Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Mass Spectrometry 238 (2004) 153–162
International Journal of Mass Spectrometry
www.elsevier.com/locate/ijms

High throughput HPLC/MS purification in support of drug discovery

W. Goetzinger, X. Zhang, G. Bi, M. Towle, D. Cherrak, J.N. Kyranos*

ArQule Inc., 19 Presidential Way, Woburn, MA 01801, USA

Received 2 February 2004; accepted 19 February 2004
Available online 27 September 2004

Abstract

High throughput purification techniques are an important part of drug discovery and provide high-quality compounds for biological screening. In this paper, we describe the purification platform developed by ArQule that is based on reverse phase high performance liquid chromatography (RP-HPLC) separation and mass directed fractionation. By strictly enforcing collection of only one fraction per sample, this purification paradigm has significantly enhanced the throughput and simplified the post-purification operation. Recovery studies have proven the reliability of this process and development of fast chromatographic separations provide enhanced throughput without additional capital investment. This approach has been used successfully to purify over half a million compounds in the past 2 years and resulted in post-purification average purity of over 97% when assessed by HPLC and low-wavelength UV.

1. Introduction

In the early 1990s, split-and-combine combinatorial chemistry emerged as an innovative synthesis approach and promised to revolutionize drug discovery by delivering mixtures of large numbers of small molecules [1,2]. The expectation was that the large number of analogs synthesized and screened against emerging novel targets provided by the genomic revolution would interrogate the biologically active chemical space much more effectively, and thus, quickly lead to new drug candidates. In general, however, the rapid generation of mixtures was not able to deliver on this promise. While many new chemical entities were synthesized, the fact that they were exposed to the targets as mixtures has led to difficulties in de-convoluting active structures and a significant need for resources to follow up on many false actives.

High throughput parallel synthesis was developed around the same time and promised to improve the productivity of drug discovery based on a different approach [3]. In parallel synthesis, individual molecules are synthesized in a spatially addressable format and these compounds can be screened similar to individual compounds made by medicinal chemists. Relying heavily on automation for leverage, this approach can generate many more compounds compared to the traditional "one at a time" approach practiced by medicinal chemists.

Although only one molecule is intended to be synthesized per reaction vessel, almost inevitably, there are impurities present at the end of the reaction, especially when multiple synthetic steps are employed to generate more complex drug- or natural product-like structures. To assess the biological results and develop meaningful structural activity relationships (SAR) that can be used to guide lead optimization, compound characterization relative to purity and quantity is imperative. There have been many different approaches in assessing the purity of parallel synthesis products. While nuclear magnetic resonance (NMR) remains the "gold standard" for traditional medicinal chemists, currently the cost and complexity of automated NMR data analysis and interpretation prevent its application as a high throughput process. Library compounds made by parallel synthesis are predominantly analyzed using high performance liquid chromatography (HPLC) with UV and/or evaporative light scattering detectors (ESLD) for purity with on-line mass spectrometry (MS) detection for structure confirmation [4–7]. This approach, however, does

* Corresponding author. Tel.: +1 781 994 0429.
E-mail address: jkyranos@arqule.com (J.N. Kyranos).

1387-3806/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2004.02.024

have limitations since none of these detectors provide a truly "universal" response for all reagents, intermediates and final products. The inherent bias of each detector such as extinction coefficients for UV, vapor pressure for ELSD or ionization efficiency for MS for each sample constituent renders a very complex analytical problem.

Many studies have attempted to address the issue of compound characterization and parallel synthesis purity evaluation [7–9]. However, to avoid disappointment resulting from inconsistent screening results due to incorrect impurity evaluation, the reaction products must be purified. Once the reaction mixtures are purified, the relative difference in compound purity assessment using various detectors is significantly reduced [10]. Screening purified compounds provides higher confidence in the initial hits with greater expectations of positive confirmation of activity. There is also significant value to drug discovery that comes from inactive compounds. The data helps scientists understand SAR more comprehensively and accelerate the lead optimization phase of drug discovery.

In an attempt to implement parallel high throughput purification within the parallel synthesis paradigm, a variety of approaches have been developed. Liquid–liquid extraction is routinely used as a work-up procedure in parallel synthesis [11] and solid phase extraction has also been adopted in a number of labs [12]. The application of scavenger resins, designed to remove impurities with specific active functional groups, is another technique that has become very popular [13]. Although all of these procedures can be used in a high throughput environment, they all suffer from a limitation in regard to their universal applicability to a wide range of compounds, even within a library, as well as their limited selectivity. To be effective, these techniques usually rely on significant differences between the target compound and potential impurities, which is not always the case in a synthesis program.

The most powerful purification technique that is suitable for parallel synthesis products is HPLC. In particular, reverse phase high performance liquid chromatography (RP-HPLC) is well suited to deal with the quantities and polarity of compounds developed for pharmaceutical interest. In the past, RP-HPLC has been widely used by medicinal chemists for the purification of natural products and individual samples from synthetic reaction mixtures [14–16], however, the field of parallel synthesis has challenged the conventional paradigms of preparative chromatography. Parallel, automated synthesis requires an equivalent high throughput approach to purification; otherwise, the typical traditional preparative chromatographic separation approach can quickly turn into the main bottleneck of the process.

The main advantages of preparative RP-HPLC are twofold. First, the selectivity of HPLC allows for better separation of closely related compounds. Second, the on-line detection of the separated components with real-time feedback to sample collectors allow for individual sample fraction collections without characterizing each separation a priori. Many groups have been using analog detectors, such as UV or ELSD, to fractionate reaction mixtures [6,17,18], followed by MS analysis, usually by flow-injection, as the most common approach to identify the fractions. This process is reliable and automated approaches to handle fractions have been described [19], however, the issue remains that the number of fractions expected for each sample cannot be predicted. Sophisticated tracking software to support this process in a high throughput environment is required. Consequently, the post-purification procedures to identify the fractions of interest and reformat these samples into a usable plate format ready for screening can be quite cumbersome. While this approach is reasonable for a typical workload of up to a few hundred compounds per week, there are serious operational issues associated with scaling this process up to thousands of samples per week.

Another obstacle that has limited the development of high throughput preparative HPLC has been the significant capital investment in instrumentation. In an attempt to limit the capital investment and increase speed, productivity and cost effectiveness, there have been a number of attempts to implement parallel approaches to preparative chromatographic separation, as well as to the analysis of collected fractions [20–22]. All of these parallel processes contribute to enhanced throughput, but they are operationally more complicated and do not address the principal problem of having to manage a different number of fractions for each sample. If the number of fractions per sample for a given library is large, then the limited deck space of the fraction collection module becomes the determinant factor of when the process has to be stopped and additional fraction tubes have to be added. This limitation severely curtails the potential unattended overnight utilization of the equipment.

Since preparative elution time can be somewhat predicted using analytical data, there have been attempts to run each individual sample on an analytical scale first to gain a better understanding of what to expect in the preparative run, and thus, adjust the collection parameters for each sample [23]. However, there are significant opportunities for improvement of the process. This approach requires analytical characterization of each sample before preparative HPLC and makes the process more time-consuming and adds complexity to data management.

While, traditionally, MS has gained widespread attention for very high sensitivity in support of bioanalytical work, it has also become an indispensable tool for high throughput parallel synthesis. Analytical groups widely rely on mass spectrometric data to confirm new product synthesis with little to no previous information about the new compounds. In recent years, mass directed fraction collection has become more popular for preparative purification of synthetic compounds [9,24] due to the high selectivity based on a specific mass to charge ratio (m/z). Generic detectors, such as UV and ELSD, respond to a wide range of samples and potential contaminants, but mass spectrometers can be set to select a specific m/z for each sample. This selectivity can be used to significantly reduce the number of fractions that are collected

from each separation. The approach is especially attractive for high throughput applications where thousands of samples per week require purification. This mass directed fractionation approach has also been implemented in a parallel fashion using a multi-column HPLC system coupled to a MUX electrospray interface [25,26].

MS is perceived as a sophisticated detector, which may not be as rugged as a UV detector. This perception and the need to inject the entire synthesis sample for purification, which may be completely lost to waste if the MS fails, has limited the development of mass directed fractionation as a high throughput technique. ArQule's vision to transform the traditional medicinal chemistry paradigm to a high throughput equivalent using the automation platform as the lever, required the development of an effective high throughput purification process. We have applied a mass directed fractionation approach for a number of years now and feel the technology has matured significantly and is as reliable and more cost-effective than any other technique [15,16,27]. In this paper, we describe our high throughput purification process that is based on mass directed fractionation. In addition to the high selectivity of the mass spectrometer, we employ features of the fraction collection software to strictly enforce a one-to-one mapping of fractions to samples (i.e., for every sample injection we collect only one fraction). We present a recovery study validation that supports the conclusion that the mass directed fractionation significantly simplifies the process and increases the throughput. Furthermore, our investigation on high-flow chromatography demonstrates potential increased purification capacity using the present capital investment.

2. Experimental

2.1. Instrumentation

Fig. 1 shows the schematic diagram of the HPLC/MS purification system. The system consists of a Shimadzu (Columbia, MD) preparative HPLC system, a Waters ZQ (Milford, MA) single quadrupole mass spectrometer, and two Gilson 215 liquid handlers (Madison, WI) for injection and collection. The Shimadzu HPLC system includes two LC-8A solvent delivery pumps, one LC-10Ai inert HPLC pump for solvent modifiers, one LC-10ADvp pump as an online dilution pump, one LC-10ADvp pump as a makeup flow for the MS, two SPD-10AVvp UV detector, and a SCL-10Avp system controller to control the HPLC gradient. Sample injection and collection are controlled by MassLynx^{TM} software. HPLC grade acetonitrile by J.T. Baker (Phillipsburg, NJ) is delivered by a custom built delivery system. Water is purified and delivered by Millipore's Milli-Q system (Bedford, MA).

2.2. HPLC and MS conditions

All preparative HPLC separations at low pH were carried out using 20 mm × 50 mm Maccel 120-10-C18 SHC18 column (The Nest Group Inc., Southborough, MA). Separations at high pH were conducted using 19 mm × 50 mm XTerra® Prep MS C18 column (Waters, Milford, MA). The standard gradient was from 5 to 95% acetonitrile in water in 3.5 min with a cycle time of 5 min. The column flow rate was 40 mL/min with 4 mL/min of 100% acetonitrile from the online dilution pump. The injection volume ranged from 400 to 2000 μL. A capillary splitter was used to divert flow after the UV detector to the mass spectrometer, which was diluted by 2 mL/min of make-up flow before entering the MS. The make-up flow consists of 90% methanol and 10% of water with 0.05% of formic acid. A second UV detector was used after the fraction collector to monitor the collection. All samples were purified by mass directed fraction collection.

2.3. Weighing and evaporation

The fraction collection tubes were tared before purification. After purification, fractions were dried and the collection tubes re-weighed. Weighing was performed using balances (d = 0.0001 g, Mettler Toledo, Columbus, OH) controlled by a Bohdan Automation weigher (Mundelein, IL). The purification blocks were bar-coded and each block contained 24 collection vials. Sample location and quantity information were stored into the company database through bar code scanning. An analytical balance (PG503-S Delta Range®, d = 0.001 g) (Mettler Toledo Inc., Columbus, OH) was used to make stock solutions of commercially available standard compounds. Solvent evaporations after purification were carried out with Mega 980 Evaporators from Genevac Inc. (Valley Cottage, NY).

2.4. Chemicals

HPLC grade acetonitrile and trifluoroacetic acid (TFA) were from J.T. Baker. Formic acid (88% A.C.S. reagent), ammonium hydroxide (A.C.S. reagent) and HPLC standards: 2-hydroxydibenzofuran (98%), 2-acetamidophenol (97%) and 3-(4-*tert*-butylphenoxy)benzaldehyde (98%) were purchased from Aldrich. HPLC-grade methylsulfoxide (DMSO) and methanol were purchased from EM Science (Gibbstown, NJ). Flavone, atenolol, chlorthalidone, cortisone, and metergoline (98%) were purchased from Sigma Chemical Co. (St. Louis, MO). Compounds used for the in-house recovery study were made using our automated high throughput synthesis platform and processes.

2.5. Recovery test procedure

Column mass loading studies were conducted using commercially available compounds as test standards. Stock solutions of atenolol, chlorthalidone, cortisone, and metergoline were purified at concentrations of 10, 20, 40, 60, 80, and 100 mM. Injection volume was 1.5 mL. Recoveries were calculated based on the dry weights of the purified compounds.

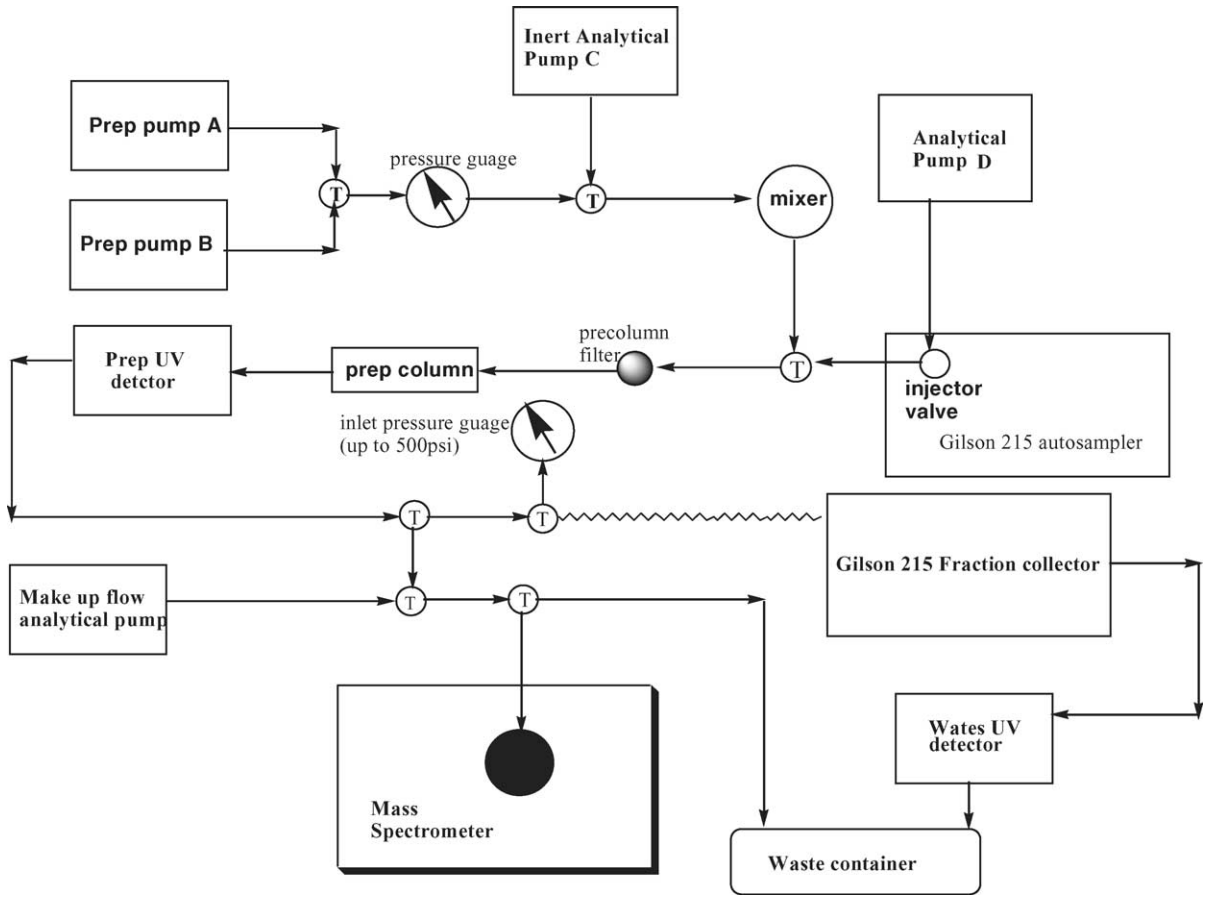

Fig. 1. Preparative HPLC/MS purification system schematic.

To extend the test compound validation to a more relevant sample set, a total of 152 ArQule compounds that had been purified under our routine process were purified a second time using the original purification conditions. To simplify the situation, only the compounds that were 100% pure by both UV and ELSD after the first purification were re-purified. Recoveries were calculated by weight data obtained after the first purification compared to the second purification.

2.6. High flow rate preparative HPLC/MS purification

Injections of a test mixture were performed at different flow rates (22, 44, 88, and 110 mL/min) using the same gradient volume. Appropriate plumbing changes were made to accommodate inlet pressure changes.

3. Results and discussion

3.1. High throughput purification

ArQule has developed a high throughput parallel synthesis platform used to increase the productivity of early drug discovery. Our approach consists of a rigorous synthesis and analytical development process in which multiple factors are considered prior to producing large or small collections of drug-like compounds. An integral and key component of the overall process is high throughput purification that ensures the highest quality compounds that go into biological screening.

The high throughput purification platform is based on the concept of mass directed fractionation and uses the signal output from a mass spectrometer to trigger collection. In addition, the system set-up includes a UV detector immediately after the HPLC column (referred to as prep UV) and a second UV detector after the fraction collector to monitor the eluent that is sent to waste (referred to as waste UV). The prep UV and the waste UV traces allow the user a direct comparison to monitor the quality of purification collection. In an ideal case, the only fraction collected is from the desired product peak, which is reflected in the waste UV trace as a missing peak relative to the prep UV trace.

Each collection tube can hold 16 mL of liquid and can accommodate collection of a peak that has a width at the base less than 21 s at a flow rate of 44 mL/min. Although the chromatography is optimized to allow collection of the large majority of the fractions, if the product peak is wider than the collection window, then part of the product peak will be seen in the waste UV trace to indicate an incomplete collection. The waste UV, thus, provides a valuable real-time

diagnosis of the purification process. Any issues in purification methods, such as inappropriate system parameter settings or poor collections caused by distorted peak shapes, will be reflected in the waste UV trace, and therefore, noticed immediately. This approach is a significant advantage over the conventional purification approach where multiple fractions are collected by the signals from analog detectors and identified off-line by MS flow-injection. Potential issues in the conventional purification process cannot be identified until the fractions are analyzed after the batch run has been completed. This sequential batch process not only delays the purification method optimization but in cases where the system is not appropriately adjusted, there is significant risk of loosing all the samples in a given batch.

In mass directed fractionation approaches, once the MS intensity of the extracted ion chromatogram from the desired product exceeds a pre-set threshold, the valve to the fraction collector opens, allowing flow to go into the collection tube. The valve closes when the MS intensity of the monitored ion drops below the threshold or the collection tube is completely filled and the flow is again diverted to waste passing first through the waste UV detector. Due to the low flow reaching the MS after the splitter relative to the high flow directed towards the collector or waste, a delay tube is placed between the prep UV and the fraction collector. The delay tube is designed to keep the eluent from reaching the collector valve too quickly, thus, allowing sufficient time for the flow directed to the MS to be analyzed and a determination to collect the sample or direct it to waste to be made.

The time difference for a peak to travel from the splitter to the MS and from the splitter to the fraction collector is further adjusted by a system dependent parameter within the MassLynx^{TM} software identified as “Split/Collector Delay” (referred as delay) [28]. A good collection is very dependent on this delay and requires routine assessment and correction. At ArQule, the delay for each purification system is quantitatively evaluated monthly at a minimum and more often if there is a potential question of recovery. Flavone, a commercially available compound, is used as a standard to assess the delay for each purification system. Flavone is a neutral compound and appears as a symmetric chromatographic peak under our standard conditions. With the typical mass/volume injection that we routinely use (35 μmol in 400 μL of DMSO), the peak width at the base is significantly narrower than the collection window. Therefore, the flavone peak disappears almost entirely (over 95%) in the waste UV trace when the correct delay is used. The presence of part of the flavone peak in the waste UV trace, immediately indicates there has been a change in the system, which requires correction. Fig. 2a represents a chromatogram obtained with a correct delay. The majority of the peak of interest is absent from the waste UV trace and indicates that collection was appropriately performed. In contrast, Fig. 2b represents a sample where the delay was inappropriately set, and as a result, a significant amount of the peak of interest has not been collected and appears in the waste UV trace. The collections are quantified by UV using a standard calibration curve as well as by weight analysis. A 95% or better recovery of flavone ensures that the system is performing to expectation.

To match the pace of sequential purification to parallel synthesis, the purification cycle time has to be minimized and post-purification processes need to be standardized. This requirement has been realized primarily through our unique approach of mass directed one-to-one fraction collection. We create a mirror image of the injection block on the collection block where the product collected from any given well from the injection block goes to the same well location in the collection block (Fig. 3). For each specific sample, only one collection is made. This collection is achieved by adjusting specific collection parameters in the Masslynx^{TM} 4.0 software. Post purification processes, including solvent evaporation, weighing and reconstitution, are substantially simplified by this one-to-one approach, as the purified sample layout remains the same as the layout of the injection block, which is tracked through proprietary in-house software.

Purification samples are submitted in custom 24-well blocks in a spatially addressable format. The pertinent sample related information is downloaded into custom software that generates the sample list for purification, and includes the expected molecular ion to trigger collection based on the intended product for each well. Since the one-to-one fraction collection is strictly enforced, both the injection deck and the collection deck can hold 17 24-well blocks that correspond to a purification batch of 408 compounds. When the batch is initiated, the analyst monitors the data quality for the first 24 samples, as well as the operation of the injection and collection systems before allowing the rest of the batch to run unattended overnight. In cases where issues are seen in the first 24 samples, purification is terminated and resumed after the problem is resolved. In the morning, the fractions are removed for evaporation and weighing and the system is available the rest of the day for attended method development or smaller purification batches. The overall process for a batch of 408 samples takes approximately 4 days from receipt of synthetic material to return of dried purified compounds ready for screening.

During the library development stages, all of the compounds generated as part of the synthetic route scouting and reagent qualification process, go through pre-purification and post-purification characterization (pre-QC and post-QC, respectively) to provide purification staff a reference as to the chromatographic and MS behaviors of these compounds. At these stages synthetic strategies and purification methods are under development and optimization. The development libraries contain small numbers of compounds that are representative of the larger libraries to be produced in future high throughput runs. For high throughput library synthesis 25% of the compounds go through pre-QC analysis. This pre-QC step serves to identify any major synthetic failures or human errors such as block switching so that mistakes can be corrected prior to purification and prevent unnecessary losses. To ensure the quality of the compounds entering biological

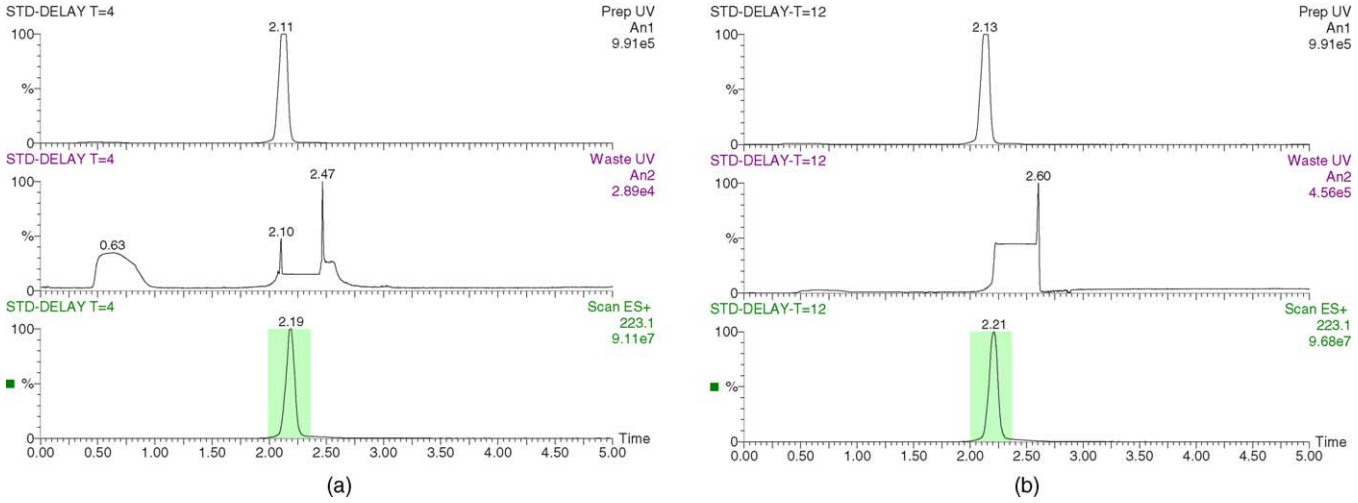

Fig. 2. Collection of flavone. From top to bottom: prep UV trace, waste UV trace and single ion chromatogram. (a) Collection performed with a correct delay (delay = 4 s). (b) Collection performed with an incorrect delay (delay = 12 s).

screening, post-QC is conducted on every compound to evaluate purity. Fig. 4 shows a typical example of pre-QC and post-QC analyses where impurities present in the pre-QC were quite effectively removed during the purification process. In the post-QC, only the desired product remains. Fig. 4 is also a good example of the difficulty in assigning purity to synthetic samples that have not been purified. The ELSD purity is significantly different than the UV purity of the pre-QC results. After purification, the differences between the two detectors are inconsequential.

3.2. Recovery study on the high throughput purification process

A major concern in purification has always been compound recovery. Although prep HPLC/MS purification has proven to be versatile and automation friendly, it too suffers inevitable compound losses. It is important to understand the recovery after purification and the consistency of the recovery on libraries that have different properties. The recovery of the preparative HPLC/MS approach is primarily influenced by two factors: physicochemical properties of the sample and column mass/volume loading. The first factor differentiates the compound's chromatographic behavior and the second factor causes potential column overloading (too much sample) or detection failure (too little sample). To validate the purification process for library compounds, we have chosen a group of commercially available compounds (referred to as test compounds) that represent a range of hydrophobicities and studied their recoveries at different mass loadings. Generic purification methods were used on these compounds depending on their acidity/basicity, as routinely done on in-house library method development, and the same purification conditions were used on any given test compound regardless of the mass loading. The purification was conducted unattended in overnight runs to mimic the high throughput library purification process. Recovery results on four test compounds are shown in Fig. 5.

Within experimental error, recoveries were consistently above 80% for all the test compounds regardless of their hydrophobicity or mass loading. These compounds were purified on different systems that were operated by various analysts, indicating that the recovery result is system and operator independent. These results support mass directed

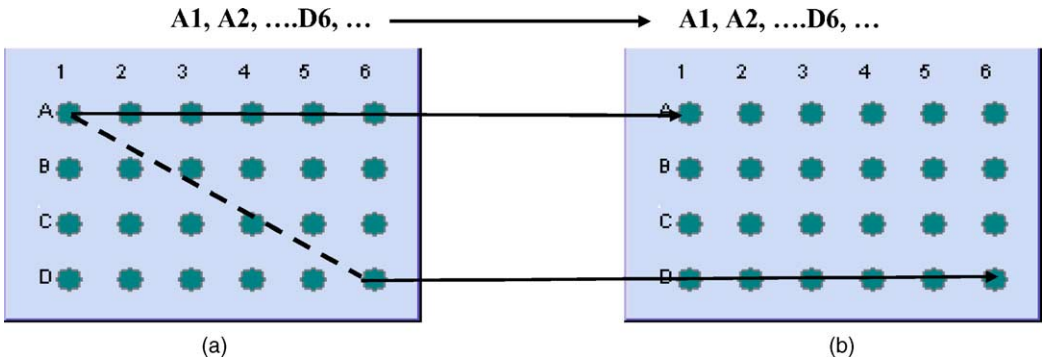

Fig. 3. Mirror image representation of the (a) injection and (b) collection blocks.

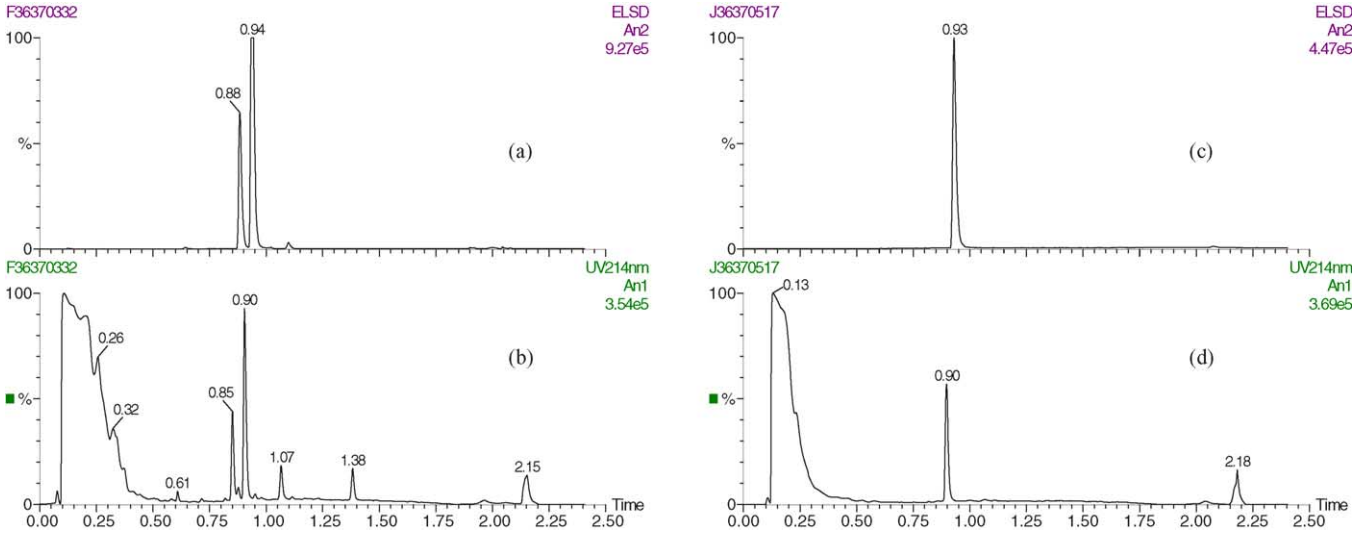

Fig. 4. Pre-QC and post-QC of a representative library compound. (a) ELSD trace of the sample before purification. (b) UV trace of the sample before purification. (c) ELSD trace of the sample after purification. (d) UV trace of the sample after purification.

fraction collection as an approach rugged enough to handle samples that come in a wide range of physicochemical properties and mass loadings.

There are two key factors in obtaining a consistently high recovery. The first is to maintain a good signal-to-noise ratio (S/N). This requires samples to have a reasonable response to the detector that triggers the fraction collection. In mass directed fractionation, low recovery can happen on samples that have low ionization efficiency. In these cases, different ionization modes (positive versus negative, electrospray versus atmospheric pressure chemical ionization (APCI)) or modifiers may be needed to enhance the signal to noise ratio. In analog detector-triggered fractionation, the same situation can happen when UV-transparent compounds are being collected based on UV signal or volatile compounds are being collected by ELSD signal. In our routine high throughput purification, most compound libraries can be effectively ionized (S/N $\gg$ 10) by positive electrospray ionization. The second key factor to ensure a good recovery is to obtain symmetric chromatographic peak shape of which the peak width at the base is narrower than the collection window. The results of this study indicate a steady expansion of the peak width as the mass loading increased, however, even at the 150 μmol mass loading, peaks can be fully collected without compromising recovery. In library purification, distorted peak shapes are sometimes the cause for low purification recovery and are therefore a major focus for method development. The standard approach to a solution for this type of issue is to optimize the separation methods by alternating gradient, modifiers, columns and mobile phase. It should be pointed out that the second factor is rather a universal concern regardless of the fraction-triggering detector. If band broadening becomes substantial, then the recovery or the purity will most likely be compromised.

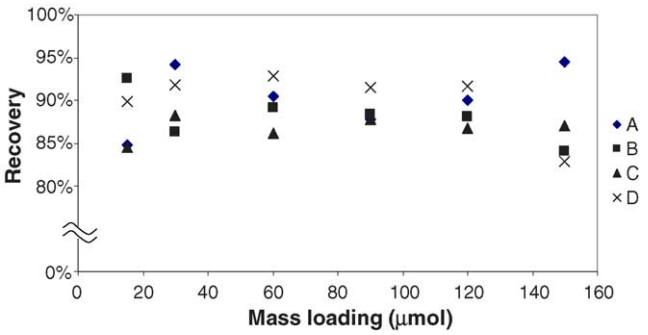

Fig. 5. Recovery data for compounds with various hydrophobicities at different mass loadings. Compound (A) chlorothalidone ($c\log P = -0.74$), (B) atenolol ($c\log P = 0.10$), (C) cortisone ($c\log P = 1.44$), and (D) metergoline ($c\log P = 0.72$).

Although commercially available test compounds are an important part of validating the high throughput purification system, there is substantial experimental information on their physicochemical properties, which is typically unavailable for novel library compounds generated in-house. Library compounds can be viewed as members of the same chemotype, with different substructures attached to the core. The resulting library has a diverse set of physicochemical properties in terms of basicity, hydrophobicity and solubility. Therefore, to fully validate our purification process, it is necessary to investigate the recovery based on compound libraries.

Since there are generally multiple impurities present in the compounds prior to purification, the weights of the intended compounds in the mixtures are difficult to determine. This challenge becomes more and more negligible as the compound purity increases. We have performed a "re-purification" test where a compound library was purified through the high throughput purification process and the

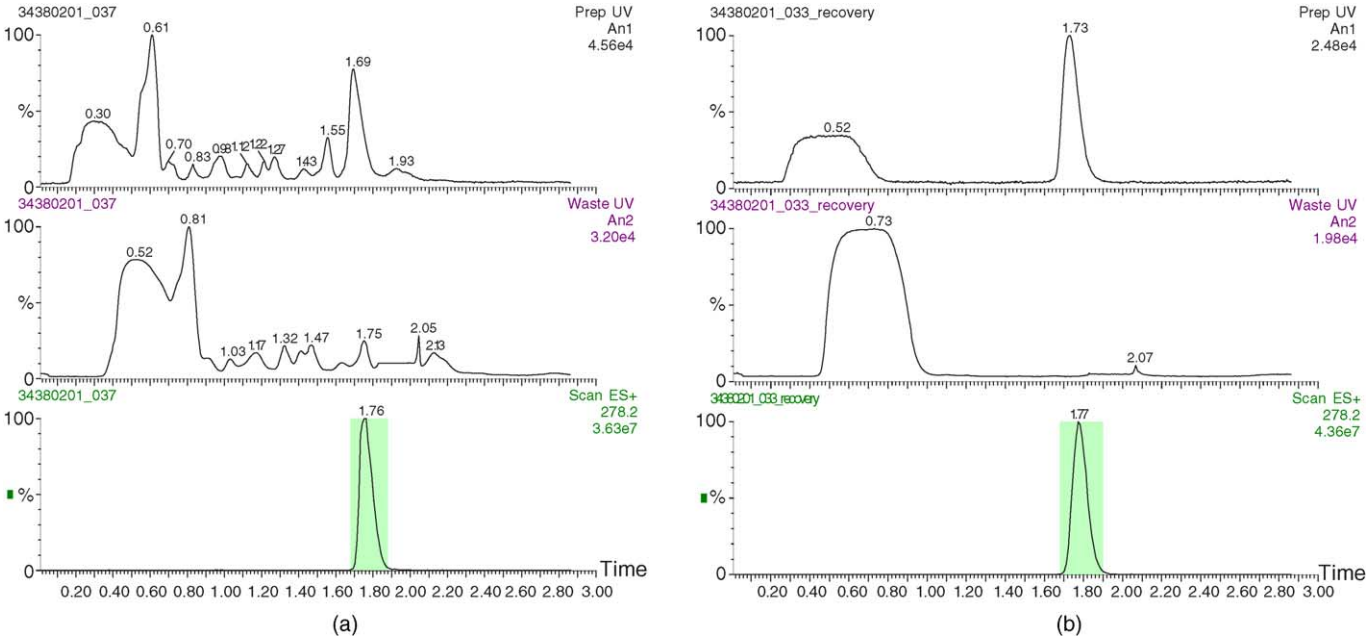

Fig. 6. Chromatograms of an ArQule compound in the first round of purification and in re-purification. From top to bottom: prep UV trace, waste UV trace, and single ion chromatogram. (a) Chromatogram of the purification of crude synthesis product. (b) Chromatogram of the re-purification on the purified fraction to evaluate system recovery.

weight of each purified compound is obtained after evaporation. The purified library was then purified again using the same method and another set of weights was generated for each collected fraction. The re-purification weights for each compound were compared to the weights obtained from the first purification to obtain the recovery.

A total of 126 compounds from the same library were used in this study. Fig. 6 shows an example of the chromatograms obtained from one of the library members after the first and second purifications, respectively. A histogram of the recovery using this re-purification approach for the library subset of ArQule compounds is shown in Fig. 7. Recoveries ranged from 75 to 102% with 93% of the samples above 75%.

3.3. High throughput purification applications

Our lab has eighteen HPLC/MS systems, which are operating continuously, and within the last 2 years more than 500,000 compounds were purified using mass directed fractionation. The results shown in Fig. 8 for high throughput synthesis in 2003 indicate the consistent quality of the process output. Over 80% of the samples have UV 214 nm purity better than 80 and 72% of the samples are 100% pure. The average synthesis purity of compounds that go into purification is 70% by ELSD and 60% by UV 214 nm, while the output average after purification is 87% by ELSD and 86% by UV 214 nm. After culling out the synthesis and process failures (0% purity) and applying minimum purity criteria of 80% by both detectors, the average purity values increase to 99% by ELSD and 97% by UV 214 nm. Fig. 9 illustrates the pre-purification purity evaluation of the 25% QC subset of the total purifications in 2003 compared to the post-purified data of the same subset. The data shows a dramatic increase in the number of 100% pure compounds after purification with an overall process loss attributed to the high throughput purification process of approximately 3%. This value is de-

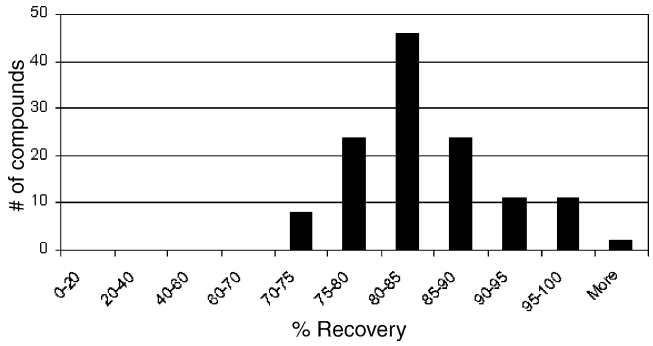

Fig. 7. Recovery results of the re-purified ArQule compounds.

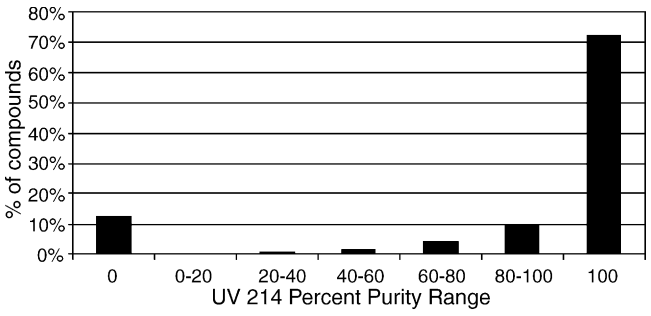

Fig. 8. Distribution of compound purity after high throughput HPLC/MS purification of over 180,000 compounds.

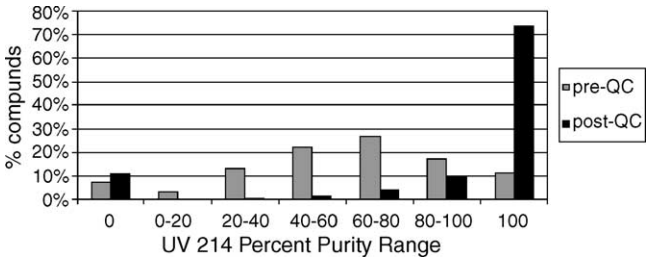

Fig. 9. Distribution of pre-purification and post-purification purities for 25% of QC sampling representing 40,000 compounds.

termined by the percent difference of the pre-QC failures that represent synthesis problems and the post-QC failures, which include both synthesis and purification process failures.

3.4. Future development on high throughput purification

Although our approach for high throughput purification using mass directed fractionation is an efficient and cost-effective process, it still requires a substantial capital investment in preparative HPLC systems and mass spectrometers. As the need for high throughput purification continues to grow, we investigated opportunities to modify the process to purify more samples per system without having to increase the equipment costs linearly. Our approach focused on improving the throughput by developing faster HPLC methods.

By operating the HPLC systems at higher linear velocities of mobile phase, similar separations can be accomplished in less time. The typical flow rate used for high throughput library purification is 44 mL/min. The limiting factor to increasing flow rate is a significant increase of the inlet pressure. By implementing minor plumbing modifications to the HPLC system, a lower inlet pressure was achieved, which allowed flow rates greater than 100 mL/min to be used. UV chromatograms of the HPLC standards obtained with a C18 20 mm × 50 mm column at different flow rates are shown in Fig. 10. The separation at 22 mL/min was achieved in 10 min,

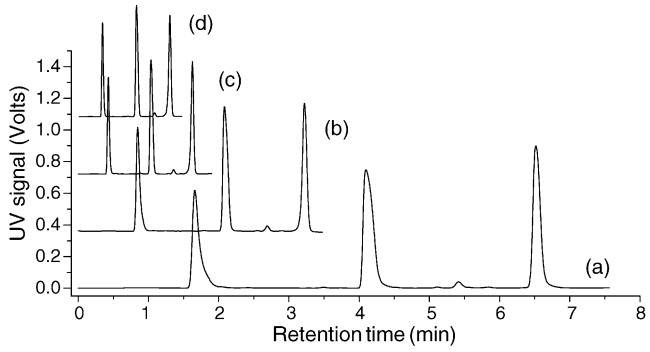

Fig. 10. Injection of a test mixture of acetamidophenol, 2-hydroxydibenzofurane and *t*-butylphenoxybenzaldehyde (20 mg/mL of each in DMSO) at different flow rate: (a) 22 mL/min (peak capacity = 33), and (b) 44 mL/min (peak capacity = 36), (c) 88 mL/min (peak capacity = 32), (d) 110 mL/min (peak capacity = 29). RP C18 column 20 mm × 50 mm. Sample injection volume is 800 μL.

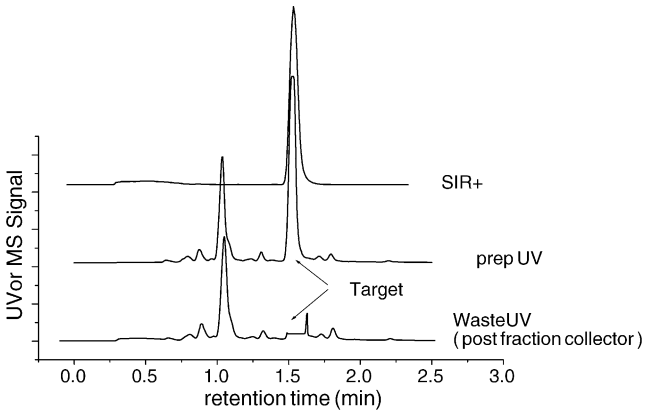

Fig. 11. Purification of crude library compound at 110 mL/min. Sample size (200 μmol) in 1.6 mL.

while a similar separation was obtained in less than 2 min at 110 mL/min. Peak capacity is inversely proportional to the peak width and varies with the flow rate of the mobile phase (in accordance with the Van Deemter equation) [29]. When the flow rate was increased from 44 to 88 mL/min, the peak capacity decreased by 10% while the purification capacity per instrument per unit time was doubled. Therefore, in cases where there is adequate resolution (impurities are well separated from the product peak), it is worthwhile to implement the high flow purification method.

Fig. 11 represents purification of a crude sample performed at 110 mL/min. Three chromatograms were recorded during the separation of the sample. Note that the major peak at approximately 1.6 min is absent from the waste UV-trace because the fraction was completely diverted to the collection vial.

4. Conclusion

We have developed and implemented a high throughput purification platform with a process based on mass directed fractionation with a one-to-one mapping of sample injection to fraction collection. A series of recovery studies validated the reliability of the mass directed fractionation coupled with preparative HPLC separation. Enforcement of one fraction per sample collection has greatly enhanced the purification efficiency and simplified the post-purification process. This approach has worked extremely well in our high throughput environment that deals with several thousand compounds per week. The purification process success rate is greater then 97% and this approach has produced several hundred thousand high quality compounds annually with an average purity over 97% by low wavelength UV detection, which is currently the most stringent quality criteria applied in a high throughput mode to synthetic compounds. In addition to the large capacity available with our present capital investment in HPLC/MS equipment, our continued development in faster chromatographic separations affords us the opportunity to

double our overall throughput without investing in additional equipment.

Acknowledgments

The authors acknowledge the contributions of colleagues Nicole Hosein, Wendy Picariello, Fushuang Liu and Tom Roark.

References

[1] M.A. Gallop, R.W. Barrett, W.J. Dower, S.P.A. Fodor, E.M. Gordon, J. Med. Chem. 37 (1994) 1233.
[2] E.M. Gordon, R.W. Barrett, W.J. Dower, S.P.A. Fodor, M.A. Gallop, J. Med. Chem. 37 (1994) 1385.
[3] C.N. Selway, N.K. Terrett, Bioorg. Med. Chem. 4 (1996) 645.
[4] M. Bauser, J. Chromatogr. Sci. 40 (2002) 292.
[5] X. Cheng, J. Hochlowski, Anal. Chem. 74 (2002) 2679.
[6] C. Edwards, D.J. Hunter, J. Comb. Chem. 5 (2003) 61.
[7] B.H. Hsu, E. Orton, S.-Y. Tang, R.A. Carlton, J. Chromatogr. B 725 (1999) 103.
[8] D.A. Yurek, D.L. Branch, M.-S. Kuo, J. Comb. Chem. 4 (2002) 138.
[9] L. Zeng, L. Burton, K. Yung, B. Shushan, D.B. Kassel, J. Chromatogr. A 794 (1998) 3.
[10] B. Yan, L. Fang, M. Irving, S. Zhang, A.M. Boldi, F. Woolard, C.R. Johnson, T. Kshirsagar, G.M. Figliozzi, C.A. Krueger, N. Collins, J. Comb. Chem. 5 (2003) 547.
[11] W.C. Ripka, G. Barker, J. Krakover, Drug Discov. Today 6 (2001) 471.
[12] P.J. Edwards, Comb. Chem. High Throughput Screen 6 (2003) 11.
[13] P. Wipf, Pharm. News 9 (2002) 157.
[14] H.N. Weller, M.G. Young, S.J. Michalczyk, G.H. Reitnauer, R.S. Cooley, P.C. Rahn, D.J. Loyd, D. Fiore, S.J. Fischman, Mol. Divers. 3 (1997) 61.
[15] J.N. Kyranos, H. Cai, B. Zhang, W.K. Goetzinger, Curr. Opin. Drug Discov. Devel. 4 (2001) 719.
[16] J.N. Kyranos, H. Cai, D. Wei, W.K. Goetzinger, Curr. Opin. Biotech. 12 (2001) 105.
[17] L. Schultz, C.D. Garr, L.M. Cameron, J. Bukowski, Bioorg. Med. Chem. Lett. 8 (1998) 2409.
[18] G.A. Nemeth, D.B. Kassel, Annu. Rep. Med. Chem. 36 (2001) 277.
[19] P.D. Lynch, D.H. Whitney, T.R. Londo, N.F. Gordon, Integrated chromatographic purification and analysis for high productivity methods development and routine purification, in: Proceedings of the 216th ACS National Meeting, Boston, MA, 1998.
[20] J.A. Ramieri, Parallel purification, Laboratory Automation in the Chemical Industries, Marcel Dekker Inc., 2001, p. 133.
[21] J. Isbell, R. Xu, Z. Cai, D.B. Kassel, J. Comb. Chem. 4 (2002) 600.
[22] G.R. Eldridge, H.C. Vervoort, C.M. Lee, P.A. Cremin, C.T. Williams, S.M. Hart, M.G. Goering, M. O'Neil-Johnson, L. Zeng, Anal. Chem. 74 (2002) 3963.
[23] http://www.discoverypartners.com/Collaborative/chem_tech_purification.html.
[24] L. Zeng, D.B. Kassel, Anal. Chem. 70 (1998) 4380.
[25] C. Edwards, J. Liu, T.J. Smith, D. Brooke, D.J. Hunter, A. Organ, P. Coffey, Rapid Commun. Mass Spectrom. 17 (2003) 2027.
[26] R. Xu, T. Wang, J. Isbell, Z. Cai, C. Sykes, A. Brailsford, D.B. Kassel, Anal. Chem. 74 (2002) 3055.
[27] H. Cai, J.P. Kiplinger, W.K. Goetzinger, R.O. Cole, K.A. Laws, M. Foster, A. Schrock, Rapid Commun. Mass Spectrom. 16 (2002) 544.
[28] C. Waters, Fractionlynx User's Guide, Milford MA, 2002.
[29] L.R. Snyder, J.J. Kirkland, J.L. Glajch, Practical HPLC Method Development, second ed., Wiley, New York, 1997.