Changes in scale in automated pharmaceutical research

Willie Harrison

The trend towards increased throughput pervades the drug discovery process. Developments in combinatorial chemistry have made available greater numbers of compounds for high-throughput screening, and advances in genomics have led to increased numbers of potential targets. These advances have been implemented by integration of sophisticated engineering and logistics. The author describes the current automated technologies used within the pharmaceutical industry and explains how continued advances in this area are necessary for maintaining growth of the industry.

In recent years, drug discovery has become increasingly dependent on high-throughput processes to identify the numbers of leads necessary to maintain commercial growth. One estimate suggests that to maintain growth, the top ten global pharmaceutical companies will need to launch at least five significant new chemical entities (NCEs) with sales potentials of more than US$350 million per year[1]. Bristol-Myers Squibb has announced plans to double the number of product launches by the year 2000 (Ref. 2). Glaxo Wellcome has publicly set itself a target of introducing three significant NCEs per year by the same deadline[3]. To put this in perspective, the top companies launched an average of only 0.45 NCEs per year between 1990 and 1994, and during the same period, only 8% of all products achieved sales of US$350 million. The trend towards increasing throughput is apparent throughout the discovery process. Developments in combinatorial chemistry have massively increased the numbers of compounds available for high-throughput screening (HTS). At the same time, the number of potential targets is being expanded through advances in genomics. Underpinning these increases are automated systems using sophisticated engineering and logistics; they do not merely represent the replacement of a manual task but are truly enabling – without them, crucial high-throughput processes could not take place.

This article reviews some of the automated systems currently in use in the pharmaceutical industry for HTS, combinatorial chemistry and genomics. The development of an industrial-scale compound library management system is also discussed as a further example of the increasing need for sophisticated automated solutions within drug discovery.

Automated screening systems

HTS is at the heart of current drug discovery, and engineering and automation have been key technologies for developing discovery processes that go far beyond human capability to manipulate, track and analyze assays. While automating a biological protocol is often no faster than performing the same manipulations manually, vast increases in scale are impossible without human error and labour becoming limiting factors.

Automation in HTS can be divided into three primary categories:

- Hand-held automation, such as automated pipettes.
- Unit automation, such as benchtop devices including liquid handling X-Y-Z robots.
- Integrated automation.

Increases in throughput require increases in automation, and in fully integrated systems, unit automation is integrated with two other aspects:

Willie Harrison, The Automation Partnership (Cambridge) Ltd, Melbourn Science Park, Melbourn, Royston, Hertfordshire, UK SG8 6HB. tel: +44 1763 262026, fax: +44 1763 262613, e-mail: wjh@autoprt.co.uk

  1359-6446/98/$19.00. PII: S1359-6446(98)01201-X

- Feeding and transfer of tubes or microplates into, between or out of individual devices, using either robots or conveyor systems.
- System-wide information handling, including planning, scheduling, sample tracking, results capture, machine control and, crucially, error management.

Beckman Instruments (Fullerton, CA, USA) has introduced the Biomek Integrated Laboratory Automation System for HTS; it combines the Biomek 2000 Laboratory Automation Workstation, the ORCA linear track robot system, SAMI software from Sagian (Indianapolis, IN, USA), detectors from Wallac Oy (Turku, Finland) and Molecular Devices (Sunnyvale, CA, USA) into one fully integrated system that can be easily upgraded or expanded. Similar systems are available from Zymark, Thurnall and Scitec, also Robocon (Vienna, Austria) who, in collaboration with Advanced ChemTech (Louisville, KY, USA), have developed the RODOLAB 9600.

Advances in all the classes of automation have enabled increases in HTS rates such that large populations of compounds can be efficiently screened *en masse*, although it should also be noted that poor implementation has undermined many attempts, and failures are common.

Combinatorial chemistry

The drive towards discovering ever larger numbers of drug candidates has been fueled by developments in organic chemistry that have vastly increased the scale of compound libraries. The underlying logic is simple: by producing larger, more diverse compound libraries, companies increase the probability of finding novel therapeutic compounds of commercial value. Automated chemistry also accelerates the process of 'working-up' lead compounds with regard to potency, toxicity and patent protectability. The advance of combinatorial chemistry has been made possible through progress in miniaturization and automation. The original focus for combinatorial chemistry was peptide and oligonucleotide library development, which are of limited therapeutic interest. However, the field changed greatly when the first combinatorial techniques for producing small organic molecules with molecular weights of up to 500 were developed by pioneers such as Sheila Hobbs de Witt at Parke-Davis Pharmaceuticals (Ann Arbor, MI, USA).

Several strategies have now been developed for combinatorial synthesis with the most crucial division traditionally being between solid- and solution-phase systems. Solid-phase synthesis greatly simplifies multistep reactions, as excess reagents can be added and then washed away after each step. Solution-phase synthesis, by contrast, has greater flexibility in terms of the range of organic reactions available.

There are now several systems that allow companies to avoid using either solution-phase or solid-phase chemistries exclusively. A combinatorial chemistry reaction block accommodating a wide variety of solvents and capable of both solid-phase and solution-phase chemistry is currently being marketed by Bohdan Automation (Mundelein, IL, USA), while several different automation systems that enable researchers to carry out solution-phase combinatorial synthesis and solid-phase peptide and peptoid synthesis have been put together by Zymark (Hopkinton, MA, USA).

Automated synthesis

The importance of automated synthesis as a major productivity booster in combinatorial chemistry was realized by several leading pharmaceutical companies who, together with The Technology Partnership, initiated the Myriad Development Programme in February 1996.

The Myriad Consortium, consisting of senior scientists from The Technology Partnership, BASF, Chiroscience Group, Merck & Co., Novartis, Pfizer, SmithKline Beecham and Takeda Chemical Industries, was formed to develop a fully automated synthesizer. The development programme included the successful testing of the Myriad Personal Synthesiser (MPS) – a modular system for automated compound synthesis using solid- or solution-phase chemistry. It is a compact system, designed to fit into a single fume cupboard. Up to 24 reaction vessels can be simultaneously incubated in the system, which is capable of carrying out automated reagent addition, sampling and other handling operations. A processing robot allows the MPS to carry out handling tasks that are beyond the capability of syringe pump-based units. An incubator, enabling accurate temperature control over the range −70°C to +150°C, and a gas-blanketed reagent storage area are located under the robot. The processing robot, incubator and reagent storage area, together with a recirculating chiller unit, are all enclosed to form the synthesizer unit.

Chemists from each of the Consortium members have also tested the Myriad Chemistry Developer (MCD) – a modular system for automated compound synthesis that is

Box 1. Contemporary automated combinatorial chemistry systems

Automated tools for combinatorial chemistry include:

- Myriad, an automated chemical synthesis technology, has been developed by The Technology Partnership (Cambridge, UK). Myriad can be fed with reagents stored and dispensed from The Automation Partnership's Haystack system. End products from Myriad can then be fed into Haystack before screening.
- Advanced ChemTech (Louisville, KY, USA) market automated systems for solution- and solid-phase synthesis and for peptide and organic synthesis.
- Tecan U.S. (Research Triangle Park, NC, USA) is marketing the CombiTec – an organic chemical synthesizer that includes reaction blocks of 8–56 chambers and a robotic sample processor.
- Chiron's (Emeryville, CA, USA) multiwash head system generates large numbers of compounds simultaneously. Third-generation units now feature all-glass reaction vessels, heating to 120°C and flexible software that allows the automation of most organic reactions.
- The OntoBLOCK, developed by Ontogen as an in-house combinatorial chemistry automation system, includes reaction blocks containing 96 reaction vessels from which compounds are transferred directly to standard 96-well microtitre plates for HTS. This system is capable of producing 1,000–2,000 small organic molecules per day by parallel array synthesis.
- A combinatorial chemistry system in the prototype stage is the Nautilus synthetic chemistry workstation that is being developed by Argonaut Technologies (San Carlos, CA, USA). The instrument, which is completely enclosed and encapsulated, with a pressurized fluid delivery system and no exposure to the atmosphere, handles a wide range of reagents with capabilities for temperature control and use of inert atmospheres.
- The CombiChem Drug Discovery Engine enables CombiChem (San Diego, CA, USA) to conduct lead optimization, lead evolution and lead generations by matching design hypotheses with novel templates from within a proprietary virtual library. CombiChem is building an integrated system of synthesis, analysis and purification to enhance rapid production of such libraries that will be capable of automating both solid-phase and solution chemistry.
- 3-Dimensional Pharmaceuticals (Exton, PA, USA) is developing a combinatorial chemistry system based on a technique called 'directed diversity'. A chemical library is generated by robotic instruments, structure-activity information is obtained on library members and data are analyzed to determine how closely the synthesized compounds match a set of desired properties.

fully compatible with the larger MPS and Myriad Automated Synthesiser (MAS).

Both the MPS and the MCD use a unique reaction vessel developed by The Technology Partnership; it enables the harvesting of reaction products either 'on-bead' or as a liquid phase after cleaving to be easily accomplished. The vessel is sealed by a long-life PTFE cap and is opened by a simple twist action. The vessel contents are protected from air and moisture ingress in both open and closed states without the use of pierced rubber septa – a feature unique to this technology. Such cap technology, coupled with full protection of reagents during storage and dispensing, allows the MPS to carry out air- and moisture-sensitive chemistries without a gas enclosure to house the machine.

Both the MAS and the MPS can be fed with reagents stored and dispensed from the Haystack system (The Automation Partnership; described in detail below); end products can then be delivered back to Haystack before screening.

A brief survey of some other contemporary automated combinatorial chemistry systems is given in Box 1.

A further application of automation in combinatorial chemistry is in the analysis of so-called encoded combinatorial libraries, such as those being used to identify aspartyl protease inhibitors at Pharmacopeia, where potential leads have emerged from libraries of about 32,000 compounds. The tags are analyzed by capillary GLC to identify the active compounds in the library. Automating the tag-reading process and combining the technology with HTS of targets of pharmaceutical interest has optimized the methodology to perform 1,000–1,500 tag decodes routinely each month using four dual-channel GLC instruments with electron capture detectors and autosamplers.

Automated compound library management

The increase in throughput made possible by HTS and combinatorial chemistry has been based to a great extent on automation. As shown above, there are currently many systems on the market that address these areas. However, the number of compounds now available has potentially shifted the bottleneck in the drug discovery process to the logistics of compound supply. In this area, the scale of automated solutions has more in common with manufacturing than with traditional benchtop pharmaceutical research. Without a sophisticated library management system using advanced engineering and data handling, the

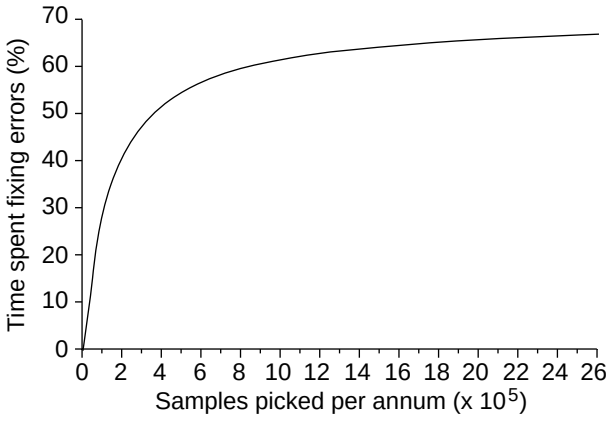

Figure 1. *Compound library management. The proportion of time spent fixing errors in a manual retrieval system with an error rate of 2%. Even at relatively low numbers of samples picked per year, over half the manual system's time can be wasted correcting errors.*

storage, tracking and retrieval of compound libraries, which may now hold millions of compounds, would lead to chaos. One simple model suggests that with a manual error rate of 2%, retrieval of 400,000 samples per year would mean half of the systems time would be spent correcting errors[4] (Figure 1). The importance of integrated compound-library management is becoming increasingly recognized and several major pharmaceutical companies have made significant investments in this area.

The Automation Partnership has designed Haystack – a modular automated preparation, storage and retrieval system to manage compound libraries for drug screening. It is now in place in several companies including Zeneca, Janssen, SmithKline Beecham, Bristol-Myers Squibb, DuPont and Proctor & Gamble (Figure 2). The installed systems are capable of handling compound libraries ranging from 100,000 to 5,000,000 samples.

The first generation Haystack system was prototyped in 1993 in a collaborative development with Burroughs-Wellcome (now Glaxo Wellcome); it was specifically engineered to meet the requirements of the company, that is, to increase assay throughput by changes to the compound library system. Because the managing and dispensing of millions of samples is a major logistical problem, a significant proportion of the Haystack development time was devoted to scheduling and data management and not just

Figure 2. *The Haystack system (The Automation Partnership) allows a compound library to be managed and dispensed from solid or liquid archives.*

to the engineering problems presented by the physical storage and retrieval of compounds. Of the total engineering development time (50 man years), 60% was devoted to software design.

The approach adopted by Burroughs-Wellcome was to use semi-targeted or 'intelligent' screens, and the compound management system was designed to maximize flexibility using a liquid library stored as aliquots in microtubes rather than a solid archive. Under these conditions, some degree of throughput rate was sacrificed to achieve maximum flexibility. Once the system was in place, representatives from 20 pharmaceutical companies were invited to visit the site and make their comments. Time was taken to re-evaluate the exact nature of the storage problems as seen by the other drug companies and to set about the adaptation of the original design to produce second generation machines. Certain requirements were addressed when setting up the next generation of compound library management systems; they included more flexibility in storage formats, compatibility with varying assay procedures and greater speed of operation.

Figure 3. *The robotic arm of the Haywain (The Automation Partnership) feeds sample vials to the balances ensuring that they are never standing idle.*

Modular systems

The second generation Haystack systems, now in place in several companies, are based on a modular design so that, without being entirely custom built, varying needs could be met according to the drug discovery approach adopted by different companies. The largest system built so far has 20 modules and the smallest just one, but there is no upper limit to the number of modules a system can utilize, giving a potentially unlimited library size. The number of samples that can be held in each module varies according to ceiling height and container size; so far the greatest number of samples that has been accommodated in one module is ~1,500,000 and the least is 100,000. The design is currently being adapted to the needs of smaller laboratories, such as those in the biotechnology sector, which have smaller, and often plate-based, compound collections. Vials are kept in racks in carousels and retrieved by robots and conveyors directed by the system software.

With the Haystack, samples are identified by bar code readers and security is enhanced by a high integrity tracking system. All container locations are recorded as compounds are moved, and the data are backed up in real time. All transactions are noted so that at any time the position of a compound and the quantity of a sample remaining in a container may be determined. The system is then capable of alerting the operator to re-order a compound when stocks diminish.

Linking retrieval to an automated compound dispensing station such as Haywain (The Automation Partnership; Figure 3) further reduces the risks of human error, as all data capture is automatic. There are four highly sensitive balances inside the Haywain cabinet that have been modified to dispense minute quantities of powder with great accuracy. Usually 1 mg samples are weighed, although sizes as low as 0.2 mg can be dispensed provided the powder is fine enough. Where appropriate, compounds can be diverted for manual processing. Haywain is capable of dispensing 50–100 dispenses per hour.

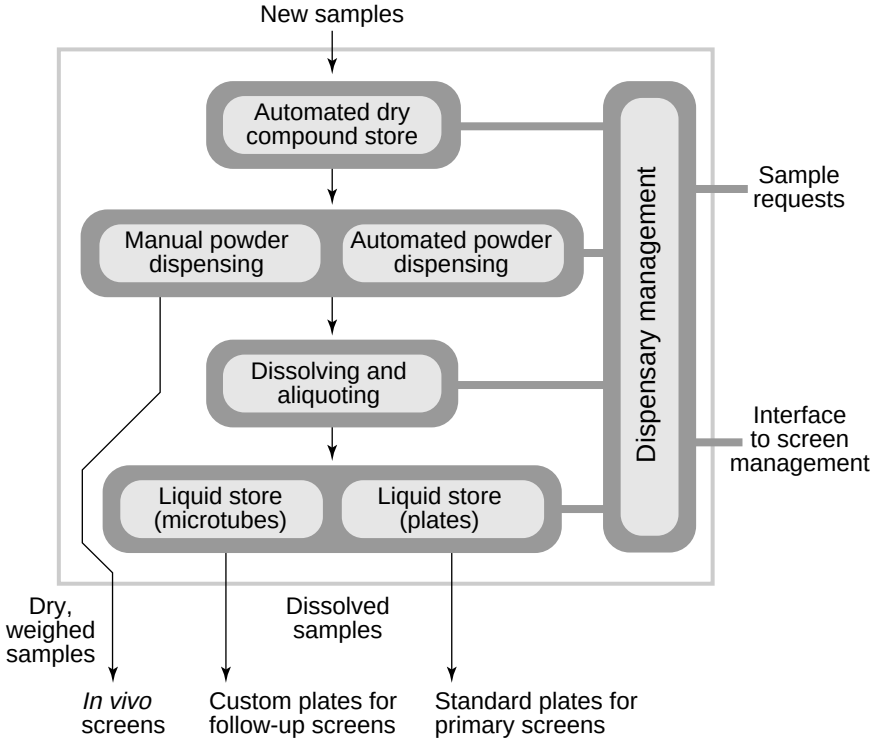

Figure 4. *An automated sample library management system allows a tiered library to be established. Compounds can be stored in solid or liquid states, depending on screening requirements.*

Once dispensed, samples are placed into prelabelled bottles or can be labelled after they are dispensed.

Depending on the discovery strategy of the company, a second tier library may be created comprising sets of compounds in mother plates and/or individual aliquots in microtubes that can be stored and retrieved automatically by additional modules (Figure 4). Haystack allows solutions of the desired molarity to be prepared for storage or for aliquotting either by hand or by further automated systems. The solid stocks may then be viewed as an archive, with drug screening routinely accessing various liquid working stocks. This use of libraries in various phases and with vast numbers of samples is made manageable only by integrated automated systems. Storing solutions in a way in which they can be readily tracked, accessed and re-used can make considerable economies possible and ensures maximum utility is derived from the original compound.

Automation advantages

Maintaining a library or libraries may demand a variety of storage criteria, and any automated system must be capable of optimal performance under differing conditions of temperature, light, atmosphere and humidity. Automated systems are better in dealing with such a range of conditions than the nearest equivalent manual approaches because robots are capable of operating in environments where the effectiveness of human workers may be limited. There are also advantages to using such systems where the compounds themselves are hazardous.

Much of the hardware used in the Haystack system is nonspecialized and well tested in other applications. Industrial robots similar to those employed in the automotive and electronics industries are used, and carousels and tracks are similar to those used in other fields. Most bottles, tubes and plates in common use in the pharmaceutical industry can be easily accommodated.

The software that supports the system is centred around a high-speed database capable of tracking the position of all compounds within the library at all times, and it was designed to be robust and error tolerant. For both generations of the system, half the software was devoted to the management of nonroutine events, such as illegible bar code labels. Only through such advanced data management can confidence be placed in scaling-up operations to utilize the vast numbers of compounds made possible by combinatorial chemistry and demanded by HTS. Such numbers are far beyond those that could realistically be managed by nonautomated systems – the effects of chaos would make them unworkable. For example, the largest Haystack system built to date is installed at Zeneca (Alderley Edge, UK) and is intended to manage efficiently archive stores of all the compounds made or acquired by the company over the last 40 years.

Genomics

As has been shown, automation is crucial in HTS and combinatorial chemistry and in linking these processes through effective compound library management. A third major area for growth in the pharmaceutical industry is genomics, and this is another field where huge numbers of samples are involved. Again, automation is critical in realizing the full potential of this science for drug discovery. The pharmaceutical industry's investment in genomics grew to more than US$1 billion in 1996 and companies estimated that by 2000, more than 60% of targets they pursue will be derived from genomics research – up from 8% in 1996 (Ref. 5). However, few companies have effectively integrated genomics into the overall discovery process; most rely on strategic partnerships with specialist genomics companies, such as Genset, Incyte or Affymetrix, and also draw upon public domain gene-sequence databases from Institutes such as the Sanger Centre. Broadly speaking, the work in genomics can be divided into two: *de novo* sequencing of the human genome and specific disease research (for example, polymorphism discovery). In either case, however, the process is characterized by similar factors:

- Dependence on very large numbers of samples and experiments
- Vulnerability to error
- Intense time pressure to produce answers quickly

It is not surprising, therefore, that successful genomics automation is usually only achieved where robust, industrial engineering is employed and, crucially, where information processing is an integral part of the system.

Future developments

The future of the drug discovery process will continue to present challenges at all stages, including HTS, combinatorial chemistry, genomics, the identification of targets and the design of assays. Largely, these challenges will have to be met not only through biology and chemistry but also through advances in engineering and logistics.

The scale of the future drug discovery operations has led some to suggest the appearance of a 'drug discovery factory' with industrial-scale engineering maintaining massive throughput rates[4]. At the same time, pressure exists to miniaturize systems, minimizing reagent usage and maximizing the possible flow of data. The trend is already apparent in the development of systems involving high-density plates and precision pipettors. Further increases in throughput are possible with assays conducted on microchips using sub-microlitre volumes and huge throughput rates. Merck & Co. have recently joined Aurora Biosciences' Ultra-High-Throughput Screening System (UHTSS) development syndicate, whose members already include Bristol-Myers Squibb, Eli Lilly & Co. and Warner Lambert's Parke-Davis Pharmaceutical Research Division. Merck will receive the licence to Aurora's fluorescent-assay technologies and Aurora will install its prototype UHTSS, which will combine the assays in a miniaturized system under computer control. More systems that integrate HTS with data analysis and decision making can be expected, allowing improved targeting of screens and greater efficiency. In addition to increasing the number of NCEs identified, the utility of leads will also be enhanced as automated, high-throughput organic syntheses become more routine. This will make the performance of exploratory chemistry even on the weakest active compounds feasible.

The maintenance of growth in the pharmaceutical industry is thus dependent on the successful continuation of an already existing trend towards sophisticated engineering to support advances in other disciplines.

REFERENCES

1 Drews, J. and Ryser, S. (1997) *Drug Discovery Today* 2, 365–372
2 Anon. (1997) *Scrip World Pharm. News* 2293, 20
3 Anon. (1996) *Nature* 384 (Suppl.), 1–5
4 Harrison, W. (1998) *Pharm. Man. Int.* 99–102
5 Banerjee, P.K. and Rosofsky, M. (1997) *Scrip Magazine* November, 35–38

In short…

In 1997 the value of the cancer market grew by over 13% to an estimated $13 billion. This growth was mainly driven by the cytotoxic segment of the market, which increased by 22% to an estimated $5 billion. According to **Datamonitor**, sales from the three new drugs, docetaxel (Taxotere; Rhône-Poulenc Rorer), gemcitabine (Gemzar; Ely Lilly) and topotecan (Hycamtin; SmithKline Beecham), account for one-third of the cytotoxic market. However, the market is still dominated by Bristol-Myers Squibb's paclitaxel (Taxol) and carboplatin (Paraplatin), and Taiho's tegfururacil (UFT). Datamonitor forecasts that the sales of the three new cytotoxic agents could peak at >$400 million. For further information see Datamonitor's report, *Market Dynamics to 2010: Global Opportunities in Cancer*.

In another report by Datamonitor, *Generic Cancer Markets: Quantitative Analysis and Forecasts*, the growth of the generic cancer market is expected to be much slower than that of the overall market. In spite of the strong growth of generic drugs in other therapeutic areas, the generic cancer market has not developed in the same way. In 1997 the market size was <$300 million in the Western developed economies, with the biggest market being in the USA and Germany. The main driver for the growth of this market is patent or exclusivity expiry – on average one cancer drug per year comes off-patent or loses its exclusivity in both the USA and Germany. This growth factor and the projected increasing requirement for anticancer drugs are not enough to match the strong growth of the total cancer market.

For more information, contact Krishna Rao, tel: +44 171 316 0001, fax: +44 171 316 0002, e-mail krao@datamonitor.com, Web: www.datamonitor.com. Reports are available from Datamonitor priced $2995 each.