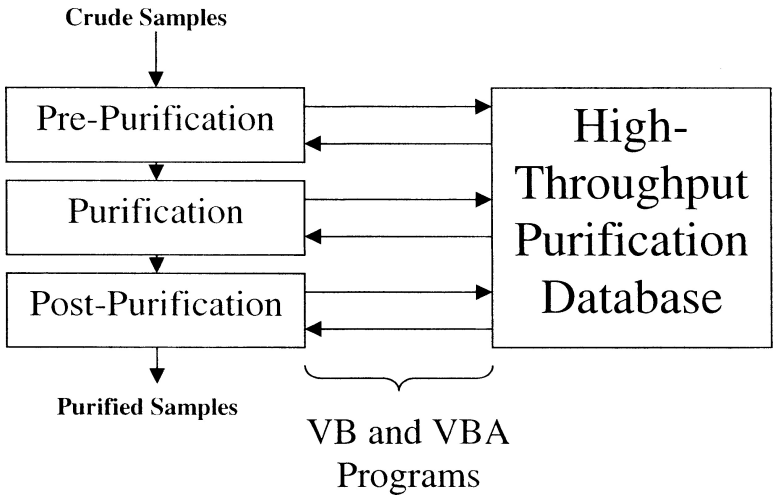
JOURNAL OF combinatorial CHEMISTRY

Article

Application of Visual Basic in High-Throughput Mass Spectrometry-Directed Purification of Combinatorial Libraries

B. Li, and E. C. Y. Chan

J. Comb. Chem., **2003**, 5 (6), 834-841• DOI: 10.1021/cc030032t • Publication Date (Web): 23 October 2003



More About This Article

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

- Supporting Information
- 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

Application of Visual Basic in High-Throughput Mass Spectrometry-Directed Purification of Combinatorial Libraries

B. Li and E. C. Y. Chan*

*S*BIO Pte Ltd., 1 Science Park Road, #05-09, The Capricorn, Singapore Science Park II, Singapore 117528*

Received March 28, 2003

We present an approach to customize the sample submission process for high-throughput purification (HTP) of combinatorial parallel libraries using preparative liquid chromatography electrospray ionization mass spectrometry. In this study, Visual Basic and Visual Basic for Applications programs were developed using Microsoft Visual Basic 6 and Microsoft Excel 2000, respectively. These programs are subsequently applied for the seamless electronic submission and handling of data for HTP. Functions were incorporated into these programs where medicinal chemists can perform on-line verification of the purification status and on-line retrieval of postpurification data. The application of these user friendly and cost effective programs in our HTP technology has greatly increased our work efficiency by reducing paper work and manual manipulation of data.

Introduction

The advent of combinatorial technique has had a tremendous impact on the way in which drug discovery research is approached. Combinatorial chemistry has been chiefly directed toward lead finding. However, its application to lead optimization has recently gained interest.[1,2] Automated parallel synthesis provides compounds suitable for early- to late-stage studies including biological activity screening, intrinsic activity comparison, physicochemical profiling, pharmacokinetic studies, toxicology studies, and screening in true disease models (in vivo testing). The main advantages of parallel synthesis (which includes solid and liquid phase) are its ability to generate a large number of discrete compounds in large quantities, its convenience to be performed in a microtiter plate format, and its ease for automation, such as the development of proprietary reactor blocks. Parallel synthesis therefore speeds up the whole process of compound production for lead discovery and lead optimization. Despite having the advantages listed above, parallel synthesis has similar shortcomings as the split and mix method, such as the expected compound not being pure or not synthesized to sufficient amounts. It is, however, important that these compounds are provided in sufficient amounts and sufficient purities to support the assay requirements. This is not unexpected since lead optimization demands the purest form of the chemical entities for reliable determination of structure−activity relationships and pharmacological properties. It is paramount to associate the activities and profiles of the chemical entity to the putative structure, instead of to an impurity or a degradant.

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. Many purification methods have recently been developed and applied to the purification of discrete compounds resulting from parallel solution or solid-phase synthesis using reverse phase chromatography coupled to mass spectrometry.[3−8] The introduction of mass spectrometry as a detection method for high-throughput purification (HTP) further improves the efficiency of the process. A liquid chromatography/mass spectrometry (LC/MS) system for automated purification provides a detect-before-collect mode of purification. This method utilizes real-time mass spectrometric ion signals to trigger fraction collection, and it eliminates the needs for excessively large fraction collection bed and postpurification analysis since only the mass of compound identified by the user in the data acquisition method is collected.

The essential properties of a commercially available LC/MS HTP system include automated injection and sampling, adequate fraction collection capacity, and accurate fraction tracking. The success of such a strategy is primarily determined by the intelligence of the software controlling the complete procedure and the coordination between software and hardware during the purification process. In our laboratory, the HTP of combinatorial parallel library involves two other important steps besides the purification process. In prepurification, the samples and their associated data (e.g., sample ID, crude amount, crude purity, exact mass, and solubility profile) are submitted by the medicinal chemists to the analytical chemists for HTP. During this phase, the analytical chemists collate the data prudently and perform random solubility testing and analytical high-performance liquid chromatography (HPLC) assays on these samples. Finally, a sample list is prepared for HTP. In postpurification, quality control HPLC assays are performed on the fractions collected and the samples are subsequently

* To whom correspondence should be addressed. Tel: +65 6827 5020. Fax: +65 6827 5005. E-mail: eric_chan@sbio.com.

10.1021/cc030032t CCC: $25.00
Published on Web 10/23/2003

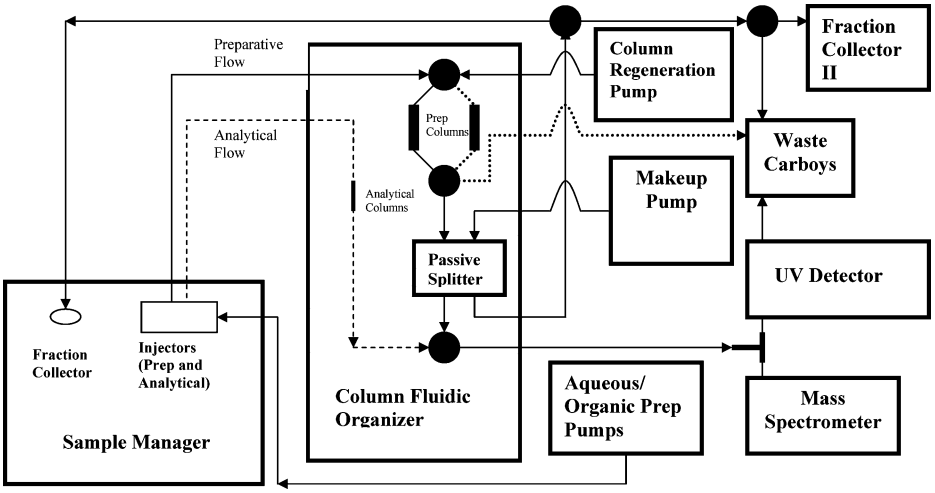


Figure 1. Configuration of Waters autopurification system.

dried. During this phase, the recovery and purity achieved for each sample are incorporated into a final report that is returned to the medicinal chemists, together with the samples. These data transfer processes involved during the pre- and postpurification steps are tedious, erroneous, and time-consuming if they are performed manually. It is not difficult to notice that the commercial software designed for the LC/MS HTP system is built mainly with the purification process in mind. It would be desirable if data submission and handling during pre- and postpurification could be controlled by the same software platform to reduce the manual transfer of data. A compound tracking system (custom designed by AXC Interactive Solutions), dedicated to tracking samples as they move through the purification process, has been reported.[9] This system, housed in an SQL server, is not commercially available, and the development cost of the program may be relatively high. To date, there is no single vendor software that addresses the needs in data handling for the pre- and postpurification steps.

In this paper, we describe the development of Visual Basic (VB) and Visual Basic for Applications (VBA) programs, using Microsoft Visual Basic 6 and Microsoft Excel 2000, respectively, and their applications for the seamless electronic submission and handling of library data. Functions were also incorporated in the programs where medicinal chemists can perform on-line verification of the purification status and on-line retrieval of postpurification data. The application of these programs in our HTP technology was found to be straightforward, user friendly, and cost effective. This approach to customize data submission and handling for HTP has greatly reduced paper work and rendered the HTP process more efficient.

Experimental Section

Materials. All reagents used were of HPLC grade. Acetonitrile and trifluoroacetic acid (TFA) were obtained from EM Science (Darmstadt, Germany) and Pierce (Rockford, IL), respectively. High-purity water was produced using a Millipore Milli-Q system (Bedford, MA). Luna C_{18} 5 μm columns were purchased from Phenomenex (Torrance, CA).

Instrumentation. Recently, a single vendor liquid chromatography electrospray ionization mass spectrometry (LC/ESI/MS) system was introduced, enabling rapid switching between analytical and preparative LC. The performance of this system for the isolation of minor impurities using LC/MS control mode was recently been evaluated and reported.[7] This LC/ESI/MS system was adopted in our HTP capability.

The autopurification system (Waters Corporation, Milford, MA) system consisted of a 2767 one bed injection−collection Sample Manager, a 2525 binary gradient module or high-pressure LC pump, a column fluidic organizer (CFO), a 2487 dual wavelength UV/vis detector, a ZQ single-quadrupole MS equipped with a Z-spray electrospray interface, a 515 pump for column regeneration, a reagent manager for the makeup flow, and a Waters fraction collector II as a waste solvent manager. The CFO also held a 1:10 000 ACURATE passive flow splitter from LC Packings (Amsterdam, The Netherlands). The mobile phases were water (0.05% TFA) and acetonitrile (0.05% TFA). Two preparative columns were used at one time for HTP using column switching. The system was controlled by Micromass MassLynx (version 4.0) and FractionLynx software. A schematic diagram of this system is illustrated in Figure 1.

The electrospray source was set at 3.0 kV capillary voltage, 30 V cone voltage, 110 °C source temperature, 175 °C desolvation temperature, 250 L/h desolvation gas flow, and 50 L/h cone gas flow. For the analyzer, the multiplier was set at 650 and 450 for analytical and preparative tune methods, respectively.

Program Development. Microsoft Visual Basic 6.0 and Microsoft Excel 2000 were used in this study to develop the HTP programs. These programs were designed to interact with MassLynx (version 4.0) software and Mettler Toledo BalanceLink version 3.01 (Greifensee, Switzerland), with the aim to enable an efficient flow of library data throughout the purification process including prepurification, purification, and postpurification. As the purification step is efficiently controlled using the vendor software, our programs were developed mainly for the pre- and postpurification steps. The VB programs developed for the prepurification step consisted of three functions: the submission of library data by medicinal chemists, the uploading of submitted data to HTP database, and the extraction of data from the same

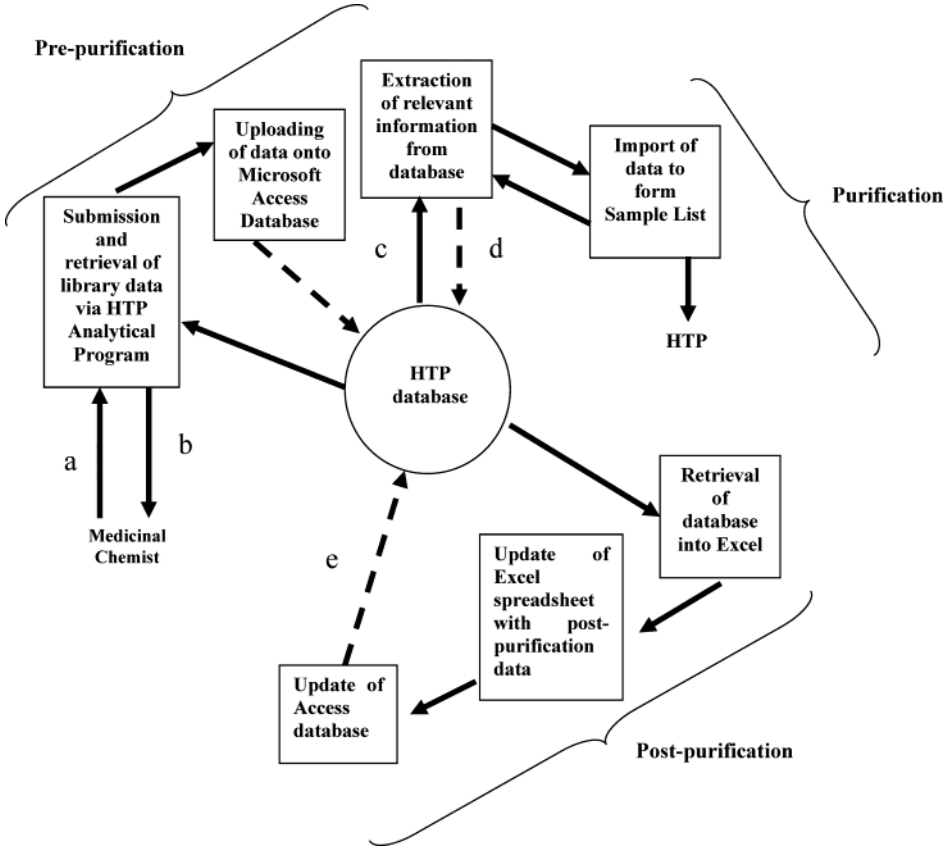


Figure 2. Schematic diagram of the flow of library data throughout HTP. (a) Submission of library data by medicinal chemist, (b) retrieval of final purification data, (c) extraction of data from HTP database to form sample list, (d) update of HTP database with instrumental and experimental data, and (e) update of HTP database with postpurification data.

database to build a sample list for HTP. The VBA programs developed for postpurification consisted of three main functions: the retrieval of the HTP database to Excel spreadsheet, the update of Excel spreadsheet with post-purification data such as purity, recovery, and mass of fraction collected, and the final update of HTP database. A schematic diagram of the flow of library data throughout HTP is illustrated in Figure 2.

Results and Discussion

Prepurification. A VB program comprised of user friendly interfaces was created using Microsoft Visual Basic 6.0 for the submission of library data. This standalone program, written using Microsoft Visual Basic 6.0, resides in a designated computer (AC2) in the analytical laboratory. The individual chemist, however, is able to assess this program on their personal computer using a shortcut key. The data submitted using these interfaces are grouped using the following fields: name of chemist, laboratory notebook number, project number, number of samples, sample identification number, crude amount (crude weight of sample), crude purity (crude purity of compound of interest measured at UV 254 nm), exact mass, solubility profile, and preference for dried or undried sample. Figure 3 illustrates some of the interfaces for library data submission. Each submission, regardless of the number of samples, is assigned automatically a unique request ID (example, HTP_03_6). Before data submission, an option is provided for the chemist to preview and/or print a table summarizing the data of the submission (Figure 3). Further amendments of data and cancellation of submission are also allowed at this stage. Once the summary table is checked and printed for documentation, the chemist selects the Submit button and automatically uploads the data to the HTP database via network communication. The HTP database in Microsoft Access 2000 resides in the hard drive of computer AC2 in the analytical laboratory. The data in the HTP database are categorized based on the unique request ID previously assigned to each submission. Additional fields are automatically created in the database to accommodate other subsequent data derivable from the HTP process such as instrumental methods and postpuriifcation data. These fields will be elaborated in the subsequent discussion.

The next step in prepurification programming involves the extraction of relevant data from the HTP database to generate a file that can be imported as the sample list in MassLynx (version 4.0). This step is controlled using a VB program that interacts with the HTP database. This standalone program, written using Microsoft Visual Basic 6.0, resides in computer AC2, but it is accessible via a shortcut key in the HTP computer where the MassLynx (version 4.0) software is housed and the autopurification system is controlled. The VB program provides a simple interface (Figure 4) where the analytical chemist keys in a unique request ID and upon activating the OK button, the relevant data identified by the request ID are transferred from the Microsoft Access database to form a Microsoft Access

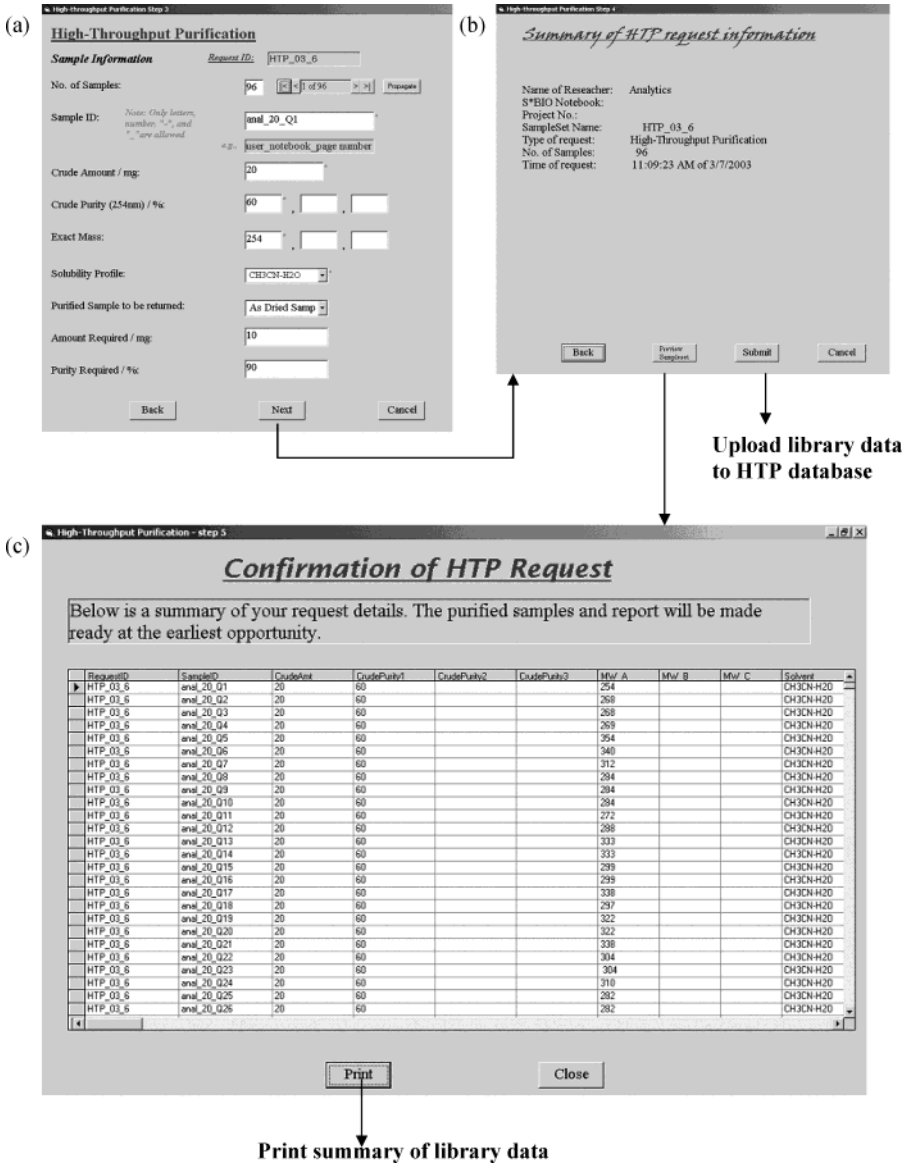


Figure 3. Selected screen captures of the library data submission process. (a) Library data submission interface, (b) brief summary of HTP request indicating the unique request ID, and (c) printable summary of library data.

Application file in the hard drive of the HTP computer. The MS-DOS name of this file is automatically designated as "request ID.mdb" where the request ID is previously assigned for each HTP library data submission. Sample lists can be created in a number of packages and imported into MassLynx (version 4.0).[10] Some of the files that are allowed to be imported include Access 97, tab and comma delimited text files, Excel 97, and OpenLynx Batch files. To import a worksheet in our application, the Menu Bar File and Import Worksheet commands are selected sequentially to invoke the Import Worksheet dialogue. The "request ID.mdb" file is selected, and the Open button is finally activated to import the worksheet. The data extracted from the HTP database include File Name (sample identification number), Mass A to C (exact masses of compounds), and Sample Location (injection position). The name and data type of these fields in the Microsoft Access Application file are matched to those in the sample list of the MassLynx (version 4.0) software so as to ensure that the worksheet is suitable for direct import. Additional fields such as Inlet File, MS File, Inject Volume, Fraction File, and Fraction Trigger (1−3) are also included in the table of our HTP database to allow subsequent update of this information. The fields finally used in the sample list in MassLynx (version 4.0) for purification are File Name, Sample Location, MS File, Inlet File, Inject Volume, Mass A, Mass B, Mass C, Fraction File, Fraction Trigger 1, Fraction Trigger 2, Fraction Trigger 3, and MS Tune File.

Purification. As elaborated, the first step in purification involves the use of the MassLynx (version 4.0) software to import the Microsoft Access Application file or worksheet, created during prepurification, to form the sample list (Figure 4). This step in importing the Microsoft Access Application file represents the seamless transfer of library data from our HTP database to the vendor software interface for HTP. The sample list in the MassLynx (version 4.0) software depicts the sample information (File Name, Sample Location, and MS File), instrumental methods (Inlet File, MS File, and Fraction File), and experimental parameters (Fraction Trigger, MS Tune File, and Inject Volume) required to command the purification. Unlike the imported sample

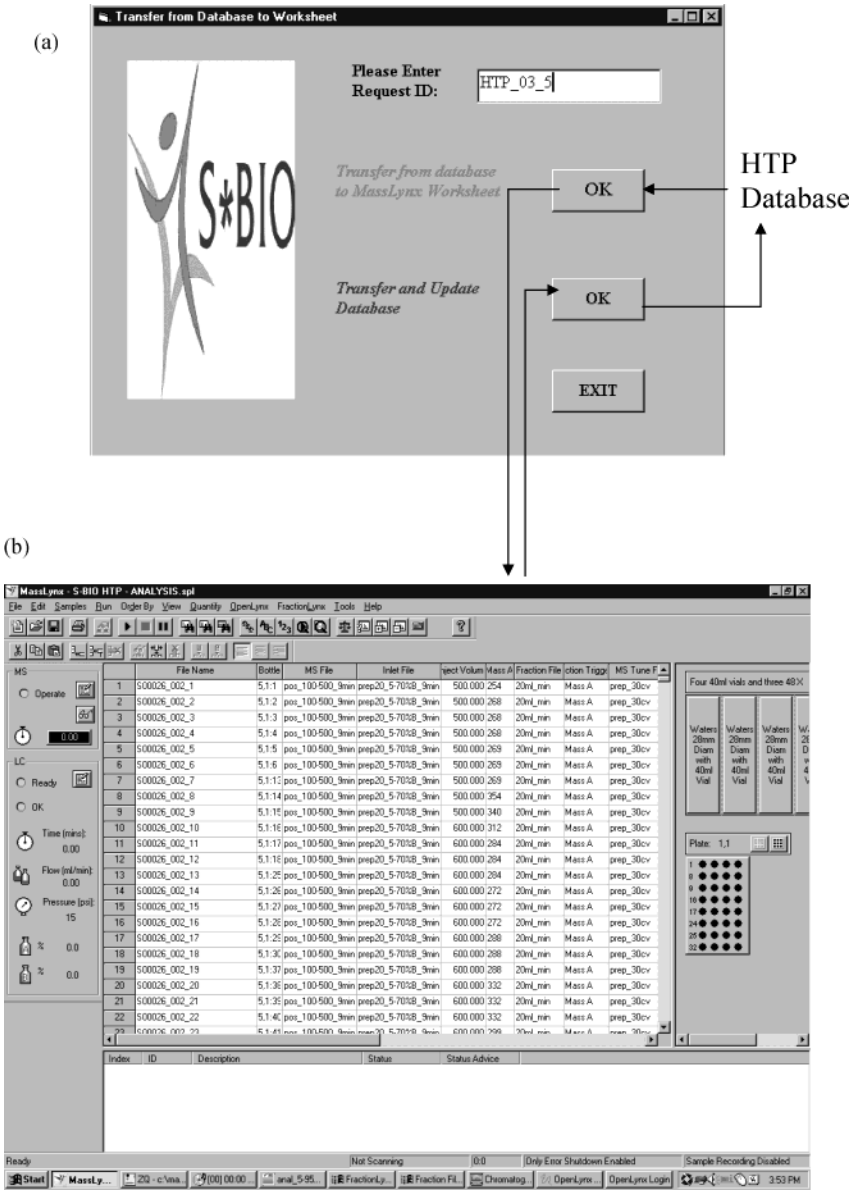


Figure 4. Screen captures of the import/export process of data via interaction with MassLynx (version 4.0) software. (a) VB program interface for the transfer of data to/from the HTP database and (b) sample list in MassLynx (version 4.0) software.

information, instrumental methods and experimental parameters are entered manually in the sample list. The optimal instrumental settings that define the HPLC, MS, and fraction collection methods for HTP are predicted based on analytical HPLC of the samples. Prior to HTP, all of the parallel library compounds are analyzed by analytical LC/ESI/MS using an open-access LC/MS instrument. Three important details are derived from these analytical data. First, the elution profile of the compound and its resolution from other components are determined and this allows the analytical chemist to predict the scale-up of analytical-to-preparative chromatography. Second, the ion intensity of each compound is determined and this helps to estimate the optimal threshold (ion intensity) for fraction collection. Third, the crude purity of the compound is determined and allows subsequent estimation of the recovery of the purification.

As mentioned, HTP is performed using the autopurification system under the control of MassLynx (version 4.0) and FractionLynx software. A diagram of the autopurification system is shown in Figure 1. Solvent gradients are formed with the 2525 Binary Gradient module capable of flow rates between 0.5 and 150 mL/min and up to 6000 psi pressure. The flow rates commonly adopted in our HTP are 20 and 40 mL/min. The makeup and column regeneration pumps are the Waters reagent manager and 515 pump, respectively. The flow rate of the reagent manager is 1 mL/min with methanol, and the flow rate of the 515 pump is 3 mL/min with a mobile phase of 95%:5% of water:acetonitrile with 0.05% TFA. Reequilibration is effected using the column regeneration pump. The reequilibration conditions and wash cycles were optimized during method development to reduce the column memory or carry-over effect to a negligible level. Sample injection/fraction collection is performed with the 2767 sample manager. The preparative injector is configured with a 5 mL sample loop, and the two columns on the preparative selection valve are 21.20 mm $\times$ 50 mm Phenomenex Luna 5 μm C_{18} columns. The analytical injector is configured with a 20 μL sample loop, and the column on

the analytical selection valve is 2.0 mm × 50 mm Phenomenex Luna 5 μm C_{18} columns. The analytical/preparative switching and preparative column switching are accomplished with the CFO. The ACURATE passive flow splitter is used to split the preparative flow between the detectors (UV and MS) and the fraction collector or waste carboy. Waters fraction collector II is set up as an optional fraction collector to collect waste solvent into individual 500 mL container per sample injection. This set up allows retrieval of compounds that missed collection and circumvents the permanent loss of precious sample into the waste carboy. The cycle time of each preparative HPLC is 9 min. The approach adopted by the autopurification system involves column switching to enable one preparative column to reequilibrate while a separation occurs in another column (Figure 1). The throughput of our HTP is 15 h for a batch of 96 samples. Although there are recent reports on higher purification throughput achieved via the use of multiple array of columns[7,11] and multiplexed ion sources,[7] the throughput that we achieved using the column switching technology is sufficient to meet our needs in combinatorial parallel library generation.

Data transfer during purification is a two way process (Figure 2). As mentioned in prepurification, additional fields dedicated to capture instrumental methods and postpurification data are created in the HTP database. Upon completing purification, a VB program is used to update these selected fields automatically with the HPLC, MS, and fraction collection methods, the injection volume, and the fraction trigger parameter. This standalone program, written using Microsoft Visual Basic 6.0, resides in computer AC2 but is accessible via a shortcut key using the HTP computer. The analytical chemist copies and transfers the data from the sample list in MassLynx (version 4.0) to the appropriate Microsoft Access Application file in the hard drive of the HTP computer. To execute the update of the HTP database, the chemist enters the Request ID, specific to each library submission, and activates the appropriate OK button on the VB program (Figure 4). The principle of this program is to update the HTP database based on the information in the Microsoft Access Application file. Another action triggered simultaneously by the above action is the update of the status of the purification from "Queue" to "Finish" in the HTP database. The chemist can check the status of their purification submission by selecting the Check Results button on the VB program (Figure 5).

Postpurification. In the first step of postpurification, Microsoft Query in Microsoft Excel 2000 was used to set up data sources in a computer designated for postpurification weighing. A query was created using the Query Wizard in Excel to retrieve data that is stored in a relational database. In our application, this relational database is the HTP database in Microsoft Access 2000. Before the HTP database was retrieved to Sheet1 ("Access Database" worksheet) in Excel, criteria were specified in the query to include only a desired set of fields so as to exclude data that are not relevant to the postpurification process. Some desired fields are Request ID (ID specific for each HTP submission), Sample ID, Crude Amount (crude weight in mg of the sample), Crude Purity1−3 (crude purity of each fraction of interest), and Mass A−C (exact mass of each fraction of interest). This query was created only once, and subsequently, the data in the "Access Database" worksheet in Excel are updated automatically by selecting the Refresh Data command in Microsoft Excel 2000.

In the next step of postpurification, a VBA program is written using Visual Basic Applications for Excel 2000 to transfer selected data in the "Access Database" worksheet to Sheet2 ("Working" worksheet) in the same Excel file where postpurification data can be appended. This process is executed by entering the specific Request ID in the Text Box and subsequently activating the Go command button in the VBA program, which is presented as an interface on the "Access Database" worksheet. In the "Working" worksheet, a number of additional fields are added for updating the purification results, namely, Amount Returned 1−3 (weight of each purified fraction in mg), Purity Returned 1−3 (purity of each fraction as determined using UV at 254 nm), Recovery 1−3 (recovery of each compound), Mass Found 1−3 (molecular weight of compound purified in each fraction), Date Returned (Date purified compound is returned to chemist), and Sample ID Return (Analytical Chemistry ID). Through the use of BalanceLink software, the AG204 balance and the LC-RS9 communication cable from Mettler Toledo (Greifensee, Switzerland), the tare weight of the 20 mL scintillation vial, and the gross weight of the vial with dried sample are easily measured. These values are transferred automatically to Sheet3 ("Weighing" worksheet) in the Excel file using simple command keys. The final weight of each purified fraction is computed using a simple mathematical formula in Excel (Gross Weight − Tare Weight), and the value is copied and transferred to the Amount Returned field of the "Working" worksheet. The final purity of each purified fraction, determined using quality control analytical HPLC/UV (254 nm), is also entered suitably in the Purity Returned field of the "Working" worksheet. The percentage of recovery is automatically generated in the "Working" worksheet using a mathematical formula written in Excel based on the following equation:

$$\text{percentage recovery} = (\text{final purity} \times \text{final weight})/(\text{crude purity} \times \text{crude weight}) \times 100$$

Both the final purity and the computed recovery of the purified compound provide valuable information to the analytical chemist in measuring the performance in batch purification of combinatorial parallel library. Finally, the analytical chemist enters the analytical chemistry ID in the Sample ID Return field of the "Working" worksheet. This analytical chemistry ID, comprising the name of chemist, notebook number, and page number, introduces traceability to the HTP experiment.

Most commonly, one fraction is collected for each preparative HPLC. However, in some unique cases, two or more fractions containing compounds of differing purity and identity may be collected. This may occur sometimes when a coeluting compound such as the starting material is desired and collected or when a peak is collected as several fractions. This gives rise to structural differences in the data and renders

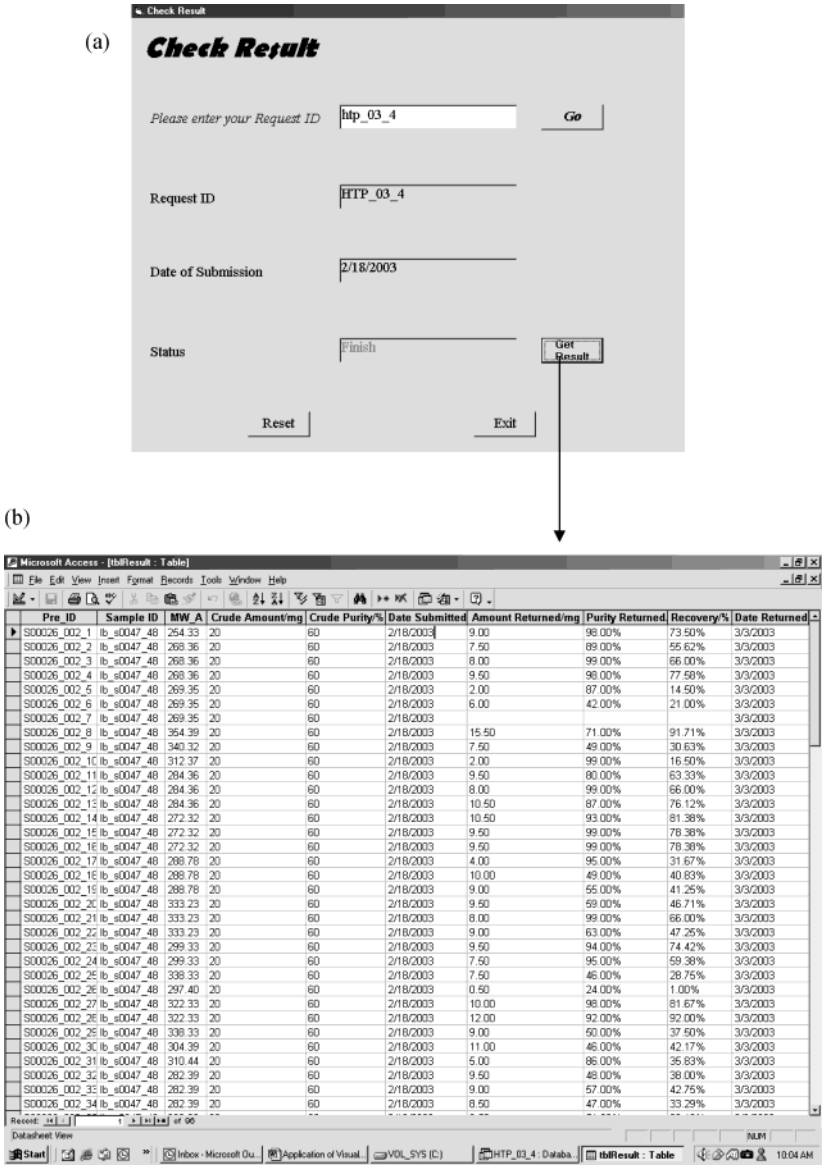


Figure 5. Screen captures of the data retrieval process. (a) VB program interface for checking status of purification and retrieving purification results and (b) an example of the result table retrieved from the HTP database.

the data handling process difficult. To address this difference in postpurification data, a VBA program is written using Visual Basic for Excel 2000 to reformat selected data in the "Working" worksheet to a desirable format. This program, activated by entering the unique Request ID and selecting the Reformat button on the "Working" worksheet, converts multiple fractions and their associated fields from their original single row sequential format to the multiple row single column format. The latter format facilitates the transfer of data (example, copy and paste of final weight of each purified fraction). This simple program reduces the manual manipulation of data structure, a method we previously adopted, that proves to be tedious and erroneous.

When the Excel spreadsheet is completed with postpurification data, we use the VBA program (residing as an interface on the "Working" worksheet) to perform two additional tasks. The first task is to transfer relevant data from the "Working" worksheet to Sheet4 ("DB Transfer" worksheet) in Excel where the information could be copied and appended to the HTP database with no further modification. The second task is optional, and it is to transfer relevant data from the "Working" worksheet to Sheet5 ("Result" worksheet) where the postpurification results can be printed as a hard copy for the chemist (submitter). Alternatively, the medicinal chemists interact with the HTP database via a VB program interface on their personal computers (Figure 5). By entering the unique request ID, each chemist then determines the status of the purification. When the status is "queue", it refers to an uncompleted HTP job. When the status is "finish", the chemist is provided with an option to download the results to the local drive of his/her computer where the report would be evaluated and printed for documentation.

Conclusions

An approach and concept using VB and VBA programs to customize the submission and handling of library data for HTP using prep/LC/ESI/MS is described. Manual transfer of data, previously adopted in our laboratory, proved to be

tedious, time-consuming, and erroneous. After incorporating these cost effective programs, the efficiency of HTP is enhanced by the improved data flow throughout the prepurification, purification, and postpurification stages. The development of these programs may appear to be complicated. However, for the medicinal and analytical chemists, the applications of these user friendly VB and VBA interfaces for the submission, handling, and retrieval of library data are relatively simple and straightforward. Furthermore, the development and maintenance costs of these VB and VBA programs are low and manageable. We have also utilized these programs for the submission of single compound or small batch compounds, synthesized routinely by our medicinal chemists, for HTP by prep/LC/ESI/MS. In the future, the applications of these programs may be expanded to encompass other analytical services in our company.

Acknowledgment. We thank S*BIO Pte Ltd. and Ms. Yap Swee Lee of Waters Asia Limited for their support in this project.

References and Notes

(1) Weller, H. N.; Young, M. G.; Michalczyk, S. J.; Reitnauer, G. H.; Cooley, R. S.; Rahn, P. C.; Loyd, D. J.; Fiore, D.; Fischman, S. J. *Mol. Diversity* **1997**, *3*, 61−70.
(2) Weller, H. N. *Mol. Diversity* **1999**, *4*, 47−52.
(3) Zeng, L.; Kassel, D. B. *Anal. Chem.* **1998**, *70*, 4380−4388.
(4) Zeng, L.; Burton, L.; Yung, K.; Shushan, B.; Kassel, D. B. *J. Chromatogr. A* **1998**, *794*, 3−13.
(5) Isbell, J.; Xu, R.; Cai, Z.; Kassel, D. B. *J. Comb. Chem.* **2002**, *4*, 600−611.
(6) Xu, R.; Wang, T.; Isbell, J.; Cai, Z.; Sykes, C.; Brailsford, A.; Kassel, D. B. *Anal. Chem.* **2002**, *74*, 3055−3062.
(7) Niessen, W. M. A.; Ling, J.; Bondoux, G. C. *J. Chromatogr. A* **2002**, *970*, 131−140.
(8) Ripka, W. C.; Barker, G.; Krakover, J. *Drug Discovery Today* **2001**, *6*, 471−477.
(9) Schultz, L.; Garr, C. D.; Cameron, L. M.; Bukowski, J. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2409−2414.
(10) Waters Corporation. *MassLynx Version 4.0 Software Manuals*; Waters Corporation: Milford, Massachusetts, 2002.
(11) Edwards, C.; Hunter, D. J. *J. Comb. Chem.* **2003**, *5*, 61−66.

CC030032T