Engineering Chemistry for the Future of Chemical Synthesis

D.E. Fitzpatrick and S.V. Ley*

Affiliation:

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom.

*Corresponding Author: Steven Ley, svl1000@cam.ac.uk

Summary:

Synthesis is changing in response to our modern resource conscious world. The principles of green chemistry are evolving as the interfaces and boundaries in science are less obvious and providing a new stimuli for future discovery. The invention and application of new chemical reactivity continues to be a primary driver since this opens up so many strategic opportunities for synthesis. However, the manual intensive efforts behind such activity inevitably lead to the need for more machine based approaches. Indeed, the engineering of chemistry delineated in this Symposium in Print seeks to collate some of the recent progress and innovation in the area with contributions from its visionary practitioners.

Keywords:

Continuous flow synthesis; Automation; Computer control; Machine-assisted synthesis; Future trends in synthesis; Lab of the future.

Graphical Abstract:
**Introduction and Preamble**

We live in a world where complex machinery impacts widely on all our activities. Even our motor cars and airplanes can now operate autonomously. The same however, is not true of a modern chemical synthesis laboratory where many of the tasks are still manual, often routine and repetitious in nature. Even the expensive equipment in this environment is often isolated and the data collected tends to be siloed and not fully utilized. Our batch-mode mentality promotes incremental development at the expense of holistic understanding and coordination. Things must change and indeed the climate is now conducive to that change.[1]

A machine assisted approach to synthesis makes sense for so many reasons,[2,3] not the least of which is improved safety, information feedback and control leading to detailed audit trails, big data management and neural networking opportunities, use of computational algorithms, especially self-optimization protocols and device communication through the Internet and ‘the cloud’. More specifically the ability to collect more data in real-time to profile reactivity and reaction optimisation is essential,[4] as is the use of more kinetic data to improve mechanistic understanding, reaction planning[5] and the discovery of new reactivity.

Much new equipment and enabling technology is becoming available to enhance the assembly of our functional molecules,[6,7] nevertheless this still needs to be better integrated with existing synthetic practice, to provide a stimulus for a new vision of how laboratories will work into the future. As with any advances, there are followers of fashion and inevitable sceptics of developing technology. However, for us to make judgments as to their value, we believe a deep understanding of the fundamental principles of synthesis, its importance, relevance and sustainability, are all a necessary requirement. Also, a proper understanding of its limitations is equally vital in guiding experimentation. Synthesis can be both an art and a science; consequently any enabling device must deliver at both ends of the spectrum and ideally respecting the Green Chemistry agenda.[8]
Given the multi-faceted nature of our research work in synthesis, a rapid and often complex decision making process arises as a result of experimental observations, which is not readily accommodated by machines. It is imperative therefore to maintain flexibility and ease of operation of any new enabling system. Accessibility and freedom to operate are increasingly important factors. Time and labour saving devices that are robust over extended time periods are now becoming readily and commercially available. Advanced computing power and more open source software opens new opportunities for everyone.

Nevertheless, as we have pointed out in many previous articles, the engineering of chemistry through control and coordination of many different pieces of equipment simultaneously and possibly across the world will lead to synergistic benefits well beyond our current capabilities.[9,10] Indeed, the ability to link and use data generated at multiple sites from diverse equipment manufacturers presents a major challenge and, while considered normal practice at the process scale, is a huge hurdle in the research environment. This said, the development of flow chemistry and continuous processing for multistep organic synthesis sequences[11] has been a game changing experience, requiring, as it did, new thinking, new equipment and indeed a new approach to how we might fundamentally assemble our molecules today.
The Early Days

Firstly we need to put our science into context. While modern developments in computers have allowed new avenues for chemical research to flourish, the idea of implementing control systems to aid in synthesis certainly is not new. For example, one of the earliest reports, dating back to 1965, details how an automated approach to solid-phase peptide synthesis enabled the preparation of two polypeptides – Bradykinin and Angiotensin II – with minimal researcher intervention.[12] In this case the automation was controlled by the “programmer,” shown to the bottom left of Figure 1, which controlled pumps and the position of two rotary valves.

![Apparatus used to automate solid-phase peptide synthesis in 1965. The “programmer” on the left was responsible for controlling reaction sequences. From [12]. Reprinted with permission from AAAS.](image)

Each rotary valve was responsible for controlling the selection of materials to be sent to the reactor column. The first, consisting of 12 steps, managed the addition of solvents and reagents necessary to increase the length of the peptide chain. The second, rotating one position for every complete revolution of the first valve, ensured that the correct amino acid was added to the chain. In this way the system operated in a simple looping action, allowing for six amino acids to be added to the base
peptide chain over a 24-hour period without manual intervention. At the exhaustion of each amino acid, the reservoirs were refilled and the process repeated with the next set of six amino acids.

The authors noted that this automation led to savings in both time and effort, a factor that would multiply in magnitude as the length of the desired peptide increased.

Automation research in these early days were focussed on controlling simple actions, such as the sequential switching of valves in the study described above, without any form of decision making required. Move forward 40 years and the desired outcomes from automation have not changed – researchers still want to reduce the time burden of repetitive tasks – but the computer technology available had progressed to such an extent that new heights of progress could be reached.

In 2005, our own laboratory conducted a study in which Grossamide, a neolignin natural product, was prepared using a variety of flow chemistry enabling techniques, including the use of polymer-supported reagents, enzymes and process automation (Figure 2a, b).[13] During the initial stages of development for this project, a computer control system facilitated the operation of a parallel-column loading and activation system enabling uninterrupted and continuous processing of an amide coupling step. By incorporating an LCMS unit into the process, we were able to gain valuable process insight (e.g. the rate of consumption of solid-supported reagents) and optimise this step to 90% conversion in a much shorter timeframe than by conventional techniques. The same LCMS unit was also used to provide feedback to researchers during the latter step of the process, in which silica-supported horseradish peroxidase was used to complete the synthesis, to form new carbon-carbon and carbon-oxygen bonds.
a) the synthesis of grossamide was enabled through the use of an automated control system. Reproduced with permission, reference [13]; b) the synthetic pathway selected to prepare Grossamide.

This machine-assisted approach to problem solving was again present in a study from our laboratories the following year in which another natural product, oxomaritidine, was successfully synthesised through the coordinated operation of a seven step flow sequence (Figure 3).[11] We
believe that this synthesis was a game changing publication at its time which has since had enormous impact. The task in this work, however, was significantly more complex than that described above; it consisted of different individual synthetic steps, required numerous input and output material streams and involved the first use of new equipment specifically designed for flow chemistry applications. The complete synthesis task was achieved within hours, a stark difference to the four days required by conventional methods.

Figure 3. The seven step synthesis of (±)-oxomaritidine.[11] Reproduced by permission of The Royal Society of Chemistry.

Later that year, we reported a further development of our computer-aided approach as applied for the preparation of 4,5-disubstituted oxazoles.[14] While this project did not apparently pose as
difficult a chemical challenge as for the previous natural products, it represented a significant increase in technical complexity since we added further features to the control process (Figure 4a, b). Feedback from UV detectors was used to adjust conditions in upstream processes, such as reactor temperatures, in order to maximise output. The process also incorporated numerous multi-directional valves, pumps, an LCMS system, in-line preparative HPLC unit and two liquid handlers to selectively manage starting materials and collect products. The use of a parallel packed-cartridge system enabled the preparation of 13 different oxazole products with excellent yields, and minimal intervention from chemists.
Figure 4. a) the level of technical integration for the synthesis of 4,5-disubstituted oxazoles was enabled through the use of a computer control system. Items outlined in red represent those with direct feedback and control by the computer; b) Example transformation carried out in a single process. Reprinted with permission from [14]. Copyright 2006 American Chemical Society.

Having gained early experience in the automation of both these synthetically and technologically challenging projects, our laboratory subsequently has reported many more applications to active pharmaceutical ingredients (APIs), natural products, porous cages, agrochemicals and perfumery products.[15,16]
The Rise of LabVIEW

In 1986, a US-based company named National Instruments released version 1.0 of their Laboratory Virtual Instrument Engineering Workbench software for the Macintosh operating system, followed six years later with versions compatible with Sun and Windows operating systems. Commonly referred to as LabVIEW,[17] this tool was designed to facilitate the collection of data from and control of laboratory and industrial equipment.

LabVIEW’s visual interface is a key feature that differentiates it from traditional text-based computer programming techniques. Users can drag-and-drop logic components, device drivers and communication modules into a central workspace for their project, connect them together by drawing lines from input and output ports on each module and thus create logical structures for a computer to follow. This process revolutionised the integration of computers across all scientific fields, as such an approach to programming opened laboratory automation to everyone. No longer did scientists require extensive programming knowledge to build their experimental control programmes; instead they could visually construct their experiment sequences.

In the late 2000’s, there was a surge of publications coming from the chemistry community that employed LabVIEW to control and automate basic processes. For this review we will focus on reports dated more recently, as they describe more advanced applications and include important modern applications of new analytical equipment.

One such piece of analytical apparatus that fitted well into the chemist’s workflow was the ReactIR, from Mettler-Toledo.[REF] By incorporating a flow-through cell it was possible to integrate this unit into continuous flow chemistry procedures, thus providing valuable real-time structural feedback to chemists and control engineers about conditions operating within the reactor product stream.[18,19]
This information was put to particularly good use by our group in 2011 when we reported taking advantage of detector feedback to alter process parameters.[20] Dispersion problems in pumped reaction streams can negatively affect flow processes, especially at mixing points where a precise stoichiometry might be required. Traditionally, downstream pumps would follow a simple binary mode of operation – either on or off. Yet in situations with significant dispersion, this would result in an undesired significant excess of the second reagent being added to the stream as the concentration of the first reagent varied. By including IR feedback (Figure 5), a LabVIEW control application was able to gradually increase reagent C’s pump flowrate to match its addition to the concentration gradient caused by dispersion. This intelligent pumping method was demonstrated for the flow synthesis of pyrazoles and during a more demanding stereoselective crotylation process.

Figure 5. Feedback from an IR spectrometer was used to control downstream pump flowrates to precisely match stoichiometry. Reproduced with permission from John Wiley and Sons, reference [2].
Having the ability to monitor accurately the composition of a reactor’s product stream, while simultaneously controlling the reactor conditions, unlocks huge potential and opportunity. If a computer can choose setpoints for equipment parameters and then gauge how well those conditions performed, subsequently it could then select new setpoints in an attempt to self-optimise the reaction – all without any manual input from the research chemist. While a number of groups have published in this area over the last six years, and we have included a few reports later in this article, we wish at this point to highlight a few notable examples here that have made use of LabVIEW.

One of the attractive features of LabVIEW is its ability to integrate with third-party software suites, enabling LabVIEW programmes to take advantage of additional functionality on offer. In the case of self-optimisation, for example, control systems need to process raw data using a variety of statistical analysis techniques which are not available by default in the standard LabVIEW installation. Luckily for non-mathematically inclined chemists, a library is included with LabVIEW that allows interfacing with MATLAB, a software product from The Mathworks Inc. that enables customisable and advanced data analysis.[21]

Poliakoff has reported numerous examples of self-optimisation, facilitated by a combination of LabVIEW and MATLAB, as mentioned in a 2011 report.[22] In this article they also described their use of the simplex algorithm[23] to optimise a number of simple etherification reactions conducted in supercritical carbon dioxide on a larger scale than that reported previously. They were able to realise a significant decrease in time taken for the optimisation approach through the use of computer automated self-optimisation.

A different study, also took advantage of the additional functionality offered by integrating MATLAB with LabVIEW for the optimisation of a Paal-Knorr microfluidic reaction.[24] A ReactIR unit was used to provide in-line analytical feedback via the controlling scripts from which reaction conversion was calculated. In this case, the system used two self-optimisation techniques, steepest descent and
conjugate gradient, to optimise production (defined in the control software as conversion divided by residence time) by varying reaction time and temperature.

Much more recently, an autonomous control system was used to explore new reaction space that had not been reported previously.[25] This study began by attempt to prevent precipitations blocking reactor tubing during a specific reaction of aniline with dimethyl carbonate over an alumina catalyst. Tetrahydrofuran (THF) was added as a co-solvent in an attempt to act as solvent for the formed reaction solids. However it was observed that this led to many unexpected by-products being formed. Instead of simply switching THF for another solvent, the authors took a different tack: they decided to utilise a LabVIEW-powered control system to systematically optimise each product and by-product, using feedback from an on-line gas chromatography system to power self-optimisation algorithms (Figure 6). By collecting characterisation data and identifying each product and by-product, potential mechanisms for their formation could be proposed.

Figure 6. Schematic showing the feedback information flow used to optimise for each of the products and byproducts observed when aniline and dimethyl carbonate were reacted the presence of THF over an alumina catalyst. Figure reproduced with permission from John Wiley and Sons, reference [25].
While there are now more efficient and practical software packages available that can be used to control reactions (see later), the continuing influence of LabVIEW cannot be understated. Indeed, it is still used widely today, as is evidenced by a report published earlier this year in which an asymmetric hydrogenation reaction was automated on a relatively large scale.[26] Using solid supported Rh-(S,S)-Ethyl-Duphos, a commercially available chiral catalyst, researchers were able to run a flow process on a 1 kg day\(^1\) scale with conversion ranging from 95% to 99.6% and an enantiomeric excess ranging from 98.6% to 98.9%, with minimal catalyst leaching.

Before moving into the next section, we felt it would be worth commenting on the influence open-source advocates have had on the uptake of LabVIEW across all disciplines. Indeed, there is a thriving online scientific community that shares equipment drivers and applications publically, thus enabling other scientists to take immediate advantage of their work without needing to rewrite their own algorithms. For example, a 2012 open-source report detailed a LabVIEW template to help with experiment automation.[27]
Machine vision

Before describing the state of computer-powered automation tools in use today, we felt it would be worth briefly outlining the world of machine vision and its application in chemical synthesis. While we have extensively reviewed literature in this area previously,[28] here we want to provide an overview of the utility of these methods.

By giving the ‘sense’ of sight to our computer control systems, we can both monitor and control experimental procedures that normally rely on visual feedback. Some typical examples include monitoring the amount of liquid in a reservoir,[10] detecting gas/liquid flow in reaction tubing[29] or observing colour changes in a reaction flask[30] – tasks that are usually carried out manually by a chemist responding to a visual stimulus.

In the realm of flow chemistry, machine vision opens up wider opportunities in the area of continuous downstream processing of crude reaction mixtures. One of the first reports of such a system describes how the combination of an inexpensive, consumer-grade webcam and open-source technologies was used in a prototype extraction unit.[31] Its operation was simple (Figure 7): an aqueous stream and organic stream were pumped together and mixed vigorously using an inline magnetic stirrer; the resulting biphasic mixture was then pumped into a transparent separation chamber where the heavier phase settled to the bottom, while the lighter phase floated on top; the position of a small coloured plastic float, sitting at the interphase between the two phases, was monitored with a webcam; using this position information, a control system written in Python adjusted the rate at which the lighter layer was removed from the column in order to maintain a set volume of the heavier phase.
We demonstrated the utility of this system with a simple condensation reaction to form hydrazones using phenylhydrazine and various benzaldehydes in the presence of pyridinium toluenesulfonate, obtaining quantitative yields in almost every case. The software system was subsequently expanded to monitor colour changes, as related to dye concentration, to monitor dispersion of a flowing stream.

This extraction process was developed further and a subsequent report described a second-generation unit that facilitated multiple extractions connected in series. In this case, the design of the extractor system was modified to include a high-pressure HPLC pump which was placed at the outlet of the heavier phase. Again, the control system used feedback from a digital camera to adjust the pump flowrate based on the position of a coloured float, but in this case provided continuous adjustment to maintain the position of the float in one location.
The flow synthesis of hydroxyvaline necessitates multiple extractions, owing to its partition coefficient of 0.74 in the binary water and ethyl acetate system. Accordingly this synthesis was selected to showcase the efficacy of the new extractor design, and three extractors were placed consecutively downstream of a reactor coil (Figure 8). The performance of this system was excellent. It was reported that several trials were carried out, each lasting longer than 24 hours and producing over 20g of product, without any input required from a chemist. These systems can especially be used to process batch generated products thereby avoiding manual liquid-liquid extraction methods.

Figure 8. Multiple extraction columns in series have been used for more complex extraction situations, especially when reaction components are soluble in both the aqueous and organic phases. Reprinted with permission from [32]. Copyright 2012 American Chemical Society.

This extraction system has been used in many different studies more recently, including one reported earlier this year which explored the bromination of enaminones with N-bromosuccinimide.[33] In this case the same layout as described above was used, and the
researchers were able to obtain highly pure products following the use of the liquid-liquid extraction unit.

**The State-of-the-Art, Today**

There has been a surge of interest in recent times to apply automation to aid with the development and synthesis of specific active pharmaceutical ingredient (API) targets, right from a discovery scale to process scale chemistry. This has been reported for both the control of telescoped steps[34] and the optimisation of reactions corresponding to individual steps.[35]

**Determination of Kinetics**

Automated control systems have also been put to excellent use to obtain kinetic information about reactions of interest. Recently Jensen has reported the use of automation to aid in the collection of kinetic data when performing continuous aminocarbonylation of a variety of aryl halides substrates.[36] In their biphasic gas-liquid system, the system screened a number of parameters including gas stoichiometry and pressure, in addition to the standard temperature and reaction time variables, to develop a model describing the kinetics of the reaction. In this case, researchers confirmed that they were able to manipulate the favoured mechanistic pathway by varying temperature.

This more recent work built on their previous reports of kinetic-based optimisation processes, including the determination of rate parameters for a Diels-Alder reaction[37] and a multi-step reaction system in which the selectivity during a nucleophilic substitution was explored using online HPLC.[38]

Researchers have also adapted this system for the determination of batch kinetics under a continuous flow regime.[39] The determination of kinetics in traditional batch processes requires
chemists to carry out a number of experiments, usually followed by manual washing, analysis and intervention to set up the next experiment. A novel approach was taken in which the flow rate of the reaction mixture travelling through a reactor was slowly decreased over time, in effect giving a varying residence time distribution. While in most continuous systems, such changes are undesired, in this case researchers were able to gather snapshots of different conditions using an inline IR detector without needing to carry out individual experiments separately. Reports noted that such an approach offered considerable time savings and importantly, reduced material consumption.

A similar approach was taken by a different team recently, in which they varied reaction residence time through the ramping of flow rate, for the determination of kinetics of a substitution reaction between pyrrolidine and 2,4-difluoronitrobenzene.[40]

Reaction optimisation

It has been reported extensively that in-line and on-line analysis techniques coupled with automated reaction control can open up new possibilities for reaction self-optimisation, in which feedback from prior reactions can be used to determine new conditions to investigate in an attempt to drive up reaction yields.[2,41,42] Consequently we will not report in detail on this, rather we will highlight a few recent examples where self-optimisation techniques have been used to aid optimisation.

While there have been many reports of self-optimisation strategies being adopted that use in-line IR spectroscopy for feedback,[43] largely owing to its ease of use, two recent reports of interest harnessed newer techniques to replace IR. In 2015, an inline NMR unit was used to optimise the reaction between 4-fluorobenzaldehyde and aniline (Figure 9).[44] Researchers developed LabVIEW scripts to control and monitor equipment parameters. Peaks corresponding to the aldehyde starting material and imine product were monitored for each experiment to provide feedback to MATLAB
optimisation functions. In this case, the system carried out 29 reactions before reaching an optimum.

Figure 9. NMR was used to provide reaction performance feedback to drive the self-optimisation of an imine formation reaction. Reproduced from [44] – published by the Royal Society of Chemistry.

More recently online mass spectrometry (MS) and HPLC were utilised to build an optimisation system for the reaction between methyl nicotinate and methylamine.[45] It was observed that at sufficiently low concentrations, the response from the MS was linear to reaction yield, as observed using online HPLC, such that after building a calibration curve the use of HPLC was no longer required. This led to a reduction in time required to process the results for each experiment iteration, making the optimisation process significantly more time efficient. The SNOBFIT self-optimisation algorithm[43] employed for this example gave similar results to a benchmark Design of Experiments (DoE) approach run in parallel.

DoE is a widely employed method for optimisation in pharmaceutical processes, as it explores a wide reactivity space and generates correlation models that can be used to predict the outcome from a reaction given particular conditions. However, as the number of reaction parameters to be optimised increases, the number of experiments required to gather sufficient information increases
greatly. Consequently, DoE optimisations require significant investments of time to conduct experiments and in the quantity of material consumed. It is also worth noting at this point that in contrast with self-optimisation algorithms, all DoE experiment set points are generated at the beginning of the optimisation process and no feedback occurs to generate new conditions.

Automation of DoE-based research programmes has been reported by a number of researchers as being able to significantly reduce the time burden required for this method.[46–48] Reported examples include multicomponent reactions such as the Suzuki-Miyaura cross-couplings between a library of starting materials,[49] which incorporated a liquid handler and a variety of ligands.

Use of computer control has also facilitated the development and adoption of novel optimisation algorithms. A study reported in 2015 describes how a polymerisation reaction was optimised using a multi-target function, optimising for both product particle size and reaction conversion simultaneously.[50] Such requirements more closely align with larger scale applications, such as the copolymerisation situation described in the paper, where a number of factors may influence how well a set of reaction conditions can be said to perform. In this example, the team conducted experiments first in silico to check the performance of their algorithm before carrying out experiments in reality. While it was found that the computed results did not closely align with actual experimentation, the researchers noted that they were able to obtain experimental conditions that satisfied their requirements without any prior knowledge of the system.

Discovery

We have written about how approaches of ‘directed serendipity’ have resulted in the discovery of new reactions previously, so we will not pursue a discussion of automation as applied for reaction discovery here. Should readers wish to find out more about this area, we would encourage them to read reference [2], where we have included a number of relevant reports.
It is our belief that autonomous systems have the power to transform the area of discovery of new API candidates, closing the gap between library generation and biological testing. Modern times are calling for a change from traditional practises – one focussed on the mass collection of data, perhaps to be used for training of machine learning algorithms and the like, from reactor systems and analytical equipment. Coupling multiple sources of information together could lead to the discovery of new or unexplored reactivity.

Our own laboratories have reported such a system that connected a synthesis process to frontal affinity chromatography (FAC) to perform inline screening tests for GABA<sub>a</sub> inhibitors[51]. In this case target proteins (Human Serum Albumin) were bound to a column through which solutions containing inhibitor analogues were pumped. By comparing the retention characteristics of the synthesised molecules and a void marker (a substance that does not interact with the bound protein), it was possible to rank interaction relatively for each of the 22 analogues synthesised and tested.

The above process was reliant on the modification of a known drug compound to generate a series of analogues. High-throughput discovery techniques are reliant on computer automation to significantly reduce the human resource required to conduct the large number of individual experiments that comprise them. In one recent report, over 1500 Pd catalysed cross-coupling experiments were carried out over the course of a day to generate a library of small molecules suitable for further biological testing.[52] Researchers harnessed automated sample dispensers, used routinely in biological applications, to rapidly dose 1536-well plates with reaction mixtures and UPLC techniques to quantify reaction performance. The utility of the system was demonstrated for route screening, reaction discovery and automated reaction optimisation (DoE).

Other reports detail automated systems that can assist with the discovery of new candidates by combining fragments in iterative processes to build new structures. In 2015, Burke reported such a system that was designed to repetitively carry out a series of deprotection and coupling reactions,
along with integrated purification between steps, to build a linear backbone in a series of small molecules.[53] The system was able to do so by incorporating a number of syringe pumps, switching valves and reaction columns (Figure 10). They also reported the use of the system for the preparation of a number of cyclic compounds, including citreofuran and oblongolide.

Figure 10. The system designed by Burke et al. consisted of a number of pumps, switching valves and reaction sections to iteratively perform deprotection, coupling and purification steps. From [53]. Reprinted with permission from AAAS.

This idea of modularity, where steps are isolated from one another and can be performed in different orders, has been adopted by Seeberger to prepare five separate APIs by switching steps at three conceptual levels.[54] In addition to the choice of starting materials, the group switched the order in which reactions occurred (in reaction ‘modules’) and modified reagents meeting starting material streams within these reaction modules. The system was demonstrated with the synthesis of five potential APIs.

*Ease of Telescoping*
Passing the output from one reactor into the inlet of a downstream process, commonly referred to as “reaction telescoping”, is an area that has been transformed by computer-controlled automation. Two of the examples listed above have incorporated some aspects of telescoping, in particular the report from Seeberger.

Other groups have developed processes where telescoping is central to their goal of small-scale production of APIs. Last year a modular system was reported that consisted of a number of different reaction and workup modules between which connections could be changed to effect different outcomes.[55] These module included reactor coils of various volumes, membrane- and gravity-based liquid-liquid separation, crystallisation units and formulation processes, all of which were designed to fit into a relatively compact volume – roughly the size of a large consumer refrigerator (Figure 11). By changing the flow paths between each unit operation in the system, and by modifying the starting material streams, researchers could produce large quantities of diphenhydramine hydrochloride, lidocaine hydrochloride, diazepam and fluoxetine hydrochloride in a continuous and telescoped manner. The researchers involved with this process had previously reported a unit for the production of a single API, including final purification and formulation steps.[56,57]
Figure 11. A compact synthesis unit capable of producing four APIs has been reported that uses common components for all target molecules. A control system switches individual unit operations into and out from the flow line to manage each synthesis. From [55]. Reprinted with permission from AAAS.

In 2015 we reported a multi-step telescoped process that included three reaction steps and four downstream processing steps (Figure 12), including continuous liquid-liquid extraction (building on our previous reports of machine vision systems) and solvent switching.[9] Automated control enabled the process to be run continuously for extended periods, which when run at steady state could produce 8 mmol.h⁻¹ of pure product.
Figure 12. Three transformation steps were telescoped using commercially available laboratory equipment, incorporating a number of downstream processing steps. Custom Python scripts were used to automate the process.
The Future of Automated Control

We believe the future of computer-powered automation is vast and has potential to impact every part of a modern scientist’s work. As we have seen in many other areas, developments in one discipline or technological sector can be applied in others with great success. Chemistry is no different.

The Cloud

Chemists will be familiar with cloud-based tools to assist them with literature searching, such as SciFinder[58] from CAS, Reaxys[59] from Elsevier and Thomson Reuter’s Web of Science,[60] and bibliography management, such as Elsevier’s Mendeley.[61] At the bench, organisations opt for electronic laboratory notebooks – many of which are accessible through an internet browser – and internet-based chemical inventory management tools.[62] It appears that several areas of a typical workday is influenced by remotely hosted software tools, yet there has been a lack in the area of using this to control “wet” chemical experimentation. This is now starting to change.

An interesting LabVIEW web application in 2013 was used to control three syringe pumps, the temperature of a reactor and display data collected from an inline UV-Vis detector.[63] This set up was designed to be used in an educational environment to allow students to conduct the synthesis of methyl orange remotely. Students signed in to the system through an internet browser, after which time they were shown an interface showing the status of each piece of equipment (Figure 13). Parameters could be changed in real-time and the effects of each change on reaction outcome could be monitored with a live camera feeds and a detector plot.
In a different study, Nottingham scientists decided to set up a remote desktop system which allowed a researcher from Nigbo, China, to remotely take control of a computer running a local LabVIEW application connected to experiment apparatus.[64] For this example they opted to try a pre-optimised etherification reaction of $n$-propanol over a catalyst of $\gamma$-alumina conducted in supercritical carbon dioxide. While this system was not cloud-based in the true sense of the word, as it operated through remote desktop software, we feel it is nevertheless interesting to highlight here as an example of how internet communication is changing the landscape for the experimentalist.

The previous two reports were developed primarily in LabVIEW, and as such were not modular and therefore not suited to a fast-paced research and development environment where chemists might conduct multiple different experiments, using different pieces of equipment, over the course of one working day. Indeed, both reports focussed more on the educational applications of such systems which, by design, consist of single experiments configured to be repeated many times over.
Earlier this year our group reported a new internet-based software platform that enables chemists to monitor and control their reactions from any location.[10] This system is entirely modular, with stand-alone driver scripts enabling the integration of reactor and analytical equipment from any manufacturer. The system can also be reconfigured in real-time, allowing researchers to hot-swap one piece of equipment for another during the development of their reactions. The system is purely based in the cloud, with user-server and server-equipment communication occurring entirely through internet (TCP/IP) protocols (Figure 14).

Figure 14. Our reported control system is fully cloud-based, with all communication occurring through TCP/IP. Users access the system through an internet browser. Reproduced from [10].

We demonstrated this system with a simple automation of a hydration reaction between 3-cyanopyridine and water over manganese dioxide, incorporating feedback from a small footprint Advion mass spectrometer (MS)[65] to provide information about conversion. We subsequently carried out two self-optimisation reactions where the system was left to find autonomously the optimal conditions for reactions.

In the first example, the system optimised the hydration reaction for three parameters (reaction temperature, residence time and concentration), again using MS data for conversion monitoring. In the second example, we decided to delve into a significantly more complex example which
optimised five different parameters for an Appel reaction (Figure 15). In this case, the system was configured to approach the problem in the same way that a chemist would; rather than just optimising for yield, it optimised for conversion, material throughput and material consumption. The cloud control system conducted 30 reactions over the space of 10 hours, finding conditions that produced 1.9 g h\(^{-1}\) of the desired product (representing 92% yield).

![EQUIPMENT SCHEMATIC FOR THE FIVE DIMENSIONAL OPTIMISATION; B) REACTION SCHEME; C) THE FIVE PARAMETERS MANIPULATED BY THE CONTROL SYSTEM DURING THE OPTIMISATION PROCESS. REPRODUCED FROM [10].](Image)

Figure 15. a) Equipment schematic for the five dimensional optimisation; b) reaction scheme; c) the five parameters manipulated by the control system during the optimisation process. Reproduced from [10].

The integration of the cloud into the experiments carried out by chemists is really only embryonic. We see the future of synthesis being made up in part by a network of sensors, each sending data to each other and having set points manipulated by a central control system. Such a system would represent a version of the ‘Internet of Things’[66] as applied in our chemistry environment, something we have described and written about recently.[67]

‘Open’ Research

Traditionally many industries are strongly protective of any internal ideas and developments for chemical synthesis, largely owing to the highly competitive nature of the sector. The changing
economic realities in the modern world are forcing changes, however. Recently 12 well-known companies including Pfizer, GSK, Merck & Co. and AstraZeneca have formed a precompetitive consortium to share their developments across a number of related areas including high throughput techniques, laboratory automation and downstream processing (crystallisation, drying and purification).[68] This group has stated their desire to publish results and new developments so that other groups are able to benefit from their collective research activities.

The process for forming such collaborative groups has been long in development. In 2010, a group from the biomedical research sector noted that expenditure in research was disproportionately large relative to the impact it was having on general human health.[69] A number of different models for precompetitive collaboration were explored to identify the most effective methods to drive multi-party R&D. Four years later, a similar workshop was held focussing on the engineering and chemistry aspects of the pharmaceutical industry,[70] from which the 12-member consortium described above developed.

The idea of open-sourcing knowledge, in which anyone is free to use, adapt and modify results or methods to suit their applications, is driving innovations in the area of 3-dimensional printing. When applied in a scientific arena, this method of creation enables chemists to rapidly prototype equipment designs, including for the design of new batch reactor systems,[71] microfluidic chips[72,73] and low cost consumable equipment.[74]

One recent report used 3D printing to not only construct a reactor framework, but also to print layers of catalysts in different sections of the system (Figure 16).[75] Once the device had been printed, the starting material reservoirs were charged with solution and the device tilted to mix the two streams in the first reaction chamber. By subsequent rotation of the device in 90° increments, researchers were able to successfully perform a 3 step synthesis. While the yield reported using this method was lower than the comparable batch methodology (32% vs. 40%), the ability to prototype
quickly new reactors rapidly opens up a number of opportunities for the sharing of chemistry developments.

Figure 16. The 3D printed vessels included a number of reaction chambers, including those that incorporated catalysts printed onto surfaces. Reproduced from [75] - published by The Royal Society of Chemistry.

The same group subsequently reported the use of a 3D printing approach for the creation of high temperature reaction vessels to produce a number of metal-organic frameworks (MOFs)[71] and for the synthesis of ibuprofen.[76]
Conclusions

In this lead article for this Symposium in Print, we have briefly reviewed some of the science that is developing not to fully automate the full repertoire of synthetic chemistry - this is unlikely to happen in the foreseeable future - but to better engineer chemistry to facilitate discovery. In a rapidly evolving scientific world, constituent elements such as chemistry must become more responsive and adapt or face becoming redundant.

Synthesis is inherently an experimental subject requiring the marshalling of deep understanding and advanced skill sets. The flexibility and diversity of the molecular assembly process provides a challenging and inspiring environment for computational and machine-based technologies. We have explored in this review how these machine-assisted and computer control systems are now performing many of the labour-intensive activities that previously required significant human resource to complete.

A holistic systems approach to our work makes sense in that the majority of our functional products require us to use multiple steps in their formation. While up-stream design and innovation still dominates our thinking, more attention to the unit operations involved in downstream processing will feature into the future.

The integration and coordination of techniques, equipment and machinery of all kinds is driving a new agenda; not one in which the synthesis chemist is replaced at the bench, for we believe skilled human input will always be required, but rather one where emphasis is placed on efficient use of time. As Prof. Robert Grubbs referred to in a recent lecture at the University of Cambridge, “you don’t have to work all the time but you have to think all the time.”

Acknowledgements

The authors gratefully acknowledge support from the Woolf Fisher Trust (DEF) and the Engineering and Physical Sciences Research Council (SVL. Grant codes: EP/K009494/1, EP/M004120/1).


The Internet of Chemical Things. Ley S, Fitzpatrick DE, Ingham RJ, Nikbin N. Beilstein Mag. 2015; DOI 10.3762/bmag.2


