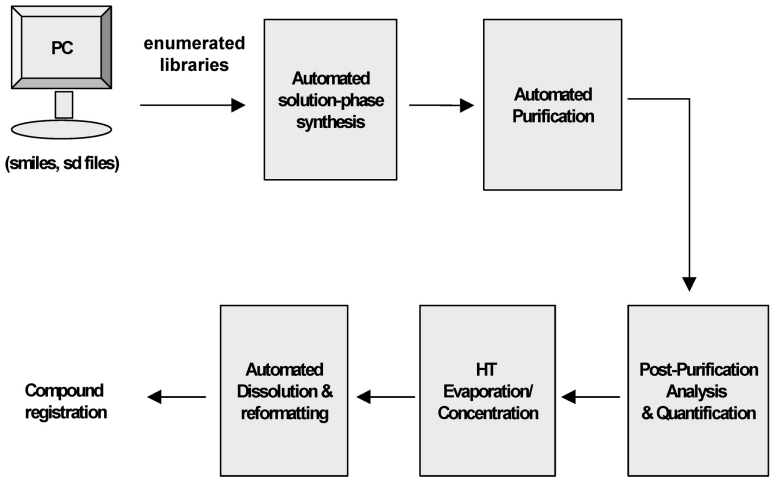
JOURNAL OF combinatorial CHEMISTRY

Article

Realities of High-Throughput Liquid Chromatography/Mass Spectrometry Purification of Large Combinatorial Libraries: A Report on Overall Sample Throughput Using Parallel Purification

J. Isbell, R. Xu, Z. Cai, and D. B. Kassel

J. Comb. Chem., **2002**, 4 (6), 600-611• DOI: 10.1021/cc0200336 • Publication Date (Web): 29 August 2002



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 4 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

Journal of Combinatorial Chemistry is published by the American Chemical Society.
1155 Sixteenth Street N.W., Washington, DC 20036

Realities of High-Throughput Liquid Chromatography/Mass Spectrometry Purification of Large Combinatorial Libraries: A Report on Overall Sample Throughput Using Parallel Purification

J. Isbell,[†] R. Xu,[‡] Z. Cai, and D. B. Kassel*

Deltagen Research Laboratories, 4570 Executive Drive, Suite 400, San Diego, California 92121

Received May 3, 2002

We report on the development of a validated, streamlined high-throughput process for the purification of parallel-synthesis-derived combinatorial libraries. The steps involved in this library purification process include dissolution of dry films of crude synthetic material, dual-column LC/MS purification, dual-column postpurification analysis, quantitation, reformatting, and submission of pure compounds for registration. Although the purification and postpurification analysis times decreased essentially linearly as a function of the number of HPLC columns employed, it was not possible to decrease the total purification process time linearly as a function of the number of columns employed in the system. This was due primarily to the fact that numerous steps in the total purification process are independent of sample analysis and purification (e.g., evaporation, reconstitution, and reformatting, etc.). Additionally, experiments were also performed to assess whether separate gradient pumps were necessary for each channel of this two-channel LC/MS or if acceptable results could be reliably obtained by splitting the flow from one set of gradient pumps between two HPLC columns. On the basis of the parallel, two-column LC/MS system employed in this work, throughput estimates were extrapolated to more massively parallel systems (e.g., four-channel LC/MS).

Introduction

Combinatorial chemistry continues to play an important role in pharmaceutical and biotechnology research.[1] The field of combinatorial chemistry has spawned a steady stream of enabling automated technologies that facilitate not only the synthesis of a large number of structurally related compounds but also the high-throughput in vitro screening against therapeutic targets. With both activity and inactivity data increasingly being used to generate SAR[2] and direct subsequent synthetic efforts, organizations are recognizing the importance of verifying the quality of compounds prior to screening, and not only those compounds for which activity was observed. To minimize false positives and false negatives,[1] it is advantageous to assay only high-quality compounds. Therefore, great effort has been devoted to the development of automated purification technology designed to keep pace with the output of high-throughput combinatorial/parallel synthesis.[1,3,4]

A number of approaches have been employed successfully to purify compounds of combinatorial libraries. Techniques including liquid/liquid extractions,[5,6] liquid/solid extractions,[7−13] fluorous extractions,[14] and scavenger "capture and release" resins[15−17] are fast and readily automated but do not consistently provide sufficiently pure final products.[13,18] Because chromatography is a more general approach for routinely isolating pure products from reaction mixtures,[1,19−28] this technique is often used for compound purification.

By use of single-channel HPLC-based purification systems, routine sample throughput of up to 192 reaction mixtures per 24-h day was reported.[26] With parallel HPLC systems, it has been reported that the theoretical throughput increases to 384 samples per day for a two-channel system and to 768 samples per day for a four-channel system. Recently, Ripka et al. reported that up to 384 reaction mixtures could be purified in a 10-h day using a proprietary four-channel supercritical fluid chromatography (SFC) system.[4]

These reports on high-throughput purification of combinatorial libraries have focused on the (theoretical) throughput achievable for the sample purification step. To our knowledge, no one has documented the time required for the entire purification process, that is, from the receipt of crude samples for purification, through purification, postpurification analysis, reformatting, dissolution, and registration of the "acceptably pure" fractions.

In this report, we describe the development, optimization, validation, and implementation of an integrated high-throughput purification process, as shown in Figure 1. A key component in our process is mass-directed fraction collection using a dual-column liquid chromatography/mass spectrometry (LC/MS) purification system. Our group developed this technology to facilitate collection of only fractions containing the desired molecular weights.[30] Two significant benefits were realized from this technology: the ability to collect only fractions likely to contain the product of interest and

* To whom correspondence should be addressed. Address: Syrrx, Inc., 10410 Science Center Drive, San Diego, CA 92121. E-mail: daniel.kassel@syrrx.com. Phone: (858) 731-3686. Fax: (858) 550-0526.

[†] Current address: Genomics Institute of the Novartis Research Foundation, 10675 John Jay Hopkins Drive, San Diego, CA 92121.

[‡] Current address: Syrrx, Inc., 10410 Science Center Drive, San Diego, CA 92121.

10.1021/cc0200336 CCC: $22.00
Published on Web 08/29/2002

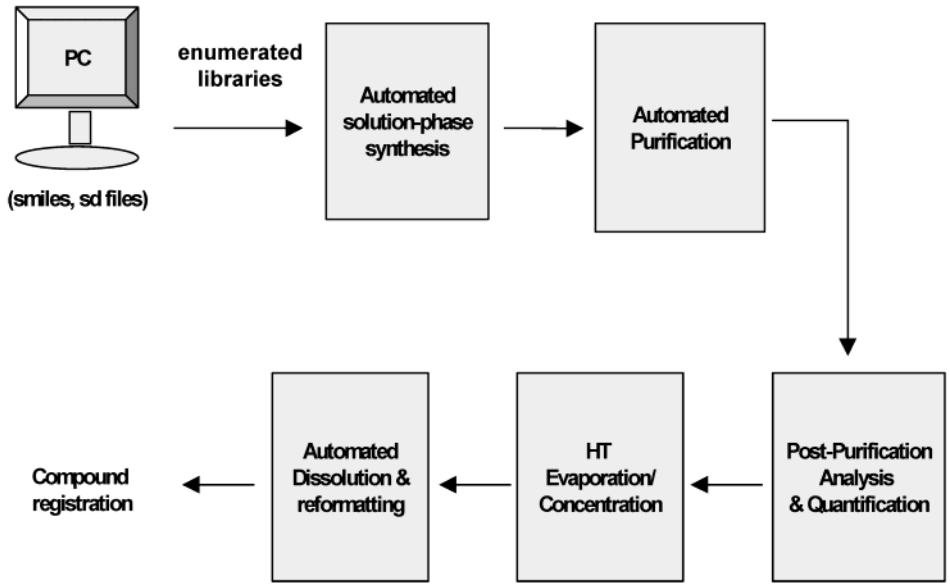


Figure 1. High-throughput purification process work flow diagram.

the ability to more easily track fractions because of the optional plate mapping feature of the system.

In this work, a custom parallel two-channel LC/MS purification system was used for all analytical and preparative LC/MS. The robustness of this dual-column LC/MS system and the fidelity of the purification process were assessed by purifying over 7800 reaction mixtures over a period of 7 months. When using technologies designed to facilitate fraction tracking and automated liquid handling, the time required to completely process samples, from receipt of sample plates for purification to submitting properly formatted plates of pure compounds for registration, is significantly more than that of the LC/MS purification step itself. Overlooking these other important steps in the overall purification process may seriously underestimate the amount of time required to purify a library and may have serious implications on plans focused on a large-scale purification effort.

To optimize the purification process, a 4400-member process optimization library, library 1, was purified, with each step examined and improved "on-the-fly." Two additional libraries were purified to assess the maximum purification throughput one scientist could achieve. On the basis of these data, we were able to estimate the impact of parallel purification technology on overall sample processing relative to a single-column system. This paper discusses the development of an integrated solution to streamline the entire purification process from library purification to final compound registration and the impact of sample handling on overall sample throughput.

Experimental Details

Materials. All the Fmoc-amino acids used as standards to evaluate system performance, Fmoc-Ala-OH, Fmoc-Thr-(tBu)-OH, Fmoc-Asn, Fmoc-Glu, Fmoc-Trp-OH, Fmoc-Try-(tBu), Fmoc-Thr-OH, Fmoc-Ile-OH, Fmoc-Val-OH, Fmoc-Pro, Fmoc-Tyr-OH, Fmoc-Asp-OH, Fmoc-Ser-OH, Fmoc-Phe-OH, Fmoc-Met-OH, and Fmoc-Aib-OH were obtained from Novabiochem (Calbiochem-Novabiochem AG, Switzerland). Eluent A was 0.05% trifluroacetic acid in deionized water and eluent B was 0.035% trifluoroacetic acid in acetonitrile. Fractions were collected into 5 mL pyramid-bottom 48-well plates. (Thompson Instrument Company, San Diego, CA).

Several libraries were obtained from our high-throughput synthesis group to evaluate and optimize the purification process. Library 1, the process optimization library, consisted of 4400 reaction mixtures, i.e., 50 plates with 88 reaction mixtures in each plate. Library 2, the process-testing library, consisted of 15 plates with 88 reaction mixtures each. More complete analysis of samples before and after purification was done for library 3 in order to determine the practical recoveries for actual library compounds purified according to the process described herein. Additional plates from other libraries were purified to assess system robustness for a range of compounds. Approximately 7800 reaction mixtures were purified in total.

LC/MS Conditions. All analytical and purification work associated with these libraries was completed using a parallel dual-column LC/MS system. The components required for LC/MS data collection were completely controlled by MassChrom 1.2 running on a Macintosh G3 computer with 256 MB RAM. A diagram of the system is shown in Figure 2. The system consists of an in-house-modified Gilson 215 autosampler that supports simultaneous sampling from the same well location across two adjacent plates (Gilson, Inc., Middleton, WI), four LC-8A HPLC pumps (Shimadzu, Tokyo, Japan), four SPD-10AD$_{vp}$ UV detectors (Shimadzu), two evaporative light-scattering detectors (ELSDs, Alltech Associates, Inc., Deerfield, IL), an SPD-10A$_{vp}$ system controller (Shimadzu), an API-150EX mass spectrometer (AB/MDL Sciex, Foster City, CA), and two Gilson 204 fraction collectors (Gilson, Inc.). Aside from the Gilson 215, all hardware components were used as received from the vendors.

Three components of the Gilson 215 were modified to support simultaneous sampling: the sample probe holder, the injection port bar,[29] and the syringe drive.[31] A special

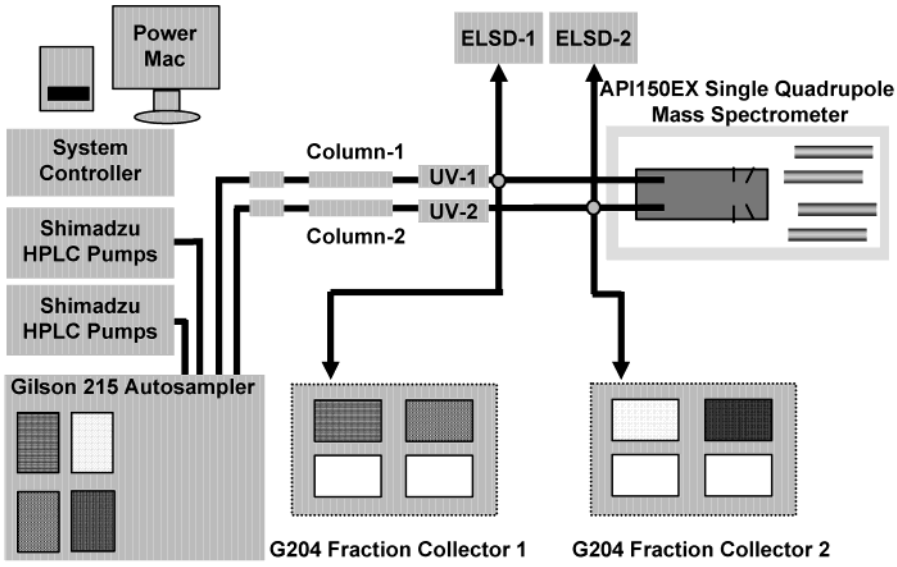


Figure 2. Block diagram of the custom parallel two-channel LC/MS used for HT purification.

sample probe holder was designed to support two standard Gilson sample probes with the space between the two probes equal to the spacing of one Gilson code 205 rack.[29] The probe spacing was precisely matched by the spacing between the injection ports of the two Gilson 819 injectors. Accurate control of the sample volume aspirated into each probe was achieved by constructing a dual syringe drive to permit simultaneous sampling from the two wells using separate syringes.[31]

The binary gradient mobile phase was delivered to each of the two Gilson 819 injectors by a pair of LC-8A HPLC pumps, with all four pumps controlled by the Shimadzu controller. Each stream passes either through a preparative column (YMC-Pack, 5 μm, 100 mm $\times$ 10 mm i.d. ODS-A column (Waters Corp, Milford, MA)) or through an analytical column (YMC 5 μm, 50 mm $\times$ 4.6 mm i.d. ODS-A column (Waters Corp.)) and then through dedicated UV detectors. To facilitate switching between preparative and analytical LC/MS, two UV detectors were used for each column, one for analytical LC/MS and one for preparative LC/MS. The effluent from each UV detector was split with a static tee splitter (ZT1C, Upchurch Scientific, Oak Harbor, WA) so that 100 μL of each column effluent was directed into the ion source through the sample and makeup flow inlets and merged at the tip of the sprayer. The balance of the effluent was directed to another static splitting tee that diverted another 100 μL from the remaining flow into an ELSD. The majority of the column effluent was sent from the second splitting tee to either a Gilson 204 fraction collector during purification or to waste during analytical LC/MS.

With the solvent flows from each of the two columns isolated from each other until they merged at the tip of the IonSpray source, the system hardware behaved as two separate, but linked, HPLC systems.

Preparative LC/MS Conditions. Since the theoretical yield of the reaction mixtures was $\leq$20 μmol, the libraries were purified using YMC-Pack, 5 μm, 100 mm $\times$ 10 mm i.d. ODS-A columns (Waters Corp, Milford, MA) with Monitor 5 μm 10 mm $\times$ 10 mm i.d. C18 guard cartridges (Column Engineering, Inc., Ontario, CA). The guard cartridges were replaced between each purification queue of 352 samples. The sample loop size was 500 μL, and the injection volume for libraries 2 and 3 was 400 μL. The injection volume for library 1 was 500 mL, as discussed below. The flow rate for each channel was 10 mL/min. On the basis of the analytical LC/MS results obtained for three plates of crude reaction mixtures, the following gradient was created and used for purification of the 4400-member process optimization library: 10% B held for 0.5 min; 10% B to 40% B in 1.0 min; 40% B to 80% B in 5.0 min; 80% B to 90% B in 0.5 min; 90% B held for 0.5 min.

Analytical LC/MS Conditions. Postpurification analytical LC/MS was performed without changing the 500 μL sample loop used for preparative LC/MS. YMC ODS-A 5 μm, 50 mm $\times$ 4.6 mm i.d. columns (Waters Corporation) were used with a flow rate of 3.0 mL/min per column. The following gradient was used: 10% B held for 0.2 min; 10% B to 90% B in 3.8 min; 90% B held for 0.5 min.

AppleScript Application To Perform Mass-Directed Parallel Purification. Mass-directed purification for the two-channel parallel LC/MS system was achieved by modifying a vendor-provided Applescript, "FC Script Combichem 1.4", so that observation of the desired ion corresponding to the molecular weights designated as "Mass 1" and "Mass 3" in the sample list would trigger collection by fraction collector 1 and those designated as "Mass 2" and "Mass 4" would trigger collection by fraction collector 2. The Applescript triggers the start and end of collection based on user-defined extracted ion chromatogram (XIC) thresholds and a maximum collect time, the latter calculated based on the time required to fill 80% of the collection vessel. Since fraction collection was triggered by the XIC for the ions corresponding to the compounds of interest exceeding a user-defined threshold, it was critical that each of the two flow streams merged at the tip of the ion source provides comparable response. Typically the response differences were less than or equal to 10% for the same compound injected onto each column.

To facilitate sample tracking and downstream processing of purified fractions, the fractions were collected in a "well-

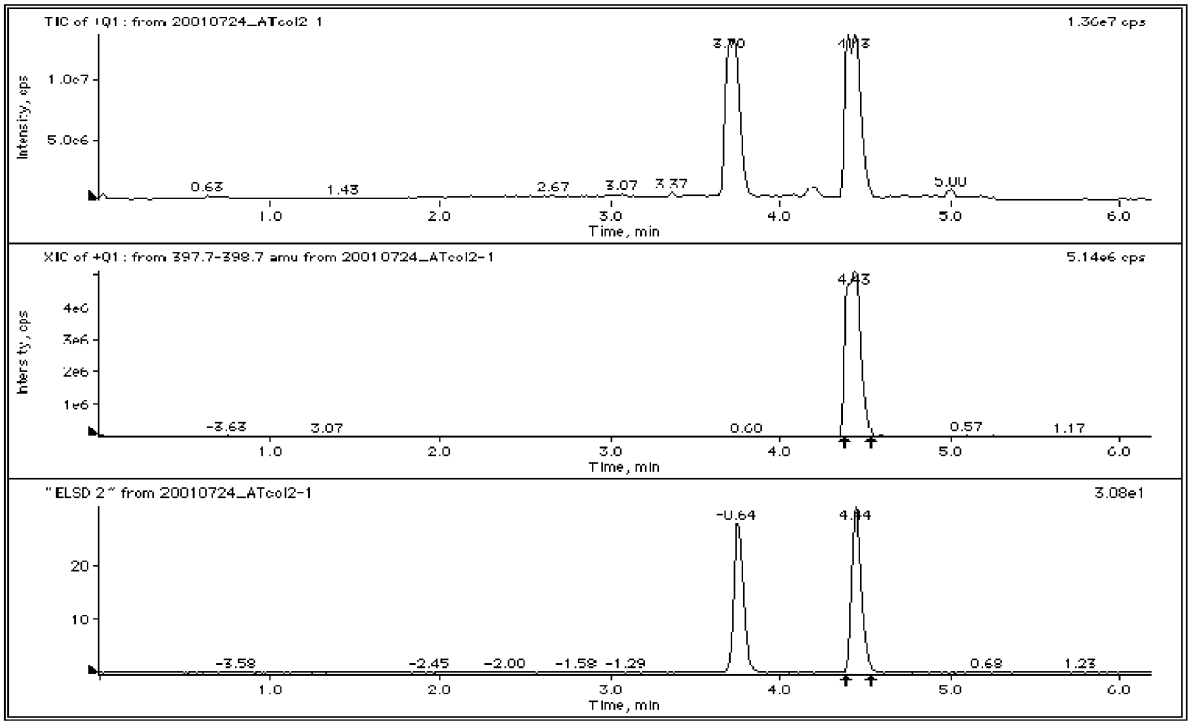


Figure 3. Recovery study showing complete collection of Fmoc-Thr(tBu)-OH.

to-well" plate-mapping mode. Although this script supports collection into any rack supported by the Gilson 204 fraction collector, microtiter plate formats were selected for ease of sample tracking and to facilitate rack transfers among all instrumentation (including evaporators/concentrators) used in the purification process.

Recovery Studies. Recovery studies were performed by injecting 200 μL of a solution containing 10.0 mg/mL each of Fmoc-Ala-OH and Fmoc-Thr(tBu)-OH. On the basis of the measured flows diverted to the MS (100 μL/min, or 1%) and to the ELSD (100 μL/min, or 1%), it was calculated that no more than 98% (1.96 mg) of each Fmoc-amino acid injected could be collected; thus, this would represent 100% recovery.

To ensure collection of the entire peak of interest, the thresholds for collection beginning and terminating were set to 5% of the expected intensity of the extracted ion chromatogram (XIC). An example of the collection of Fmoc-Thr using these preset ion intensity threshold values is shown in Figure 3. The results suggest that the entire Fmoc-Thr peak was collected, as indicated by the start and end collection arrows shown directly below the mass chromatogram (*m*/*z* 398) for Fmoc-Thr. To calculate the percent recovery, the volume of the collected fraction was measured and the expected concentration was calculated for 100% recovery. The collection plate was sealed and shaken. A sample of this solution was injected and the ELSD peak area compared with that obtained for an aliquot from the same 10 mg/mL solution diluted to the expected concentration. Recovery studies were done before each purification queue of 352 samples to verify system performance. If the recovery before purification fell below our minimum acceptable recovery of 85%, the problem was corrected and the recovery study repeated prior to purification. Comparison of several recovery studies done before and after a purification batch generally indicated there was little (~5%) or no decrease in collection efficiency between the start of the queue and the end of the queue.

Preparation of Library Plates for Purification. Since many of the wells of the source plates from library 1 contained precipitate that did not dissolve in 600 μL of DMSO, the library plates were centrifuged prior to purification. After centrifuge, the plates were carefully transferred to the autosampler and 500 μL was aspirated from each well for injection. By use of this technique, injector clogging was observed just once during the purification of library 1. Even though the other libraries purified contained a small percentage of wells with visible precipitate, no clogging was observed when all the material in the well was injected for purification.

Generation of Postpurification QAQC Plates for Quality and Purity Assessment. Upon completion of the purification sample queue, the collected fractions were evaporated using a Genevac HT-12 (Genevac, Ipswich, England). Since variable volumes were collected for each compound purified and since some unknown amount of evaporation occurs during the purification queue, reliable postpurification quantification could not be done directly using aliquots from the collection plates. By use of a Genesis RSP 200 (Tecan Group, Ltd., Zurich, Switzerland), the dried plates were redissolved in 1.0 mL of DMSO and shaken. A 100 μL aliquot was transferred from each well of two 48-well collection plates into one 96-well plate (QAQC plate) for postpurification LC/MS analysis. The QAQC plate generated was a mirror of the purification source plate, thereby facilitating tracking of postpurification data with both the source plate and the collection plates. To prepare the collection plates for final dissolution and reformatting, the

remaining 900 μL in the collection plates was evaporated, again using a Genevac HT-12 with 60% lamp power.

Postpurification Quantification and Purity Assessment. Immediately prior to LC/MS analysis of the postpurification QAQC plates, ELSD calibration curves were generated for each column using Fmoc-Ala-OH and Fmoc-Thr(tBu)-OH as calibration standards. The calibration curve spanned the range of at least from 6 to 60 μg. All collected fractions were analyzed using the parallel two-channel LC/MS system, and the collected data were processed using a customized Applescript that extracted purity information and ELSD area information for each fraction. This script has options that enable the script to estimate purity and quantity for multiple isomers in a fraction and/or to save a "snapshot" of the processed data as a GIF file that can be uploaded into a tracking system on a PC platform.

Each of the compounds selected for library reformatting had a purity requirement of at least 85%, as determined using the average of the purities from two detectors, ELSD and UV 220. A National Instruments NI-488 board facilitated the acquisition of more than two analog signal inputs. Using this NI board, the MassChrom 1.2 software was capable of monitoring a maximum of 16 analog inputs.

Preparation of Purified Library Master Archive Plates. A custom database was used to link the purity and quantity of purified fractions with the expected structures, molecular masses, and theoretical quantity of crude sample synthesized. Although this custom tracking system was a significant improvement over sample tracking using spreadsheets, it was designed initially for tracking of small batches of 96−384 samples. Since this system proved somewhat cumbersome for uploading information for thousands of fractions, modifications were made to this system to further facilitate tracking of samples in 48-well collection plates and to permit batch uploading of data. These modifications decreased tracking time by over 70%.

The database is integrated with a Tecan Genesis liquid handling system to automate redissolution of individual samples to a user-defined concentration. Furthermore, this system is capable of reformatting the new solutions to master archive plates using Boolean logic and based on user-defined thresholds for quantity, purity, and/or other user-defined criteria.

Results and Discussion

To achieve a highly efficient purification process, the integrity of each component needed to be assessed and each task needed to be optimized and integrated into the overall process. Since efficient high-throughput (HT) purification entails more than LC/MS purification, we were interested in determining the resources necessary for one person to process a large library.

Assessment of Process Integrity. Before embarking on a large-scale purification effort, it was critical to determine the amount of material lost while traversing the complete purification process and if such losses are acceptable in exchange for obtaining highly purified compounds. Since the parallel LC/MS purification system had demonstrated the ability to routinely collect at least 85% of the injected standards and system performance was evaluated before each purification run, the remaining sources for potential losses could be encountered only during postpurification sample handling.

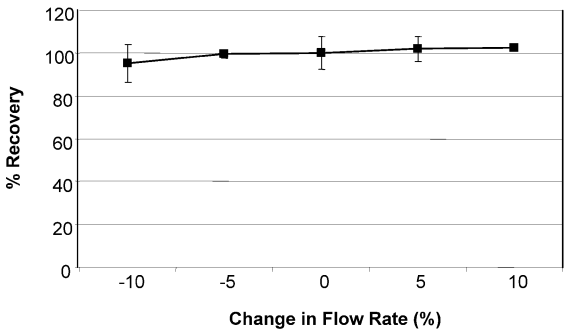


Figure 4. Prep flow rate experiments (±10% changes in flow rate do not have a statistically significant impact in recoveries relative to standard (10.0 mL/min) flow rate).

To determine if the significant sample losses were occurring during the postpurification sample processing, a known quantity of Fmoc-Thr(tBu)-OH was placed in an empty well of a collection plate. Following routine plate processing, the quantity of Fmoc-Thr(tBu)-OH in the spiked well was quantified against an aliquot of the original solution diluted to that same expected concentration. It was found that the losses in postpurification sample handling were not significant, i.e., <10%. Thus, with purification recoveries greater than or equal to 85% and postpurification sample handling resulting in no more than 10% lossess, overall yields of approximately 70% would be expected when using this production-scale purification process with solutions free of insoluble materials. This 70% purification yield is supported by the experimental data obtained for library 3.

Pump Flow Splitting. Although our group has reported successfully purifying libraries using just one set of gradient pumps and splitting the flow to the two injectors using a static splitting tee,[30,31] there has been no definitive report comparing this technique (dubbed the "split-and-pray" approach) to the more traditional method of using one set of pumps per column. With one set of gradient pumps, the relative flow rates between the two columns are regulated by the relative column backpressures, which may be affected by the library chemistry and/or differential column degradation. If the flow rate fluctuates in the course of a purification queue, the critical peak detection and fraction collection timing may be affected and may result in potentially greater losses during purification.

To assess the general need for highly accurate flow control during purification, mass-directed fraction collection was done at 9.0, 9.5, 10.0, 10.5, and 11.0 mL/min. As shown in Figure 4, there is no statistically significant difference in recoveries over these flow rates. It was noted that a 10% change in flow rate resulted in a 30 bar change in column backpressure. With two sets of gradient pumps, it was possible to monitor the relative column backpressures before and after each purification batch. For library 1, a library with visible precipitate in many wells, a 30 bar backpressure change between the two columns occurred at some time during the purification of ∼15% of the process testing library. Such differences in relative pressures were not observed

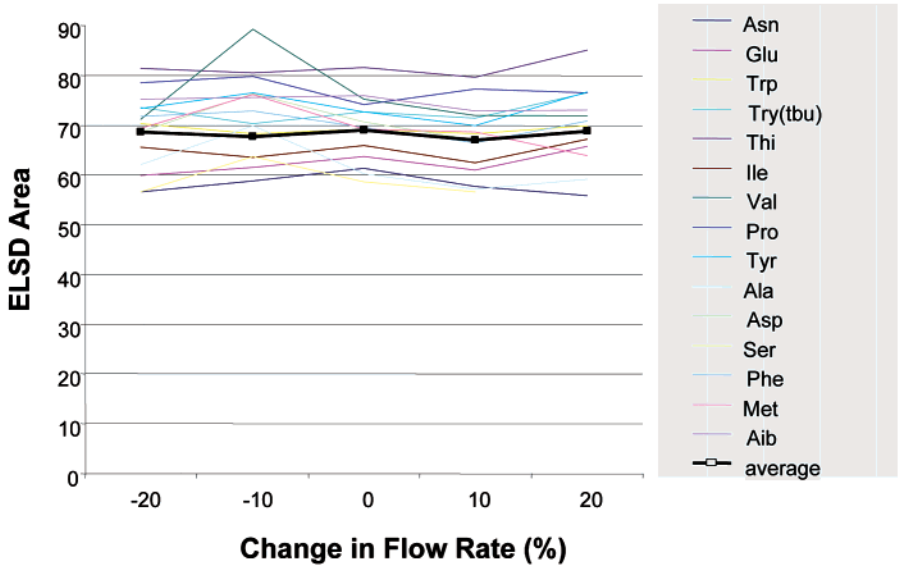


Figure 5. ELSD response vs flow rate, indicating impact of flow rate changes of up to ±20% relative to standard analytical flow rate of 3 mL/min on the ELSD response.

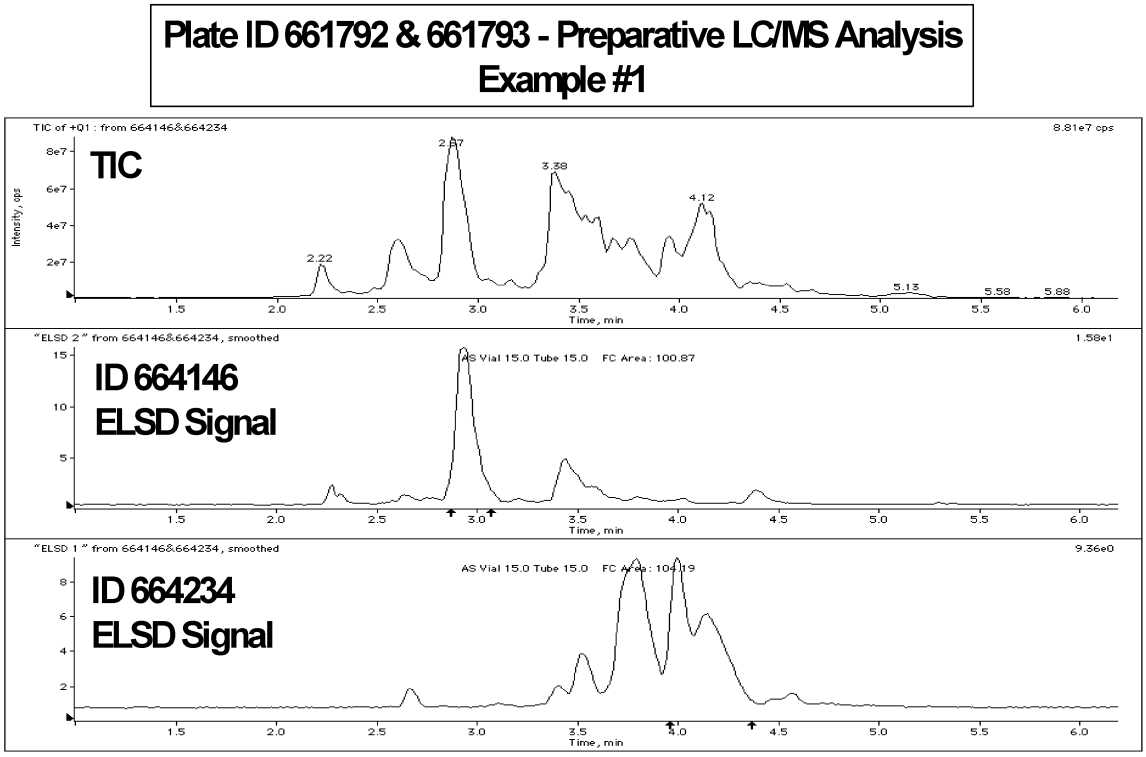


Figure 6. Dual-column LC/MS purification of two wells from library 1. The composite total ion current chromatogram (TIC) shown in the top panel and the ELSD chromatograms for columns 1 and 2 (middle and lower panels) illustrate the complexity of the library mixtures. The target mass was collected by mass-directed fraction collection. The start and end collection times are indicated by the arrows superimposed on the ELSD traces.

during the purification of the last 39 plates for which samples were free of visible precipitate. Our data demonstrate that two sets of gradient pumps are not essential for parallel purification, *provided the samples are free of precipitate or no precipitate is injected.* However, the use of four gradient pumps had a positive impact on troubleshooting and instrument repair during library purification.

The need for highly accurate flow control was assessed also for analytical LC/MS and for quantification. The ELSD areas for 0.5 mg/mL solutions of 15 Fmoc-amino acids were obtained at 2.4, 2.7, 3.0, 3.3, and 3.6 mL/min, thus representing flow rate changes of up to ±20% of the target flow rate of 3.0 mL/min. The data, shown in Figure 5, suggest that such significant changes in analytical flow rate have only a modest impact of ≤20% on ELSD areas. The average deviation of ELSD area for all the amino acids examined was <5%. On the basis of these data, the additional pumps are considered superfluous for our analytical LC/MS needs.

Estimating Library Throughput with an Optimized Purification Process. From the 4400 reaction mixtures in

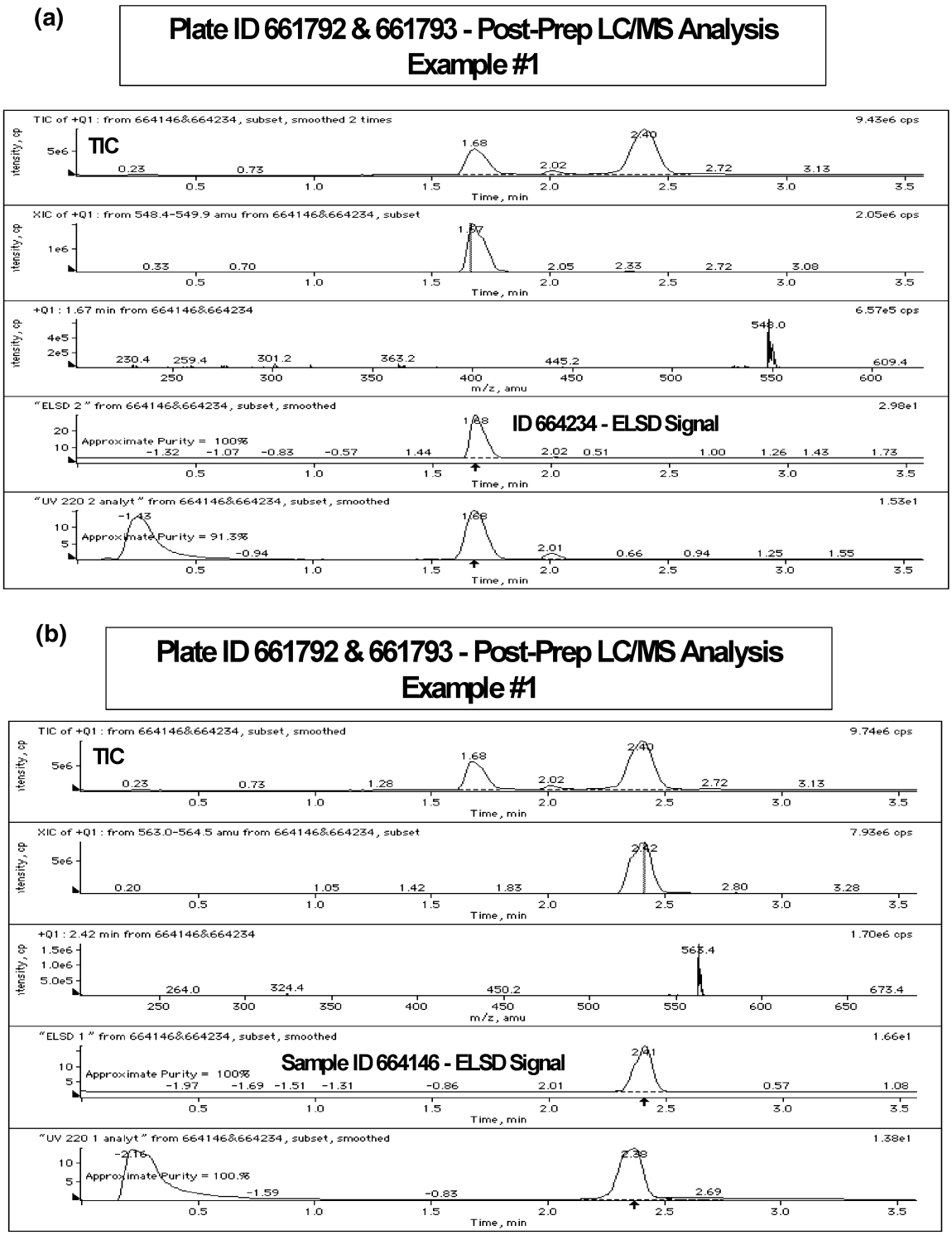


Figure 7. (a) Dual-column LC/MS postpurification analysis of two wells from library 1 shows that the isolated product from column 1 following dual-column LC/MS purification is of high purity. The composite TIC (top panel) shows two peaks because, as alluded to earlier, it is a mixture of the signals arising from the two columns, since the source does not incorporate an "indexed" or "MUX" interface. The extracted ion current chromatogram (XIC) for the compound isolated from column 1, the mass spectrum at the apex of XIC peak, the ELSD trace, and the UV chromatogram (second, third, fourth, and bottom panels, respectively) indicate that the compound was isolated to near homogeneity (greater than 95% pure). (b) Dual-column LC/MS postpurification analysis of two wells from library 1 shows that the isolated product from column 2 following dual-column LC/MS purification is pure. The XIC for the compound isolated from column 2, the mass spectrum at the apex of the XIC peak, the ELSD trace, and the UV chromatogram (second, third, fourth, and bottom panels, respectively) indicate that the compound was isolated to near homogeneity (greater than 95% pure).

the process optimization library (library 1), 2678 compounds (61%) were isolated in quantities that met or exceeded the minimum required amount. Importantly, 2531 (95%) of these compounds, which met or exceeded the minimum required amount, also met the minimum 85% purity criteria. These data are supported in the example of the dual-column LC/MS purification of two wells of library 1 shown in Figure 6. The *composite* total ion current (TIC) chromatogram (i.e.,

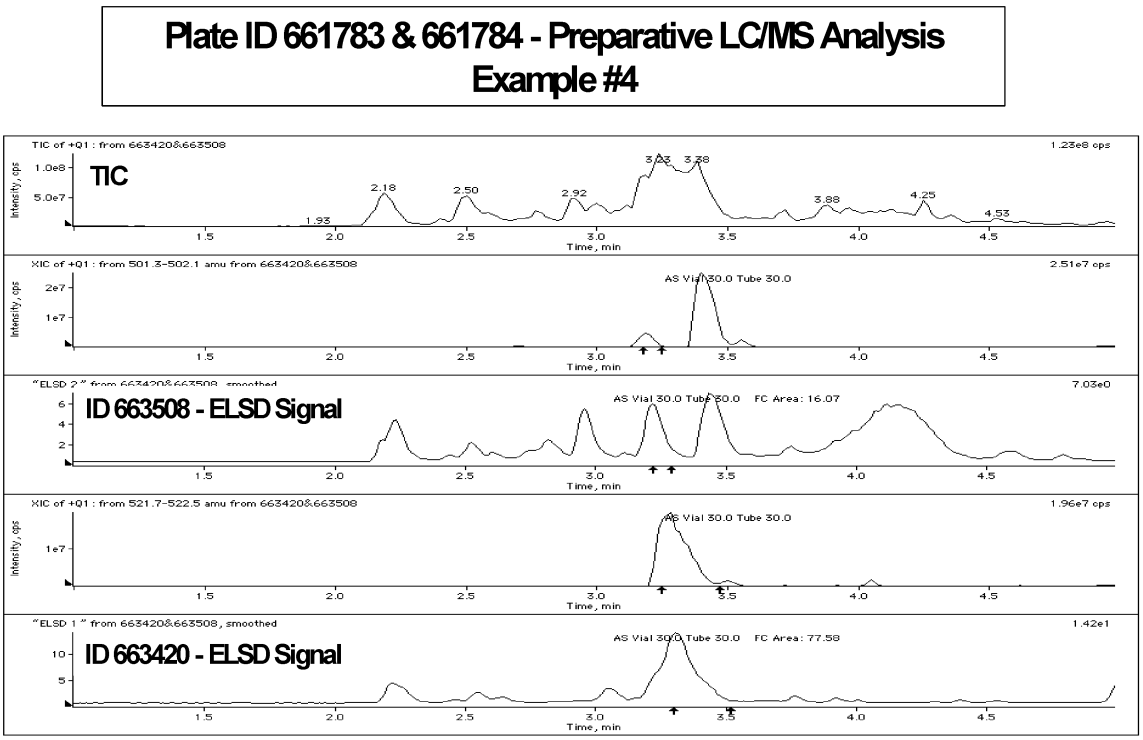


Figure 8. Dual-column LC/MS purification of two wells from library 2. The composite TIC chromatogram shown in the top panel and the ELSD chromatograms for columns 1 and 2 (middle and lower panels) illustrate the complexity of the library mixtures. The target mass was collected by mass-directed fraction collection. The start and end collection times are indicated by the arrows superimposed on the ELSD traces.

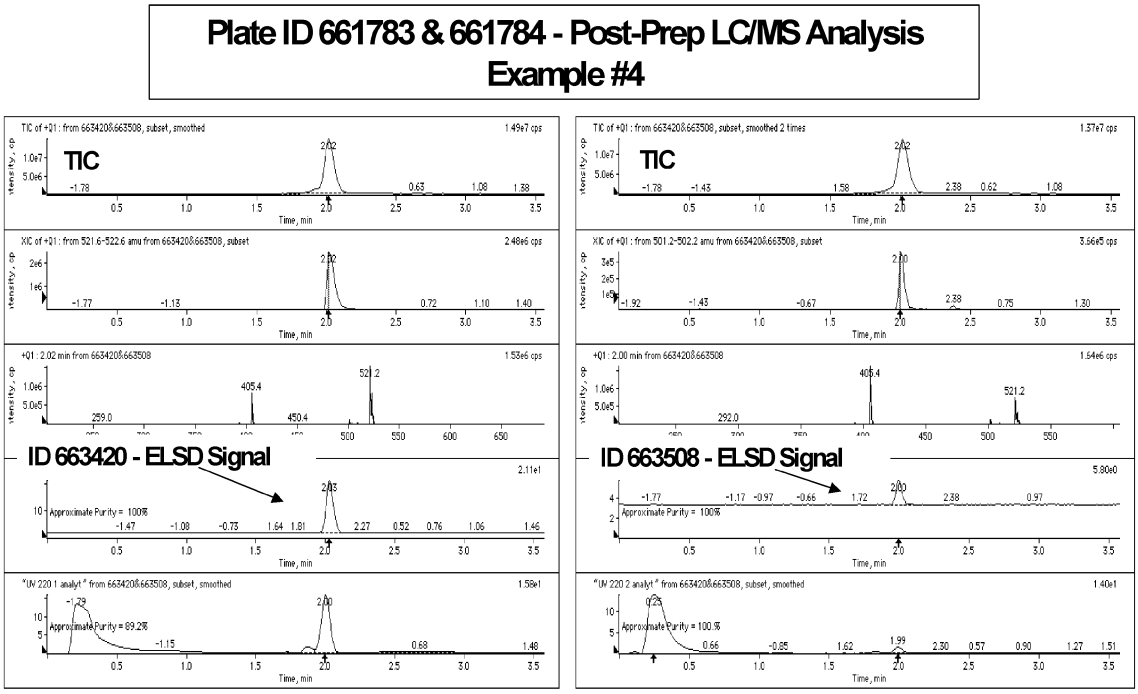


Figure 9. Dual-column LC/MS postpurification analysis of two wells from library 2 shows that the isolated products from columns 1 and 2, following dual-column LC/MS purification, are highly pure. The XICs for the compounds isolated from columns 1 and 2, their corresponding mass spectra, the ELSD traces, and the UV chromatograms (second, third, fourth, and bottom panels, respectively) indicate that both compounds were isolated to near homogeneity (greater than 95% pure).

the combined ion signals from columns 1 and 2 observed in the ion source, since no "indexing" or "MUX" interface was used) and the ELSD chromatograms for these two samples illustrate the complexity of the library synthesis. These wells were in fact very much representative of the quality of the library 1 synthesis. Figure 7 shows the postpurification analysis results. The TIC, UV, and ELSD traces for both isolated products show, upon postpurification analysis, that they were isolated to near homogeneity. Taken together, the data clearly demonstrate that even when the library mixtures

are very complex, highly purified products can be obtained from this dual-column LC/MS purification procedure. The results on these compounds as well as the remaining members of the 4400-component reaction library 1 mixture indicate that this process is capable of generating reasonable quantities of very pure compounds.

Library 2 showed similarly complex mixtures and equally high-quality purification results. Figure 8 shows an example of two components of library 2 purified by dual-column LC/MS. Again, the composite TIC chromatogram and the ELSD chromatograms for columns 1 and 2 indicate that both wells contained very crude mixtures and that the product of interest is only a small component of the entire mixture. Mass-directed purification, however, enables the expected product to be isolated to near homogeneity, as illustrated in Figure 9. Similar to library 1, the postpurification analysis for these two library 2 components shows that the expected products are greater than 95% "pure" (based on UV220 and ELSD integration) following postpurification analysis.

Library 2, consisting of 1208 reaction mixtures, was used as a model to determine the maximum purification throughput achievable by one person. If the only task associated with library purification were mass-directed fraction collection, then purification of this library would have been completed in just 96 h. However, as shown in Figure 10, the work associated with library purification does not begin and end with the purification, since there are additional tasks associated with sample preparation and sample handling of purified fractions.

Samples for purification are delivered preferably as dry films and are then redissolved using the minimum volume of DMSO required. Once the compounds are dissolved, an aliquot is taken for prepurification analytical LC/MS to determine the appropriate gradient and XIC thresholds for purification. On the basis of the analytical results obtained for a subset of the samples (5−10%), a method is created. For each sample analyzed, the retention times of the desired component and up to two of the closest-eluting side products are recorded. The purification gradient is calculated on the basis of the average of the retention times for the desired compounds, the average elution time difference between the nearest eluting side products and the product of interest. The acetonitrile composition of the average retention time for the desired products was calculated. The method was then generalized for the library by extending the gradient to encompass acetonitrile compositions of ±20% of the average composition calculated for eluting the compounds of interest.

For the determination of the appropriate purification method for the library, a recovery study is done to validate performance of the instrument. This evaluation is done before each set queue of 352 compounds to ensure acceptable purification recoveries and to enable the assessment of instrument robustness. Recovery studies done after several of the purification queues were typically comparable to the prepurification recovery studies, indicating slight or no degradation in system performance. Thus, it is only from the use of a nonindexed ion source that sample loss can be anticipated. With the flow streams merging at the tip of the ion source, losses can occur due to ion suppression and/or

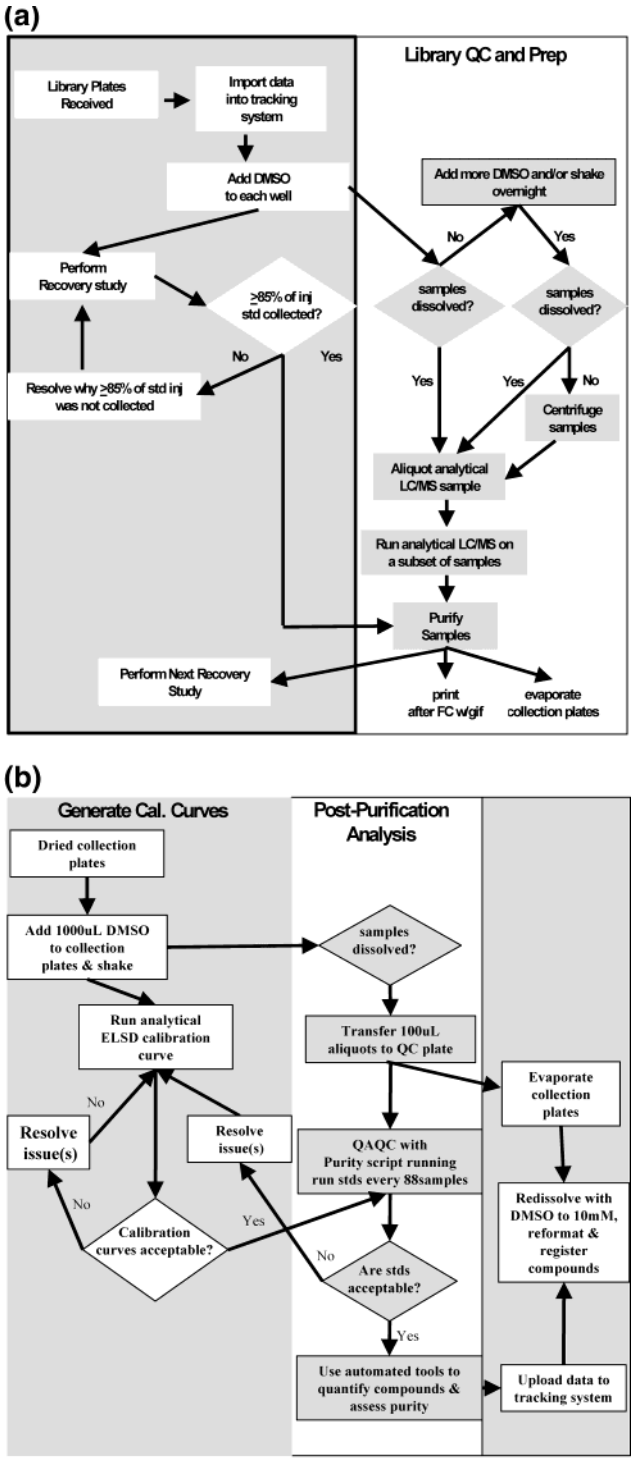


Figure 10. (a) Detailed flowchart for sample preparation and purification. (b) Detailed flowchart for postpurification sample processing.

due to the inability of the system to intelligently trigger the appropriate fraction collector when the two channels both have compounds with the same nominal molecular mass injected. The former has seldom been observed, and the latter was observed for just two or three pairs of wells per pair of plates purified. When the expected masses are the same for both columns, the fraction collection software is unable to reliably assign a peak of the desired ion to a correct channel, which may cause the loss of the desired compound. Such

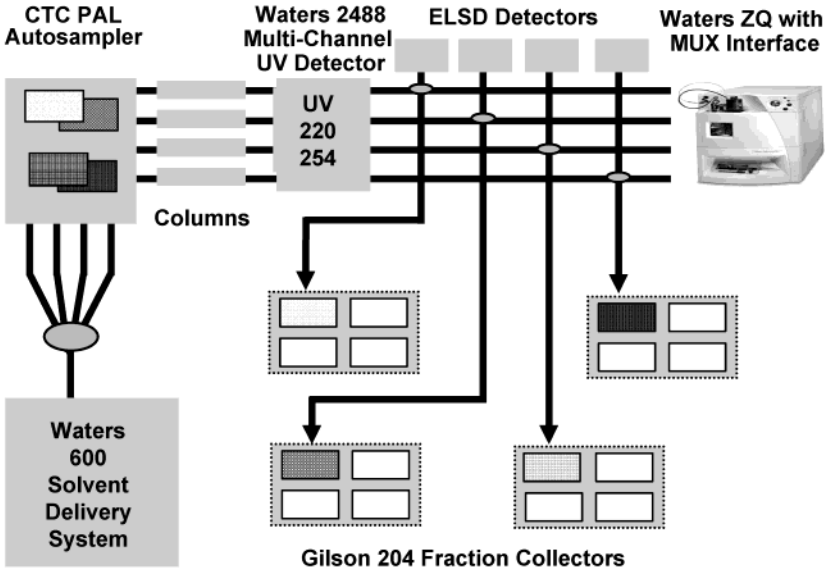


Figure 11. Schematic diagram of a four-channel parallel LC/MS purification system.

minor losses are acceptable given the significant throughput advantage of this parallel LC/MS system.

With fraction collection done using a plate mapping mode to facilitate sample tracking, each 96-well source plate purified generates either two or four 48-well collection plates. When one mass collection is performed, injections are made from a 96-well plate and the target mass is collected into the corresponding well of one of two 48-well fraction collector plates. Wells 1−48 of the source plate are collected into the first 48-well fraction collector plate, and wells 49−96 of the source plate are collected into the second 48-well fraction collector plate. When two fractions (or two target masses) are collected for each mixture injected, a total of four 48-well fraction collector plates are incorporated. The Gilson 204 fraction collector bed easily accommodates these four 48-well fraction collector (or destination) plates.

In many instances, it may be advantageous to collect the same target mass twice during the course of the purification. This is of particular interest when it is known that there are potential interferences in the library mixtures (i.e., isobars of the target species) or if the synthesis is expected (or known) to generate diastereomers that may be separated readily under RP-HPLC conditions. For two target masses (whether they are the same mass or unique masses), the first mass is collected into the corresponding well of the first two sets of 48-well fraction collector plates and the second target mass is collected into the corresponding well of one of the two remaining 48-well collection plates. There are a number of empty wells, so prior to biological screening, those fractions meeting or exceeding the purity and quantity requirements should be reformatted or condensed into master library plates for screening and the contents of these master library plates registered in a corporate database.

Because one cannot be certain that all collected compounds are in fact pure, the purity of each collected fraction needs to be assessed. The evaporation and redissolution steps after purification are to permit simultaneous purity assessment and postpurification ELSD quantification. Although the quantity of compound collected can be estimated using "on-the-fly" ELSD quantification, postpurification ELSD quantification was used because it directly estimates the material actually present in the collection well while the "on-the-fly" methodology indirectly estimates what should have been collected.

Impact of Number of Purification Channels on Overall Process Throughput. It was found that 11 days were required to completely process a library of 1208 reaction mixtures, significantly more time than the 96 h needed to purify the library by dual-column LC/MS. By use of this information, it was possible to make throughput estimates using a commercially available single-column LC/MS purification system.[26,27] In addition, the four-channel LC/(MUX)/MS system shown in Figure 11 has recently been developed in collaboration with Waters.[32,33] The comparison, shown in Figure 12, suggests that the data and liquid handling steps of the process impact the enhanced throughput realized using parallel LC/MS purification and analysis. For a library of 1408 reaction mixtures, it is estimated that doubling the number of LC/MS channels enhances total throughput achievable by one person with one instrument by approximately 30% and that throughput would approximately double by increasing the total number of LC/MS channels from one to four. The estimates in Figure 13 suggest it is possible for one person to use the purification process described herein with one four-column LC/MS system to process 4400 reaction mixtures from one or more libraries synthesized on a ≤ 20 μmol scale in 1 month.

It is expected that parallel LC/MS purification will become a mainstay in many labs particularly because of the availability of vendor-supported solutions[33] that may alleviate the need for custom components.

Conclusions

A high-throughput purification process was developed, validated, and implemented based on mass-directed fraction collection using a custom parallel two-channel LC/MS system. The complete process included all sample-handling steps after delivery of crude samples as dry films to final

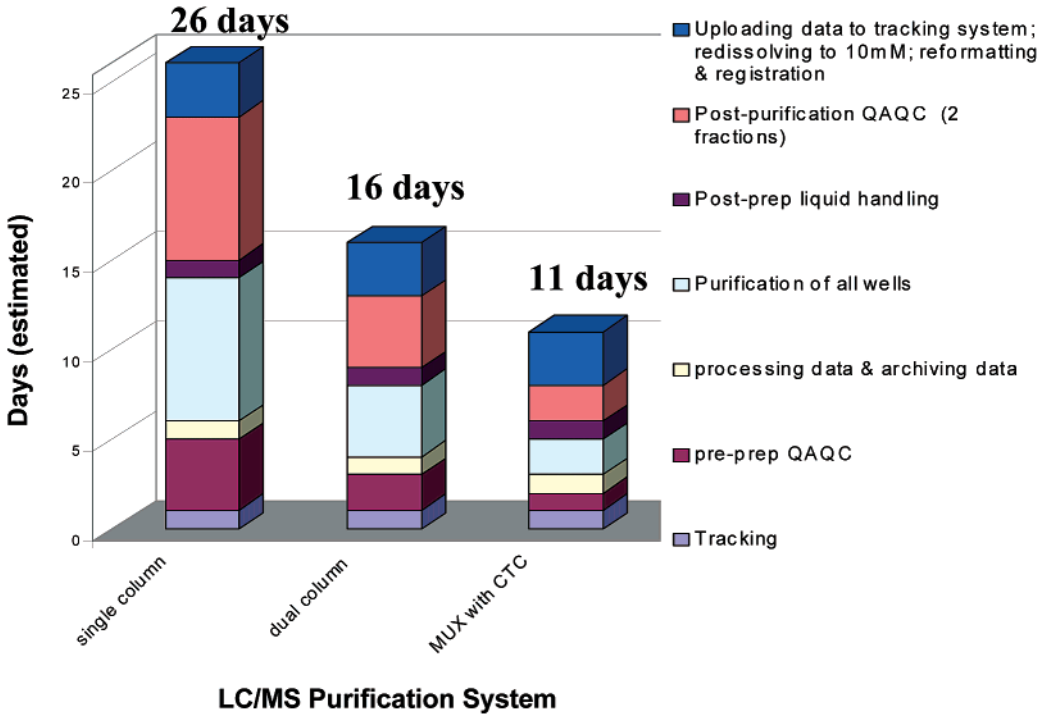


Figure 12. Estimated time for processing a 16-plate combinatorial library using a single-channel, a two-channel, and a four-channel LC/MS purification system.

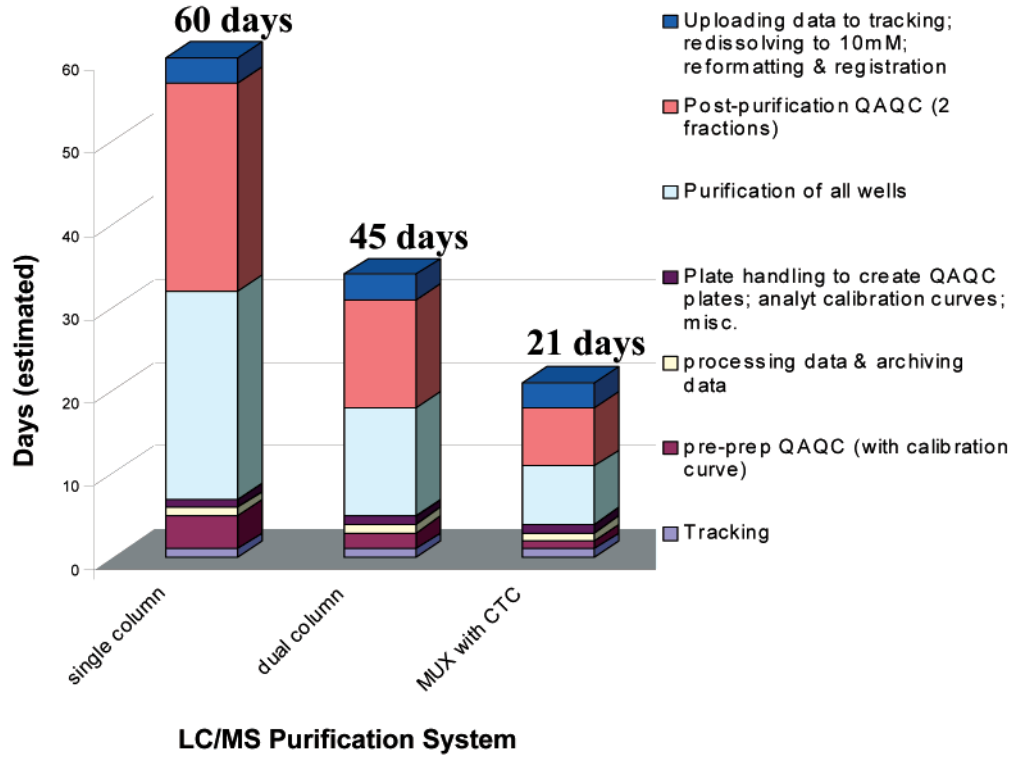


Figure 13. Estimated time for processing a 50-plate combinatorial library using a single-channel, a two-channel, and a four-channel LC/MS purification system.

compound reformatting and submission of the reformatted plate with the electronic files necessary for automated compound registration. With this system one scientist was able to completely process a 15-plate library in 11 days. Purification of this library was completed within 96 h, and the remaining time was required for sample handling. As one would expect, doubling the number of purification channels does double the throughput in the purification step, but other aspects in the process resulted in realized throughput gains of approximately 30%. Several areas for improvement were identified for postpurification sample handling, namely, the manner in which postpurification QC plates for quantification and purity assessment was generated and the resources allotted for routine plate reformatting.

It was demonstrated that a two-channel LC/MS system generally does not require a separate set of gradient pumps for each channel. We found that flow rate fluctuations of 10 % of the optimal flow rate (10 mL/min) had no statistically significant effect on sample recovery. Furthermore, the HPLC pressure changes associated with these flow rate deviations were not observed during the purification of 39 plates free of significant visible precipitate. Flow rate changes of up to 20% in analytical LC/MS were found to have a small but acceptable impact on ELSD quantification results.

Acknowledgment. The authors are grateful to Leo Raftogianis of AB/MDS SCIEX for providing us with a demo API-150EX. We also thank Andrew P. Combs, Tasir Haque, Albert Robichaud, and Loren Thompson for providing the 89 96-well plates for process validation and throughput determination. The authors acknowledge the assistance provided by Scott Harris, Ben Owen, Christine Rualo, and Bao Mai in automating sample-transfer steps. We also thank Tao Wang, Ingo Hardt, and Mark Suto for helpful discussions.

References and Notes

(1) Kassel, D. B. *Chem. Rev.* **2001**, *101*, 255−267.
(2) Myers, P. L.; Greene, J. W.; Saunders: J.; Teig, S. L. *Today's Chemist Work* **1997**, *6*, 46−53.
(3) Weller, H. N. *Mol. Diversity* **1999**, *4*, 47−52.
(4) Ripka, W. C.; Barker, G.; Krakover, J. *Drug Discovery Today* **2001**, *6*, 471−477.
(5) Peng, S. X.; Henson, C.; Strojnowski, M. J.; Golebiowski, A.; Klopfenstein, S. R. *Anal. Chem.* **2000**, *72*, 261−266.
(6) Cheng, S.; Comer, D. D.; Williams, J. P.; Meyers, P. L.; Boger, D. L. *J. Am. Chem. Soc.* **1996**, *118*, 2567−2573.
(7) Flynn, D. L.; Crich, J. Z.; Devraj, R. V.; Hockerman, S. L.; Parlow, J. J.; South, M. S.; Woodward, S. *J. Am. Chem. Soc.* **1997**, *119*, 4874−4881.
(8) Booth, R. J.; Hodges, J. C. *J. Am. Chem. Soc.* **1997**, *119*, 4882−4886.
(9) Booth, R. J.; Hodges, J. C. *Acc. Chem. Res.* **1999**, *32*, 18−26.
(10) Parlow, J. J.; Flynn, D. L. *Tetrahedron* **1998**, *54*, 4013−4031.
(11) Rabinowitz, M.; Seneci, P.; Rossi, T.; Dal Cin, M.; Deal, M.; Terstappen, G. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1007−1010.
(12) Johnson, C. R.; Zhang, B.; Fantauzzi, P.; Hocker, M.; Yager, K. M. *Tetrahedron* **1998**, 4097−4106.
(13) Boger, D. L.; Goldberg, J.; Andersson, C.-M. *J. Org. Chem.* **1999**, *64*, 2422−2427.
(14) Studer, A.; Hadida, S.; Ferritto, R.; Kim, S.-Y.; Jeger, P.; Wipf, P.; Curran, D. P. *Science* **1997**, *275*, 823−826.
(15) Baxendale, I. R.; Ley, S. V. *Bioorg. Med. Chem. Lett.* **2000**, *10* (17), 1983−1986.
(16) Bookser, B. C.; Zhu, S. *J. Comb. Chem.* **2001**, *3*, 205−215.
(17) Bhat, A. S.; Whetstone, J. L.; Brueggemeier, R. W. *J. Comb. Chem.* **2000**, *2*, 597−599.
(18) Breitenbucher, J. G.; Arienti, K. L.; McClure, K. J. *J. Comb. Chem.* **2001**, *3*, 528−533.
(19) Wang, T.; Hardt, I.; Kassel, D. B. *Rapid Commun. Mass Spectrom*. **2001**, *15*, 2067−2075.
(20) Berger, T. A.; Fogleman, K.; Staats, T.; Bente, P.; Crocket, I.; Farrell, W.; Osonubi, M. *J. Biochem. Biophys. Methods* **2000**, *43*, 87−111.
(21) Griffey, R. H.; An, H.; Cummins, L. L.; Gaus, H. J.; Haly, B.; Herrmann, R.; Cook, P. D. *Tetrahedron* **1998**, *54*, 4067−4076.
(22) Weller, H. N.; Young, M. G.; Michalczyk, S. J.; Reitnauer, G. H.; Cooley, R. S.; Rahn, P. C. Loyd, D. J.; Fiore, D. I.; Fischman, S. J. *Mol. Diversity* **1997**, *3*, 61−70.
(23) Schultz, L.; Garr, C. D.; Cameron, L. M.; Bukowski, J. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2409−2414.
(24) Hughes, I. *JALA* **2000**, *5* (2), 69−71.
(25) Kibbey, C. E. *Lab. Rob. Autom.* **1997**, *9* (6), 309−321.
(26) Zeng, L.; Wang, X.; Wang, T.; Kassel, D. B. *Comb. Chem. High Throughput Screening* **1998**, *1*, 101−111.
(27) Zeng, L.; Burton, L.; Yung, K.; Shushan, B.; Kassel, D. B. *J. Chromatogr. A* **1998**, *794*, 3−13.
(28) Kiplinger, J. P.; Cole, R. O.; Robinson, S.; Roskamp, E. J.; O'Connell, H. J.; Ware, R. S.; Brailsford, A.; Batt, J. *Rapid Commun. Mass Spectrom.* **1998**, *12* (10), 658−664.
(29) Zeng, L.; Wang, X.; Wang, T.; Kassel, D. B. Design of a Fully Automated Parallel Chromatography/MS System for the Analysis and Purification of Combinatorial Libraries. In *Proceedings of the 46th ASMS Conference on Mass Spectrometry and Allied Topics*, Orlando, FL, May 31 to June 4, 1998.
(30) Zeng, L.; Kassel, D. B. *Anal. Chem.* **1998**, *70*, 4380−4388.
(31) Isbell, J. J.; Xu, R.; Nemes, C.; Kassel, D. B. A Fully Integrated and Automated Approach to the Design, Synthesis, and Characterization of "Drug-Like" Libraries. In *Proceedings of the 48th ASMS Conference on Mass Spectrometry and Allied Topics*, Orlando, Florida, June 11−15, 2000.
(32) Kassel, D. B.; Xu, R.; Wang, T.; Brailsford, A.; Smith, B. To "MUX" or not to "MUX," That Is the Question. In *Proceedings of the 49th ASMS Conference on Mass Spectrometry and Allied Topics*, Chicago, Illinois, May 27−31, 2001.
(33) Xu, R.; Wang, T.; Isbell, J.; Cai, Z.; Sykes, C.; Brailsford, A.; Kassel, D. B. *Anal. Chem.* **2002**, *74*, 3055−3062.

CC0200336