Chemistry—Methods

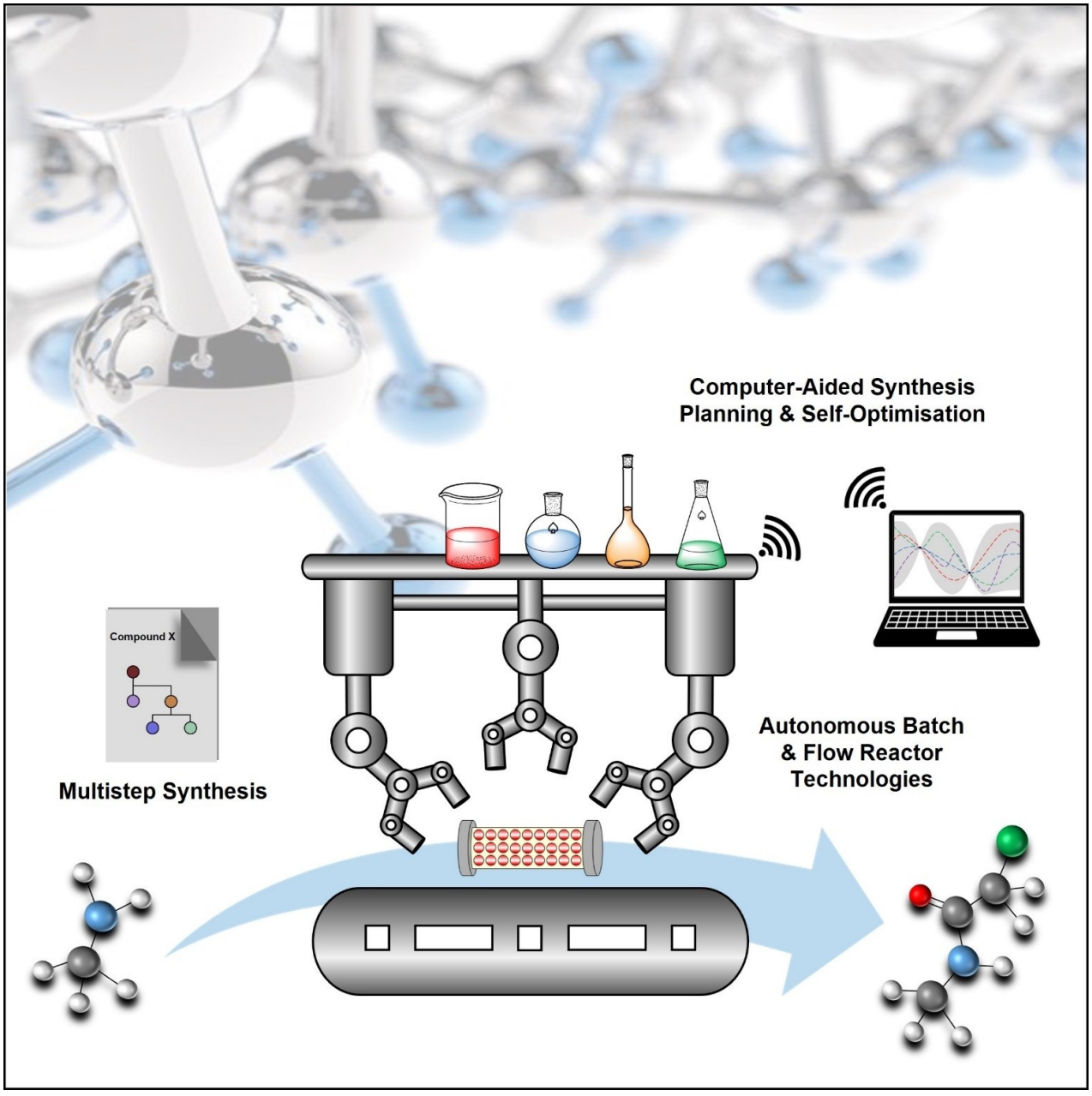
Review
doi.org/10.1002/cmtd.202300021

Chemistry Europe
European Chemical Societies Publishing

www.chemistrymethods.org

Recent Developments in Reactor Automation for Multistep Chemical Synthesis

Adam D. Clayton*[a]



Reactor automation is revolutionising the way new chemical processes are discovered and developed. Assigning repetitive aspects of chemical synthesis to machines, such as experimental execution and data collection, provides more time for researchers to focus on critical interpretation and creative problem solving. The ability to autonomously prepare late-stage intermediates and complex products, rather than just simple starting materials, will play a central role in applications such as the efficient exploration of chemical space and responsive manufacturing. However, translating automated technologies from specific single-step tasks to more general multistep syntheses remains a significant challenge, owing to high structural diversity and chemical/physical interdependencies between the steps. Robotic batch and continuous flow platforms are gradually becoming more universal, providing access to a wider range of chemistries required to achieve autonomous multistep synthesis. Advances in process analytical technologies have enhanced our ability to monitor interconnected reactions in real-time, thus accelerating data collection and giving greater process control for ensuring a high standard of safety and product quality. Integration of these tools with control software creates a feedback loop, which can be harnessed for adaptive and flexible multistep screening or holistic self-optimisation. This review presents recent developments in the application of automated reactor technologies for multistep chemical synthesis, including batch and continuous flow platforms. Specifically, this review highlights how the integration of control software with advanced process analytical technologies and machine learning algorithms are accelerating the synthesis of complex molecules.

1. Introduction

The synthesis of complex molecules that form many indispensable medicines and materials are comprised of multiple transformations from readily available starting materials, which requires precise control over both chemical reactivity and reaction conditions. Competing side reactions between chemical species involved in each individual step are often avoided by conducting intermediary work-ups and purifications, which require significant input of energy and materials (e.g., solvents). In contrast, directly concatenating reactions (known as 'one-pot processes' or 'reaction telescoping') can reduce waste generation, thus better aligning with the net zero carbon vision of the future.[1] In reality, most total syntheses will utilise a combination of these approaches where appropriate, and therefore encompass a variety of complex development challenges which requisite holistic consideration.[2]

Efficient multistep processes must be developed and optimised in an ever-shortening timeframe, owing to a shift in recent years from large batch production to more responsive manufacturing. The importance of being able to rapidly synthesise complex molecules in response to changes in supply and demand was highlighted during the COVID-19 pandemic; a concept which extends to many aspects of modern society. For example: (i) as the supply of traditional resources are depleted, existing processes must be changed to utilise more sustainable starting materials and reagents; (ii) improving treatments with personalised medicine results in a continually fluctuating demand for different active pharmaceutical ingredients.

[a] *Dr. A. D. Clayton*
Institute of Process Research and Development
Schools of Chemistry & Chemical and Process Engineering
University of Leeds
Leeds, LS2 9JT (UK)
E-mail: A.D.Clayton@leeds.ac.uk

The ongoing digitalisation of chemistry has started to transform the way scientists and engineers discover and develop new synthetic processes.[3] For example, the application of machine learning to literature datasets has enabled the prediction of retrosynthetic pathways and reaction outcomes,[4] and the design of novel materials.[5] Notably, cyber-physical systems have accelerated the on-demand development of improved syntheses, through automation and real-time data generation for enhanced process control and advanced modelling,[6] often outperforming expert decision-making. These technologies were initially developed for single-stage reactions, however translating these benefits to more useful multistep processes has the potential to revolutionise the next generation of chemical synthesis. For example, compound library generation often relies on late-stage functionalisation of a precursor, owing to the laborious nature of manually conducting numerous multistep syntheses. However, automation of the repetitive aspects of multistep synthesis would overcome this barrier, making early-stage diversification a more realistic strategy for accessing and exploring a larger portion of chemical space. In this review, we present some recent (i.e., 2018 onwards) representative examples of applying automated experimental platforms (batch and continuous flow) towards multistep syntheses.

2. Discussion

2.1. Batch platforms

Automation of single tasks in traditional synthetic chemistry, such as column chromatography, were amongst the first examples of lab digitalisation. The main advantage of these was an increased efficiency, as it freed researchers from manual and repetitive experimental procedures. Naturally, these systems evolved to automate entire workflows, such as the solid-phase synthesis of peptides,[7] oligonucleotides,[8] and oligosaccharides.[9] In all these cases, building blocks are connected in a multistep synthesis involving sequential iterations of a single coupling reaction. More recently, automated

Review
doi.org/10.1002/cmtd.202300021
Chemistry Europe
European Chemical Societies Publishing

methodologies for iterative C–C bond formation have started to emerge for the controlled synthesis of organic molecules with greater structural diversity.

A notable example was the orchestrated formation of C(sp^2)–C(sp^2) bonds via sequential Suzuki–Miyaura cross-couplings of commercially available *N*-methyliminodiacetic acid (MIDA) boronates.[10] Key to this process was the ability to iteratively cycle deprotections and couplings enabled by the MIDA ligand, which prevented undesired oligomerisation. In addition, MIDA boronate intermediates could be autonomously purified by silica gel chromatography utilising a catch-and-release protocol. However, MIDA boronates are incompatible with the aqueous basic conditions or nucleophilic reagents utilised in C(sp^3)–C bond forming reactions, making them unsuitable for automated synthesis of valuable non-planar and stereogenic small molecules.

Therefore, Burke et al. reported the development of hyperstable tetramethyl *N*-methyliminodiacetic acid (TIDA) boronates **8** (Figure 1a), which enabled automated stereospecific C(sp^3)–C(sp^2) and C(sp^3)–C(sp^3) bond formation via Suzuki–Miyaura cross-couplings and 1,2-metallate rearrangements respectively.[11] A bespoke iterative synthesis platform was developed, which was designed around a single syringe pump with an 8-position valve. This was used to control the fluid movement around the different unit operations for the iterative deprotection-coupling-purification cycles (e.g., silica gel column for catch-and-release purification of TIDA boronate intermediates). This technology was leveraged towards the automated total synthesis of two natural products: ieodomycin C **7** and macrolactone sch725674. The carbon skeletons of the products were constructed using two C(sp^3)–C(sp^2) bond formations for ieodomycin C **7**, and two C(sp^3)–C(sp^3) bond formations for macrolactone sch725674.

However, this work also highlighted the requirement for very different reaction conditions (catalyst, ligand, base, solvent, time) for Suzuki–Miyaura cross-couplings depending on the substrates. Indeed, the reliance of iterative synthesis platforms on either generalised or manually predetermined reaction conditions remains a significant challenge. The ability to autonomously re-optimise reaction conditions for different substrates, or predict suitable conditions using machine learning models, will help overcome these limitations in the future and drive towards a more fully autonomous workflow.

Dr Adam Clayton is currently a University Academic Fellow in Process Research and Development at the University of Leeds. He graduated with a degree in Chemisry from the University of Huddersfield in 2016, and a PhD in Chemical and Process Engineering from the University of Leeds in 2020. He was then awarded a Royal Academy of Engineering Research Fellowship in 2022. His research interests focus on sustainable process development, including multistep flow chemistry, integrated catalysis and autonomous optimisation.

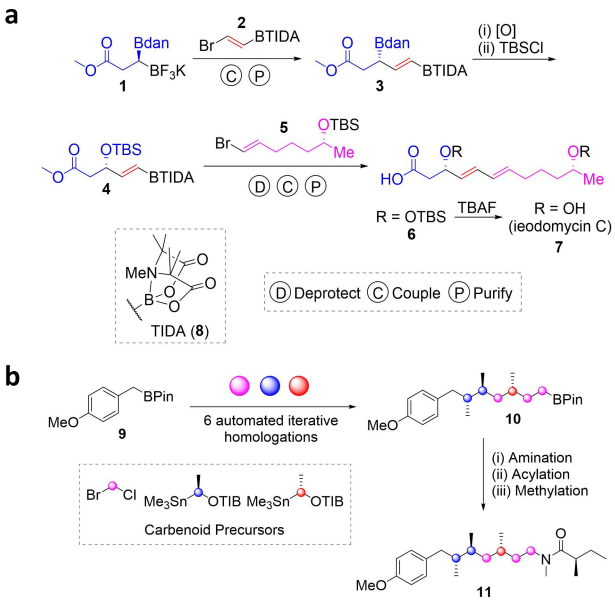


Figure 1. Automated assembly of small molecules using iterative organoboron chemistry. (a) Synthesis of ieodomycin C via C(sp^3)–C(sp^2) Suzuki–Miyaura cross-couplings enabed by TIDA boronates. (b) Stereocontrolled synthesis of (+)-kalkitoxic precursor **11** via six sequential carbenoid homologations.

In similar work, Aggarwal et al. reported the stereocontrolled formation of C(sp^3)–C(sp^3) bonds via iterative boron homologations with chiral carbenoid building blocks (Figure 1b).[12] In contrast to the previous example, this approach enables carbon chains to be built one-carbon-at-a-time in an assembly-line fashion, rather than coupling larger fragments. This methodology is well-suited for automation, as it employs a single set of reaction conditions using a small set of common repeat building blocks, and does not require additional deprotection steps. However, the use of air sensitive organometallic reagents and thermally unstable carbenoids represent significant challenges. To overcome these, a commercially available robotic platform (Chemspeed Swing Platform) was used, which is capable of performing reactions in an inert environment and at low temperatures.[13]

Automated protocols for iterative Matteson homologations and chiral carbenoid homologations were developed, which required adjustments to be made to the standard laboratory procedures. These included vigorous shaking during addition of *n*-BuLi, and the replacement of volatile diethyl ether for anhydrous *t*-butyl methyl ether (TBME) as the solvent. The solid-phase extraction module was utilised for automated silica-plug filtration to remove the lithium salt by-products, followed by concentration under reduced pressure. However, for chiral carbenoid homologations using lithiated benzoate esters, formation of a viscous reaction mixture forced a solvent switch to 1,2-dichloroethane (DCE) prior to work-up. Indeed, the requirement for homogeneous solutions remains a limitation for chemical robotics. The developed approach was demonstrated for the automated assembly-line synthesis of a late-stage intermediate **11** towards neurotoxin (+)-kalkitoxin. Six iterative homologations were used to construct the carbon

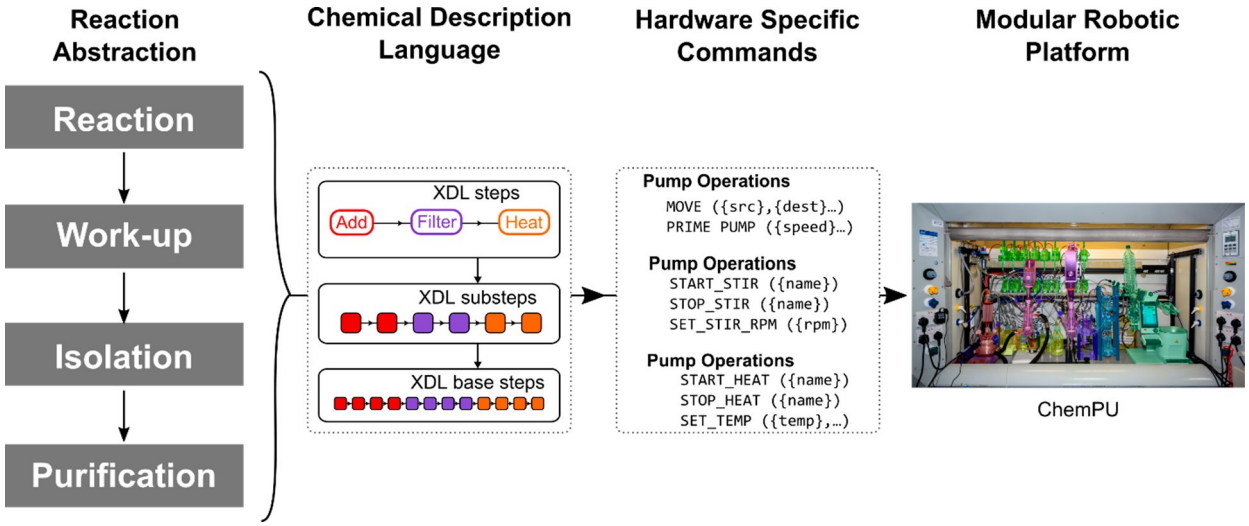


Figure 2. Abstraction and scripting of complex molecule synthesis into steps that can be run using physical hardware. Reproduced from Ref. [16] under the terms of CC BY 4.0. Copyright the Authors.

chain with the desired stereochemistry, followed by an amination-acylation-methylation sequence using the same robotic platform.

Although significant advances have been made towards automated assembly of small molecules, these are still mostly limited to successive iterations of similar reactions. Due to the complexity and variety of organic molecules, a platform capable of carrying out a diverse range of multistep processes is desirable. However, this would require a universal standard to enable automation of chemical synthesis more generally. To achieve this, Cronin et al. developed the Chemputer, a system capable of interpreting standardised synthetic protocols and conducting them autonomously on a modular robotic platform (Figure 2).[14] The Chempiler is a programme that was developed to produce low-level instructions for the physical system. Information about the system is compiled in GraphML format and combined with a chemical assembly (ChASM) scripting language. Interpretation of synthetic procedures are achieved by formalisation of a written synthetic scheme using a chemical descriptive language (XDL). Hence, it is possible to abstract and run reported syntheses without reconfiguration of the platform.[15]

To maximise the compatibility of the system with existing chemical literature, the platform was designed to conduct reactions using conventional bench chemistry. Hence, reactions were performed in round-bottomed flasks, and had access to subsequent work-up (liquid/liquid separations), isolation (filtration) and purification (evaporation) modules. The modules were connected to a fluidic backbone, comprised of a series of pumps used to transfer reaction components/mixtures between different physical operations, as well as conduct washing procedures between steps. The capability of the system was demonstrated for the synthesis of three active pharmaceutical ingredients (APIs) with varying multistep complexity: (i) rufinamide – 2 reactors/1 filtration; (ii) sildenafil – 2 reactors/3 filtrations/1 separation/1 evaporation; (iii) nytol – 4 reactors/2 filtrations/5 separations/3 evaporations.

In the absence of reaction monitoring, even the most well-equipped platforms are limited to conducting standardised protocols designed for the synthesis of a single molecular target. However, chemical discovery often relies on the synthesis and screening of small-molecule libraries, where the application of a rigid procedure may limit the breadth of the substrate scope. To account for changing reactivities of different substrates, Hein et. al. developed an adaptive auto-synthesiser with online process analytical technologies (PAT) for real-time feedback (Figure 3).[17] Online HPLC and FTIR were combined as orthogonal PAT tools for accurate quantification and detection of non-UV active or unstable species respectively, where sample aliquots were delivered to the HPLC using a series of pumps and valves. Notably, an autonomous peak

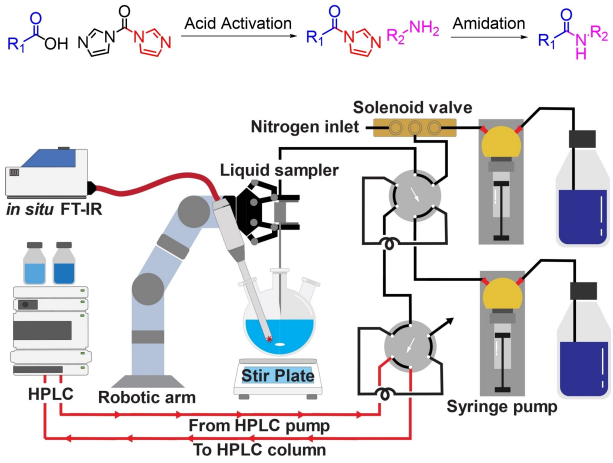


Figure 3. Schematic of the Auto-Synthesiser with online PAT feedback used to flexibly perform one-pot two-step CDI-mediated amidations. Reproduced and adapted from Ref. [17] under the terms of CC BY-NC 4.0 with permission from the Authors.

identification process was designed to assign each peak as an added chemical, substrate, or product ready for data processing. Hence, the platform was capable of performing flexibly-timed procedures by adjusting the timing of critical actions based on information feedback loops.

This approach was implemented for one-pot two-step CDI amidations, where the first step involves acid activation to form an acyl imidazolide intermediate, which then reacts upon subsequent addition of the amine. This process required three flexibly-timed critical actions for optimal operation: (i) unreacted CDI from the first step required an aqueous quench to avoid undesired side reactions with the amine; (ii) CO_2 by-product must be purged after the CDI quench as it negatively affects the rate of the amide bond formation; (iii) amine must be added immediately after the CDI quench and CO_2 purge, as the acyl imidazolide intermediate can hydrolyse back to the starting carboxylic acid. HPLC monitoring of the starting materials, intermediate and product, coupled with IR monitoring of CDI and CO_2, enabled adaptive synthesis of six amides using substrates with different reactivity profiles.

2.2. Continuous flow platforms

In general, multistep continuous flow syntheses involve linear sequences of interconnected reactors where more than one transformation occurs. In contrast to one-pot batch reactions, continuous flow systems enable inline purification and facile addition of reagents at precise points in the sequence.[18] Furthermore, steps which require different reaction conditions can be compartmentalised in separate reactors, enabling more optimal operation and efficient processes. These advantages create the possibility of uninterrupted reaction networks, which have shorter production times and are less at risk of potential supply chain disruptions. This, combined with the design of modular and reconfigurable continuous flow platforms, enables the on-demand synthesis of different complex molecules.[19] However, limitations of this approach include solvent/reagent incompatibility and a mismatch of timescales between consecutive steps, which can lead to inefficient downstream reactions. Nevertheless, there has been a drive in the application of continuous flow as an enabling technology for multistep synthesis, where ease of integration with online/inline PAT enables these platforms to be more readily automated.

Although most examples of automated multistep continuous flow synthesis are for small molecules (e.g., APIs), there is an increasing application of this technology towards materials science, such as the development of self-driving polymer labs.[20] In this context, automation enables systematic variation of polymer compositions and properties through precise control of stoichiometry between monomers and the initiator. Park and Waymouth et al. utilised this for the computer controlled high-throughput synthesis of polyester and polycarbonate libraries.[21] A homopolymer library of 41 poly($_L$-LA)'s, with degrees of polymerisation from 10 to 50 in increments of 1, were synthesised in 6 minutes by programmatically varying the flow rates of THF and monomer solutions through a single stage reactor.

The system was then adapted for the more challenging multistep synthesis of diblock copolymers (Figure 4), where the varying reactivity profiles of monomers often necessitates different residence times. To overcome this challenge, a catalyst switching approach was developed by exploiting the acidity-dependant activity of urea anions. After polymerisation of the first monomer, a second urea is added which undergoes proton transfer with the first urea anion. This quenches the first catalyst and generates a new urea anion which has matched reactivity with the second monomer. As previously, this approach was used to programmatically synthesise a library of 100 AB diblock poly(VL)-*block*-poly($_L$-LA) copolymers, with degrees of polymerisation from 10 to 46 in increments of 4, in just 9 minutes. The narrow average molecular weight distribution achieved ($Đ$ = 1.13) demonstrated the high degree of control provided by this method.

In addition to the challenge of disparate reactivity profiles, autonomous continuous flow synthesis of small molecules must also consider reagent compatibility between the steps. Hence, recent approaches have focused on developing platforms which removes any interaction between the individual synthetic steps. For example, Khan, Wu et al. automated the synthesis of non-peptide APIs by merging solid-phase synthesis (SPS) and continuous flow (Figure 5).[22] In this method, the starting substrates are bound to a solid support within a column reactor, and the target molecule grown through a series of transformations by treatment with various reagent solutions. As the reagents remain in the mobile phase, they are kept independ-

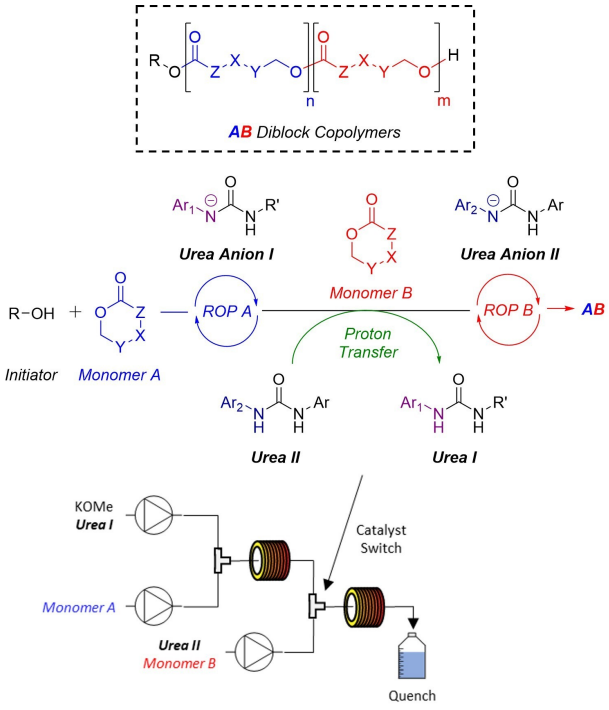


Figure 4. Reactor set-up for AB diblock copolymer synthesis utiliisng a catalyst switch based on proton transfer.

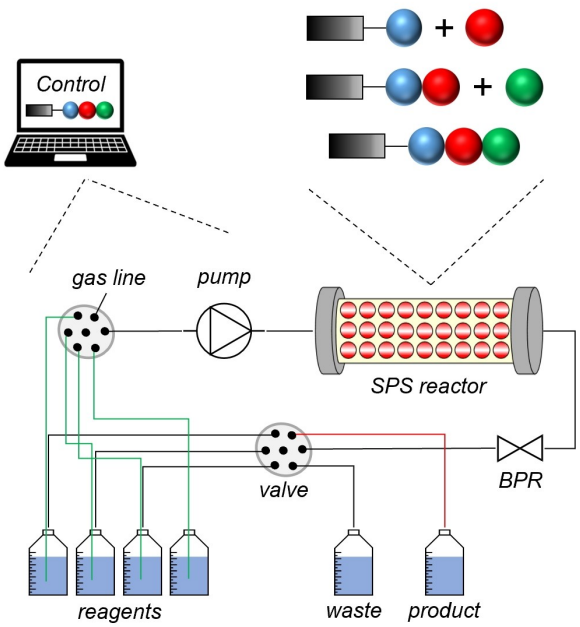


Figure 5. Schematic of the automated SPS flow synthesiser.

ent from one another, thus overcoming the limitation of reagent incompatibility, and enabling the automation of different reaction types in a single synthetic sequence. At the end of the sequence, a cleavage step can be performed to remove the desired product from the support.

The platform utilises multiposition valves to route reagents and solvents through the column reactor for each step, and to either loop the flow back through the reactor to achieve the desired residence time, or direct it to waste or product collection. To demonstrate the versatility of the platform, a six-step synthesis of Prexasertib was performed, which included: protection, S_N2, Claisen condensation, hydrazine condensation, S_NAr and cleavage. In addition, a library of 23 derivatives were synthesised by exchanging a single step in the sequence with no interference on the other reactions. Different reactions were performed for the fifth step (amide coupling, reductive amination, *N*-triflation) and the second step (Mitsunobu, click chemistry), highlighting the ability to conduct both late-stage and early-stage diversification respectively. However, this approach currently requires manual batch optimisation of the synthetic route, cannot utilise solid reagents (e.g., heterogeneous catalysis), and is limited to the synthesis of molecules with a developed linker strategy.

Convergent syntheses are shorter and more efficient strategies for multistep synthesis than their linear counterparts. Despite this, the majority of automated multistep syntheses rely on linear processes, which are limited by physical interdependencies and often result in equipment redundancy. To overcome these limitations, Gilmore et al. developed the radial synthesiser; an automated platform with continuous flow modules arranged around a central switching station.[23] Similar to the SPS-flow approach, this allows reactions to be performed sequentially rather than simultaneously, addressing both chemical incompatibilities and the requirement for different residence times between steps.

The radial synthesiser (Figure 6) is composed of four modules: reagent delivery system (RDS), central switching station (CSS), standby module (SM) and collection vessels (C). The entire system is pressurised under nitrogen gas, enabling flow rates and reagent additions to be controlled by venting and mass flow controllers. The RDS stores and delivers solvents, reagents, and synthesised intermediates into sample loops, which are then combined in a connector before entering the CSS. The reaction mixture can be directed to a reactor for the desired residence time, or to a FlowNMR for inline analysis. The mixture then exits the CSS via a FlowIR and, for a multistep synthesis, is directed to either a storage vessel in the RDS (for later use in convergent synthesis) or a sample loop in the standby module (for use in a subsequent step). The automated synthesis of three essential medicines were achieved using this system, including the two-step synthesis of local anaesthetic lidocaine.[24]

The major benefit of this approach is that a single reactor can be used for multistep syntheses involving reactions requiring different conditions, including any combination of long and short residence times. Hence, different synthetic routes can be evaluated on a single platform without the need for reconfiguration. This was demonstrated for the comparison of a convergent and linear three-step synthesis of rufinamide.[23] Both syntheses required the same reactions in a different sequence: (i) alkyl azide synthesis via nucleophilic substitution; (ii) aminolysis; (iii) Cu-catalysed cycloaddition. In this case, the convergent route provided a higher isolated yield compared to the linear route (70% *cf.* 45%). Both routes were used to

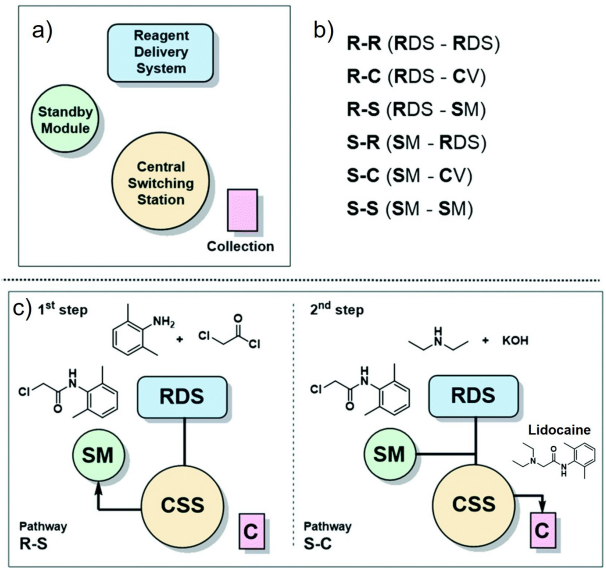


Figure 6. Schematic of the radial synthesiser. (a) Four modules of the instrument: reagent delivery system (RDS), central switching station (CSS), standby module (SM) and collection vessels (C). (b) Different pathways of solution through the instrument described by starting and finishing locations. (c) Reagents and pathways for the multistep radial synthesis of lidocaine. Reproduced and adapted from Ref. [24] under the terms of CC BY 3.0. Copyright the Authors.

synthesise a library of 12 derivatives, depending on which route required the least re-optimisation for the change in substrate. The versatility of the system was also demonstrated by expanding the CSS module to include a photochemical reactor, which was used to synthesise aniline derivatives via dual nickel/photoredox C–N cross-couplings. The biggest drawback of this platform is the mitigation of some of the advantages of flow processes. For example, the need to store solutions between reactions restricts the chemistries to those with stable and non-hazardous intermediates.

The advantages of multistep synthesis in continuous flow can also be harnessed for the automated screening of catalytic performance. The ability to synthesise and test catalysts in an uninterrupted sequence can accelerate the discovery and optimisation of a wider range of new catalysts, including those with poor stability in air. For example, Willans et al. developed a multistep electrochemical flow platform for the automated screening of metal-NHC catalysts.[25] In contrast to chemical reagents, electrons are not converted into side products, and can therefore be classified as clean reagents. Hence, the electrochemical synthesis of metal-NHC complexes enables direct addition into subsequent steps without purification, which is ideal for telescoped continuous flow processes.

In this case, Cu-NHC complexes were electrogenerated from azolium starting materials and Cu electrodes under mild conditions, with hydrogen gas as the only by-product. Their activity was then tested in a downstream click reaction between benzyl azide and phenyl acetylene, which was monitored using online HPLC. To enable screening of the catalysts, an autosampler and sample loop were integrated into the platform to inject and deliver the ligand precursors. In total, 11 azolium salts and acetic acid (to make CuOAc) were screened, where it was found that less sterically demanding NHCs were favourable, and the electronic nature of the backbone was less important.

Although autonomous high-throughput screening platforms enable various parameters and outcomes to be assessed, their 'brute-force' approach to optimisation requires significant resources, including both experimental time and cost. These are further confounded for multistep processes, owing to an increase in the number of reaction variables and the introduction of complex interactions between the steps. To reduce this experimental burden, self-optimisation platforms were developed to intelligently explore reaction design spaces.[26]

In contrast to screening platforms, self-optimisation integrates machine learning algorithms with automated reactors and online analytics to create a feedback loop, where the data from the previous experiments is used to inform the search (Figure 7a).[27] This technology has been widely shown to accelerate the development of single reactions, where current state-of-the-art systems are capable of simultaneously optimising continuous (e.g., time, temperature, stoichiometry) and categorical (e.g., catalyst, ligand, solvent) variables across multiple objectives.[28–30] Intelligent automation also enables substrate-specific optimisations to be readily performed during compound screening, rather than relying on the application of pre-optimised conditions for a model substrate, which often limits the success of library generation.[31]

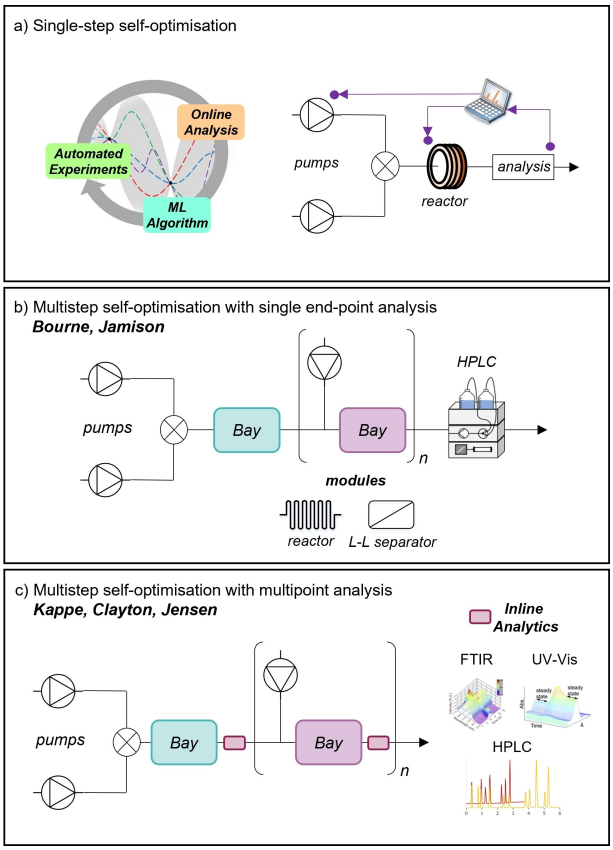


Figure 7. Development of automated continuous flow optimisation platforms. (a) Self-optimisation of a single reaction step utilising a feedback loop connecting experiments, analysis and machine learning algorithms. (b) Simultaneous self-optimisation of multiple unit operations, including reactor and separator modules which can be reconfigured in each bay. Optimisation is driven by a single analytical measurement at the end of the process. (c) Holistic self-optimisation of telescoped processes. Integration of multipoint analytics enhances process understanding and can be used to optimise the output of each step.

Not only does this technology benefit those with direct access to these systems within their labs, but integration with Cloud-based servers enables remote operation from anywhere in the world. This facilitates collaboration and promotes the standardisation of data and protocols, which in turn will improve the quality of machine learning models in the future. Ley et al. demonstrated the ability to work across international borders by using servers based in Japan, equipment based in Cambridge, UK and operators based in Los Angeles, USA.[32] Using this approach, they were able to self-optimise the individual steps for the multistep syntheses of two APIs: lidocaine and bupropion. Although only one automated system was used in this work, the method could be extended in the future for parallel optimisation of multiple steps at different locations, thus enhancing researcher capabilities by providing access to a wider range of reactor technologies.

In the case of bupropion, the individually optimised steps were then combined with appropriate work-up operations to ensure compatibility. The previously unexplored route was optimised and combined into a continuous multistep process in four working days, producing bupropion at an average rate of

2.88 g h^{-1}. However, integrating multiple reactions into continuous flow sequences creates interdependencies between the steps, where combining the individual reaction optima does not necessarily lead to the global system optimum.[33] Therefore, to avoid having to re-optimise steps, it is beneficial to optimise all variables in the overall process simultaneously.

This holistic self-optimisation approach was initially applied by taking a single analytical measurement downstream of the overall process (Figure 7b). For example, Bourne et al. simultaneously optimised a biphasic Claisen–Schmidt condensation reaction with subsequent liquid-liquid separation.[34] Adjusting flow rates and ratios at the reactor inlet effected the residence time and reagent stoichiometry of the reaction, whilst simultaneously changing the solvent ratio, and therefore the partitioning of chemical species between the organic and aqueous phases. Multiobjective optimisation was performed to maximise the purity, space-time yield (STY) and reaction mass efficiency (RME) of the process, which enabled the impact of the downstream work-up on economic and environmental objectives to be simultaneously evaluated.

Similarly, Jamison et al. developed an easily reconfigurable flow platform for the optimisation of diverse multistep processes.[35] The platform was designed with five universal bays which could host any type of module, including different reactors and a liquid-liquid separator. The backbone of the platform consisted of an array of six reagent feeds and pumps, enabling addition of reagents and solvents at the inlet of each module. The "black-box" Stable Noisy Optimisation by Branch and Fit (SNOBFIT) algorithm was utilised to provide a flexible and general approach for identifying the global optimum without requiring prior knowledge of the system. This approach was used to optimise the yield of five diverse multistep reactions and flow sequences: (i) Buchwald–Hartwig cross-coupling with liquid-liquid separation; (ii) two-step Horner–Wadsworth olefination; (iii) two-step reductive amination; (iv) photoredox iminium generation with nucleophilic trapping; (v) ketene generation with 2+2 cycloaddition. Although these approaches successfully provide a global optimum for the specified optimisation problem, the amount of process understanding gained is limited, as the output of each step is convoluted into a single HPLC measurement. Hence, it is impossible to correlate the effect of each individual step and factor on the overall response.

For multistep continuous flow chemistry, being able to quantify the formation and consumption of products, intermediates, and impurities in real-time at multiple points along the synthetic pathway can enhance process understanding and control. In this vein, Kappe et al. integrated four complementary process analytical technologies (NMR, UV/Vis, IR and UHPLC) with a continuous flow reactor, all monitored and controlled through a single supervisory control and data acquisition (SCADA) software.[36]

The capabilities were demonstrated for a multistep synthesis of API mesalazine **15** (Figure 8a): (i) aromatic nitration of **12** with acid/base extraction; (ii) hydrolysis of aryl chloride **13**; (iii) hydrogenation of nitroarene **14**. Step (i) was monitored using inline NMR with a 12 second measurement frequency. Due to peak overlap in the spectra, an indirect hard model (IHM) was built, which fits Gaussian/Lorentzian peaks to the signals. This enabled quantification of the starting material and two regioisomeric products. Inline UV/Vis was used to monitor step (ii) in 2 second intervals. As no distinct spectral features for each compound could be observed, a neural network was constructed by combining the UV/Vis measurements with the input concentrations determined by NMR from step (i). Finally, step (iii) was monitored using inline IR with spectra been acquired every 15 seconds. In this case, a partial least squares (PLS) regression model was developed to facilitate concentration predictions.

UHPLC was integrated at the end of the process to provide precise quantification of all nine chemical species in a method time of 7.5 minutes, which also served as validation for the IR PLS model. The ability of the developed approach to detect and monitor process deviations in real-time was demonstrated under dynamic operation, where the temperature of each reaction step was varied. Although the IR PLS model did not always make accurate predictions, each PAT was able to identify the correct trends in concentrations with changing temperature. In a manufacturing setting, this would enable correction of reaction conditions in a timely manner.

A similar real-time multipoint PAT approach was applied to the self-optimisation of a two-step synthesis of API edaravone **19** (Figure 8b), via formation of imine **18** and subsequent cyclisation.[37] IHM calibrated inline NMR was used to monitor the condensation of ketoester **17** with hydrazine **16**, whilst inline FTIR used a PLS model to quantify all species in the subsequent imine cyclisation step. The fast-sampling time of these techniques enabled simultaneous optimisation of both steps using a Thompson Sampling Efficient Multiobjective Optimisation (TSEMO) algorithm.[38] A total of 85 experiments were conducted over a period of 26 hours, revealing the trade-off between the yield of the first step, the STY of the second step, and the overall equivalents of reagents used.

Nevertheless, the application of real-time PAT for multistep self-optimisation does have significant drawbacks. High equipment costs, combined with the requirement for advanced data processing techniques, often makes this technology inaccessible for the standard synthetic lab. In addition, a substantial amount of time is spent training chemometric models for quantification on a case-by-case basis, which can still typically have an error of up to 5%. Although correct trends can still be identified, more precise quantification of low-level impurities is often required to meet high regulatory standards.

With the aim of increasing the accessibility of multistep self-optimisation, Clayton et al. developed a simple approach for using a single HPLC instrument for automated multipoint sampling across a multistep continuous flow process (Figure 8c).[39] This was achieved by daisy-chaining sample valves positioned at the outlet of each reactor stage, and coding them to trigger sequentially within the optimisation program. This enabled accurate quantification of compounds at each step within a telescoped process, thus increasing the process understanding gained throughout the optimisation.

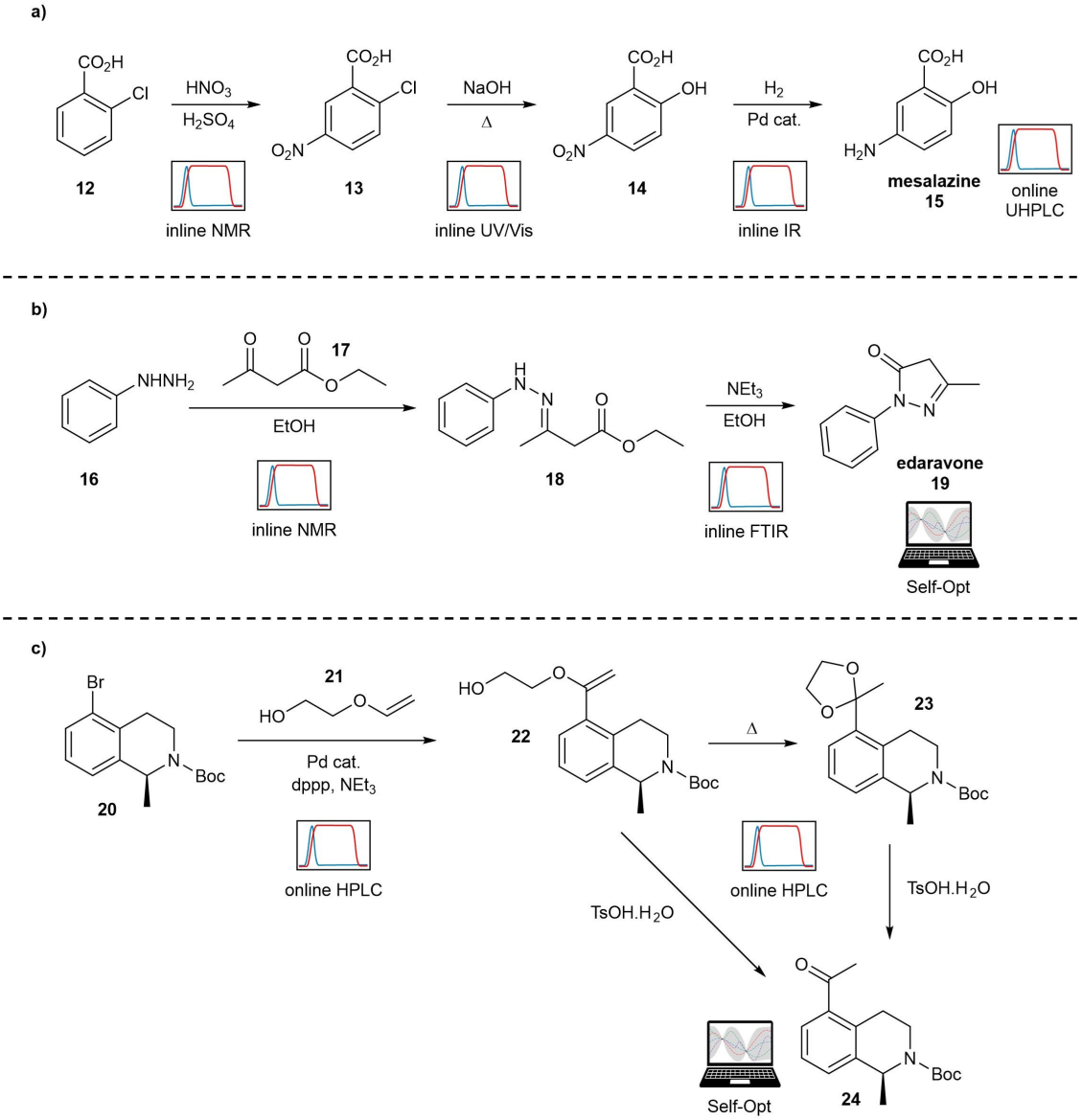


Figure 8. Examples of multistep process development and control in continuous flow using multipoint sampling. (a) Three-step synthetic process of mesalazine **15**, where each step was monitored by a different PAT tool. (b) Self-optimisation of the two-step synthesis of edaravone **19** using different PAT tools. (c) Self-optimisation of the Heck-cyclisation-deprotection reaction sequence for the synthesis of methyltetrahydroisoquinoline **24**. Multipoint analysis with a single online HPLC instrument was used to quantify the output of each reactor.

Due to the longer method times of HPLC compared to PAT (minutes *cf.* seconds), it was important to integrate a state-of-the-art Bayesian algorithm to minimise the number of experiments required, and thus ensure the optimisation remained practically viable. Bayesian algorithms, which balance the exploration of areas of uncertainty with the exploitation of available information, have been applied as a tool for chemical reaction optimisation in recent years.[40] In this case, Bayesian Optimisation with Adaptive Expected Improvement (BOAEI) was applied, which dynamically controls the explore/exploit trade-off, and thus reduces the risk of inefficient global or excessive local searching.[41]

The developed approach was demonstrated on the self-optimisation of a Heck-cyclisation-deprotection reaction sequence for the synthesis of a pharmaceutically relevant methyltetrahydroisoquinoline **24**. A total of 32 experiments were conducted over 45 hours, however the optimum was identified after only 13 experiments and 14 hours, just 4 experiments after the initial Latin hypercube. This efficiency is comparable to previously reported single step optimisations, despite the additional complexities of telescoped processes. In addition, the global optimum found was different to that expected by combination of individual optimisations, due to the identification of a favourable competing pathway.

Despite the advances in automated experimental platforms highlighted above, these approaches still require a significant input of researcher time for the initial formulation of the multistep synthesis. Computer-aided synthesis planning (CASP)

tools can now propose forward reaction conditions for the multistep synthesis of organic compounds.[42] However, predictions made from these tools are based on limited and general data, therefore optimisation of approximate reaction conditions is still required.

To accelerate the planning and development workflow of multistep flow synthesis, Jensen et al. reported the integration of CASP tools with an automated robotic platform (Figure 9).[43] The open source CASP software ASKCOS was used to propose synthetic routes for sonidegib. The proposed high-ranking routes were manually assessed, and a feasible three-step synthesis selected and refined: (i) nucleophilic aromatic substitution; (ii) multiphasic Pd-catalysed nitro reduction with subsequent phase separation; (iii) amide coupling. The synthetic route was optimised on a modular, robotically reconfigurable, continuous flow synthesis platform. The platform was equipped with online LC–MS and inline FTIR analysis for reaction monitoring, which were combined with a multiobjective Bayesian algorithm (Dragonfly) for feedback optimisation.[44] The platform was capable of optimising both continuous and categorical variables, including the halonitropyridine starting material for the S_NAr reaction and the activating agent for the amide coupling. The reconfigurable capabilities of the platform also enabled reactor volumes to be varied, which allowed different residence times in the amide coupling step to be explored.

Notably, attempts to optimise this as a fully telescoped process were unsuccessful. The inline monitoring capabilities at each step revealed a nonintuitive chemical incompatibility, where a by-product from the S_NAr reaction caused catalyst deactivation in the subsequent nitro reduction step. As a result, the S_NAr step was optimised separately from the telescoped nitro reduction-amide coupling process. Nevertheless, both processes were autonomously optimised for different objectives (yield, productivity, cost) in a combined number of just 45 experiments. This study demonstrated how autonomous experimental optimisation can be used in combination with CASP to conduct more of the repetitive aspects of multistep process development. This enables researchers to focus on areas where human input is clearly still required, such as critical interpretation and creative problem solving.

3. Outlook

It is clear that advances in automated reactor technologies in recent years are continuing to transform the way chemists approach synthesis. Although the application of automation for multistep synthesis is still in its infancy, the drive towards more flexible and responsive manufacturing of complex products is expected to cause this field to develop rapidly. The design of new iterative bond forming strategies and modular platforms has enabled the automated synthesis of molecules with greater structural diversity. Indeed, further development of universal synthesis platforms will continue to expand the toolbox of reactions that can be performed autonomously, giving on-demand access to a wider range of products.

Integration of complimentary analytical technologies has enabled real-time monitoring of each step in a one-pot or telescoped process, which when combined with a feedback loop, provides unprecedented levels of adaptive control and flexibility for multistep procedures. In the future, the combination of advanced process analytics with dynamic trajectories in flow will enable data-rich exploration of multistep reaction design spaces,[45] and autonomous determination of the reaction models for each step.[46]

Many of the reported examples of autonomous multistep synthesis have utilised continuous flow as an enabling technology. When combining two or more steps together in continuous flow, the number and possible combinations of variables

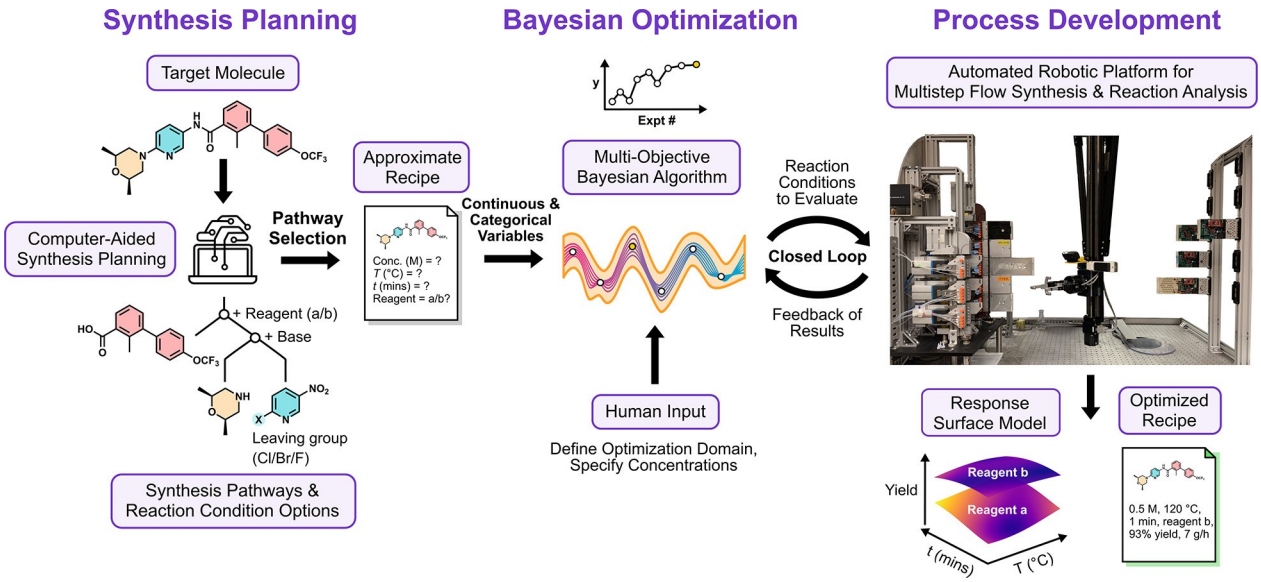


Figure 9. Machine-assisted multistep synthesis planning and process development. Reproduced from Ref. [43] under the terms of CC BY-NC-ND 4.0 with permission from the Authors.

increase exponentially. Hence, self-optimisation platforms have been further advanced for application towards multistep synthesis, including integration with CASP software, multipoint analysis, and advanced algorithms. By establishing international collaborations, networking autonomous synthesis and optimisation platforms will enhance the accessibility of these technologies and generate extensive validated datasets to advance machine learning methods in multistep synthesis planning.

Nevertheless, high material consumption remains a significant limitation for continuous self-optimisation. The development of droplet-flow microfluidic platforms has started to address this challenge for single step reactions,[47,48] and their extension to multistep self-optimisation will likely form the basis of future work.[49] Radial synthesis and SPS-flow approaches have also been developed to overcome limitations of linear flow sequences, including chemical incompatibility and physical interdependencies between the steps. Integration of feedback loops with these platforms also presents an exciting opportunity to expand the technologies available for autonomous multistep synthesis.

Acknowledgements

A.D.C. thanks the Royal Academy of Engineering for support under the Research Fellowships scheme.

Conflict of Interests

The authors declare no conflict of interest.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Keywords: Continuous Flow · Digitalisation · Machine Learning · Reactor Automation · Self-Optimisation

[1] Y. Hayashi, *Chem. Sci.* **2016**, *7*, 866–880.
[2] R. B. Leng, M. V. M. Emonds, C. T. Hamilton, J. W. Ringer, *Org. Process Res. Dev.* **2012**, *16*, 415–424.
[3] I. W. Davies, *Nature* **2019**, *570*, 175–181.
[4] C. Coley, W. H. Green, K. F. Jensen, *Acc. Chem. Res.* **2018**, *51*, 1281–1289.
[5] S. M. Moosavi, K. M. Jablonka, B. Smit, *J. Am. Chem. Soc.* **2020**, *142*, 20273–20287.
[6] F. Häse, L. M. Roch, A. Aspuru-Guzik, *Trends Chem.* **2019**, *1*, 282–291.
[7] R. B. Merrifield, *Science* **1965**, *150*, 178–185.
[8] M. H. Caruthers, *Science* **1985**, *230*, 281–285.
[9] O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Science* **2001**, *291*, 1523–1527.
[10] J. Li, S. G. Ballmer, E. P. Gillis, S. Fujii, M. J. Schmidt, A. M. E. Palazzolo, J. W. Lehmann, G. F. Morehouse, M. D. Burke, *Science* **2015**, *347*, 1221–1226.
[11] D. J. Blair, S. Chitti, M. Trobe, D. M. Kostyra, H. M. S. Haley, R. L. Hansen, S. G. Ballmer, T. J. Woods, W. Wang, V. Mubayi, M. J. Schmidt, R. W. Pipal, G. F. Morehouse, A. M. E. P. Ray, D. L. Gray, A. L. Gill, M. D. Burke, *Nature* **2022**, *604*, 92–97.
[12] D. Leonori, V. K. Aggarwal, *Acc. Chem. Res.* **2014**, *47*, 3174–3183.
[13] V. Fasano, R. C. Mykura, J. M. Fordham, J. J. Rogers, B. Banecki, A. Noble, V. K. Aggarwal, *Nat. Synth.* **2022**, *1*, 902–907.
[14] S. Steiner, J. Wolf, S. Glatzel, A. Andreou, J. M. Granda, G. Keenan, T. Hinkley, G. Aragon-Camarasa, P. J. Kitson, D. Angelone, L. Cronin, *Science* **2019**, *363*, eaav2211.
[15] P. S. Gromski, J. M. Granda, L. Cronin, *Trends Chem.* **2020**, *2*, 4–12.
[16] A. J. S. Hammer, A. I. Leonov, N. L. Bell, L. Cronin, *JACS Au* **2021**, *1*, 1572–1587.
[17] J. Liu, Y. Sato, F. Yang, A. J. Kukor, J. E. Hein, *Chemistry-Methods* **2022**, *2*, e202200009.
[18] J. Britton, C. L. Raston, *Chem. Soc. Rev.* **2017**, *46*, 1250–1271.
[19] A. Adamo, R. L. Beingessner, M. Behnam, J. Chen, T. F. Jamison, K. F. Jensen, J.-C. M. Monbaliu, A. S. Myerson, E. M. Revalor, D. R. Snead, T. Stelzer, N. Weeranoppanant, S. Y. Wong, P. Zhang, *Science* **2016**, *352*, 61–67.
[20] M. Abolhasani, E. Kumacheva, *Nat. Synth.* **2023**, https://doi.org/10.1038/s44160-022-00231-0.
[21] B. Lin, J. L. Hedrick, N. H. Park, R. M. Waymouth, *J. Am. Chem. Soc.* **2019**, *141*, 8921–8927.
[22] C. Liu, J. Xie, W. Wu, M. Wang, W. Chen, S. B. Idres, J. Rong, L.-W. Deng, S. A. Khan, J. Wu, *Nat. Chem.* **2021**, *13*, 451–457.
[23] S. Chatterjee, M. Guidi, P. H. Seeberger, K. Gilmore, *Nature* **2020**, *579*, 379–384.
[24] M. Guidi, S. Moon, L. Anghileri, D. Cambié, P. H. Seeberger, K. Gilmore, *React. Chem. Eng.* **2021**, *6*, 220–224.
[25] C. Schotten, J. Manson, T. W. Chamberlain, R. A. Bourne, B. N. Nguyen, N. Kapur, C. E. Willans, *Catal. Sci. Technol.* **2022**, *12*, 4266–4272.
[26] C. J. Taylor, A. Pomberger, K. C. Felton, R. Grainger, M. Barecka, T. W. Chamberlain, R. A. Bourne, C. N. Johnson, A. A. Lapkin, *Chem. Rev.* **2023**, *123*, 3089–3126.
[27] A. D. Clayton, J. A. Manson, C. J. Taylor, T. W. Chamberlain, B. A. Taylor, G. Clemens, R. A. Bourne, *React. Chem. Eng.* **2019**, *4*, 1545–1554.
[28] O. J. Kershaw, A. D. Clayton, J. A. Manson, A. Barthelme, J. Pavey, P. Peach, J. Mustakis, R. M. Howard, T. W. Chamberlain, N. J. Warren, R. A. Bourne, *Chem. Eng. J.* **2023**, *451*, 138443.
[29] K. Y. Nandiwale, T. Hart, A. F. Zahrt, A. M. K. Nambiar, P. T. Mahesh, Y. Mo, M. J. Nieves-Remacha, M. D. Johnson, P. García-Losada, C. Mateos, J. A. Rincón, K. F. Jensen, *React. Chem. Eng.* **2022**, *7*, 1315–1327.
[30] L. M. Baumgartner, C. W. Coley, B. J. Reizman, K. W. Gao, K. F. Jensen, *React. Chem. Eng.* **2018**, *3*, 301–311.
[31] P. Mueller, A. Vriza, A. D. Clayton, O. S. May, N. Govan, S. V. Ley, T. W. Chamberlain, R. A. Bourne, *React. Chem. Eng.* **2023**, *8*, 538–542.
[32] D. E. Fitzpatrick, T. Maujean, A. C. Evans, S. V. Ley, *Angew. Chem. Int. Ed.* **2018**, *57*, 15128–15132.
[33] T. Dietz, K. Klamroth, K. Kraus, S. Ruzika, L. E. Schäfer, B. Schulze, M. Stiglmayr, M. M. Wiecek, *Eur. J. Oper. Res.* **2020**, *280*, 581–596.
[34] A. D. Clayton, A. M. Schweidtmann, G. Clemens, J. A. Manson, C. J. Taylor, C. G. Niño, T. W. Chamberlain, N. Kapur, A. J. Blacker, A. A. Lapkin, R. A. Bourne, *Chem. Eng. J.* **2020**, *384*, 123340.
[35] A.-C. Bédard, A. Adamo, K. C. Aroh, M. G. Russell, A. A. Bedermann, J. Torosian, B. Yue, K. F. Jensen, T. F. Jamison, *Science* **2018**, *361*, 1220–1225.
[36] P. Sagmeister, R. Lebl, I. Castillo, J. Rehrl, J. Kruisz, M. Sipek, M. Horn, S. Sacher, D. Cantillo, J. D. Williams, C. O. Kappe, *Angew. Chem. Int. Ed.* **2021**, *60*, 8139–8148.
[37] P. Sagmeister, F. F. Ort, C. E. Jusner, D. Hebrault, T. Tampone, F. G. Buono, J. D. Williams, C. O. Kappe, *Adv. Sci.* **2022**, *9*, 2105547.
[38] E. Bradford, A. M. Schweidtmann, A. Lapkin, *J. Glob. Optim.* **2018**, *71*, 407–438.
[39] A. D. Clayton, E. O. Pyzer-Knapp, M. Purdie, M. F. Jones, A. Barthelme, J. Pavey, N. Kapur, T. W. Chamberlain, A. J. Blacker, R. A. Bourne, *Angew. Chem. Int. Ed.* **2023**, e202214511.
[40] B. J. Shields, J. Stevens, J. Li, M. Parasram, F. Damani, J. I. M. Alvarado, J. M. Janey, R. P. Adams, A. G. Doyle, *Nature* **2021**, *590*, 89–96.
[41] D. Jasrasaria, E. O. Pyzer-Knapp, **2018**, *arXiv Preprint* DOI: 10.48550/arXiv.1807.01279.
[42] C. W. Coley, D. A. Thomas III, J. A. M. Lummiss, J. N. Jaworski, C. P. Breen, V. Schultz, T. Hart, J. S. Fishman, L. Rogers, H. Gao, R. W. Hicklin, P. P. Plehiers, J. Byington, J. S. Piotti, W. H. Green, A. J. Hart, T. F. Jamison, K. F. Jensen, *Science* **2019**, *365*, eaax1566.
[43] A. M. K. Nambiar, C. P. Breen, T. Hart, T. Kulesza, T. F. Jamison, K. F. Jensen, *ACS Cent. Sci.* **2022**, *8*, 825–836.
[44] K. Kandasamy, K. R. Vysyaraju, W. Neiswanger, B. Paria, C. R. Collins, J. Schneider, B. Póczos, E. P. Xing, *J. Mach. Learn. Res.* **2020**, *21*, 1–25.

[45] F. Florit, A. M. K. Nambiar, C. P. Breen, T. F. Jamison, K. F. Jensen, *React. Chem. Eng.* **2021**, *6*, 2306–2314.
[46] C. J. Taylor, M. Booth, J. A. Manson, M. J. Willis, G. Clemens, B. A. Taylor, T. W. Chamberlain, R. A. Bourne, *Chem. Eng. J.* **2021**, *413*, 127017.
[47] L. M. Baumgartner, J. M. Dennis, N. A. White, S. L. Buchwald, K. F. Jensen, *Org. Process Res. Dev.* **2019**, *23*, 1594–1601.
[48] C. Avila, C. Cassani, T. Kogej, J. Mazuela, S. Sarda, A. D. Clayton, M. Kossenjans, C. P. Green, R. A. Bourne, *Chem. Sci.* **2022**, *13*, 12087–12099.
[49] A. A. Volk, R. W. Epps, D. T. Yonemoto, B. S. Masters, F. N. Castellano, K. G. Reyes, M. Abolhasani, *Nat. Commun.* **2023**, *14*, 1403.

Manuscript received: April 10, 2023
Version of record online: June 22, 2023