time, which is beyond the mere savings of instrument time and has made meaningful impact to the decision-making of the drug discovery projects.

## 4. Conclusions

An automated, streamlined workflow has been reported for high throughput metabolic stability screening. A custom Visual Basic program (IMSAPS) serves as the architect or central processing unit, overseeing the entire process. Key components of the system are a novel pooling strategy based on *c* Log *D*, incorporation of high resolution and high sensitivity UPLC/MS/MS, intelligent data processing and intelligent re-analysis. These features ensure fast turnaround while ensuring high data quality. This workflow has enabled metabolic stability assessment of all newly synthesized in-house compounds for lead optimization projects. Compounds for metabolic stability screening are plated at the same time as those for biological screening, enabling the chemists to simultaneously optimize the structure–activity relationship and structure–ADME relationship to accelerate their drug discovery efforts.

## Acknowledgments

The authors would like to thank Kheng B. Lim and Shaokun Pang for initial liquid handler programming for microsomal stability assay, James Thompson for conducting part of the sample incubations, Dan Hascall for writing the Compound Selection program module, and Jeff Geibel, Christine Clayton, and Matthew Pustelnik for providing informatics support.

## References

[1] I. Kola, J. Landis, Nat. Rev. Drug Discov. 3 (2004) 711.
[2] E.H. Kerns, L. Di, Drug-like Properties: Concepts, Structure Design and Methods from ADME to Toxicity Optimization, Academic Press, 2008.
[3] W.A. Korfmacher, Drug Discov. Today 10 (2005) 1357.
[4] D.B. Kassel, Curr. Opin. Chem. Biol. 8 (2004) 339.
[5] B.L. Ackermann, M.J. Berna, A.T. Murphy, in: M.S. Lee (Ed.), Integrated Strategies for Drug Discovery Using Mass Spectrometry, Wiley, 2008, p. 315.
[6] S.K. Balani, G.T. Miwa, L.S. Gan, J.T. Wu, F.W. Lee, Curr. Top. Med. Chem. 5 (2005) 1033.
[7] R.S. Obach, Drug Metab. Dispos. 27 (1999) 1350.
[8] L. Di, K.H. Edward, X.J. Ma, Y. Huang, G.T. Carter, Comb. Chem. High Through. Screen. 11 (2008) 469.
[9] J.H. Ansede, D.R. Thakker, J. Pharm. Sci. 93 (2004) 239.
[10] P.V. Kaplita, H. Hu, L. Liu, T.M. Farrell, H. Grbic, D.M. Spero, J. Assoc. Lab. Autom. 10 (2005) 140.
[11] M. Fonsi, M.V. Orsale, E. Monteagudo, J. Biomol. Screen. 13 (2008) 862.
[12] J. Wang, L. Urban, Drug Discov. World 5 (2004) 73.
[13] L. Di, E.H. Kerns, S.Q. Li, S.L. Petusky, Int. J. Pharm. 317 (2006) 54.
[14] K.M. Jenkins, R. Angeles, M.T. Quintos, R. Xu, D.B. Kassel, R.A. Rourick, J. Pharm. Biomed. Anal. 34 (2004) 989.
[15] E.H. Kerns, L. Di, J. Assoc. Lab. Autom. 10 (2005) 114.
[16] X.S. Tong, S. Xu, S. Zheng, J.V. Pivnichny, J. Martin, C. Dufresne, J. Chromatogr. B 833 (2006) 165.
[17] J. Ayrton, G.J. Dear, W.J. Leavens, D.N. Mallett, R.S. Plumb, J. Chromatogr. A 828 (1998) 199.
[18] R.J. Scott, J. Palmer, I.A. Lewis, S. Pleasance, Rapid Commun. Mass Spectrom. 13 (1999) 2305.
[19] L. Romanyshyn, P.R. Tiller, E.C.A. Hop, Rapid Commun. Mass Spectrom. 14 (2000) 1662.
[20] L. Romanyshyn, P.R. Tiller, R. Alvaro, A. Pereira, E.C.A. Hop, Rapid Commun. Mass Spectrom. 15 (2001) 313.
[21] R. Xu, C. Nemes, K.M. Jenkins, R.A. Rourick, D.B. Kassel, C.Z.C. Liu, J. Am. Soc. Mass Spectrom. 13 (2002) 155.
[22] G. Hopfgartner, E. Varesio, TrAC Trends Anal. Chem. 24 (2005) 583.
[23] E.R. Wickremsinhe, B.L. Ackermann, A.K. Chaudhary, Rapid Commun. Mass Spectrom. 19 (2005) 47.
[24] D.A. Volmer, L. Sleno, K. Bateman, C. Sturino, R. Oballa, T. Mauriala, J. Corr, Anal. Chem. 79 (2007) 9000.
[25] J. Gobey, M. Cole, J. Janiszewski, T. Covey, T. Chau, P. Kovarik, J. Corr, Anal. Chem. 77 (2005) 5643.
[26] R. Xu, T.M. Quintos, M. Manuel, J. Cramlett, K.B. Lim, D.B. Kassel, Proc. 56th ASMS Conf. Mass Spectrometry and Allied Topics, Denver, CO, 2008.
[27] S.X. Zhao, D. Forman, N. Wallace, B.J. Smith, D. Meyer, D. Kazolias, F. Gao, J. Soglia, M. Cole, D. Nettleton, J. Pharm. Sci. 94 (2005) 38.
[28] J.S. Halladay, S. Wong, S.M. Jaffer, A.K. Sinhababu, S.C. Khojasteh-Bakht, Drug Metab. Lett. 1 (2007) 67.
[29] R. Xu, K.B. Lim, M. Manuel, S. Pang, J. Cramlett, D. Hascall, D.B. Kassel, Proc. 54th ASMS Conf. Mass Spectrometry and Allied Topics, Seattle, WA, 2006.
[30] R. Xu, M. Manuel, J. Cramlett, K.B. Lim, S. Pang, D. Hascall, D.B. Kassel, Proc. 55th ASMS Conf. Mass Spectrometry and Allied Topics, Indianapolis, IN, 2007.
[31] K. Valko, in: R.T. Borchardt, E.H. Kerns, C.A. Lipinski, D.R. Thakker, B. Wang (Eds.), Pharmaceutical Profiling in Drug Discovery for Lead Selection, AAPS Press, 2008, p. 129.