CHIMIA 2000, 54, No. 1/2

Chimia 54 (2000) 37–40

ISSN 0009–4293

Focal Point: Medicinal Chemistry

ILMAC Congress: October 14, 1999

High-Throughput Analysis, Purification, and Quantification of Combinatorial Libraries of Single Compounds*

Keywords: Combinatorial chemistry · High-throughput analysis · ILMAC · Medicinal chemistry

The possibility of rapid analysis and purification of combinatorial libraries of individual compounds in multi-milligram quantity has opened new perspectives for the effective application of combinatorial technology in medicinal chemistry for lead optimization, new lead finding and compound archiving.

High-throughput analysis and purification offers medicinal chemists the possibility to evaluate precisely the success of their combinatorial synthesis, *i.e.* in terms of purity and yields! This might also contribute to the prevention of the painful experience of false positive hits and less evident issues with false negative hits. Finally the necessity to draw structure-activity relationships is, for many biological assays, only realistic with compounds of quality and adjusted concentration!

The objective of the mini-symposium was to illustrate the state-of-the-art of high-throughput analysis and purification methods, and to outline ongoing and near future developments of this technology, which has made impressive progress over the past few years. Three pioneers in this technology had the opportunity to present and discuss their views, strategies, and set-ups for carrying out high-throughput analysis and purification of large arrays of compounds.

Combined Use of HPLC, MS, and NMR for High-Throughput Analysis, Purification, and Quantification

Derek P. Reynolds[a], Keith Brinded[a], John Hollerton[a], Stephen J. Lane[a], Andrea Seffler[b], and Kenneth Lewis[b]

[a]Physical Sciences Unit, Medicinal Sciences Division, GlaxoWellcome R&D, Stevenage, Hertfordshire, UK

[b]Analytical Chemistry Department, GlaxoWellcome Inc., Research Triangle Park, North Carolina, USA

For the analytical chemist the development of parallel synthesis continues to pose new challenges. In the past, analytical characterization of combinatorial libraries was often cursory and confined to analysis of part of the library by mass spectrometry. Today in GlaxoWellcome it is appreciated that it is essential to have more rigorous quality control and it is standard practice to analyze all libraries of single compounds by LC/UV/MS. In lead optimization the need for reliable structure activity relationships has led to the introduction of additional techniques for compound quantification performed in parallel to the LC/MS measurement (*e.g.* NMR or CLND). To improve sample quality, libraries can now be conveniently purified using commercially available automated HPLC and LC/MS systems that accommodate microtiter plate-based sample formats and operate in an automated mode with computerized fraction tracking.

At the present time bottlenecks in the overall synthetic process often occur before robotic synthesis (*i.e.* in library rehearsal) or in post-synthesis purification. We have therefore tried to develop a flexible 'toolkit' of high-throughput methods that exploit 'generic' gradient HPLC in combination with various detectors (MS, UV, ELSD, CLND) and complementary parallel techniques. By making as many techniques as possible 'open-access' to synthetic chemists, we enable them to have immediate access to the data they need for rapid reaction scanning. Compatible preparative HPLC and HPLC/MS systems allow them to automatically translate the results from analytical HPLC into conditions for automated library purification.

A common feature of the open-access systems deployed at GW in Stevenage and RTP is a simple PC interface that assumes no knowledge on the part of the chemist of the instrument hardware. Each user types in his name and sample identifiers, and selects the required methods from a menu. The computer specifies the correct positions in the autosampler for placement of his sample tubes or plates in a queue for analysis. After acquisition, the data is automatically archived over the computer network and a report is automatically printed or e-mailed back for viewing on a normal desktop PC. Current open-access systems include NMR, HPLC/UV/MS, GC/MS, HPLC/UV, and IR. For many of these systems it is now possible to view instrument and sample status over the Intranet using a standard Web Browser.

Fast 'generic' gradient HPLC methods are key to the success of this approach. It is important to realize that there is always a compromise between speed and resolution in HPLC. At present, cycle times of about 5 min are typical, and shorter run times will only produce adequate separations if short high efficiency columns are available and used at the highest practicable flow rates with modern HPLC equipment with very low mixing volumes and low dead volumes. We operate on various scales and have a number of variations available, but the same stationary phase/mobile phase combination is used whenever possible to assist method transfer.

LC/UV/MS is a powerful combination, however, it suffers from the problem that relative response factors are not to be relied upon, and therefore %purity figures are at best approximate and

*Organized by:
Dr. Rudolf Giger
Novartis Pharma AG
S-507.4.05
CH-4002 Basel
Tel.: +41 61 324 81 95
Fax: +41 61 324 42 36
E-Mail: rudolf.giger@pharma.novartis.com

CHIMIA 2000, 54, No. 1/2

accurate solution concentration can only be determined if standards are available. The situation can be rectified if a quantitative technique is run in parallel to the LC/MS. We have implemented a system where flow-inject NMR is combined with HPLC/UV/MS and the concentration of the samples is determined from the NMR data. An alternative for nitrogen-containing compounds is the Chemiluminescent Nitrogen Detector (CLND). An HPLC-compatible version (methanol eluent) is available from Antek and produces a response proportional to the amount of nitrogen in a compound. This detector can also be deployed in parallel with LC/UV/MS and, like NMR, is capable of absolute quantification on a routine basis to an accuracy of about ±5%. NMR has lower sensitivity than CLND but has the advantage of providing complementary structural information.

With all these methods when applied to large libraries, data interpretation becomes a time-consuming task and paper reports are unwieldy and inefficient. Data visualization software is now available to all our chemists by download from the Intranet. We have chosen as far as possible to use commercially available packages, *e.g.* Diversity Browser from Micromass for HPLC/UV/MS data, and SpecMan from ACD for NMR. There is a real need for improvements here, especially for software that can assist in the interpretation of data from multiple techniques (*e.g.* MS, NMR, HPLC, CLND). In-house we have developed a database application (Spectrum Finder) that on input of a sample i.d. will retrieve archived spectra together with chemical structures from the corporate database.

Library purification can be performed when required using automated preparative HPLC. The Fraction Lynx system from Micromass allows mass-directed purification so that an 80-well plate can be directed to an 80-well output format based on molecular ion detection and all unwanted peaks are rejected automatically. This approach is feasible for small-scale purification (1–10 mg). Most problems arise from inadequate compound solubility and the success rate is improved markedly by the use of 1:1 DMSO/MeOH as injection solvent (500 µl) and by selection of a preparative gradient based on the analytical chromatogram that uses as high a percentage of organic mobile phase as possible for the starting solvent. Our standard is a 10 min gradient, but higher throughputs are possible, and more generic gradients can be used if the scale is reduced. Larger quantities are generally separated using preparative LC/UV systems. The Gilson systems with Autopoint software have fraction tracking facilities and can be used for purification of small libraries, but fraction management and identification of the correct compounds is tedious for large numbers because all peaks detected by UV are collected.

The current focus of our work is more on improving the quality of libraries and the quality of our analytical information than on increasing sample throughput. However, the productivity of synthetic chemists continues to increase, and new advances in HPLC (*e.g.* high pressure pumps, new 1–2 µm HPLC packings) and mass spectrometry (*e.g.* TOF MS, parallel electrospray interfaces) will inevitably lead to faster runtimes and higher throughputs. As this happens there will be an ever increasing need for 'intelligent' software to assist with interpretation of data from a multiplicity of complementary techniques.

Application of Preparative HPLC to Purification of Automated Synthesis Products in Drug Discovery

Harold N. Weller
Combinatorial Discovery and Instrumental Development Group, Bristol-Myers Squibb Pharmaceuticals Research Institute, Princeton, NJ 08543, USA

The discovery and development of a new drug, from inception to market, is a process that can take ten or more years including drug discovery, development, clinical trials, and regulatory approval [1]. The period of market exclusivity, meanwhile, is limited by patent protection which extends for 20 years from the time of patent application in major markets such as the United States. Since patents are filed early in the drug discovery process, any delay in the timeline from discovery to market will result in a corresponding loss of exclusive marketing for that drug. Blockbuster drugs today commonly have gross sales in excess of $1 billion per year, thus the cost (or benefit) of an increase (or decrease) in time to market can be valued similarly.

The earliest part of the overall process is *drug discovery*. During this phase, an initial biochemical concept is tested using many thousands of new compounds. Once initial biochemical activity is observed for a 'lead compound', its structure is optimized through an iterative process for both intrinsic activity and secondary activities until a drug candidate is nominated to proceed into development. This overall discovery process can take from two to ten years, and requires the synthesis of thousands of new compounds for testing. Traditionally, these new compounds have come from individual manual synthesis at a rate of 25 to 50 new compounds per chemist per year. In the last five years, automation technologies and parallel synthesis have enabled chemists to synthesize new compounds at a rate in excess of 1000 compounds per chemist per year. To achieve this rate of synthesis, a suite of integrated tools for design, synthesis, analysis, purification, and information handling is required. At BMS we have developed and/or deployed a suite of tools including the MiniBlock™ Synthesizer and the IRORI Accutag™ system for synthesis, the Discovery-VP™ Chromatography system for high-throughput analysis and purification, and the BMS Synthesis Workbook for synthesis information handling.

In developing tools to facilitate synthesis, analysis, and purification, it quickly became apparent that post-synthesis processing will be most efficient if done identically regardless of the synthesis method. This, in turn, requires that products be delivered into the same vessel and format regardless of synthesis method. To facilitate this streamlined post-synthesis processing, we developed the MiniBlock™ Synthesizer. The MiniBlock is a 48-position reaction block that can be used for both solution- and solid-phase synthesis under a wide variety of conditions, including high and low temperature and under inert atmosphere. The MiniBlock delivers products into racks of 96 custom microtubes containing up to 2.5 ml volume. Since we also use the IRORI AccuTag™ system [2], a modified AutoSort™ device was developed to deliver MicroKans into the MiniBlock for cleavage. Thus, products are delivered into the same format (96-well microtubes) regardless of synthesis method using the MiniBlock. The MiniBlock is now commercially available from Bohdan Automation [3] under license from BMS.

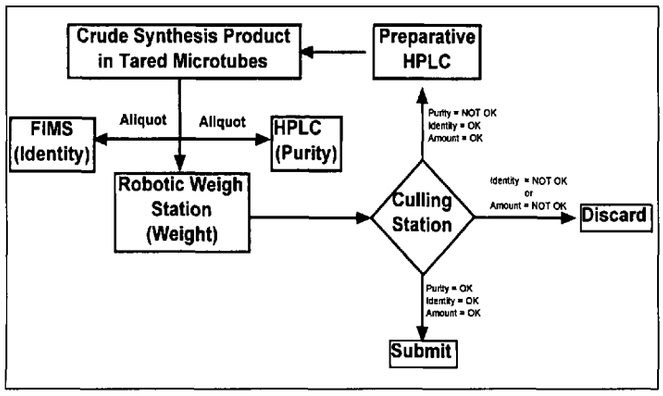
Following synthesis, we wish to estimate three parameters for each compound: purity, structure confirmation, and mass. Since products are delivered into individual microtubes, mass can be obtained by simply weighing the tubes using a robotic weighing device such as the one we developed in collaboration with

Bohdan Automation. Our weighing station can weigh up to 30 racks of 96 tubes (2880 tubes) in 24 h, with precision of 0.1 mg under ideal conditions. The system features queue-based software to facilitate shared use by many different chemists. Structure of all products is confirmed by high-throughput flow injection mass spectrometry using the commercially available MassLynx Diversity system from Micromass Ltd. [4]. Systems such as this are capable of routinely analyzing up to 1000 samples per day.

Purity of products is estimated by high-throughput HPLC using proprietary software. High throughput is obtained using short columns and high flow rates [5]. Custom software has been developed to facilitate 'open access' HPLC in a shared environment by untrained users. This system, now marketed by Shimadzu Scientific Inc. under the name *Discovery-VP*, provides a simple user interface, automated column selection and preparation, and e-mail notification to users and administrators. For automated synthesis applications, a browser has been added to allow import of plate data prior to analysis and to visualize purity data following analysis.

The Discovery-VP system has been extended to include mass spectrometer control. In this case, the Discovery-VP software serves as the single user interface. Discovery-VP directly controls all HPLC components and indirectly controls the mass spectrometer *via* communication with the Micromass MassLynx software. With this architecture, conventional HPLC and LCMS are implemented using the same user interface. This has significant implications for usability and training in an open access environment.

After analysis, the three pieces of data (purity, identity, amount) are used to make a decision about each compound synthesized. Compounds that meet all three criteria are selected for submission to biological assay. Compounds whose structure is not confirmed or for which we have synthesized insufficient amount for testing are discarded. Compounds with confirmed structure and sufficient amount, but insufficient purity, are cycled to preparative HPLC for further purification and reanalysis. Physical sorting of products according to these criteria is accomplished using a pick and place robot we call the 'culling station'. This robot picks individual microtubes and sorts them into new racks according to category. Delivery of products after synthesis into individual containers (custom microtubes) rather than into blocks or plates, allows efficient culling of products after analysis. The entire post-synthesis workflow is diagramed in the Figure.



Compounds whose structure is confirmed and for which there is sufficient amount, but where the purity is inadequate, are recycled for further purification. We have four levels of purification available, with increasing purification power at the cost of decreased throughput. Our highest throughput method is SPE purification using the MiniBlock as an SPE vacuum box apparatus [6]. This technique allows us to purify 96 samples in parallel in under one hour, but with minimal flow control and no online detection. SPE with greater flow control, but still no online detection, is conducted using a custom eight-channel parallel SPE robot designed in collaboration with Bohdan Automation [6]. Using this device, we can purify 96 samples in about two hours. For greater purification power, we turn to preparative HPLC – once again using the Discovery-VP software.

The Discovery-VP preparative HPLC system consists of two Shimadzu LC-8a pumps, a modified SIL-10A autosampler, a variable wavelength UV detector, and up to six fraction collectors (in addition to various valves), all controlled through a modified SCL-10Avp system controller which is, in turn, controlled by the Discovery-VP software. Multiplexing of up to six fraction collectors provides for sample capacity, while online column selection allows for scalability. We use columns ranging from 10 to 50 mm diameter to allow purification from 2 mg up to 200 mg or more on the same system. Automatic column selection allows multiple users to purify widely varying amounts without regard to what the others are doing in the same queue on the same system. Fractions are collected based on peaks in the UV chromatogram, with each peak directed to a different fraction and areas between peaks discarded. The software provides the same simple user interface as the Discovery-VP analytical HPLC system and the Discovery-VP analytical LCMS system, thus facilitating training and user comfort with the system.

In its most complete version, the Discovery-VP Preparative LCMS offers features similar to both analytical LCMS and preparative HPLC, with the additional option of fractionation based on either the UV chromatogram or the single ion chromatogram from the mass spectrometer. In either case, an analogue voltage signal provides input to fraction collector firmware which works in combination with parameters provided by Discovery-VP to direct fractionation. Mass-directed fractionation is most useful for large numbers of samples where fraction collector capacity and post-purification fraction handling are significant issues. The tools described above were designed to work together as an integrated system. Integration is accomplished *via* an information handling tool called the Synthesis Workbook. This tool is a relational database with a workflow oriented client application for access. It serves as the hub in a hub and spoke data network among various lab instruments, and assists tracking all information associated with high-throughput synthesis.

The availability of these tools for high-throughput synthesis is reflected in a large increase in the number of new compounds synthesized over the last four years. More importantly, the number of development candidates has increased while the number of chemists per candidate and the average discovery project duration have significantly decreased. Given the importance of time to market in the pharmaceutical industry, these data suggest that investment in high-throughput technologies provides a significant return.

[1] New York Times, May 16, **1999**.
[2] *www.irori.com*
[3] *www.Bohdan.com*
[4] *www.Waters.com*
[5] 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, 'High Throughput Analysis and Purification in Support of Automated Parallel Synthesis', *Molecular Diversity* **1997**, *3*, 61–70.
[6] M.G. Young, H.N. Weller, J. Roberge, W. Ruediger, 'Solid Phase Extraction as an Efficient Method of Purification for Automated Syn-

thesis', Proceedings of the Sixteenth International Symposium on Laboratory Automation and Robotics (ISLAR), Boston MA, October 18–21, **1998**.

From High-Throughput Parallel Synthesis to High-Throughput Parallel Analysis and Purification: Existing and Emerging Tools to Address the Analytical Bottleneck in Combinatorial Chemistry

D. B. Kassel, L. Zeng, T. Wang, R. Xu
CombiChem, Inc. 9050 Camino Santa Fe, San Diego, CA 92121, USA

Rapid advances in both high-throughput parallel synthesis and high-throughput screening of combinatorial libraries have led to a dramatic reduction in the time taken to identify and optimize lead compounds for drug discovery. Unfortunately, advances in these two disciplines have only shifted the bottlenecks in the drug discovery process to other disciplines. One such discipline that is acutely affected is analytical chemistry. Previously, chemists would make a statistical sampling of their libraries to assess their quality/purity, *etc.* With the overwhelming trend towards single compounds per well (instead of the traditional split-mix mixture approach) and more importantly, *pure* compounds per well, a tremendous burden has been placed on the shoulders of the analytical community to develop techniques which permit rapid characterization, purification, and profiling of these libraries.
In our laboratory we have adopted a fully integrated and modular approach to accelerating drug discovery, placing an emphasis on providing seamless 'links' between each cog of the discovery processes from computational design → synthesis → characterization → purification → screening. We have developed computational tools to dissect both structure activity and inactivity relationships into their principle components. The automated platform for high throughput, a robotic synthesizer, is linked directly to an automated purification workstation where the library of compounds are fully characterized, purified, and quantified in a single system. In so doing, we have been able to maintain at all times a highly efficient and easily trackable system. Each system enables us to characterize, purify and quantify *routinely* 192 compounds per day (*i.e.* two plates per day) at the low mg level in a format immediately available for biological testing. Improvements in our synthesis platform necessitated we push the analysis/purification throughput even further and consequently we implemented a parallel LC/MS workstation concept. A modification to a commercially available Gilson 215 autosampler was made to permit simultaneous sampling across two microtiter plates. We customized a fraction collector script so that sample A1, injected from microtiter plate 1 is isolated into the corresponding well of a deep-well microtiter plate of fraction collector 1. Sample 2, simultaneously injected from well A1 of microtiter plate 2 is isolated into the corresponding well of a deep-well microtiter plate positioned in fraction collector 2. In this way, we are able to purify a maximum of 384 compounds per instrument per day using a dual-column parallel workstation. The amount of isolated fraction contained within the deep wells of the fraction collector microtiter plates is measured by a proprietary method developed in-house that measures the amount of isolated material 'on-the-fly'. More recently, we have established a proof-of-concept for a multi-parallel column workstation, capable of increasing our purification throughput to 768 compounds per instrument per day. This system incorporates a multiple probe autosampler which delivers four samples to an array of either four preparative or four analytical columns coupled to four dual wavelength detectors and a single mass spectrometer incorporating either a standard ESI interface or a multi-parallel interface developed in-house.
In addition, we have taken the concept of parallel LC/MS and extended it to the ultra-high throughput analysis and rapid profiling of combinatorial libraries. The system consists of a multiple probe autosampler to deliver eight compounds at a time to an array of analytical scale columns, an array of eight dual wavelength UV detectors and a single mass spectrometer. In addition, the PE SCIEX software has been customized to permit acquisition of all 16 UV channels.
This system is now employed routinely for two principal purposes. For libraries purified into microtiter plates (as was described earlier by Zeng *et al.*, Comb. Chem. High Throughput Screening (1998)), we are able to perform rapid post-purification analyses to verify compound purity prior to biological screening. Samples are processed eight at a time using our standard 5 min high-speed HPLC/MS gradient for compound purity assessment. Because the samples are processed eight at a time, this leads to an effective analysis time per purified compound of less than 28 sec. The second application of this ultra-high throughput technology is in the area of compound profiling for ADME properties (including models for cell permeability and microsomal stability). We have successfully demonstrated that eight samples can be screened in as little as 1.0 to 1.5 min, leading to an effective analysis time of < 8–12 sec per sample (R. Xu *et al.*, manuscript in preparation and ASMS book of Abstracts, 1999).

CHIMIA 2000, 54, No. 1/2

EUROPEAN COOPERATION
IN THE FIELD OF SCIENTIFIC
AND TECHNICAL RESEARCH

Second Swiss COST Chemistry Symposium

Friday, October 15, 1999

ILMAC 99
Convention Center of the Messe Basel, Basel, Switzerland

Reports and Abstracts by the Authors

Organizer:

Prof. Philippe Renaud
Université de Fribourg
Institut de Chimie Organique
CH–1700 Fribourg, Switzerland
Phone: +41 (0)26 300 87 77
Fax: +41 (0)26 300 97 39
E-Mail: philippe.renaud@unifr.ch

Sponsored by:

- Federal Office for Education and Science
- New Swiss Chemical Society
- Swiss National Science Foundation
- F. Hoffmann-La Roche AG
- Lonza AG
- Novartis Pharma AG