Journal of Cellular Biochemistry Supplement 37:22–27 (2001)

Life in the Fast Lane: High-Throughput Chemistry for Lead Generation and Optimisation

David Hunter*

Discovery Research, High-Throughput Chemistry, GlaxoSmithKline, New Frontiers Science Park, Third Avenue, Harlow, Essex, CM19 5AW, United Kingdom

Abstract The pharmaceutical industry has come under increasing pressure due to regulatory restrictions on the marketing and pricing of drugs, competition, and the escalating costs of developing new drugs. These forces can be addressed by the identification of novel targets, reductions in the development time of new drugs, and increased productivity. Emphasis has been placed on identifying and validating new targets and on lead generation: the response from industry has been very evident in genomics and high throughput screening, where new technologies have been applied, usually coupled with a high degree of automation. The combination of numerous new potential biological targets and the ability to screen large numbers of compounds against many of these targets has generated the need for large diverse compound collections. To address this requirement, high-throughput chemistry has become an integral part of the drug discovery process. J. Cell. Biochem. Suppl. 37: 22–27, 2001. © 2002 Wiley-Liss, Inc.

Key words: high-throughput chemistry; automation; lead generation

In recent years, the pharmaceutical industry has come under increasing pressure to reduce cycle times and rates of attrition in drug development whilst maintaining high standards of safety and tolerability of the new medicines it produces. Increased regulatory restrictions and intense competition have at the same time contributed to the rapidly escalating cost of developing new drugs. Emphasis has been placed on identifying and validating new targets from the human genome and on generating quality leads for these. The response from industry has been evident in genomics and high-throughput screening, where along with combinatorial chemistry, technology has dramatically changed the dynamics of the drug discovery process [Harrison, 1998]. Nowhere are these changes more apparent than in medicinal chemistry, where the application of high-throughput automated synthesis is now well established [Patel and Gordon, 1996, Brooking et al., 1999]. The early emphasis on mixtures of compounds and very large numbers has largely given way to a more measured approach based on arrays of single, well characterised compounds, albeit still in significant numbers. However, the purpose of high-throughput compound production is not simply to facilitate expansion of the bank of compounds which is available for screening. Central is the need to increase efficiency and enhance the ability of chemists to respond rapidly to hits from high-throughput screening, and this response should include the generation of broader structure-activity relationships (SAR) earlier in the discovery process. There is still a need to extend the application of high-throughput technologies and efficient automated methods more widely throughout medicinal chemistry [Merritt, 1998; Coates et al., 2000]. Key to the success of all these objectives is the development of a flexible process and some degree of integration with medicinal chemistry and medicinal chemists.

There are many issues that need to be considered for the practical implementation of high-throughput compound production into medicinal chemistry. The purpose of the compounds, whether for lead generation or lead

*Correspondence to: David Hunter, Discovery Research, High-throughput Chemistry, GlaxoSmithKline, New Frontiers Science Park, Third Avenue, Harlow, Essex, CM19 5AW, UK. E-mail: david_hunter-1@gsk.com

Received 3 August 2001; Accepted 3 August 2001

DOI 10.1002/jcb.10062

optimisation, will influence many of the choices that need to be made between the synthesis of large libraries, as mixtures or singles, or the synthesis of smaller, more focused libraries as singles; solid phase or solution phase synthesis; rigorous analysis and purification of all compounds or the screening of impure reaction products; individual chemist orientated equipment utilising manual or semi-automated methods or sophisticated automation employing expensive robotics in a core facility [Hird, 1999]. Laboratory design may be a luxury for many who have to adapt existing laboratories to accommodate robotic equipment, but there are clearly different requirements, even if only that work areas, such as benches and fume hoods, invariably need to be bigger. In addition, a powerful integrated informatics system is required to handle the huge quantity of data generated by the design, synthesis, analysis and purification of compounds in a high-throughput process [Warr, 1997; Cargill and MacCuish, 1998].

Clearly, there is no "best" method for the implementation of high-throughput compound production, rather the "best" approach is that which is most appropriate to a particular project. Consequently, an integrated and flexible process has to be developed that meets the needs of both lead generation and lead optimisation, and which includes application of solid phase and solution phase chemistry, automation and manual methodology. Both medicinal chemists and the core combinatorial group need to be involved. In our own consideration of high-throughput technologies, we opted for a significant open access component supported by a specialist group, as we felt that this was the most effective way to disseminate the technology, gain the support of the medicinal chemists, and thus leverage wider benefits in efficiency.

The array synthesis process can be broken down into a number of components: reagent management, synthesis, post-synthesis processing, quality control, registration and entry into screening. The number of post-synthesis subsidiary processes: work-up, sampling, evaporation, weighing and purification, emphasises the amount of processing that follows synthesis and the requirement for equipment to complete this efficiently with the throughput matching the rate of synthesis. Integral to this process at GlaxoSmithKline is the in-house software package RADICAL (Registration And Design Interface for Combinatorial and Array Libraries) which facilitates the entire process from design through synthesis and submission to screening [Calvert, 1999]. While arrays can be prepared without such software, it soon becomes impossible as the numbers increase. RADICAL provides facile transfer of files to and from equipment such as weighing, QC and purification stations, and of course registration and screening processes.

The synthesis tools utilised in the process include a range of equipment allowing maximum flexibility. Thus, a Zymate solution synthesiser (Hopkinton, MA) has moderate chemical capability but is able to provide a few grams of each product as reagents for use in other robotic and manual systems. The Bohdan RAM™ (Bohdan Automation, Vernon Hills, IL) is also a solution synthesiser, which operates on a smaller scale but has better chemical capabilities and is particularly useful for SAR work. Extensive use is made of the Robbins FlexChem™ system (Robbins Scientific, Sunnyvale, CA), particularly for high-throughput "simple" solution phase chemistry incorporating the application of polymer-supported reagents and scavenger resins in a microtitre plate format. The Bohdan MiniBlock Synthesiser is a useful addition to this type of technology and can also be used for solid phase chemistry.

A variety of ACT systems (Louisville, KY), suitable for solid and solution phase synthesis, are available for library production and the preparation of directed arrays. While somewhat limited in their chemical capability and by their non-modular format, they have proved particularly useful in recent years as part of a high-throughput solution synthesis training kit. This short intensive training programme, open to all chemists, serves the dual purpose of utilising proprietary reagents for the synthesis of libraries and exposing medicinal chemists to automated methods and the high-throughput synthesis process. Equipment utilised includes the ACT 496 for solution synthesis, purification by solid phase extraction, Tecan Miniprep (Hombrechtikon, Switzerland) or Hamilton MicroLab for preparation of analytical samples, evaporators such as the Genevac (Ipswich, Suffolk, UK), and a Bohdan Automated Weighing Station. A browser is used where necessary to check the automated QC output from the Micromass MUX™ LC/MS (Manchester, UK) and of course RADICAL is used throughout.

The Irori technology (San Diego, CA) is ideally suited to the solid phase multi-step synthesis of large lead generation libraries [Nicolaou, 1995]. This technology exploits the parallel reaction advantage of "mix and split" synthesis, but to generate arrays of single compounds the use of individual mini-reactors, each with its own unique radiofrequency tag, enables each component to be tracked through the synthesis by the "mix and sort" method. This methodology has been utilised for the preparation of diversity arrays for screening but also for generic and targeted arrays within medicinal chemistry groups. Chemistry is usually conducted in normal laboratory equipment and it can be an entirely manual operation, but as array size increases and/or more synthetic steps are involved, then automation of the most labour intensive steps is important.

The Myriad Core System (MCS^{TM}; Mettler-Toledo Myriad, Melbourn, Herfordshire, UK), an example of the new generation of synthesiser, has been developed specifically for high-throughput and high performance chemistry (solution and solid phase). As well as the modular design, the key to the instrument is the reactor which comprises a fritted glass U-tube with a novel twist-cap seal to allow septum free access in an inert atmosphere. In addition, reagent handling is carried out by positive displacement using disposable fluoropolymer pipettes, enabling reliable use of corrosive reagents as well as eliminating the need for time consuming wash steps to prevent cross contamination. The MCS operates with 192 reaction vessels which are independently processed as 4 sets of 48 vessels, and further special purpose modules for rapid resin washing and liquid–liquid extraction have been developed to extend the range of synthetic tasks that can be performed on the MCS. Complimentary to the MCS are the Myriad Personal Synthesisers (PS) which use common hardware and software. These 24 reaction vessel systems have been distributed throughout Medicinal Chemistry and, along with analysis, evaporation and purification modules, are a key component of the Automated Medicinal Chemistry Workstations. In addition to their utility in the synthesis of small targeted arrays, their common hardware and software features allows chemistry from the PS to be readily transferred to the MCS. A wide range of solid and solution phase chemistry has been demonstrated in-house at GSK on both the MCS and PS, including the use of corrosive, air and moisture sensitive reagents, dropwise addition of sensitive reagents and the generation of reactive intermediates.

A useful recent addition to the tools available for synthesis is that of microwave equipment designed specifically with chemistry in mind. Equipment such as the Smith Personal Synthesiser^{TM} (Personal Chemistry, Uppsala, Sweden) allows precise control of temperature and pressure with a high degree of reproducibility. Numerous reports are beginning to appear on the application of microwaves in high throughput chemistry [Brain and Brunton, 2001, Larhed and Hallberg, 2001].

The range of synthesis tools that are available allows for flexibility in the choice of the appropriate high-throughput synthesis technology. Overall, the intention is to develop a process that provides a high performance, high-throughput capability in both solution and solid phase chemistry and to ensure that chemistry is not the limiting factor. Clearly, it is important to choose the technology that has the capability to handle the particular chemistry and produce the required quantities and numbers of compounds—i.e., to work within the capabilities of the technology. Instruments also need to be matched to tasks to be used efficiently, and usage must be co-ordinated, especially when the users are from a large pool of chemists. Standardised processes need to be employed whenever possible, irrespective of the synthesis technique, because all of the processes must converge sooner or later. Essentially, the process is the same for one compound as it is for a thousand, but the logistics problems of processing large numbers of single compounds dictates a degree of automation and the concept of a "factory" environment becomes increasingly appropriate.

Although a great deal of effort has been directed to high-throughput synthesis, there is an obvious requirement that all the post-synthesis tasks require a degree of automation to fully enable high-throughput compound production. Thus instruments for work-up, purification, sample reformatting, evaporation and analysis have been integrated into the process. A key component of any high-throughput production process is quality control, and automated analytical instrumentation has advanced rapidly to meet these needs [Hughes

and Hunter, 2001]. All of the products from high-throughput chemistry, irrespective of the synthesis technique, are analysed by LC/MS using the high-throughput MUX™ LC/MS technology which allows *on-line* LC/MS analysis in parallel with simultaneous detection on one electrospray LC-TOF-MS detector. This technology offers an eightfold increase in throughput over a conventional LC/MS [de Biasi et al., 1999]. NMR spectra are also obtained for representative samples of large lead generation arrays. Targeted SAR arrays for Medicinal Chemistry, where a consistently higher quality product is required, are subjected to more rigorous analysis, similar to that traditionally applied to single compounds. After analysis, products are quantified by mass using tare and gross weights obtained on a Bohdan Automated Weighing Station.

The initial QC analysis results are imported into RADICAL where the plate manager component of the data management system is used to categorise the products into those that pass QC, those that require purification and those that must be discarded. The recognition that it is not always possible to prepare pure compounds by synthesis alone has led to the development of a number of purification techniques and automated purification systems [Weller, 1998, Hughes and Hunter, 2001, Kassel, 2001]. Purification may simply be a matter of liquid–liquid extraction or the use of scavenger resins to remove impurities or excess reagents. The growing importance of solution phase chemistry in high-throughput synthesis has stimulated interest in using solid phase reagents as synthesis tools [Ley et al., 2000]. If these techniques prove inadequate then several automated flash chromatography systems, in addition to higher throughput preparative HPLC systems, are now available.

The FlashMaster II, from Jones Chromatography (Mid Glamorgan, UK), is an automated flash chromatography system with the capacity to purify 10 compounds sequentially, using UV triggered fractionation, and is ideally suited to the purification of reagents and intermediates required for array production. The Biotage Quad 3 (Charlottesville, VA) is a parallel 12-column instrument in which fractions are automatically collected by time intervals using a fraction collector. Ideally suited to the purification of small focused arrays, these instruments are used extensively throughout Medicinal Chemistry. The Biotage Parallex™ is a fully automated, high-throughput, preparative HPLC system running four columns in parallel with UV detection for fractionation [Schultz et al., 1998]. Scale is typically 10–30 mg (though it can be larger) with a throughput of 200–300 samples/day. Although largely used for the purification of large lead generation libraries that would otherwise fail to pass QC criteria, it has also found application in the purification of targeted SAR arrays for Medicinal Chemistry. The key to its successful application is the ability to reduce the burden of post-purification fraction processing and to streamline the liquid handling required for reformatting. To this end, "intelligent fraction collection" is used to reduce the number of fractions collected by adjusting the fraction collection criteria: slope, threshold, and UV monitoring wavelength. In addition, the initial analytical data for the crude samples can be used to guide fraction collection. Subsequently, "Winnow" [Hughes, 2000], an Excel-based application developed in-house, generates an interactive display of the fractions collected from each sample in the form of a chromatogram. The key purpose is to enable "intelligent fraction selection", based on absorbance, volume and initial QC data, to be used to reduce the number of fractions selected for final QC. The post purification process requires the integration of the Parallex, the Mux™ LC/MS and a Tecan Genesis (combined with a Zymark plate hotel) that prepares analytical plates from the fractions selected for QC and subsequently transfers the bulk solutions, based on the QC results, to vials for evaporation. The data transfers required for these operations are an additional functionality of "Winnow". As an extreme example, 4320 compounds from an 11-step solid phase synthesis were purified; typically only 3–4 fractions were collected per sample, and of these only 1–2 were selected for analysis by LC/MS. 4012 compounds were subsequently submitted for screening (a success rate of 93%), more than 90% of the compounds were greater than 95% pure.

The final stages in the process, evaporation and weighing, are completed efficiently using Genevacs and the Bohdan AWS. All of the compound data: structure, analytical results and weight, are integrated within RADICAL, along with registration of the compounds into the GSK compound database and the data

transfer to the high-throughput screening group.

To judge the success of the process, the overall aims of high-throughput compound production at GSK must be considered. High-throughput technologies are now an integral part of all Medicinal Chemistry groups. The productivity of these groups has undoubtedly been enhanced, while the quality of the compounds produced has been maintained. The Myriad personal synthesisers have been deployed throughout Medicinal Chemistry and, along with associated analytical, purification and evaporation equipment, form the basis of Medicinal Chemistry Workstations. Focused SAR arrays are being prepared with this technology and the chemistry utilised has successfully been transferred to the Myriad Core System. High-throughput synthesis has had a substantial impact on the progress of research programmes, with numerous examples of "instant SAR", reduced cycle time and more potent leads. In addition, increased serendipity generated from new combinations of existing building blocks has generated novel leads for other unrelated targets. The number of compounds in the GSK screening collection has been increased dramatically with the major contribution coming from the core Combinatorial Group, but with a significant contribution from Medicinal Chemistry. Perhaps most importantly, 50% of the leads identified in 2000 originated from high-throughput chemistry.

High-throughput compound production will without doubt continue to play an integral part in drug discovery and the use of automation in Medicinal Chemistry will see further expansion. While the open access nature of much of the equipment at GSK has been emphasised, this is not the only possible approach, and the development of a more focused production type environment may also be appropriate. Certainly, some of the smaller, independent companies have already made some progress towards this. The typical organic synthesis laboratory is an expensive, highly serviced area, with fixed fume hoods, lacking flexibility and based on traditional strategies of single compound synthesis. This contrasts with a factory or production facility, which is usually a cheaper building, less highly serviced and with open work areas and drop down services for in-built flexibility, designed for efficient production. So for optimum high-throughput compound production, integration of high level synthesisers, work-up stations, evaporators purification and QC systems in a purpose designed facility, "The Compound Factory", would be desirable. This type of concentrated production unit would require co-ordination from reagent input to product output and on into screening, which would be the next work group on the production line producing quality leads as the products. The design and construction of just such a dedicated facility is currently in progress at GSK.

REFERENCES

Brain CT, Brunton SA. 2001. Synthesis of 1,3,4-Oxadiazoles using polymer-supported reagents. Synlett 3:382–384.

Brooking P, Crawshaw M, Hird NW, Jones C, MacLachlan WS, Readshaw SA, Wilding S. 1999. The Development of a solid phase Tsuge reaction and its application in high-throughput robotic synthesis. Synthesis 11:1986–1992.

Calvert S. 1999. Managing the combinatorial chemistry process and information. MDL US User Conference, San Francisco, CA, USA.

Cargill JF, MacCuish NE. 1998. Object-relational databases: The next wave in pharmaceutical data management. Drug Discov Today 3:547–551.

Coates WJ, Hunter DJ, MacLachlan WS. 2000. Successful implementation of automation in medicinal chemistry. Drug Discov Today 5:521–527.

de Biasi V, Haskins N, Organ A, Bateman R, Giles K, Jarvis S. 1999. High throughput liquid chromatography/mass spectrometric analyses using a novel multiplexed electrospray interface. Rapid Commun Mass Spec 13:1165–1168.

Harrison W. 1998. Changes in scale in automated pharmaceutical research. Drug Discov Today 3:343–349.

Hird NW. 1999. Automated synthesis: New tools for the organic chemist. Drug Discov Today 4:265–274.

Hughes I. 2000. Separating the wheat from the chaff: High-throughput purification of chemical libraries. JALA 5: 69–71.

Hughes I, Hunter D. 2001. Techniques for analysis and purification in high-throughput chemistry. Curr Opin Chem Biol 5:243–247.

Kassel DB. 2001. Combinatorial chemistry and mass spectrometry in the 21st century drug discovery laboratory. Chem Rev 101:255–267.

Larhed M, Hallberg A. 2001. Microwave-assisted high-speed chemistry: A new technique in drug discovery. Drug Discov Today 6:406–416.

Ley SV, Baxendale IR, Bream RN, Jackson PS, Leach AG, Longbottom DA, Nesi M, Scott JS, Storer RI, Taylor SJ. 2000. Multi-step organic synthesis using solid-supported reagents and scavengers: A new paradigm in chemical library generation. J Chem Soc Perkin Trans 1:3815–4195.

Merritt AT. 1998. Uptake of new technology in lead optimisation for drug discovery. Drug Discov Today 3: 505–510.

Nicolaou KC. 1995. Radiofrequency encoded combinatorial chemistry. Angew Chem Int Ed Engl 34:2289–2291.

Patel DV, Gordon EM. 1996. Applications of small-molecule combinatorial chemistry to drug discovery. Drug Discov Today 1:134–144.

Schultz L, Garr CD, Cameron LM, Bukowski J. 1998. High-throughput purification of combinatorial libraries. Bioorg Med Chem Lett 8:2409–2414.

Warr WA. 1997. Combinatorial Chemistry and molecular diversity. An overview. J Chem Inf Comput Sci 37:134–140.

Weller HN. 1998. Purification of combinatorial libraries. Mol Diversity 4:47–52.