# Monitoring, Modelling and Optimisation of Continuous Flow Reactions Using On-line Mass Spectrometry

Christopher Stefan Horbaczewskyj

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

University of Leeds

School of Chemistry

March 2019

The candidate confirms that the work submitted is their own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The work in Chapter 1 of the thesis has appeared in the publication as follows:

Chapter 12, Green Chemical Engineering, The Handbook of Green Chemistry, Nov 2018, C. S. Horbaczewskyj, A. A. Lapkin, R. A. Bourne, C. E. Willans.

I was responsible for compiling and writing sections 12.1, 12.2, 12.3, 12.4, 12.5.2-12.5.6 and 12.6.

The contribution of the other authors was writing section 12.5.1., proof reading and editing the document.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Christopher Stefan Horbaczewskyj to be identified as Author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

© 2019 University of Leeds and Christopher Stefan Horbaczewskyj.

## Acknowledgements

I would firstly like to thank my PhD supervisors, Dr Richard Bourne and Dr Charlotte Willans, for giving me the chance to work in such diverse and exciting research groups, and for their support and advice throughout all projects undertaken.

I would like to thank all the iPRD research group members past and present who have helped me over the past 4 years: Chris, Mary, Will, Katie, Alastair, Mike, Nick, Lisa, Maria, James, Illias, Adam, Connor, Calum. I would specifically like to thank Chris for his initial support and teaching during my studies. Also, my thanks go to all technicians and support staff within the department; Martin Huscroft for LCMS, Simon Barrett for NMR, Matthew Broadbent in the mechanical workshop and the staff in stores.

My thanks go to the University of Leeds, EPSRC and AstraZeneca for providing funding for this PhD CASE Studentship. And thank you to AstraZeneca, for also providing the opportunity to carry out a 3-month placement at their Macclesfield site during my studies. In particular I would like to thank my industrial supervisor, Tony Bristow as well as Graeme Clemens, Mubina Mohammed and Brian Taylor for their help throughout the placement and their help with the AZD5634 and CPACT projects. Regarding the work completed in chapter 5, I would like to thank both Francis Singer and Michael Chapman for their help in synthesising all imidazolium ligand precursors and pure catalytic samples.

I would also like to thank all the friends I met throughout my PhD who helped keep my work-life balance in equilibrium. Thank you to those who I have met though union societies, in particular those I have met through the Ju Jitsu society, climbing society and the caving society and continue to help support and develop my skill sets.

I would like to thank my family and friends for their support over many years of studying and stressing not only during my PhD but throughout my time getting to this point. Alison for helping me with applications and support through the initial stages. Most importantly my Mum and Dad for their continued support of me partaking in science and allowing me to do the subject which interests me the most.

Finally, I would like to thank Rosi for making the last few years of my PhD full of travelling, exploration, new activities and the great outdoors. Thank you for all your current and continued support of everything that I do.

# Abstract

An on-line mass spectrometry method has been developed to monitor, model and optimise continuous flow reactions. This method makes use of dual-piston pumps, tubular reactor block, Vici sample actuator, an Advion Compact Mass Spectrometer (CMS) and other analytical systems to investigate a variety of chemical systems based on their need for process improvement. Full reaction automation employed MATLAB, the Snobfit algorithm, along with Modde DoE software.

On-line mass spectrometry has advantages over other analytical techniques as it has shorter acquisition times (2-60 s), low chemical sensitivity ($\sim10^8$ mol%) and chemical identity as well as the potential to provide quantitative information. In this work, reaction quantitation has been explored using four chemical systems, where each of them was monitored by a variety of analytical techniques, with the overall aim being to examine if on-line mass spectrometry can be used for quantitative analysis.

For all cases investigated, process improvements were made whilst also determining optimal operating conditions to improve conversions, yields or selectivities as well as looking at reaction waste reduction. Flow chemistry and the work conducted has shown how waste can be reduced for certain reactions when compared to more traditional approaches. This method relies on machine learning, full process automation and quick process analytical technology to determine optimum conditions as well as build large reaction data sets. Large data sets were created using a hybrid DoE-kinetic composite circumscribed orthogonal design.

Mass spectrometry provided valuable reaction information and has the potential for reaction quantitation depending on the required application, reaction system and ionisation settings. Compound thermal stability can be problematic in APCI+ mode whilst ion suppression is problematic in ESI+ mode. Still a versatile analytical tool, on-line mass spectrometry was found to be inherently quantitative.

The continuous-flow-on-line-MS-self-optimisation platform was used to investigate a variety of different reactions to show versatility of the MS system. These reactions are summarized below.

1) An N-Boc deprotection of AZD5634 (Scheme 1) for optimisation and process scale-up, with achieved conversions >95% and scale-up to pilot and commercial scale using on-line mass spectrometry).
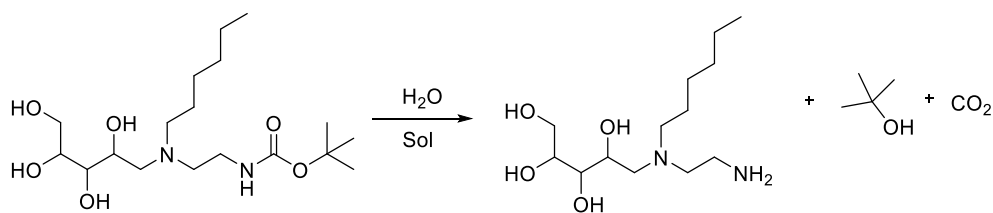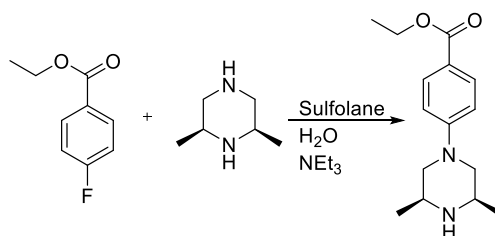


*Scheme 1 Reaction scheme for the deprotection of AZD5634 using $H_2O$ as reagent. By-products of $^t$butanol (as $H_2O$ is in large excess) and carbon dioxide are also formed.*

2) An N-Boc deprotection reaction using a hybrid DoE-kinetic model for optimisation and large data set generation, with achieved conversion >90% (Scheme 2).
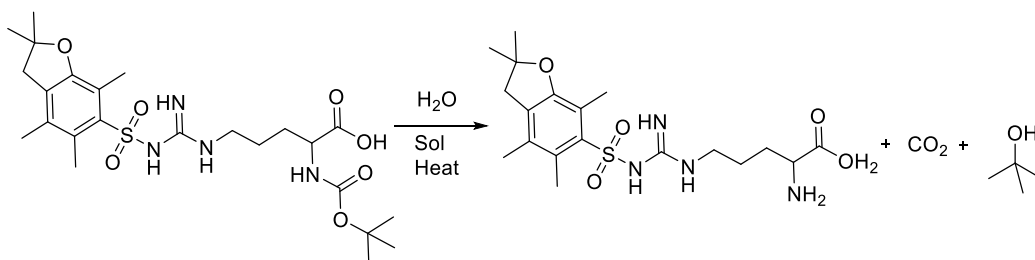


*Scheme 2 Reaction scheme for the deprotection of Boc-Arg(Pbf)-OH using $H_2O$ as reagent. By-products of $^t$butanol (as $H_2O$ is in large excess) and carbon dioxide are also formed.*

3) An $S_NAr$ reaction of AZD4547 for product selectivity and yield improvement, with achieved conversion of ~38% and DP yield of ~30% (Scheme 3).



*Scheme 3 Reaction scheme for the $S_NAr$ reaction to form AZD4547 in continuous flow.*

4) The synthesis and optimisation of Fe-N-heterocyclic carbene complexes using an electrochemical method for use in a C-H hydroxylation reaction (Scheme 4). Optimum electrochemical conditions of either 7 V and 4 minutes residence time, or 2.5 V and 15 minutes residence were achieved.
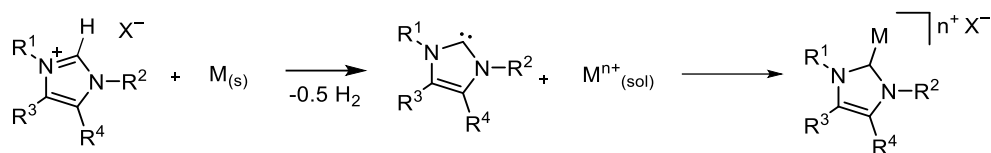


*Scheme 4 Electrochemical reaction performed in continuous flow to generate Fe-NHC complexes, monitored by on-line mass spectrometry.*

**Table of Contents**

**List of figures**

**List of Schemes**

**List of Tables**

## List of Abbreviations

| | |
|---|---|
| 2D | 2 Dimensional |
| 3D | 3 Dimensional |
| 4D | 4 Dimensional |
| APCI | Atmospheric Pressure Chemical Ionisation |
| API | Active Pharmaceutical Ingredient |
| AZ | AstraZeneca |
| Boc | $^{t}$Butyloxycarbonyl Protecting Group |
| BPR | Back Pressure Regulator |
| CCC | Cubic Centred Circumscribed |
| CCD | Charge Coupled Device |
| CCF | Face Centred Cubic |
| COSY | Correlation Spectroscopy |
| CPACT | Centre for Process Analytics and Control Technology |
| CSTR | Continuous Stirred Tank Reactor |
| DC | Dissatisfaction Coefficient |
| DCM | Dichloromethane |
| DoE | Design of Experiments |
| DP | Desired Product |
| ESI | Electrospray Ionisation |
| FD | Factorial Design |
| FFD | Fractional Factorial Design |
| FT-IR | Fourier Transform Infrared Spectroscopy |
| GC | Gas Chromatography |
| GLC | Gas Liquid Chromatography |
| HETCOR | Heteronuclear Correlation Spectroscopy |
| HPLC | High Pressure Liquid Chromatography |
| HSQC | Heteronuclear Single Quantum Coherence |
| IPA | Isopropyl Alcohol |
| IS | Internal Standard |
| LCMS | Liquid Chromatography Mass Spectrometry |
| LSL | Large-Scale Lab |

| | |
|---|---|
| MATLAB | Matrix Laboratory |
| MeCN | Acetonitrile |
| MeOH | Methanol |
| MeTHF | Methyl Tetrahydrofuran |
| MIT | Massachusetts Institute of Technology |
| MODDE | Software used for DoE modelling |
| MRA | Mass Rate Attenuator |
| MS | Mass spectrometry/ Mass Spectrum |
| N-Boc | Nitrogen protected by ᵗButyloxycarbonyl group |
| Nd:YAG | Neodymium-doped Yttrium Aluminium Garnet |
| NMR | Nuclear Magnetic Resonance |
| NMSIM | Nelder-Mead Simplex |
| OVAT | One Variable at a Time |
| PAT | Process Analytical Technology |
| PEG-PCL | Biodegradable Diblock Copolymers |
| RSM | Response Surface Methodology |
| RTD | Residence Time Distribution |
| SM | Starting Material |
| SMSIM | Super Modified Simplex |
| SNOBFIT | Stable Noisy Optimisation by Branch and Fit |
| SP | Side Product |
| STY | Space-Time-Yield |
| TFA | Tri Fluoro Acetic acid |
| TIC | Total Ion Chromatogram |
| UV-Vis | Ultra Violet Visible Spectroscopy |

Chapter 1

**Introduction and Concepts of Continuous Flow Reactions, On-line Analytics, Process Optimisation and Process Scale-up**

## 1. Introduction

The work presented in this thesis demonstrates how new technologies, in particular, continuous flow techniques and on-line analysis and optimisation approaches, can be used to improve traditional chemical syntheses. Current synthetic techniques can be slow, inefficient, hazardous and problematic when scaling-up. The focus of this work was to monitor a variety of reactions, completed in continuous flow, using on-line mass spectrometry (MS), modelling the data to form a more complete reaction profile and ultimately optimise the process. On-line MS as a quantitative tool for reaction optimisation has been developed.

This chapter introduces the essential concepts that are involved in performing chemical transformations in continuous flow processes. Namely, the set-up of continuous flow reactors, the on-line monitoring through process analytical technology, different methods in which a flow reaction may be optimised and finally how a continuous flow process can be applied to industry for addressing the issues of scale-up associated with bath processes.

## 1.1. Concepts of Continuous Flow Methodologies
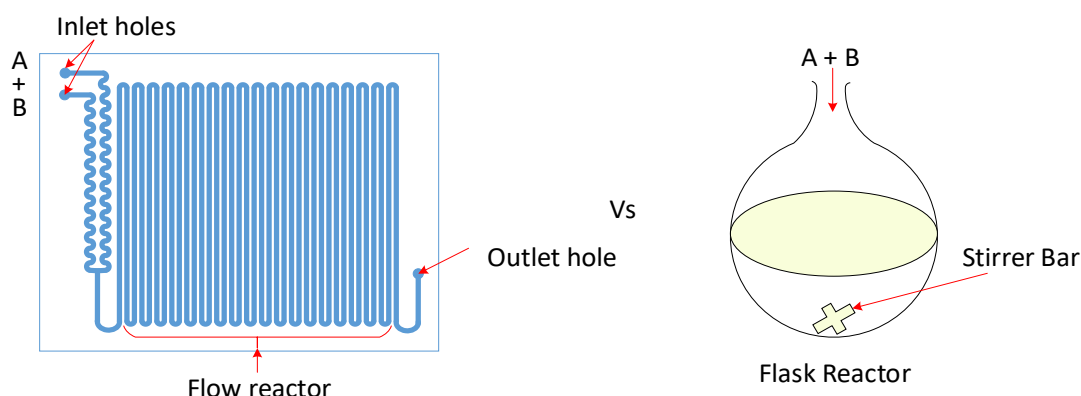


*Figure 1 Left: example of a continuous flow reactor; A + B continuously pumped separately, mix and react within the thin reactor tubing. Right: flask reactor traditionally used in synthesis; A + B mix and react in the same reactor one batch at a time.*

Performing chemistry in a continuous fashion is a relatively new technique to the fine chemical and pharmaceutical industries. Continuous flow chemistry can be defined as performing chemical transformations in tubes/pipes with constant addition of reagents *via* pumps through a reactor, (Figure 1).[1] This method of synthesis has been coined flow

chemistry. These assemblies can become more advanced, employing sample switching devices and process analytical technology to make sampling quicker and easier. Traditional sampling methods are usually carried out manually by allowing the exit tube to collect the stream in a sample vial before analysis can take place off-line. The sample loop allows automatic sampling to be carried out and is automatically analysed by any process analytical techniques employed. The continuous flow system presented in this thesis made use of several separate pieces of equipment in a modular fashion allowing apparatus to be swapped out if the intended use necessitates a unique piece of apparatus.[2,3]

### 1.1.1. Fluid Control Systems for Continuous Flow Reactors

Precise fluid control is important for an automated continuous flow process as it regulates the length of time the stream resides within the reactor (residence time) and controls the stoichiometry of two or more reagent streams which are to be combined. In most examples, this control is provided by pumps which create a pressure at the inlet end of the system allowing reagents to be hydrodynamically pumped to the outlet which is at a lower pressure.[4] This means the pump must be able to achieve a pressure which is always above that of the outlet even under the employment of a pressure regulation tool *e.g.* back-pressure regulator.[5]

There are several types of units which may be used for liquid delivery; syringe, peristaltic, piston, gear or mass-flow controllers (Table 1).[6] Each type of pump may be connected to a computer system, using an RS232 control board, to help automate the system. Direct and precise control allows flow rates to be changed either by hand *via* the computer or using algorithms designed to take control of the chemical system. Being able to control the pumps in this way is the first step in creating a fully closed-loop system which can self-optimise through a feedback loop.

*Table 1 Summary of different types of pumps which can be used for continuous flow chemistry. Examples of each pump type are provided below each head given in brackets.*

| Pump Type | Advantages | Disadvantages | Flow Rate Accuracy |
|---|---|---|---|
| **Syringe**[7] <br> **(Harvard A HPLC pump1 SyrDos 2)** | • Lower flow rates therefore use for microreactors. <br> • Mechanical components away from reaction. <br> • Multiple syringe pumps to pump large volumes continuously. <br> • Improved safety due to no hazardous reagent accumulation. | • No/little back pressure usage. <br> • Single syringe pumps lower volume. | • Mechanical parts, reliable flow rates below 0.05 mL min$^{-1}$. |
| **Piston**[8] <br> **(Jasco PU980)** | • Higher pressure systems (0-500 bar). <br> • Single piston; Cheap and simple. <br> • Dual piston; more stable flow profile, low cavitation, greater reproducibility of flow rate. | • Blockage with precipitates or slurries. <br> • Lower lifespan of mechanical components. <br> • Single piston; pressure spikes & flow pulses. <br> • Dual Piston; moving parts, greater chance of failure. | • Electronic flow rates from pressure transducer. Can loose accuracy at low rates, below 0.05 mL min$^{-1}$. Reliable at higher flow rates. |
| **Peristaltic**[9] <br> **(VapourTec E-Series)** | • Pumping slurries without blockage and fouling of the pump. <br> • Wide configuration variety. <br> • Self-priming. <br> • Precise pumping measurements. | • Lower lifespan of mechanical components. <br> • Lower back pressure usage. | • Mechanical parts but relies on |
| **Gear**[10] <br> **(Gorman-Rupp G Series)** | • Can handle a wide range of viscosities. <br> • Less sensitive to cavitation. <br> • Easy to maintain and rebuild. <br> • Self-priming. | • Cannot pump solids or slurries. <br> • High cost. <br> • Lower lifespan of mechanical components. <br> • Can be noisy. | • Very reliable flow rates over wide range. |
| **Mass-Flow Controllers**[11] <br> **(Brooks Coriolis QMB Series)** | • High accuracy and reliability. <br> • Pump liquids and gases. | • High cost. <br> • Problems with low density gases. <br> • Not for solids or slurries. | • Very reliable flow rates over a wide range. |

### 1.1.2. Continuous Flow Reactor Types

Traditional batch chemistry exploits widely available reaction vessels to house reactants, reagents and solvents in a single, separate reactor at each stage of the process. Once the reaction has been completed the desired product is removed, by crystallisation if a solid, liquid separation if in the liquid state or directly transferred to the next stage in the synthesis route. If the product requires further transformation then a new vessel containing reactants, reagents and solvents is prepared and the procedure repeated until the final desired product created. Batch reactor vessels can occupy volumes from a few millilitres, allowing synthesis of milligrams of product, to ten thousand litre tank reactors capable of producing tonnes of product.[12,13] Although these types of reactor vessels have been used for hundreds of years they have not always been ideal for certain reactions which may evolve gases, generate reactive intermediates[14] or cause excess heat energy to be dispelled rapidly.[15] Large scale reactor vessels typically contain a pressure relief valve and a rupture disk which bursts at a specific pressure rating.[16] If the volume of gas is too great for the relief valve to handle (depressurisation is too slow) then the rupture disk bursts to relieve the pressure. The latter two issues are, generally, more problematic at larger scale as continuous stirred tank reactor (CSTR) vessels have a much higher volume compared to the surface area of the vessel. The nature of these reactor vessels is to rapidly stir the contents to ensure complete mixing, so the reaction can progress to completion in the lowest time possible whilst maximising the final yield of desired product. The high mixing potential can have an adverse effect if a reactive intermediate is generated causing a cascade reaction to occur incredibly fast throughout the vessel and will often generate a large exotherm leading to reaction runaway. Other problems may also arise with adding chemicals to the tanks through inlet ports at the top of the vessel. Semi-batch reactors are much like typical batch reactors, using a single stirred tank for a chemical reaction, but are modified to allow reagent addition and/or product removal at any point in time.[17,18] Semi-batch processes are designed to increase the safety (*via* handling of hazardous chemicals) associated with large scale industrial processes by charging the tank vessel from a safe distance using large addition vessels attached to the tank. However, when adding reactive species to a reaction in one go or smaller portions, macromixing can have a significant effect on the reaction uniformity in the vessel. A build-up of reactive species may occur in isolated places, again potentially dispelling heat and causing reaction runaway.

New technologies may have a remarkable effect on improving large scale chemical processes in terms of safety, efficiency and sustainability.[19] In its simplest form continuous flow chemistry uses tubes as the 'vessel' and pumps to continuously push reactants, reagents and solvents through at a set rate. This means that separate pumps can be used to ensure reagents are kept separate until mixed in the tubular reactor and creates a level of modularity to the system.

A team from Novartis-MIT[20] demonstrated use of new flow technology where an end-to-end synthesis of the pharmaceutical drug, aliskiren hemifumarate (Figure 2), was created combining 3 reaction stages, quenching, workup, isolation and purification. Subsequently, this led to the reaction being scaled-up producing 100 g hr$^{-1}$ and delivering the final stage as tablets containing 112 mg of the synthesised compound, aliskiren hemifumarate. This continuous flow method requires a 136 L reactor, 1 hr reaction time and only 48 hr processing time. If carried out as a batch process this series of reactions would use a 1500 L reactor, each batch would take 48 hr, processed for 300 hr using 21 different operations.



**1.1**
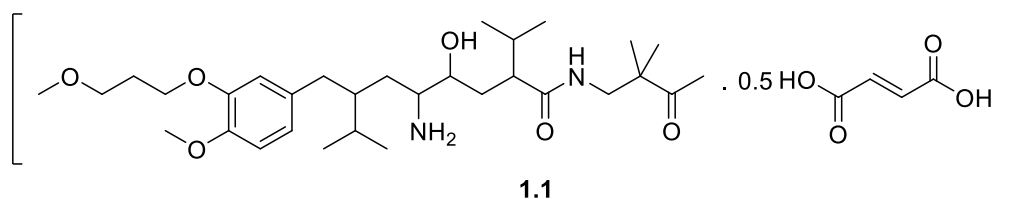
*Figure 2 Aliskiren hemifumarate drug compound synthesised by the Novartis-MIT group using an end-to-end continuous flow process.[20]*

Similarly, a single step synthesis with quench and separation was used for the production of diphenhydramine hydrochloride (Figure 3) in continuous flow with a final output of 2.4 g hr$^{-1}$.[21] The process was developed to minimise waste, reduce purification steps and reduce the overall production time. This API is widely used as an antihistamine for allergy relief and is made in quantities of 100,000 kg yr$^{-1}$ worldwide. Although this reaction was completed on a small scale, 720 µL, it shows the diversity and applicability of continuous flow to the pharmaceutical industry.
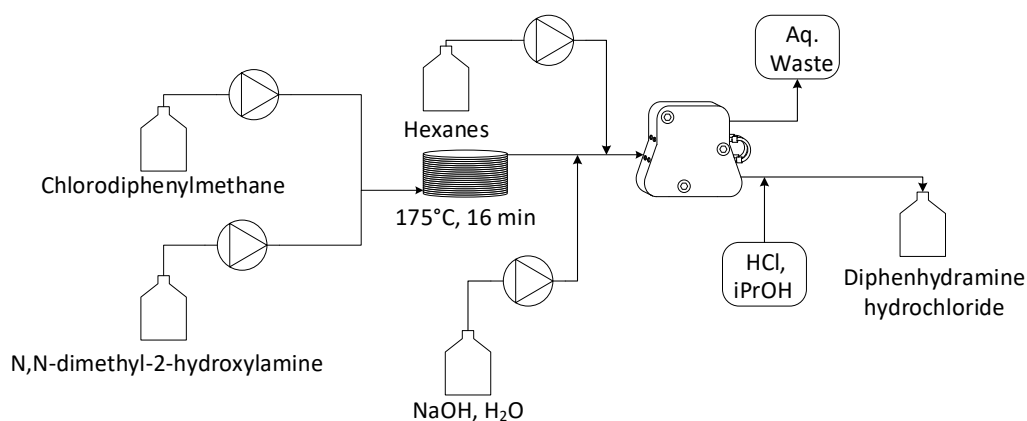
*Figure 3 Continuous flow synthesis of diphenhydramine hydrochloride.*[21]

Continuous flow methods can be described as many small batch reactions (plug flow)[22] being carried out simultaneously within the device (Figure 4). These devices are much smaller in volume but still allow sufficient time for the chemical reaction to occur in the reactor system.[23–25] The residence time, the time the reactants spend in the reactor (equivalent to reaction time), is a function of the total flow rates of all pumps used and the reactor volume.[26]
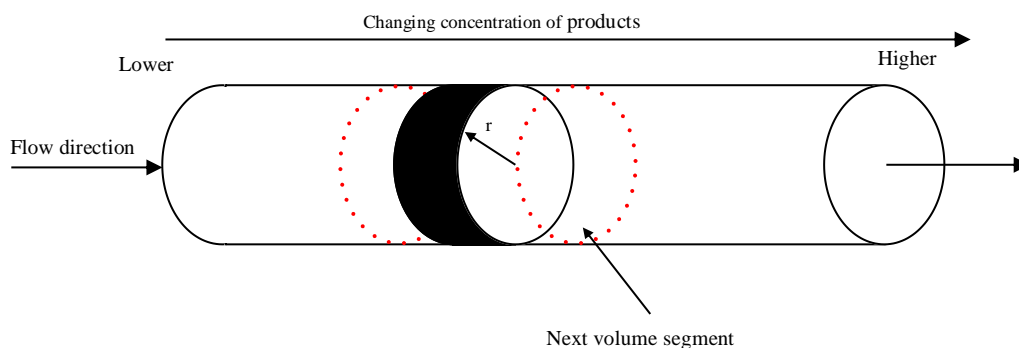


*Figure 4 Diagram of a plug flow system showing how segments are independent from each other.*

$$\tau = \frac{m}{f} \tag{1}$$

Where $\tau$ is the residence time, $m$ is the overall quantity of material to pass through the reactor (reactor volume) and $f$ is the overall flow rate of material at steady state. Using this method ensures a tailored residence time for each specific reaction which greatly increases reaction control. Steady state must be guaranteed in a continuous flow reaction

as this dictates when the reaction is in a state of none varying conditions in order to achieve reliable results (Figure 5).
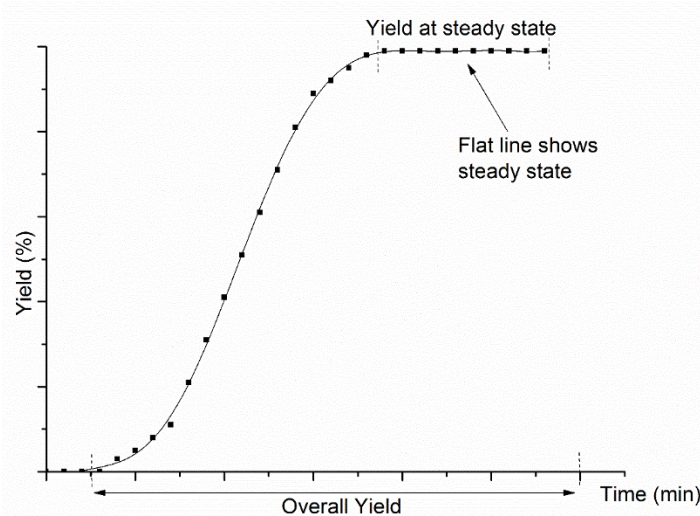


*Figure 5 Steady state of a system in a continuous flow reactor and the difference between steady state and overall conversion.*

### 1.1.3. Mixing in Flow Reactors

Smaller reactor volumes reduce diffusion distances for reactants, which allows mixing to be better controlled (potentially including a static mixer) and occur perpendicular to the flow direction. Smaller diameter channels result in larger surface area (SA) to volume (V) ratios which is shown in the comparison below for a flask reactor, tank reactor and flow reactor.

- Surface area for a 100 mL round bottomed flask reactor (d = 65 mm) *ca.* 132.7 cm$^2$, 8:6 surface area to volume ratio,
- Surface area for a 1 L tank reactor (d = 11 cm, h = 11 cm) *ca.* 570.2 cm$^2$, 4:7 surface area to volume ratio,
- Surface area for a 10 mL flow reactor (1/32" ID, h = 20 m) *ca.* 0.198 cm$^2$, 1:50 surface area to volume ratio.

A larger SA:V indicates better heat transfer rates, enabling more efficient heating and cooling therefore easier process intensification.[1,23,25] Macromixing and accumulation of reactive intermediates are limited in flow reactors, better heat transfer allows better

control of highly exothermic reactions minimising reaction runaway and greatly improving the safety of certain processes.[1,25] Mixing within a micro reactor is generally more rapid than a batch reactor because of their small channel size. The plug flow nature of flow reactors means that mixing longitudinally (in the direction of flow) is minimized permitting radial mixing to dominate. As radial mixing is dominant, the subsequent 'plug' and previous 'plug' react completely independently of one another and have no effect on each other (Figure 4).[1,27]

Extent of mixing, given by Reynolds numbers; $R_e$, can be changed and tailored using various mixers. Static mixers (Figure 6) join different fluids together whether mixing two liquids, two gases, dispersing gas into a liquid or blending two immiscible liquids together.
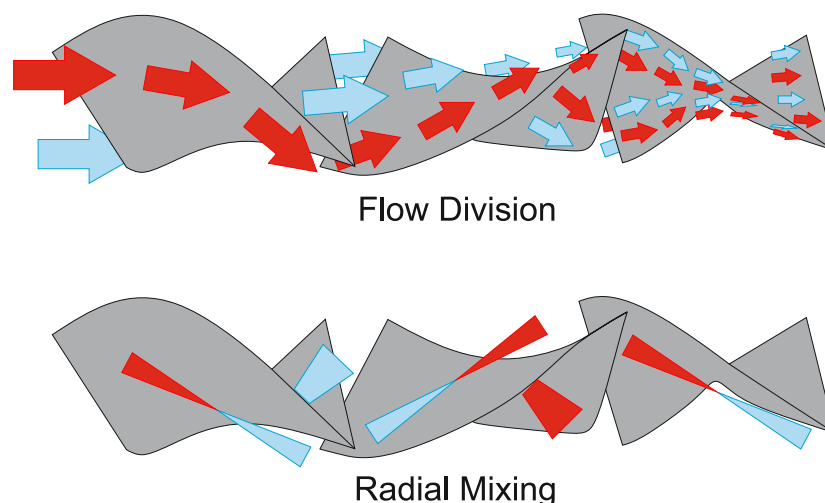


*Figure 6 Example of a static mixer used for continuous flow processes. Left: indicates a mixer which splits the flow stream in half several times before reforming the streams in a different order. Further radial mixing (red triangles) helps the create Eddy currents and ensure a more homogeneous solution.*[28]

Mixing can be separated into three different regimes depending on their Reynolds number: laminar (<2000), transitional (2000-4000) and turbulent (>4000). Tubular continuous flow systems, which only use T or Y mixing pieces, exhibit laminar flow regimes and those with static mixers typically can be found in the transitional regime. Another parameter which plays a large role in the mixing within a microreactor is the Damköhler number (*Da*). The Damköhler number is a dimensionless number which is described by the ratio of the rate of reaction to the rate of mass transfer by diffusion. If *Da* < 1 then mixing can be said to be >95% homogenous and can be fulfilled before the

reaction occurs. Conversely, if $Da > 1$, mass transfer dominates the flow stream forming concentration gradients within the reactor. This may have greater effects on the overall selectivity of reaction if there is a competitive side reaction which is highly influenced by mixing.[6,29]

The performance and extent of mixing can be visualised by a residence time distribution (RTD). RTDs are useful as molecules spend different times in the reactor which affects the product distribution at the end of the reactor. An RTD (for a flow reactor) is measured using a tracer, *e.g.* a UV-active substance, which is injected into the reactor at time, t=0, flowed through the reactor and through a detector at the end where it is detected by UV-Vis spectroscopy. At t=0 the tracer will show a sharp peak but once passed through the detector, a peak which is longer and shallower, but still with the same area, will be observed (Figure 7).
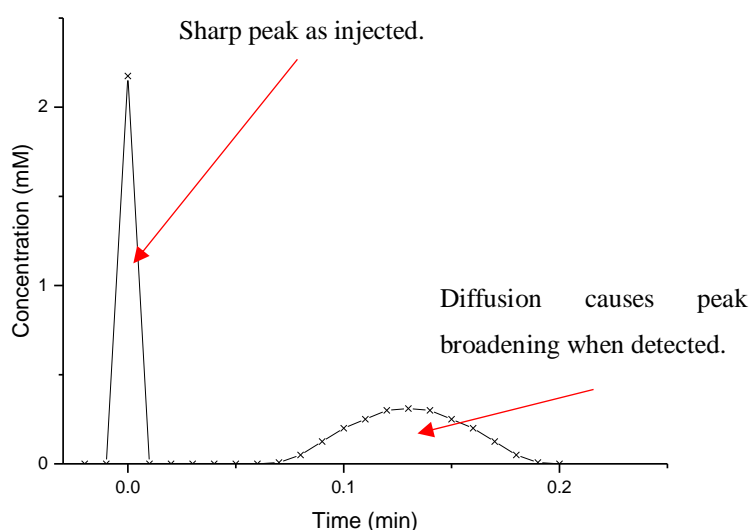


*Figure 7 Residence time distribution example. At t=0 the tracer example will show a sharp peak. Upon detection after passing though the reactor the tracer will spread out creating the curve seen on the right.*

This pattern shows ideal behaviour within the flow reactor, any deviation from this curve suggests the reactor is behaving beyond ideal. Peak tailing suggests the reactor contains a dead volume. The lengthening of the peak is indicative of the stream front moving faster than the edges of the flow stream, due to frictional forces on the liquid from the reactor walls. The deviation from an ideal RTD for a flow reactor can be due to

turbulence, diffusion or a non-uniform velocity profile. If more than one peak occurs, then it may suggest a parallel pathway to the exit exists or strong reactor circulation.[30]

### 1.1.4. Using Flow for Elevated Temperature & Pressure

In general, reactions carried out in continuous flow have been shown to reduce the overall waste, increase the energy efficiency of a process and give a greater output of new compounds. In addition, it enables reactions to operate at higher pressures and temperatures, above the boiling point of the solvent, or even at much cooler temperatures.[22–25] As previously mentioned, reaction runaway is a potential problem in larger scale batch reactions where the SA:V ratio is small. The miniaturised dimensions of micro-flow reactors improve the performance of the reactions through more efficient heat transfer reducing the likelihood of runaway. Any excess build-up of heat will be transferred from the internal volume of the system to the atmosphere rapidly dictated by the heat transfer coefficient ($U$) which allows the calculation of heat transfer rates ($q$). Heat transfer rate is directly proportional to the surface area of the reactor, and so heat dissipation is faster when the surface area is larger. This relationship can also be used when cooling the reaction. Much faster cooling will occur when the surface area is large meaning any excess heat can be rapidly balanced through rapid cooling of the reactor.[31,32] Many reactions carried out in the lab are typically carried out at ambient conditions and never go beyond the boiling point of the solvent as high temperatures and pressures can greatly increase the hazards of the process. Being able to operate a reaction above the boiling point of the solvent in use, and therefore at greater than atmospheric pressure, can have beneficial effects on the reaction by reducing the overall time of the process. It is possible to run reactions in a batch reactor at higher temperatures but necessitates higher boiling solvents which may complicate future process stages such as workup and product purification. Fine tuning of reaction pressure can easily be done in flow processes using a back-pressure regulator (BPR). Constant pressure (spring) BPRs keep the entirety of the upstream system under a set pressure and cannot be changed whilst variable BPRs can be employed to change the pressure of the system

### 1.1.5.   **Transferring Sample from Reactor to Analytical Tool**

Reactions can easily be automated using pumps and flow techniques but are much easier to optimise when directly linked to an analytical technique. Reaction samples are taken, and a feedback loop generated using a sample switching device (actuator). The sample actuator has two parallel flow streams acting as an interface between reaction and analysis. The reaction passes into a set internal volume where the device is then triggered (at a set rate) and a spectrum taken using contact closure *via* connection to a computer (Figure 8). Contact closure automatically starts the on-line technique keeping the cycle flowing as the algorithm progresses. This sample can be transferred to HPLC or GC for separation, then to MS for peak *m/z* characterisation or directly into the MS for direct characterisation in real time.
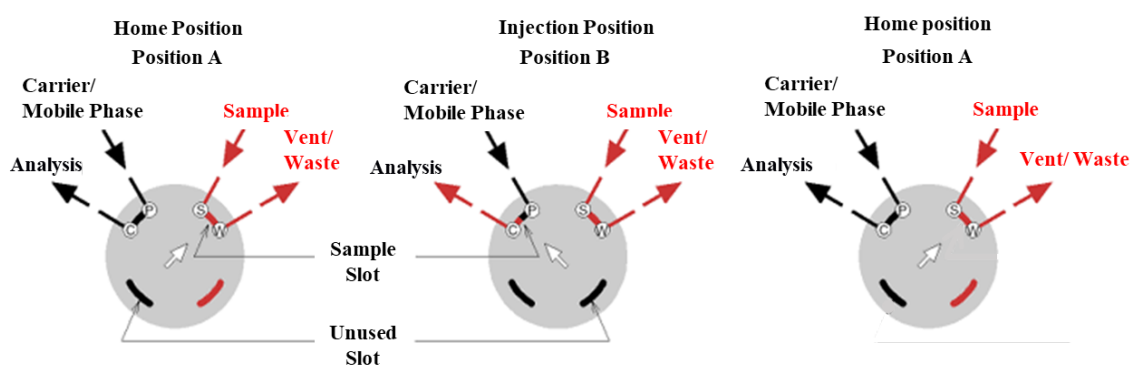


*Figure 8 Sample loop diagram showing the two parallel flow streams before sampling (left, position A, home position) and after the triggering of the sample loop (right, position B, injection position). The black line indicates the analytical technique mobile phase and the red line the sample stream. Once injected the valve turns back to position A, home postion.*

### 1.2.   **Continuous Flow Process Analytical Technologies (PATs)**

Automating or optimising a process requires algorithms to generate a feedback loop. This loop relies on the output function from process analytical technology without the need to manually transfer it for off-line analysis. On-line analysis allows a process to be continuously monitored by extracting a portion of the flow volume whilst in-line analysis

monitors the entire flow stream directly. Each technique offers its own advantages and disadvantages depending on the intended application. PATs aid optimisation as reactants, products, intermediates and impurities can be seen in real-time allowing live yield calculations, kinetic data and the search for optimum conditions to be determined. Many techniques are available and provide unique insights into continuous processes. A summary of each technique is shown in Table 2.[33]

Each chosen analytical method has specific aims which must be met for it to be effective either for qualitative data, quantitative data or both:

(i)     To minimise the delay between the sample taken and results being available in the correct form which can be read and evaluated by the employed algorithm against the set experiment objective.

(ii)    To increase and maximise the observability of different system states. For chemical synthesis these different states correspond to qualitative identification of molecules (analysis of their structure) and quantitation of reaction composition.

Point (i) ensures the experimental system is not idle and a steady stream of data is being collected and evaluated. Performing large numbers of experiments with very small quantities of reaction sample is key to differentiating between automated continuous flow reactions compared to traditional development and optimisation strategies. Point (ii) can be much more difficult to achieve accurately but is critical for algorithmic discovery and reaction optimisation procedures.[33]

*Table 2 Outline of process analytical techniques used for continuous flow process giving quantitative detail on reaction composition and qualitative information on substituents. Each technique gives varying detail in different acquisition times i.e. some can give data on steady state whilst others detail of concentration.*[34]

| Technique | Information Type | Sensitivity / mol% | Acquisition Speed / s | Limitations |
|---|---|---|---|---|
| **Mid-IR** | Chemical identity, concentration. Solid, liquid or gas samples. | $\sim 10^{-1}$ | $\sim 1$ | Short fibres. Intolerant to water. |
| **Near-IR** | Chemical identity, concentration. | $\sim 10^{-1}$ | $\sim 1$ | Less informative than Mid-IR, tolerant to water. |
| **Raman** | Chemical identity, crystal structure, concentration. Solid or liquid samples. | $\sim 10^{-1}$ Potentially to individual molecules in the case of SER(R)S | $\sim 1$-100 | Fluorescence masking Raman signal |
| **UV-Vis** | Chemical identity, concentration. | $\sim 10^{-4}$ | <1 | Limited number of species. |
| **NMR** | Molecular structure, I.D. of unknown compounds, concentration. | $\sim 10^{-3}$ | $\sim 10$ | At present flow method is limited in sensitivity and resolution due to low field. |
| **GC** | Concentration. | $\sim 10^{-6}$ | 10-1500 | Typically – slow. Cannot I.D. unknown compounds. Difficult to automate. |
| **HPLC** | Concentration. | $\sim 10^{-6}$ | 200-1500 | Long method development times. Must be combined with MS for proof of molecular identity. |
| **MS or MS/MS** | Chemical identity, concentration. | $\sim 10^{-8}$ | 5-20 | Requires cheminformatics expertise. MS/MS is more informative, but few process instruments on the market. Difficult method development on more advanced systems. |
| **Process Sensors** | N/A | Pressure Temperature<br>pH<br>Conductivity<br>Viscosity<br>Dynamic light scattering<br>Ultrasound | | Sensor fouling<br>Sensor fouling<br>Requires specific in-line cell<br>Difficult for in-line, requires dilution<br>Underdeveloped |

### 1.2.1. Spectroscopic Analytical Techniques

Spectroscopic analysis is a non-destructive technique which allows reactions to be screened in-line. Spectroscopic techniques provide energy to the sample *via* photons of different wavelength light causing bonds in the analyte to bend, stretch, or nuclei to spin.

#### *1.2.1.1.    Fourier Transform Infrared Spectroscopy*

FT-IR is a spectroscopic technique which is non-invasive and non-destructive because the infrared interface does not interfere with the reaction.[35] FT-IR offers very rapid acquisition times and can quantify the concentrations of different compounds in the reaction. Concentration can be mapped accurately because peak intensity is proportional to the concentration of the analyte, though requires calibration using a standard. Infrared spectroscopy offers advantages to optimisation if the reaction does not contain large levels of impurities or by-products and if the peaks are de-convoluted.
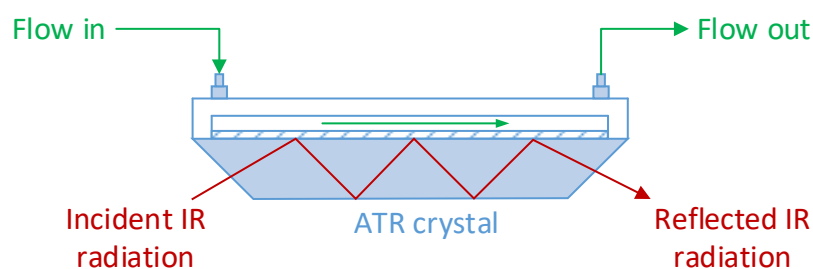


*Figure 9 Diagram outlining the flow cell of a Fourier transform infrared in-line flow detector. The flow stream enters at one side, fills the cell and makes direct contact with the ATR crystal before exiting. The incident IR radiation passes through the cell and interacts with the edge of the flow stream.*

Skilton *et al.*[36] used FT-IR to optimise the reaction of 1-pentanol, **22**, with dimethyl carbonate to form pentylmethylether, **23**, (previous work used GLC). FT-IR analysis significantly reduced the analysis time from 35 minutes to 3.2 minutes due to the increased sampling rate of the FT-IR. This technique allowed measurement of chemical yields (>99%) when the reactor was at steady state. It was also necessary to calibrate the FT-IR instrument using on-line GLC which reduced the initial speed advantage of FT-IR. However, with the use of FT-IR the volume of chemicals used was significantly reduced and a much wider image of the experimental space was permitted.
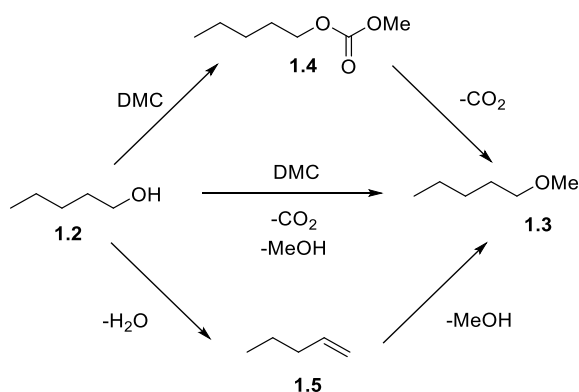
*Scheme 5 The reaction between dimethyl carbonate (DMC) and 1-pentanol, **22**, along with competing side reactions which occur at different temperatures.*

Extensive work has been completed in the Ley group using in-line IR as an analytical tool for continuous flow chemical processing. This technique has been used to monitor fluorinations[37], chlorosulfonylations[38], prepare benzamide derivatives for the pharmaceutical sector[39], measure CO concentrations during methoxycarbonylations[40], prepare arylmagnesium reagents in flow[41] among others.

### *1.2.1.2.    Raman Spectroscopy*

Raman spectroscopy (Figure 10) relies on the manipulation of bonds within the molecule where energy from a laser beam causes them to bend and stretch. Unlike infrared, Raman monitors light scattering as all molecules give unique patterns. Raman spectroscopy can be used to monitor reaction profile or changes in morphology of crystalline materials.[42,43]
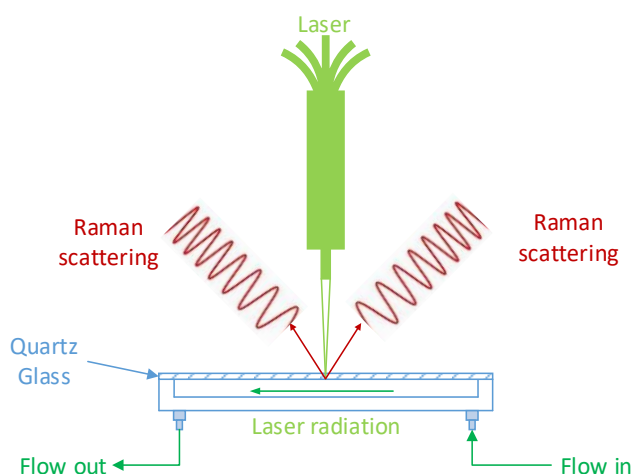


*Figure 10 Schematic of a flow Raman spectrometer. The flow stream passes into the flow cell beneath a quartz glass pane. The Raman laser beam passes into the stream where the light is scattered and detected.*

Work completed by Šahnić *et al.*[44] shows how Raman spectroscopy can be used to monitor the synthesis and crystallisation of Omeprazole. The final reaction stage, oxidation of the sulphide, was monitored for the formation of the desired product and side-product sulphone (Scheme 6). Omeprazole is widely used for several conditions affecting the gastrointestinal tract.



| 1.6 | | Omeprazole 1.7 | | Sulphone side-product 1.8 |
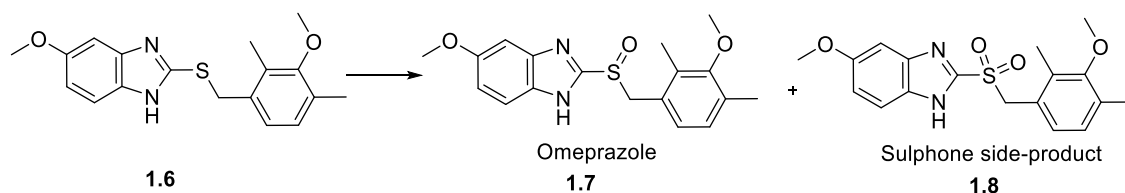
*Scheme 6 Synthesis route for the formation of Omeprazole using in-line Raman spectroscopy with a sulphone side-product produced.*[44]

The quantitative ability of Raman spectroscopy was tested using different extraction wavelengths (785 & 1064 nm) for each reaction species. Formation of the sulphone directly influences the purity of the Omeprazole after isolation, therefore it was important to develop a sufficient and reliable method, using in-line Raman. Current methodologies use off-line HPLC for reaction quantitation, but this is far slower than in-line Raman. Comparative off-line measurements by LC-SPE/NMR we completed on aliquots of the reaction mixture. Their method allowed maximum yield and satisfactory purity of the desired isolated product, but lower yields were observed using longer reaction times as increased sulphone levels were detected. Success during this work saw the procedure scaled-up and transferred to the pilot plant for kilo-scale synthesis whilst using in-line Raman rather than off-line HPLC.

### 1.2.1.3. *UV-Vis Spectroscopy*

Miniaturisation of UV-Vis spectrometers has allowed adaption use in flow. Light source, fibreoptic cables, flow cell and detector comprise a typical setup similar to a traditional UV-Vis instrument. The flow cell is a machined, stainless steel block with entry and exit for the flow stream along with flow entry for the source (Figure 11). The light path allows the incident beam (from the lamp transported using fibre optic cables) to pass across the flow stream at a set distance and into the detector where sample absorbance or transmittance can be measured ultimately transferring data on sample concentration as given by the Beer-Lambert law.[45]

This equation shows absorbance as a function of the lights initial intensity divided by its final intensity (equation A).

$$A = log_{10}\frac{I_0}{I} \quad (equ.\ \boldsymbol{A})$$

Absorbance can also be calculated from the sample using path length (l), concentration (c) and molar absorptivity (ε), equation B.

$$A = \varepsilon \cdot C \cdot l \qquad (equ.\ \boldsymbol{B})$$
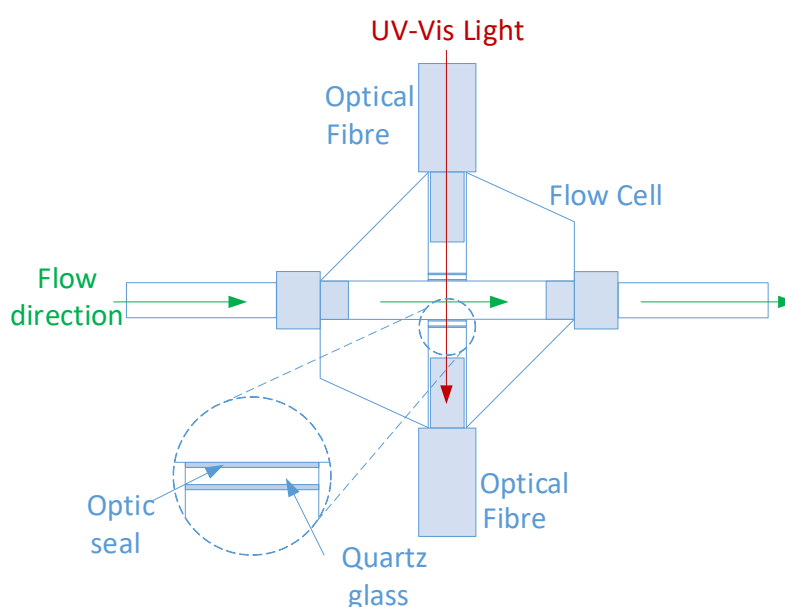


*Figure 11 Diagram of a UV-Vis flow cell used for in-line UV-Vis spectroscopy. The flow stream moves perpendicular to the light beam which passes through fibre optic cables across the reaction stream and to the detector.*

Günther *et al.*[46] monitored the formation of monomeric azo dyes using a segmented flow microflow reactor with UV-Vis spectroscopy from 350 – 850 nm (visible light region only). In-line analysis allowed each dye formation to be monitored as well as indicating the concentration of each. This technique proved useful for the characterisation of segment properties, such as size and shape. Different sized segments possessed different residence times and therefore different absorbances. If a low maximum absorbance was observed, it suggested that the segment size was smaller than the diameter of tubing used.
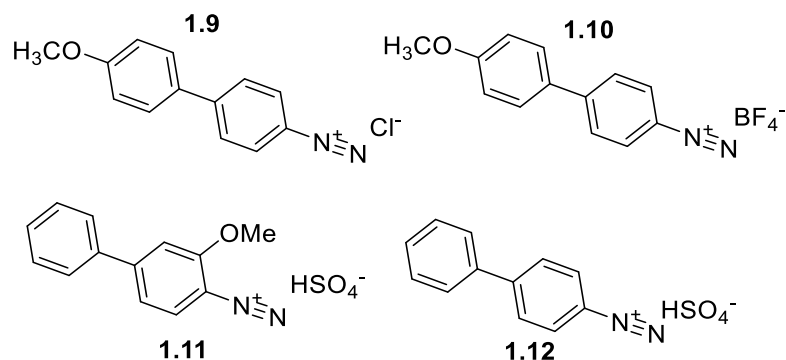
*Scheme 7 Diazonium salts used to form azo dyes in segmented flow using in-line UV-Vis to monitor their formation.*[46]

In-line UV-Vis spectra are highly dependent on detector environment where temperature fluctuations can cause differences in the electronics within the detector changing background noise and potentially reducing accuracy and sensitivity of the analysis. The concentration of the flow stream may also affect the quality of spectra; highly absorbent samples easily exceed the detector linearity giving inaccurate data. This may be fixed by changing flow cell path length or use of a dilution pump before the flow cell. Calibration samples must be taken before reaction monitoring to get reliable reaction concentrations of all components in the system. This is not always ideal if the reaction produces complex reaction mixtures and response factors are unknown.

### *1.2.1.4. Nuclear Magnetic Resonance Spectroscopy*

Nuclear magnetic resonance (NMR) spectroscopy is a powerful tool for chemists in the determination of species synthesised in reactions. Bench-top NMR instruments can be integrated into continuous processes offering real-time data by allowing the flow stream to enter and exit the instrument (Figure 12). Application of magnetic fields induce spin interactions between nuclei in the compound where they can absorb or emit radiation. Nuclei relax via emission of energy and relaxation depends on the NMR frequency giving a signal proportional to the concentration of the analyte useful for conversion and yield calculations without calibration. NMR spectroscopy supplies data quickly and constantly as the reaction stream passes through the core of the instrument and gives an accurate identification of different compounds.[25,47,48]
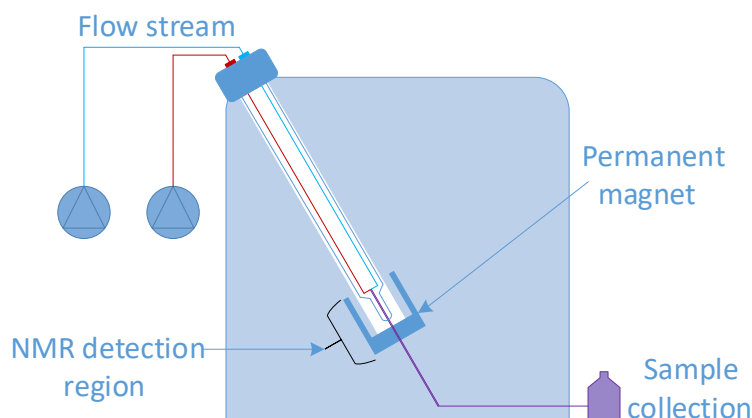
*Figure 12 Schematic of a bench-top flow NMR device. The samples mix and flow through the flow cell to the permanent magnet where the analyte, impurities or side-products are detected. The flow stream flows out of the spectrometer to collection.*

New bench-top instruments have limited resolution and a more time and investment to set-up, typically a minimum of 24 hrs for the instrument to equilbrate. These problems mean spectra will be complex, providing less useful information about reagent signals with large levels of peak overlap.

Sans *et al.*[25] used a Spinsolve NMR spectrometer to optimise different reactions in flow using various NMR techniques before attempting a full self-optimisation. Initially, a condensation reaction was monitored to form an imine, **32**, (Scheme 8) using reagent flowrate control. 2D NMR techniques such as HSQC, HECTOR and COSY were used to see spectra detail from a continuous flow process. Self-optimisation for the formation of **32** using the Nelder-Mead (modified) simplex, maximised the output yield, with a maximum of 79 %.
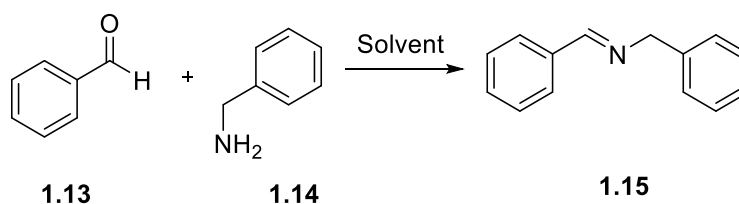


*Scheme 8 The use of in-line NMR spectroscopy to probe the formation of an imine.*[25]

Currently, bench top NMR spectrometers are not only limited by their resolution but are far less time and cost efficient for pharmaceutical and fine chemical industry use. However, use during a continuous flow process certainly reduces the time taken to sample compared to off-line sampling and also cuts cost compared to large volume transfer lines to higher frequency NMR spectrometers. Current low field systems may

not benefit industry as a very high concentration of sample is needed to overcome the low sensitivity and resolution of the instruments. At 60 - 83 MHz many signals will overlap with ones which may be of interest for the reaction carried out (Figure 13).
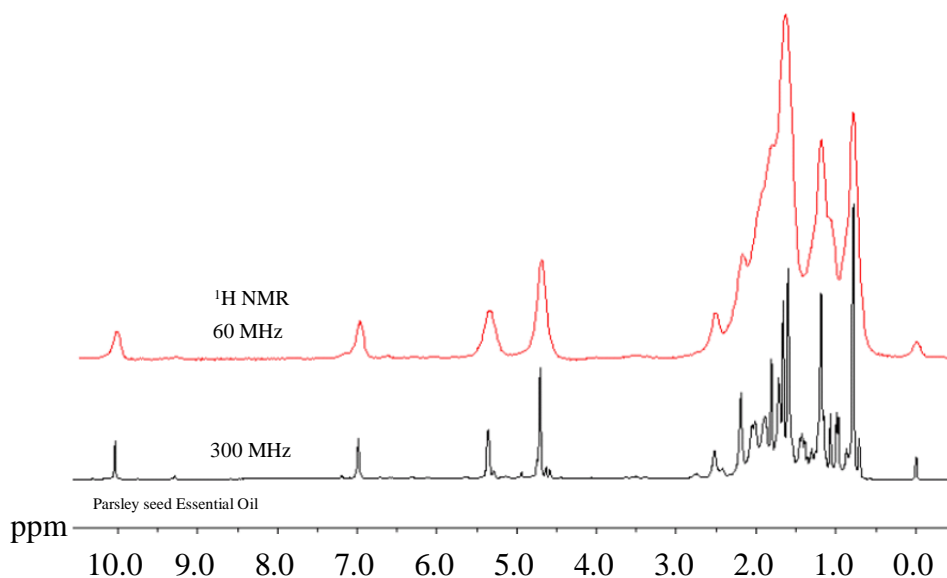


*Figure 13 A comparison of a 60 MHz bench-top NMR spectrometer and a 300 MHz NMR spectrometer in the $^1$H NMR spectra of parsley seed essential oil.*[49]

New advances have been able to adapt existing NMR spectrometers with frequencies of, 300-900 MHz, to use with continuous flow reactors by addition of a flow probe. This enables flow reactions to be carried out if high resolution or sensitivity is needed. The downside to this method is the capital investment in the flow probe and taking the flow reactor to the NMR spectrometer. Few laboratories are equipped to have the relevant safety procedures necessary for a flow reaction to be performed close to an NMR instrument. The type of reaction may also limit the use of a flow NMR device as only liquid samples may be analysed efficiently as gaseous media may not be analysed if insufficient mixing has occurred.

### 1.2.2. Separation Analytical Techniques

#### 1.2.2.1. Gas Chromatography

Gas chromatography (GC) can be used as an on-line technique to monitor continuous flow processes (Figure 14). GC allows analysis of volatile compounds and gaseous species from reactions, suited to species that vaporise above 350 °C.

Identical to operation off-line, a portion of the flow stream is injected and passes through column using a mobile phase. The sample separates, and each individual peak is detected

by the chosen detector. With use in a self-optimisation process the peak area is used to determine the yield.
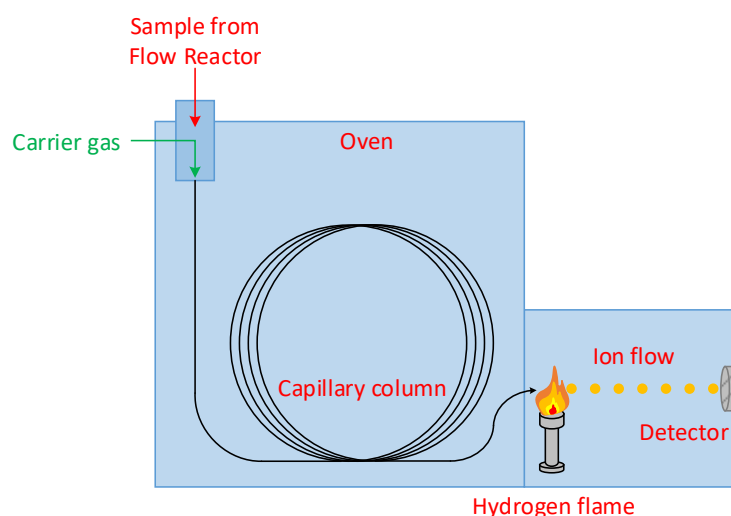


*Figure 14 Diagram of an on-line gas chromatogram (GC). The sample from the flow reaction is injected into the GC from a sample switching device before being separated on the heated capillary column, ionised and detected.*

Parrott *et al.*[50] ran an optimisation of a bimolecular dehydration reaction of ethanol, **1.16**, to diethyl ether, **1.18**, over γ-alumina as a catalyst (Scheme 9) using in-line GC to monitor the reaction. GC results had to be read and converted by MATLAB before a yield could be determined. The high volatility of the reagents helped facilitate in-line sample transfer from the in-line sample loop to the injector and oven. Maximum yield totalled 68 % (0.4 mL min$^{-1}$, 212 bar, 340 °C) from a minimum of 0 %. On-line GC also helped in finding a small plateau region in which all factors produced a yield $\geq$ 67 %.
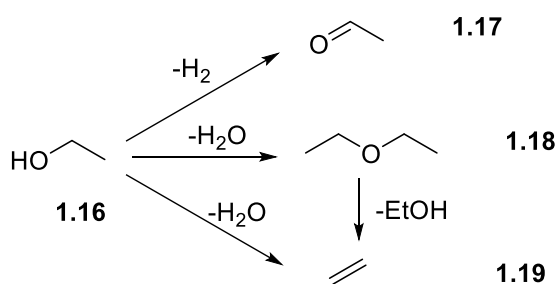


*Scheme 9 The use of γ-alumina in the dehydration of ethanol, **1.16** to diethyl ether **1.18**.[50]*

GC separates volatile compounds, and a detector is needed in order to quantify the final amount of analyte. Many of these are compound specific but may include; flame ionisation, electrolytic or thermal conductivity, electron capture *etc.* Similarly to HPLC,

GC can make use of other analytical techniques for structural characterisation *e.g.* mass spectrometry.

### *1.2.2.2.    High Performance Liquid Chromatography*

High pressure liquid chromatography (HPLC) is an analytical technique offering numerous advantages for optimising reactions (Figure 15). HPLC generally has much slower acquisition times than spectroscopic methods but can be used during self-optimisation. Chromatograms show distinctive and quantifiable peaks for each compound eluted shown by retention time with each peak relating to each compound. However, like GC, HPLC must have a detector to be qualitative so compounds can be identified making it more advantageous with prior knowledge of the reaction. Calibration ensures quantifiable results to assist calculation of concentration of peaks. UV-detection is typically used necessitating a chromophore in the compound being detected.
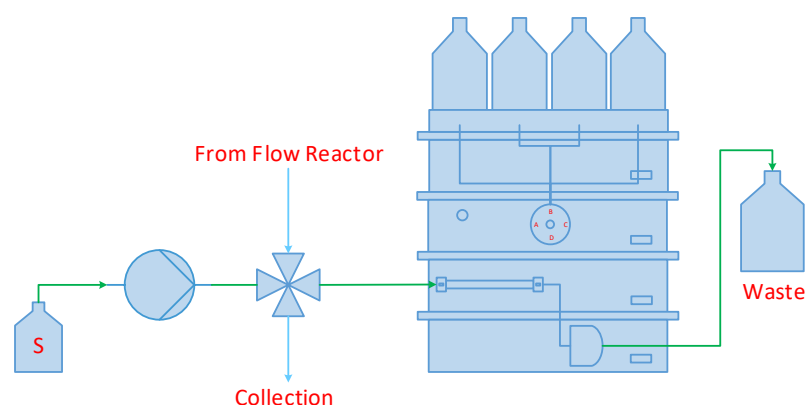


*Figure 15 Diagram of an on-line high-performance liquid chromatogram (HPLC). The sample from the flow reaction is injected into the HPLC from a sample switching device before being separated on the HPLC column and detected via a UV detector.*

McMullen and Jensen[51,52] studied various reactions using on-line HPLC to aid optimisation. HPLC can distinguish between regio- and stereo-chemically different compounds and slight differences between reagent and product structures which may be missed *via* spectroscopic analysis. Low concentration reactions may be completed and sampled in-line to directly separate each component, or on-line (for more concentrated samples) which requires the use of a sample actuator (**1.1.5**) to switch a portion of the reaction into a parallel stream to flow into the HPLC for separation. On-line measurements are not seen as 'real time' as method lengths tend to vary anywhere

between 2 and 20 minutes depending on the number of components needing to be separated. Due to this it may be necessary to use a technique which has a quick acquisition time and both low sensitivity and high resolution.

### 1.2.3. Mass Spectrometry as a Process Analytical Tool

Mass spectrometry (MS) and continuous flow have, only recently, been used together to monitor reactions in real-time. Bench-top MS relies on compound polarity for ionisation to occur; the most common techniques are electrospray ionisation (ESI, **1.2.3.2**) and atmospheric pressure chemical ionisation (APCI, **1.2.3.3**). For a compound to be detected it must be ionised to either a positive or negative state. Detection measures the mass-to-charge ratio (*m/z*) of gaseous ions in the system to make a plot of signal intensity against *m/z*. The technique scans over a range of *m/z* to isolate each separately.

#### *1.2.3.1.    General Theory of Mass Spectrometry*

MS, in general, delivers unique analyte specificity because each species is structurally based upon its detected *m/z* giving accurate identification of unknown compounds. There are two modes of MS, ESI and APCI, where the sample is first nebulised, then desolvated before being ionised, separated and detected (Figure 16).
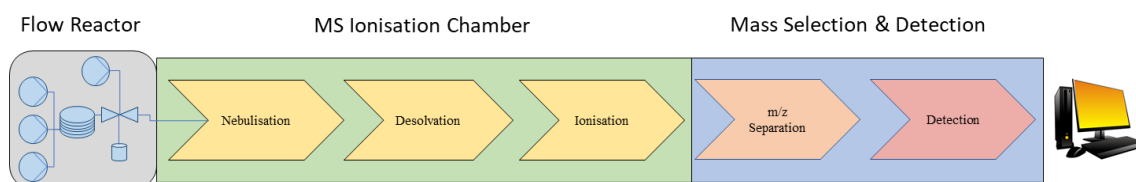


*Figure 16 General process schematic for the ionisation and detection of species, from the reaction system (left), by a mass spectrometer. The output total ion chromatogram (TIC) and mass spectra are shown directly on the computer in real time (right).*

The power of MS detection is only enabled if the correct ionisation technique is used to get the reaction solution into ionised particles. Both ionisation modes were used during this PhD project and both can give a positive or negative charge. Deciding on which mode to use largely relies on the polarity of the analyte and the composition of the matrix.

#### *1.2.3.2.    Electrospray Ionisation*

Ion generation from electrospray ionisation (ESI) begins with the production of charged eluent droplets at the very tip of the ESI needle. The flow stream is fed by a mobile phase eluent which is sprayed directly into the ionisation chamber of the MS at atmospheric

pressure. Either positive or negative ion generation can be selected, depending on the analyte of interest. In positive mode, a voltage is induced between the capillary needle (anode) and aperture plate (cathode). This causes a build-up of positively charged eluent at the tip of the needle which is periodically repelled and attracted to the oppositely charged aperture plate. In negative mode, the opposite process occurs. The capillary needle acts as the cathode whilst the aperture plate is the anode. Negatively charged eluent particles dominate the needle tip and travel electrophoretically towards the aperture.[53,54]

Upon application of the voltage, between the needle and aperture, the build-up of ions occurs (either positive or negative) to the point where repulsive forces match the forces of the surface tension of the eluent. The balance between these forces is known as the Rayleigh instability limit. When the repulsive forces exceed the surface tension of the eluent a cone shaped meniscus is formed called the Taylor cone. This cone allows a continuous spray of charged droplets to migrate from the surface towards the aperture, hence the term electrospray (Figure 17).  Subsequently, between the capillary tip and the aperture, desolvation of the charged analyte species occurs. As the solvent evaporates its effective radius decreases whilst its charge stays constant. This brings into effect the Rayleigh instability limit as the repulsive forces between like-charged particles overcome the surface tension of the eluent and undergo a Columbic cascade explosion forming much smaller charged droplets. Ions in the charged droplet are then able to evaporate from within the solvating media to form gaseous, charged analyte ions.[53–55]
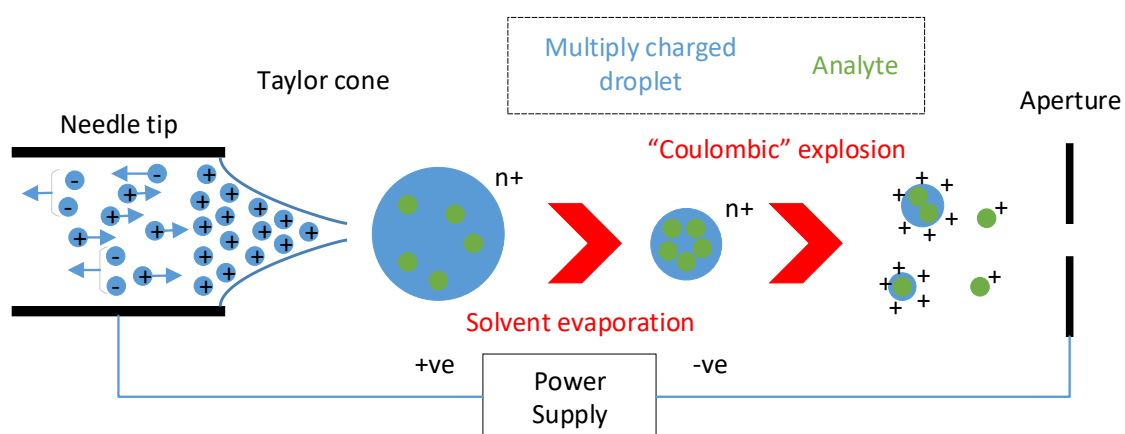


*Figure 17 A schematic showing the mechanism of ion formation in electrospray ionisation.*[54]

ESI is a highly versatile method of analyte ionisation and is very useful for a wide range of molecular weight compounds and can ionise anything which will form an ion in

solution. Ionisation is preferential to molecules that contain groups which can accept a charge, typically heteroatoms like N, S and O atoms, molecules containing these can be analysed by ESI. This also means that any molecules that can accept a charge *via* charge induction may also be analysed with ESI. Ion formation using ESI relies greatly on the pH of the mobile phase and the pKa value of the analyte and follows principles of acid-base theory. However, being sensitive to pKa means ESI is more sensitive to contaminants such as alkali metals or basic compounds and it is not a good ionisation technique for uncharged, non-polar compounds.[53,55]

### *1.2.3.3.*     *Atmospheric Pressure Chemical Ionisation*

Atmospheric pressure chemical ionisation (APCI), utilises a corona discharge to form ions relying on the formation of a mobile phase plasma. Charge transfer processes dominate ion generation which reduces the need to pre-form charged eluent species. The sample is nebulised before being heated into the gaseous phase to form an analyte solvent vapour. This vapour (the bulk being solvent) is then ionised *via* corona discharge through proton addition (positive mode) or *via* proton abstraction or electron capture (negative mode). Charged solvent molecules proceed to chemically ionise the gaseous analyte *via* either protonation or deprotonation to produce the required charged species ready for detection (Figure 18). The flow of ions is transported to the aperture via the carrier gas and does not rely on a potential difference between the sample feed and the MS sample cone.[53,56]
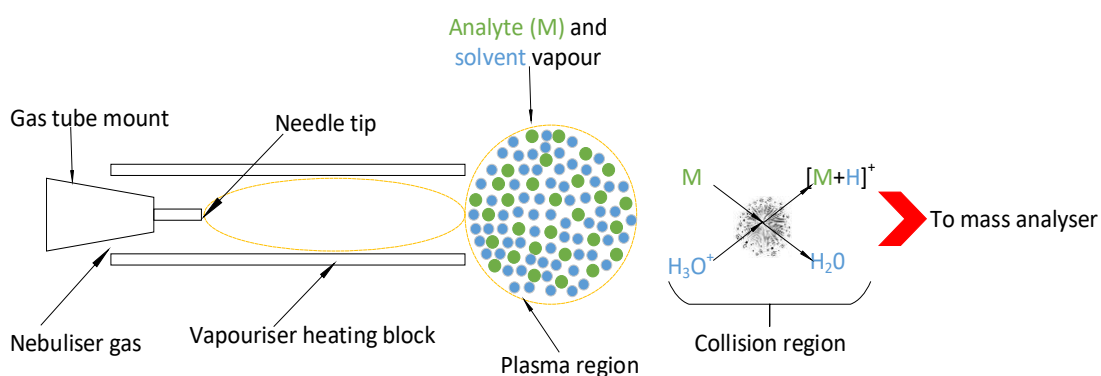


*Figure 18 Diagram showing the ionisation process of solvent and analyte using an APCI source.[57]*

In positive ion mode, APCI ionises gaseous analyte molecules *via* proton transfer or charge exchange. Equation **C** outlines the ionisation of species if analyte (M) has a larger proton affinity than solvent (S).

$$M + [S + H]^+ \rightarrow [M + H]^+ + S \qquad (equ.\ \textbf{C})$$
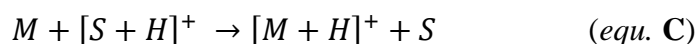
Equation **D** shows how ionic species are formed if both M and S have similar proton affinities.

$$M + [S + H]^+ \rightarrow [M + S + H]^+ \qquad (equ.\ \textbf{D})$$
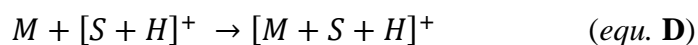
If a charge transfer process occurs, then ionic species are formed as shown in equation **E**.

$$M + S^{\cdot+} \rightarrow M^{\cdot+} + S \qquad (equ.\ \textbf{E})$$

In negative ion mode, charged species are produced by either proton extraction or electron capture. Equation **F** outlines the ionisation of species if solvent (S) has a larger proton affinity than analyte (M).

$$M + [S - H]^- \rightarrow [M - H]^- + S \qquad (equ.\ \textbf{F})$$
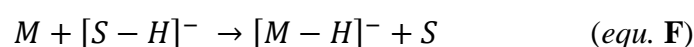
Anions produced in MS from APCI can also help form other negative species through anion (A⁻) attachment, equation **G**.

$$M + A^- \rightarrow [M + A]^- \qquad (equ.\ \textbf{G})$$

In negative APCI mode, gaseous molecular analyte species may be ionised through an electron capture process as shown in equation **H**.

$$M + e^- \rightarrow M^- \qquad (equ.\ \textbf{H})$$

APCI is a useful ionisation technique and is widely used in industry due to its lower sensitivity to chemical interferences making it a rugged, reliable and efficient ionisation process, particularly compared to ESI. The benefit of carrying out ionisations at atmospheric pressure is that all gases in the ionisation chamber undergo collisions with the surrounding carrier gas ($N_2$) inducing a cascade reaction and charging more of the sample matrix. This method of ionisation is most efficient for molecules which have lower polarity, low molecular weight and any other molecule which is unsuitable to be ionised by ESI. As APCI uses high temperatures to vaporise samples it is unsuited to thermally unstable analytes. APCI is a soft ionisation technique producing very few fragments of analyte but instead an intense molecular ion signal.[53,56]

On-line MS requires minimal sample from the reaction and is able to provide detailed structural information and isotopic information about reagents, intermediates, products and impurities. Sampling time is very rapid, and data are constantly recorded as both a total ion chromatogram and mass spectrum. However, MS is prone to ion suppression as analyte ions compete for ionisation with side products, impurities or molecules in the mobile phase. Although MS gives detailed information it cannot differentiate between

regioisomers, relying on compound mass, so further techniques should be employed *e.g.* HPLC or MS/MS. Mass spectra can be simple but have the potential to be complex if species are prone to fragmentation. Gaseous acid-base reactions can cause groups to dissociate, interact or cluster higher mass adducts; such as $Na^+$ or $K^+$ ions or solvate molecules. Different combinations of these phenomena may also occur *e.g.* a fragmentation reaction followed by formation of a $Na^+$ adduct.

On-line MS (4000 MiD, Mosaic Systems) was employed by Bristow *et al.*[58] to monitor a Hofmann rearrangement (Scheme 10) which is widely used in the pharmaceutical industry.
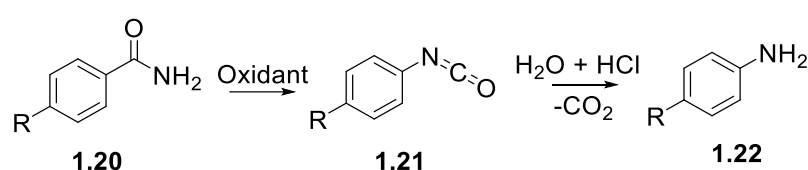


*Scheme 10 A generic reaction scheme for the Hofmann rearrangement from a benzamide, **1.20**, to an isocyanate intermediate, **1.21**, before the desired aniline, **1.22**, is formed.*[58]

The current batch process for the Hoffmann rearrangement reaction produces large quantities of impurity, increasing upon scale-up and reducing overall process robustness. The group hoped to improve the selectivity, increase the yield and provide a more robust process in flow compared to batch. Bench-top, on-line MS overcame sample transfer problems as the reaction was directly connected *via* a mass-rate-attenuator (MRA) to continuously transfer sample for analysis. Dilution *via* make-up flow improved sample compatibility with electrospray ionisation (ESI). ESI+ was used and mass spectra were collected from 100-800 *m/z*. The group aimed to monitor the consumption of reagents and the formation of impurities and products over several temperatures. A small-scale reactor chip (100 μL) was used with syringe pumps to introduce reactants to the reactor to explore the effect of the sampling make-up flow and how it affected the MS data output. Varying compositions of make-up flow (MeCN and $H_2O$) were studied with different additives (formic acid, TFA and a mixture of formic acid and sulfamethazine) in varying amounts (Table 3).

*Table 3 An outline of the make-up flow compositions explored by Bristow et al.[58]*

|   | Ratio MeCN:H$_2$O | Additive | Concentration/ % |
|---|---|---|---|
| **1** | 50:50 | Formic acid | 0.1 |
| **2** | 20:80 | Formic acid | 0.1 |
| **3** | 50:50 | TFA | 0.05 |
| **4** | 20:80 | HCOOH & sulfmethazine | 0.1 & 5 µgmL$^{-1}$ |

Reproducibility was evaluated with flushing of the reactor multiple times between each run using make-up flow in row 4 of Table 3. From 20 – 50 °C, the size of reactant peak, **1.20**, was observed to decrease whilst the isocyanate intermediate, **1.21**, increased. Subsequently, amine formation was monitored, **1.22**, with the addition of 2 M HCl as well as the production of chlorinated impurities in the reaction.[58] This work outlined that MS can be successfully coupled to a flow reactor to understand reactions and to gain knowledge of side product and impurity formation. If the reaction was too concentrated an underestimate of the sample matrix is made due to inefficiencies during ionisation and due to detector saturation. This experimental process only provides information on how to use MS systems qualitatively and fails to show its quantitative ability.[58]

Browne *et al.* investigated using a miniature MS instrument, a 3500 MiD miniature MS (Microsaic Systems), as an on-line reaction monitoring tool using a Uniqsis FlowSyn system (20 mL reactor volume) with a sample loop (5 µL) which switched between the reactor flow and the MS make-up flow (50:50 MeCN : H$_2$O with 0.1 % formic acid).[59] The group aimed to show how rapid MS data acquisition is, how it can screen starting materials, intermediates and products and overall to optimise the formation of benzyne *via* diazotisation of anthranilic acid and cycloaddition to furan (Scheme 11). Flow conditions were chosen as accumulation of one reactive intermediate presents an explosion risk particularly when crystallised.
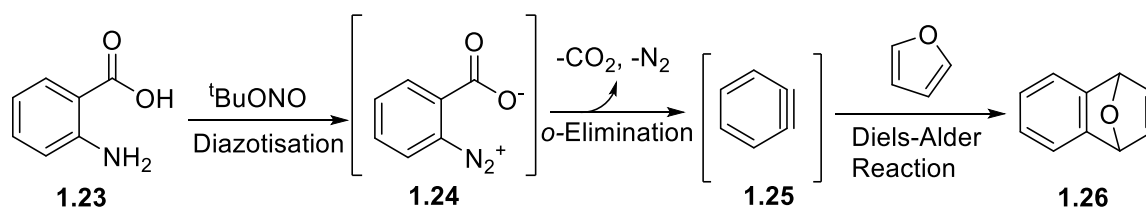


*Scheme 11 Reaction scheme leading to the formation of 1,4-endoxide-1,4-dihydronaphthalene via diazotisation, ortho-elimination and Diels-Alder reactions.*

Initial studies aimed to characterise each reagent and product. It was determined that a collision-induced loss of water forming a positive ion with lower *m/z* was dominant rather than the product *m/z*. Runs of the reagents at 25 °C and 50 °C were performed to see the difference in overall MS composition. Other peaks from induced losses of species such as $N_2$ and other reactive intermediates occurred during ionisation. Optimisation of flow parameters were made by detailing a cumulative spectrum of selected peak intensities. Observations showed that starting material was minimised at a temperature of 50 °C with maximization of the product peak. Samples of the crude flow mixture at both 25 °C and 50 °C, were taken and analysed using off-line LCMS, giving $1.4 \times 10^7$ and $2.5 \times 10^7$ respectively showing the product yield is 79% greater than at the lower temperature. The procedure was repeated using 2,5-dimethylfuran on a 5 mL scale to reduce the amounts of reagents used. This system was again scoped to see reagent and product peaks at different temperatures to see the growth of the product peak at the higher limit of 50 °C.

On-line MS in this system created a picture of starting materials, products and reactive intermediates essential in the reaction from direct and rapid analysis. The peak height of products in the MS corroborate with calculated yields from off-line LCMS analysis although no further work to fully calibrate the system to calculate yields from MS spectra was carried out. Rapid analysis of MS was not combined with algorithms to optimise and understand the chemical system. Further research is still needed in order to fully develop continuous flow processes with on-line MS that is calibrated and can be applied to multiple reaction pathways.

### 1.3.    Optimisation Techniques Used for Continuous Flow Reactions

Process optimisation is paramount to academia and industry where reactions consist of many factors, often requiring maximum possible values to be found; product yield and selectivity or conversion of SM, to name a few. Parameter optimisation through minimisation may also be completed if important to the process *e.g.* side-product yield, E-factor or cost.

There are numerous process optimisation methods, from the simplest one-variable-at-a-time, to the more complex experimental design and even the use of computer-controlled algorithms. The use of self-learning algorithms, a recent approach in the optimisation of chemical syntheses, has been made possible by new flow

technologies and miniaturised process analytical technologies.[60] Methods of self-learning supersede traditional one-variable-at-a-time approaches, making this optimisation a method many chemists are likely to rely on in the future.

### 1.3.1. One-Variable-at-A-Time

The 'one-variable-at-a-time' (OVAT) optimisation approach is the most well-known and exploited method to optimise synthetic reactions and uses, as the name suggests, single variables (whilst keeping all other variables constant) changed iteratively giving a series of reactions. Basic patterns and trends can be seen relating to the changed variable potentially finding conditions for the highest yield. However, there are several drawbacks of this approach. OVAT necessitates many more runs to ensure similar levels of precision as other optimisation techniques. OVAT also struggles to form clear estimates between variable interactions and because OVAT relies on singles factor changes, there may be factors that are overlooked which can lead to true optimal settings being missed.[61]

*Table 4 Factors and levels for the hypothetical optimisation of a reaction method.*[62]

| Factor | Levels |
|---|---|
| **Reactor Temperature/ °C** | 100 and 250 |
| **Reaction Concentration/ M** | 0.2 and 0.4 |
| **Residence Time/ min** | 2 and 280 |

Table 4 outlines a theoretical experimental set-up to illustrate OVAT further. Optimisation begins by performing two experiments using both temperatures (100 & 250 °C) whilst keeping the concentration and residence time constant. Subsequent data from both experiments will show the "optimum" reaction temperature which is set at this value for the next experiments. This process is repeated (changing either concentration or residence time) using the highest response from temperature experiments. It is impossible to know whether a global optimum has been found as factor combinations are missed giving a lack of process information. Temperature may have a profound effect on the outcome response over several residence times. If different starting points are chosen, subsequent experiments may lead to a completely different conclusion for the optimal combination of factors (Figure 19).
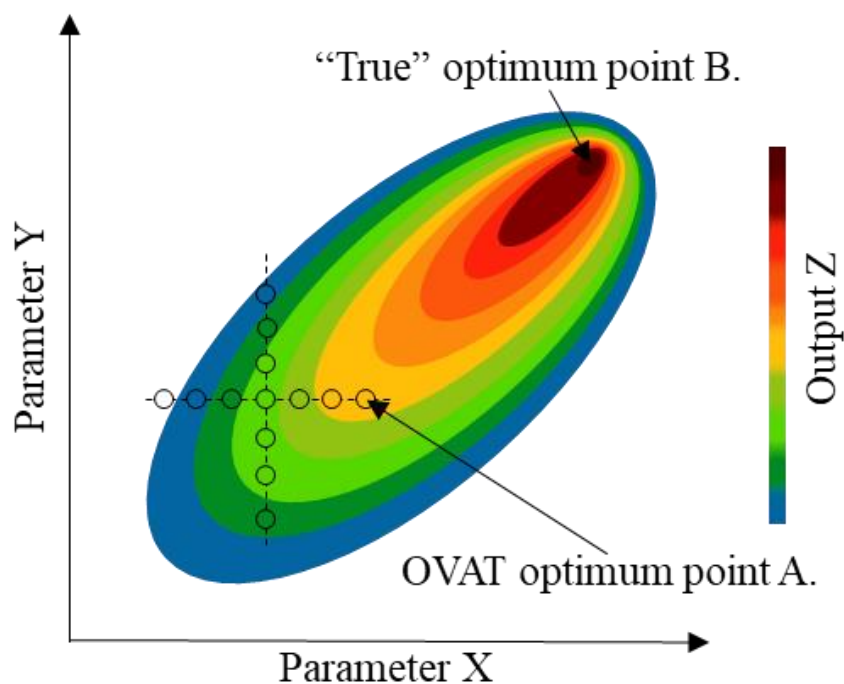
*Figure 19 Optimisation by probing changes in variables "x" & "y" over a range of values (circle points) to see how they affect output "z". The small span of the conditions to find local maxima "A" via the OVAT study is still distant from the global maximum "B".*

Reddy *et al.*[63] investigated how reaction variables influenced the resolution of a racemic mixture of sodium omeprazole, **15**, (Scheme 12).[63] Initially, water content during resolution was studied where it was found that successful batches contained higher volumes of water (1-2 mol). This led to an investigation of water content conducted at the higher set of temperatures (35-40 °C) where no resolution was seen due to precipitation of the $TiO_2$ and none at the minimum content of water. Subsequently, it was discovered that there was no resolution at 0 °C or above 40 °C as many impurities were formed. Greatest selectivity (>99% ee) was found to be between 35-40 °C. Screening experiments where varying moles of $Ti(O^iPr)_4$ and D-(-)-DET showed that the best ratio to use was 0.5:1. This approach included 3 factors and took a minimum of 8 reactions to find the most basic reaction information. Other optimisation approaches can be used to give better quality information, in fewer experiments, as well as give information about variable interactions and detailed reaction information plots.
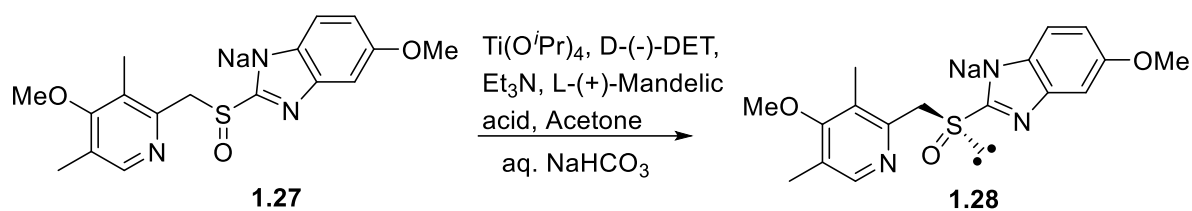
*Scheme 12 Resolution of sodium omeprazole to form (S)-(–)-omeprazole.*[63]

### 1.3.2. Design of Experiments

Chemical processes may be better understood using design of experiments (DoE). This method uses pre-defined experimental conditions generated by statistical design software to help generate a process and reaction model. It can detail variable effects from specified outcome criteria, experimental reproducibility, data predictability and how each variable may interact with one another *i.e.* detailed process understanding. Other statistical parameters are generated alongside models relaying far more detail than traditional OVAT *e.g.* $R^2$, $Q^2$, residuals and reproducibility. Experimental results are used to predict other data points and find an optimum quicker and easier but may also hinder the process if the data contains outliers or if values are missing.

DoE consists of four stages each providing different information on the reaction, for example, a three-dimensional reaction *i.e.* three variables. Set points are displayed throughout the space representing different condition combinations to perform and maximise data output from the design.[64]

Figure 20 outlines a sequential experimental design workflow which is used to identify optimal operating conditions and provide process knowledge and understanding.

- Scope: A fractional factorial design, requires prior process knowledge to choose factors important to the reaction. If unknown, then the scope can disregard factors keeping important ones based on relative error from reactions performed.

- Screen: Full factorial design, determines the main effects and variable interactions from all condition combinations and centre points. $R^2$ and $Q^2$, model validity and reproducibility are determined. Variable linearity or curvature can be estimated but not determined.

- Optimisation: Several different designs *e.g.* FCC, CCC (Figure 20) and CCO, applies quadratics models to fit a response surface. Measures the curvature of the process to find maxima and minima. Expands the original "screen" by considering more design space for better accuracy.

- Robustness: A fraction factorial design over four levels surrounding the found optimum. Ensures full validation of the optimum, testing its ability to be targeted repeatedly.
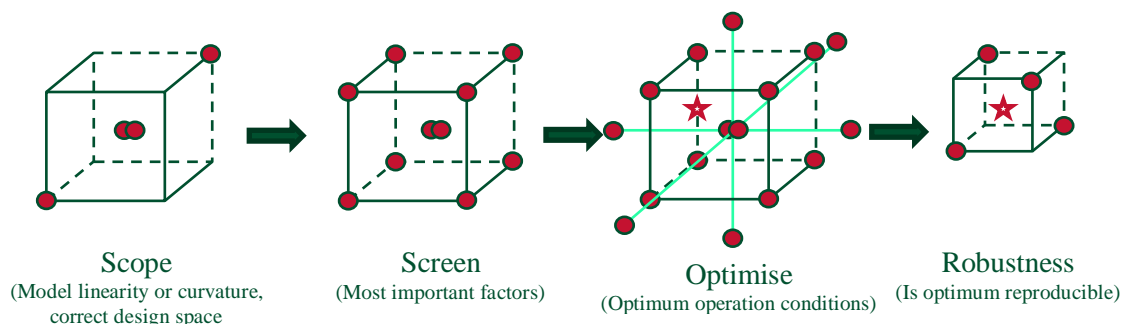


*Figure 20 Sequential workflow of a design of experiments process to find the optimum operating region for a chemical system. Below each cube is the level label which outlines the section of the DoE and what that sections main outcomes are.*

### 1.3.2.1.Factorial Designs

Calculating the number of experiments for factorial designs uses $m^n$, where n is the number of chosen variables and m is the level (number of conditions) for each parameter. A 3-parameter system each with 2 levels, will need $2^3 = 8$ experiments. Experimentation results form a unique system model relating each response to the experimental conditions *via* a mathematical expression (equation **I**). These results are then used to determine system coefficients to calculate values for $x_1$, $x_2$ and $x_3$ to solve the equation and find the value of "Y" the desired output value. The factorial design (FD) equation theoretically represents the entire design space.

$$Y = b^0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{1,2}x_1x_2 + b_{1,3}x_1x_3 + b_{2,3}x_2x_3 + b_{1,2,3}x_1x_2x_3 \quad (equ. \textbf{I})$$

Both model and equation show each factor in combination with one another from single parameter effects to dual parameter effects and finally triplet parameter effects, if any. A factor is significant to the process if its overall effect is greater than the standard deviation, otherwise experimental error dominates, and the factor can be disregarded.[65] It is important to include all independent variables as missing variables may cause unknown variation in the model causing undesired effects in the output optimum.

For large numbers of variables, a fractional factorial design method (FFD) can be used. This method is almost identical to FD but reduces the number of conditions combinations and therefore experiments needed for model generation. For example, if an FD approach

with 7 parameters, each having two conditions, was completed then there would be $2^7$ experiments or 128 in total.[66]

FFD uses the rule $m^{n-p}$ where $m$ is the number of values, $n$ is the actual number of reaction conditions and $p$ is the number of "ignored" variables e.g. 2 values for 4 conditions will give a total of 8 experiments ($2^5 = 32$, $2^{4-1} = 8$ runs).[66] Model coefficient aliasing or confounding can sometimes be problematic as less information is provided upon setup thus fewer factor level combinations are created combining more complex terms together. Complex terms therefore cannot be estimated independently in the design space. Simpler terms are estimated independently just like when using an FD method so can be assumed to be major contributors (Figure 21).[65]
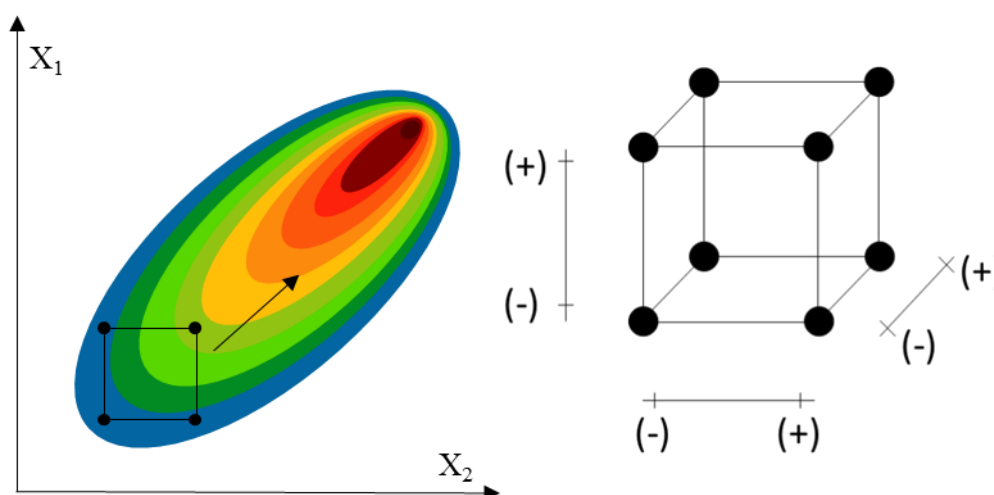


*Figure 21 Left: Factorial design approach showing the simultaneous variation of factors $x_1$ and $x_2$ towards the global maximum.[67] Right: A $2^3$ factorial design representation outlining each corner experimental point.[65]*

### 1.3.2.2. Central Composite Designs

Central composite designs are advanced versions of FDs, adopting several types to optimise parameters. Initial experiments follow those created in FDs, $2^n$ with corners and centre points, but include further experiments localised around the face of the cube (for 3 factors). Composite face centred cubic (CCF) methods include experiments directly on the face of the cube (Figure 22 - left). This design is helpful when the relationships of factors in chemical processes are known to be quadratics (equation **J**).

$$Y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 + b_{11} x^2{}_1 + b_{22} x^2{}_2 \ (equ. \ \textbf{J})$$

A central composite circumscribed (CCC) design has points surrounding the central point positioned at equal length to that of the corner points from the centrum. This creates a spherical design space and offers a much greater understanding of the experimental area due to its larger volume (Figure 22 - right). The CCC approach is advantageous over a CCF design as each axis has 5 different magnitudes of each variable. This allows the design to offer a greater understanding for cubic factor interactions as it can map out more detailed information on stronger graphical curvature.
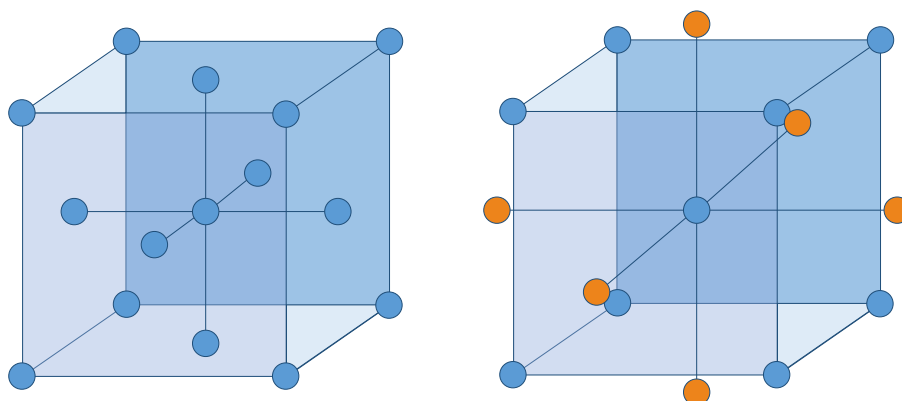
*Figure 22 Left: A Composite Face Centred cubic design representation outlining $2^n$ experiments from the corners of the cube and 2 x* n *experiments on the faces of the cube surrounding the central experimental point. Right: A Central Composite Circumscribed design representation outlining $2^n$ experiments from the corners of the cube and 2 x* n *experiments surrounding the central experimental point. The points coloured orange represent the "star" points, those outside the cube.*[65]

### 1.3.3. Kinetic Profiling

Physical organic approaches and kinetic parameters can be used to optimise chemical processes. System models can be created, and time-series data compiled from calculable kinetic parameters, such as activation energies, rate constants and equilibrium constants.[22] Availability of kinetic data allows this approach to be useful as multiple factors can be optimised simultaneously, however, prior knowledge must be sought to help understand the procedure before optimisation, ultimately increasing the time and cost of this method. Automation is possible as computer programs can be used to run experiments – through a feedback loop - and automatically analyse the reaction, although the program will not 'speak' directly to any on-line analytical techniques, which might be used. This method requires data calculations and spectra interpretation to be completed manually and input into the optimisation software before the reaction is fully

optimised. This approach is often considered an open loop approach as each data point needs to be calculated by hand before the change in the reaction system can be made. It is possible to automate this data response, but an initial experimental design needs to be inputted followed by best fit literature data and a kinetic model.[33]

To aid the design of a chemical reactor, several kinetic parameters need to be considered; heat, mass and component balances, heats of reaction and kinetics for each reaction required to build a complete model.[31,68] The variables mentioned each allow kinetic parameters, activation energies and rate coefficients to be estimated using well known theories; either transition state theory or kinetic theory. Wherever possible, use of direct experimental data is paramount to finding these parameters as accurately and efficiently as possible. The kinetic models gained not only mean the possibility of scale-up to another, larger flow reactor but also other types of reactors which are still often used in industry e.g. large-scale batch CSTR (continuous stirred tank reactor).[69]

The simplest way of finding kinetic data is to carry out reactions manually at different temperatures and residence times (similarly to OVAT) to span a large kinetic area and to maximise the data output for a model. Efficiency can be improved by automation of the reaction conditions (using computer-controlled systems to automatically run a set of predefined reactions) although it may need manual data mining, interpretation and model fitting. New techniques are allowing chemists to adopt algorithms to automate the reaction, and use data provided from real-time analysis, to interpret and model fit without the need of extracting data manually.[42]

Relevant and quantitative composition data from an analytical technique is essential to perform a kinetic study on a reaction but can often be challenging to acquire. As with other types of reaction optimisation there are a few different types of analysis which can be used to get the data needed for a profile; extractive (at-line), in-line and non-invasive. Each one offers a different angle for analysis of the chemical reaction and provides a unique data set, but any can be used to give the information required to outline a kinetic profile and other reaction parameters, such as activation energies and rate constants. Extractive techniques (GC, HPLC and on-line MS) can be placed at any point on the microreactor, utilizing a sample injector, with the potential of providing information before, during and after the reaction to maximise the data output. Reactor design can be a limiting factor when it comes to forming a kinetic profile or optimising a reaction as certain analyses can only be completed at specific points on the reactor. Some reactors

allow sampling to be made at multiple points; before the reaction, on the reactor or after the reactor, permitting a reaction profile to be formed where kinetic data can be derived.[42,69]

Kinetic data is widely used to understand a reaction mechanistically and gauge an understanding of specific parameters of a system which helps with understanding the formation of by-products and side-products. This is useful in lab-scale synthesis as it helps to expand current knowledge for new reactions and find details that may have previously been missed due to limited resources or technology. New techniques for determining this kinetic data allow parameters to be found and models fit more quickly and more efficiently than traditional means. Continuous flow allows complex reaction systems to be de-convoluted and optimised using less material and more forcing conditions for a larger modelling area increasing the accuracy and reducing the uncertainty of data. Kinetic data can reduce the extent of side reactions under specific conditions with a higher degree of operation certainty. It is this data that is essential when scaling up a chemical process and new techniques, or the improvement of old techniques, allow it to be sought much quicker and more accurately than in the past.[69–71]

### 1.3.4. Algorithm Optimisation Techniques

Optimising a process using computer algorithms may be more beneficial than kinetic means, namely algorithmic optimisation. Computer algorithms use a set of rules followed by calculations to optimise the desired output using user defined variable boundaries *e.g.* a specific temperature range. These rules allow flow reactions to be automated *i.e.* automatically generate experiments, carry out each reaction, analyse and perform calculations without much human involvement. Automation requires equipment to be connected to a computer, so the algorithm can make changes to each equipment's parameters *e.g.* flow rate or temperature. As previously mentioned, this occurs through generation of a feedback loop (**1.2**) and improves optimisation speed and efficiency because the algorithm drives the system away from a minimum output towards a global maximum.[34]

Algorithms can also be adapted to the need of the reaction if certain requirements are to be met. The algorithm will only progress if all constraints and requirements are satisfied *e.g.* maximisation of reaction yield but also limiting reagent equivalency if, for example, the reagent is expensive. There are various algorithms that can be used to self-optimise reactions but the most common are; Simplex,[72] Nelder-Mead simplex (NMSIM),[73] super

modified simplex (SMSIM)[74] and the stable noisy branch and fit algorithm (SNOBFIT).[75]

### *1.3.4.1. Simplex and Modified Simplex Algorithms*

Simplex algorithms first developed by Spendley *et al.*[72] are the most simplistic used for response surface exploration. A surface equal to the number of conditions, *n*, is created along with a polygon of $n + 1$ points.[73] Each response is then calculated and ranked from lowest to highest. The midpoint between the lowest and the next ranked points is calculated before a true reflection, away from the lowest ranking point, is performed; this provides the next conditions for further experiments and the procedure is repeated until an optimum is found (Figure 23).[72] Simplex is efficient for optimsation but does not explore much of the design space and may get stuck in local maxima if there are any in the experimental area.
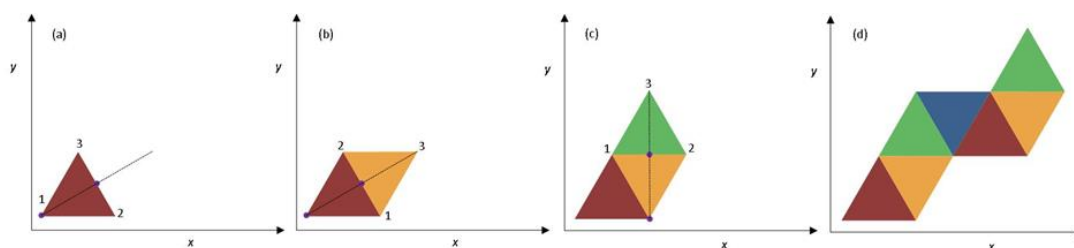


*Figure 23 Simplex progression of a 2-factor system – temperature and pressure. (a) Formation of the first simplex with ranked points - 1 (lowest), 2 (middle) & 3 (highest) for the greatest output - with the mid-point set for reflection. (b) Reflection of the first simplex and formation of the second simplex with ranked points. (c) & (d) reflection to form the third and fourth simplexes through the vertices of the previous ones with ranked points regarding total output. Each colour represents a newly generated simplex showing movements towards optimum conditions.*[76]

Adaption of the Simplex algorithm led to improvements by Nelder and Mead[73], introducing a variable sized simplex, offering many advantages. The generated polygon could change size as the algorithm progressed "uphill", it could contract and re-expand in different directions and could contract repeatedly around an optimum point (Figure 24).[77,78] These new modifications led to further improvements by Routh *et al.*[74] which became known as the super modified simplex (SMSIM).

The SMSIM creates a new simplex by determining the worst response of the polygon, the centre point of the shape and the reflected vertices, then the algorithm fits the data responses to either a second-order polynomial curve or a Gaussian curve. The new curve is also further extrapolated by a percentage factor of the polygon to give two different

curves; a concave up curve or concave down curve. Either curve is evaluated and allows the program to continue towards an optimum value. This procedure allows simplexes to have a much greater freedom of position on the experimental surface.[74,78]



Reflection    Expansion    Outside Contraction    Inside Contraction    Shrink
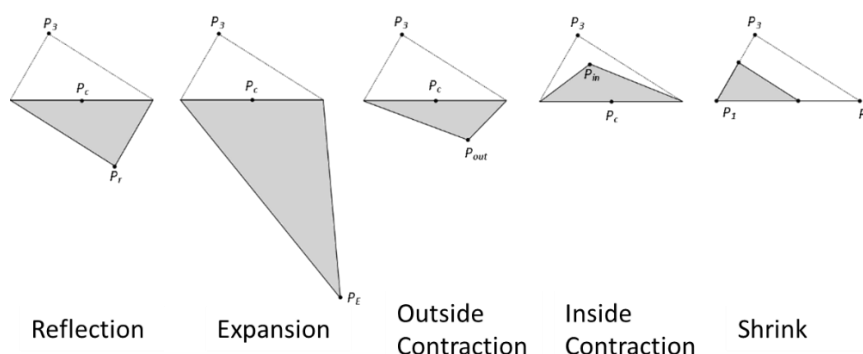
*Figure 24 Diagram showing the different modes the Nelder-Mead simplex (NMSIM) can adopt when searching for an optimum output. $P_3$ is the lowest ranking point in the simplex, P is the mid-point of the previous simplex.*[73]

### 1.3.4.2. Stable Noisy Branch and Fit

Another algorithm which can be used to self-optimise processes is: stable noisy optimisation by branch and fit (SNOBFIT). Coded in MATLAB by W. Huyer and A. Neumaier it is designed to optimise constrained, expensive functions by combining global and local searches by branch and fit methods.[75]

User-defined values are set in the program interface before the optimum search begins. The algorithm evaluates the number of points specified during each step before progressing onto the next and creating successive surveys of each box within the experimental grid. After this, the program fits polynomials to the calculated response which allows a value for the output function to be determined. This process continues until the global maximum has been found. Each cycle of experiments has *n* points associated with it to boost the algorithms robustness;[76] an optimum point, a secondary optimum, a point in close proximity to the optimum (to help improve the polynomial fit), an explanatory point and a random scatter point (Figure 25).

The SNOBFIT algorithm allows a large variety of different factors to be varied simultaneously based upon the previous output feedback of the on-line analytical technique, and as random points are scouted, gives a much greater image of the reaction field. Krishnadasan *et al.* used the SNOBFIT algorithm for the self-optimisation of fluorescent CdSe quantum dot nanoparticles using in-line 335 nm Nd:YAG laser and

charge-coupled device (CCD) spectrometer. They optimised their reaction for a function of wavelength, converting the detected wavelength into a dimensionless quantity – 'dissatisfaction coefficient' (DC), in order for the SNOBFIT program to find a minimum value for the output.[60]

This algorithm is very useful at gaining detailed information about an experimental surface as well as finding an optimum set of conditions *i.e.* it can be called both exploitive and explorative, also known as branch-without-bound.[79] Normally algorithms proceed like a branch and tree *i.e.* main pathway of the algorithm is the primary branch of the "tree" and any other exploratory steps are secondary branches. SNOBFIT can proceed with the primary branch without formally creating a secondary branch from the current point of interest and can randomly search other areas of the design space and see if the main branch is one containing the "true" path to optimality. However, SNOBFIT is still prone to several problems which can make it harder to deal with.

Each subsequent experiment is based on the ones previous meaning if there is one "bad" experiment it can throw the algorithm into the wrong area. This can typically be overcome by making the algorithm more explorative so it can find its way again but slows down the optimisation speed. The SNOBFIT algorithm also does not have a termination criteria built-in to the software meaning it will continue until the user stops the optimisation. This has the potential to waste material when optimising a chemical reaction. If an optimum is found sooner than expected or during an overnight run the program will continue.

SNOBFIT is useful for any function where constraints are placed upon them during optimisation *e.g.* boundary conditions or a maximum yield output. There is the possibility to restart an optimization using previously gained information if the conditions need to be expanded *i.e.* no loss of old information.
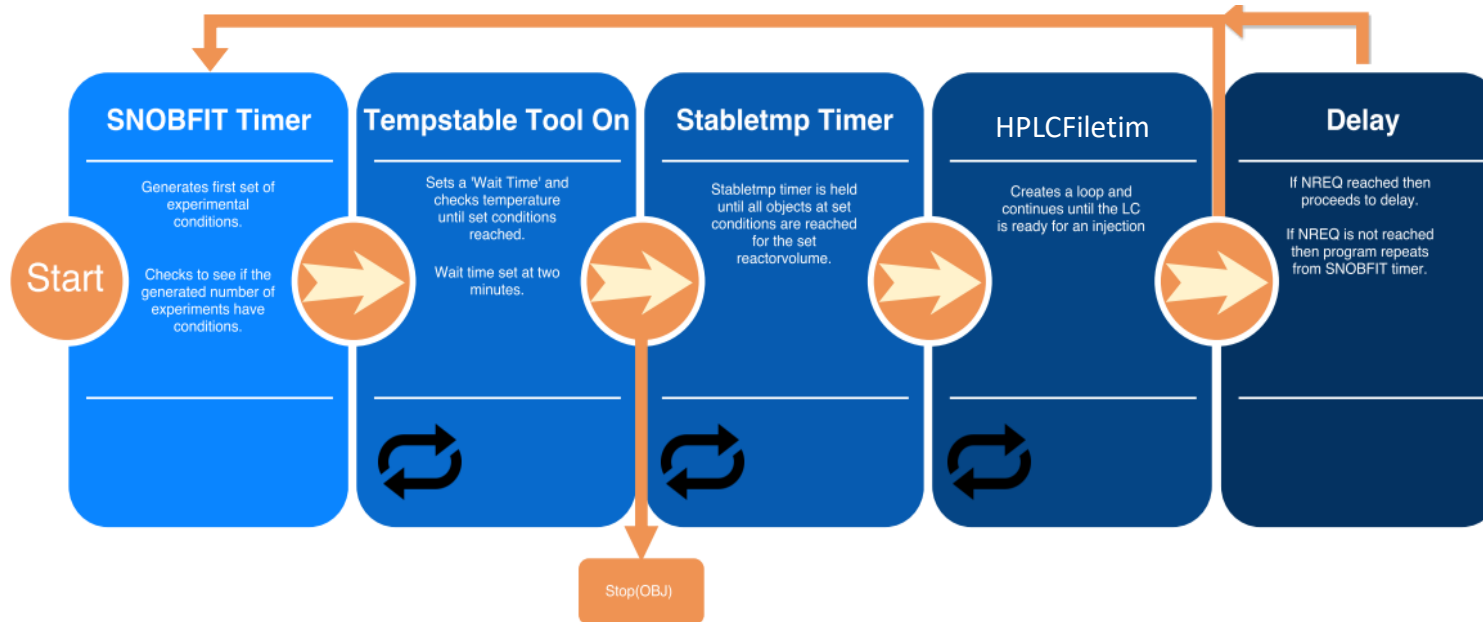
*Figure 25 Flow diagram outlining the timers involved in the progression of the SNOBFIT algorithm or adaptions. Once the SNOBFIT timer begins the program generates conditions and progresses to the next timer where call back functions take over and monitor the temperature. The next timer is similar but monitors the flow rates and checks the reactor has pumped for the correct volume/time before triggering the next timer to wait for the HPLC. The final timer occurs once the HPLC is triggered and causes the program to wait to HPLC method length before cycling back to the beginning.*

## 1.4. Industrial Uses and Scale-up of Continuous Flow Reactors

A typical batch reactor may offer difficulties in scaling up reactions from lab to plant volumes meaning certain steps and rules must be followed in order to ensure high levels of process safety for workers and all risks are reduced to a minimum.

During the development of a pharmaceutical compound it can take at least ten years for a new treatment to become available on the market from initial discovery (Figure 26).[80]

- Initial discovery of target (natural products) generates a large catalogue of molecules. Initially 5 to 10 mg are needed for primary tests followed by 500 mg to 1 g for activity, metabolism, pharmacokinetic and selectivity studies before lead targets need a toxicological profile (1 – 2 g). Detailed toxicology studies must be carried out and need between 5 and 30 g for *in vivo* tests.

- Early process optimisation of a lead compound requires further scale-up (30-500 g) for further toxicity and side product studies,

- Process scale-up for pilot and commercial production requires a far larger scale for use in clinical trial ready for approval and long term commercial synthesis (kilograms to tonnes).[27]



*Figure 26 An outline of industrial synthesis scale. Vial 1 shows a mass of 5-10 mg, vials 2 & 3 show 500 mg – 1 g, vial 4 1-2 g, vial 5 between 5 and 30 g, vials 6 & 7 show 1 kg and 5 kg respectively. Vials 1 – 4 outline the discovery section of a pharmaceutical's timeline whilst 5-7 larger scale synthesis working towards a commercial process.*

The general synthesis of a compound changes for each scale. A discovery synthesis typically makes use of highly reactive and hazardous reagents *e.g.* strong bases, non-environmentally friendly solvents *e.g.* dichloromethane and reactants that may be costly to buy in bulk. When creating a process synthesis, it is essential to use cheap, readily available, stable and non-hazardous reactants, reagents and solvents as well as taking all the relevant precautions for the process itself.[81]

Currently, industry favours batch processes because of equipment availability (a continuous flow reactor needs a large investment) meaning they are limited in their process variability. Continuous processing can overcome many of the associated disadvantages with batch processes in the pharmaceutical industry, providing advantages from the start during the medicinal chemistry stage, where a flow route can be used with minimal optimisation. This flow route can then be used in later stages with more focus on the optimisation of the product yields.

As safety is such a concern in industry, lab research can be useful to create safe processes prior to scale-up. Lab scale processes typically use flask reactors (100 - 500 mL), large scale 25-50 L reactors and commercial 1000 - 6000 L tank reactors.[82] All reagents are added *via* a manual inlet or from a reservoir and stirred for the reaction duration. Heating in batch involves a heated jacket surrounding the tank or a heated coil within the reactor. A flow process consists of all reagents in separate reservoirs connected to pumps *via* convergent tubes that enter the reactor. A large-scale flow device is heated as with smaller-scale devices; in a heated oven or heating the plates directly. Each microreactor can be kept at a different temperature depending on the reaction.[83]

Automated flow reactors allow new ways to quench reactions chemically or thermally improving reaction selectivity leading to integrated work-up procedures directly after the reaction improving, simplifying and reducing costs of work-up stages.[84] Often work-up is not needed if reaction stages can be linked together. Direct transfer from one stage to the next also simplifies the process route. This can be completed using an in-line membrane separator to resolve biphasic reactions or by solvent reduction and subsequent reaction.

Borukhova *et al.*[85] used a membrane separator to separate parallel reactions during the formation of an API, cinnarizine and three other APIs. The reaction used benzhydrol and HCl, before being quenched with NaOH, separated, basified and reacted with piperazine and reacted to form cinnarizine (Figure 27). The membrane separators worked to remove

the aqueous phase and side-products before the sequential reaction could occur. Use of the separator made the reaction stages easier to handle and telescope subsequent reactions. In-line separation also allowed the collection of pure products and their immediate consumption to gain good yields of 82% and above.
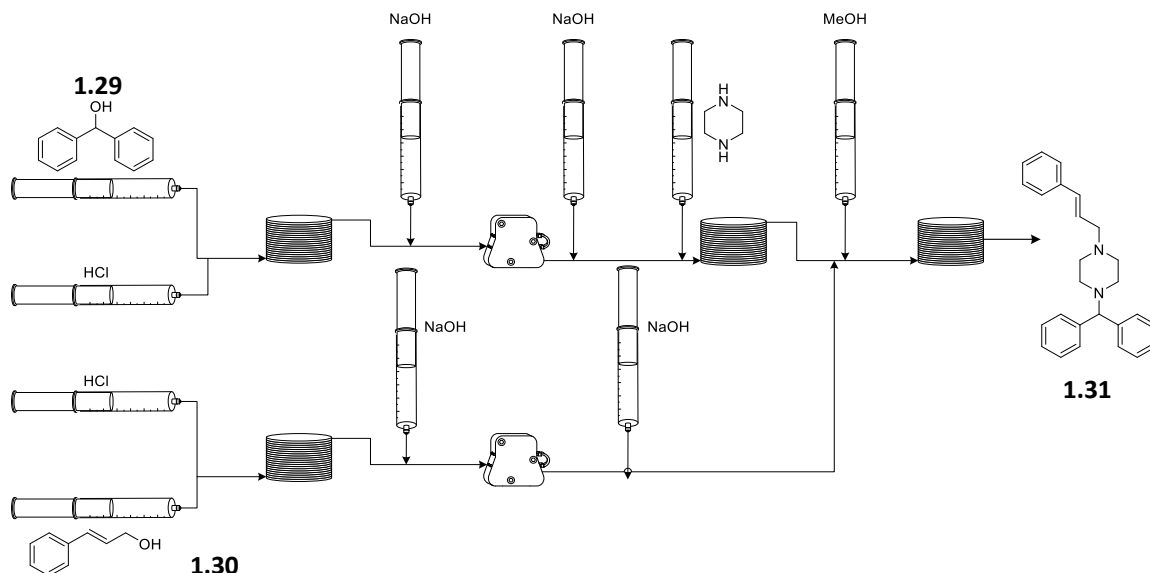


*Figure 27 Continuous flow synthesis of cinnarizine and three other APIs where four transformations convert the starting alcohol to their respective APIs.*[85]

Industrial processes are built up on several different factors that help the company maximise the processes efficiency in terms of cost and productivity. It has been reported that a full production of an active pharmaceutical ingredient (API) typically contains an average of eight steps with each intermediate becoming the key reactant of the next step.[86] For a multistep synthesis, fine tuning the process or carrying it out in flow can help improve overall reaction yields. Often, large scale processes are more dilute particularly later on in the process. More dilute solutions can help to reduce normally harsh reaction conditions and can favour overall yield and selectivity when compared to the high concentration reactions.[86] Continuous flow benefits reactions with dangerous or harmful intermediates by reducing accumulation. Small channel sizes reduce the amounts of dangerous materials forming a critical build-up limiting their harm.

A significant part of large-scale industrial developments is based upon the total cost of the process which is highly dependent on labour. Multistep batch syntheses take a much higher work force to maintain and perform each step in the process. A continuous process is much more advantageous as many steps can be cut out and performed automatically meaning an overall reduction to labour costs. Norbert Kockmann outlines a scale-up of

a lithium exchange reaction followed by a coupling reaction from the laboratory to pilot plant.[82] The scale-up to the pilot plant is initially operated by technicians and shift workers. Progression sees the plant solely operated by the shift workers until the duration of the experiments had elapsed. Subsequently, work was continued to see how the main sections of the reaction would proceed when they were carried out in microreactors in the laboratory (Figure 28).[82]
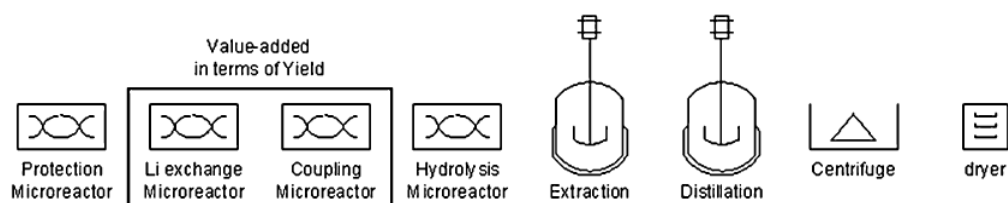


*Figure 28 Diagram to show the first four stages of a lithiation reaction carried out in microreactors, followed by product extraction, distillation, centrifugation and drying in batch reactors.[82]*

Time is often a limiting factor in the fine chemical and pharmaceutical industries, particularly for process optimisation of a compound; a flow process enables maximum optimisation in minimal time. Continuous processes are highly automatable using process analytical techniques and are simple methods to finding the best yield for the desired compound efficiently. They can also provide large volumes of useful information including details about mechanisms, intermediates and by-products improving efficiency, data richness and safety. High levels of control of reaction parameters allows optimum conditions for reagent flow rates, reactor temperature and product yield can be found within 24 h using minimal reagents and be implemented straight away without direct human interference.[87] Optimum conditions are generally transferable so quick scalability of small scale processes – 10s of grams – to large scale – 100s of grams – can happen on a case by case basis.[27]

There are several ways a continuous flow process can be scaled to increase the output of desired product. First, the total internal volume of the reactor can be increased to allow a greater amount of reactant to be fed through for reaction. This allows scale-up to be successful only if the dimensions of the reactor are kept constant so mass and heat transfer are almost identical, and as long as the residence time is kept constant at the found optimum. With the increase in reactor volume the total flow rate is increased maintaining residence time to improve reaction productivity also known as space-time-yield (STY). STY is a measure of reactor productivity where space-time is

the time required to process a single reactor volume based on the entrance conditions of the reactor.[88]

The second method named "Numbering Up" uses multiple identical reactors to run the reaction on the same scale, in parallel, using found optimum conditions.[89] This is beneficial as the reactor has already been used and modelled but also increases productivity of the reaction. The third final method is to scale-out[90,91] the reaction by running the process for longer in the same reactor again improving productivity. Improving reaction productivity reduces process costs but generates more product to be sold. However, capital investments are still necessary for a fleet of identical flow reactors (numbering up) or new larger flow reactor (scaling-up) but both former methods have a higher STY than reaction scale-out. Scaling-up may be the most attractive as it only involves a single, larger flow reactor, which although will still need to be characterised, will be easier to run long-term than many smaller reactors and will use less tubing and other fittings.

Any scale-up method (or chemical reactor) allows the reaction to be monitored by process analytical technology to ensure reaction progression without the trouble of sample workup, purification and off-line analysis.

## 1.5. Summary

- Industrial development of reactions has a huge interest in reducing the material use, and time and costs of processes where development is often limited by methodological capabilities.

- Typical chemical process development is carried out using batch processes.

- Both industry and academia have begun to adopt continuous flow as a processing technique to help push process intensification, improve safety, sustainability and product quality and reduce energy usage, costs and material consumption.

- Continuous flow devices are highly modular and can be designed for any reaction. They can be cheap, improve reaction speed and help with process intensification.

- Most reactions in academia or industry are optimised primarily using one-variable-at-a-time (OVAT) or design of experiments (DoE).

- New optimisation techniques using computer based algorithms are designed to be used alongside continuous flow processes for rapid optimisation.

- These optimisation algorithms can be specifically designed for different processes and have the potential to optimise multiple conditions simultaneously.

- Each experiment carried out is designed to be as information dense as possible to limit the number carried out whilst gaining as much information about the reaction as possible.

- Process analytical techniques (PATs) can provide data on the chemical reaction taking place without the need for off-line sampling.

- Certain PATs can offer real-time reaction data of a continuous flow process allowing rapid reaction optimisation as well as large data sets.

## 1.6. Key Aims and Objectives

The aim of the work presented in this thesis is to investigate if on-line mass spectrometry can be used to monitor continuous flow chemical reactions, thus reducing analysis time and providing valuable qualitative and quantitative information. This will be completed using the continuous flow platform housed at the University of Leeds along with an Advion CMS and Agilent HPLC using MATLAB software and the Snobfit optimisation algorithm.

Continuous flow technology allows processes to be intensified and automated quicker and easier. Process improvement *via* waste reduction from traditional means and optimisation will be completed using the modular system. Optimisation approaches will be used (fully automated statistical DoE and automated self-optimisation) and will vary from study to study. The optimisation algorithm used, Snobfit, is both an exploitive and exploratory algorithm to help improve the reliability of reaction optimisation.

On-line MS was chosen as it has improvements of current on-line analytical methods, short acquisition times, low sensitivity therefore small sample volumes used, chemical identity and the potential to provide concentration and quantitative data. Comparisons will be made to HPLC (a known quantitative analytical technique) for verification that MS can also be quantitative.

Together this technique will be used on a variety of reactions to examine its variability and usefulness for organic transformations and inorganic catalysis to develop the continuous-flow-on-line-MS-self-optimisation platform. This work will also demonstrate how the gathered data is useful to the pharmaceutical industry, where it can be applied to compounds of interest, optimised and scaled-up.

Chapter 2

**Deprotection of an N-Boc Protected Amine Active Pharmaceutical Ingredient:
Monitoring and Optimisation Using On-line Mass Spectrometry**

## 2. Deprotection of a N-Boc Protected Amine in Continuous Flow Monitored by On-line Mass Spectrometry

Chapter 2 will outline work conducted using an API from AstraZeneca, named AZD5634. This compound is currently in early stage development where the process is being fine-tuned for large scale production before clinical trials. The work presented outlines how the project came to be, problems encountered during development and results from both work completed at the University of Leeds and at the AstraZeneca site itself. A brief introduction will discuss the general concepts of amine protection using the N-Boc group before moving onto traditional approaches of removal after manipulation of other functional groups in the molecule. Subsequently, this will lead onto newer methods of N-Boc removal and how they may be used for large scale synthesis. The full molecular structure has been included with the main transformation in the molecule only involving the terminal amine group and the attached N-Boc protection group.

### 2.1 Introduction

One of the most widely used groups for amine protection is the *tert*-butyloxycarbonyl (Boc) group where this groups use is paramount as a large proportion of organic reactions contain amine functional groups within their structure. Current batch protection reactions are quick and simple and do not form harmful by-products; carbon dioxide, *tert*-butanol and the desired protected amine (Scheme 13) but $CO_2$ is not ideal being a greenhouse gas.
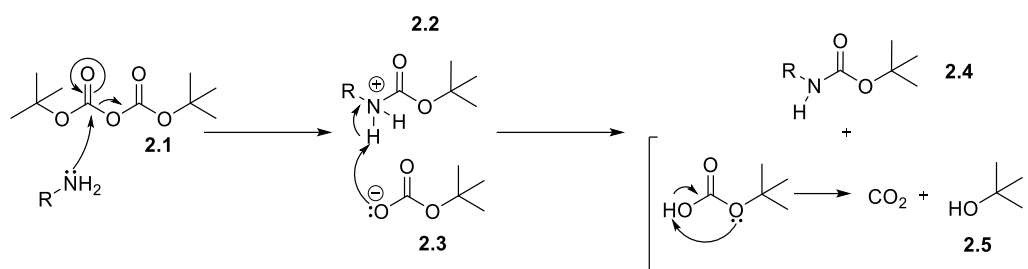
*Scheme 13  A general reaction procedure for the protection of an amine using Boc₂O. The amine attacks the carbonyl group and causes the C-O single bond between the central oxygen and carbonyl to break. The excluded group then attacks one proton on the amine to negate its own negative charge and the positive charge on the amine. By-products of $CO_2$ and tert-butanol are formed.*

The N-Boc protecting group (or any protecting group) is a temporary unit used to ensure the protected group is stable under the conditions needed for the chemical transformation allowing other groups to be reacted. Removal often follows to form either the free amine or ammonium salt as the Boc group is, usually, not needed in final products of APIs. Deprotection is carried out in batch favouring long reaction times and strong mineral acids ($H_2SO_4$,[92] $H_3PO_3$,[93] $HCl$[94] or TFA[95]) in $CH_2Cl_2$ to help the reaction to completion (Scheme 14).
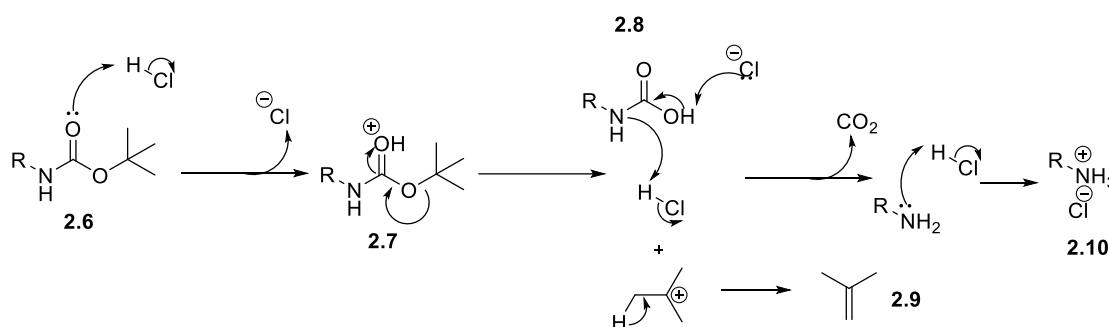


*Scheme 14 The deprotection of a tert-butyl carbamate using HCl producing the chloride salt of desired product and $CO_2$ and tert-butyl cation as by-products.*

Scheme 14 shows the deprotection mechanism of an N-Boc protected amine using acid. The acid attacks the carbonyl oxygen forming a double bonded OH group with a positive charge. The remainder of the N-Boc group breaks off to balance the charge on the OH group forming a *tert*-carbocation which releases a proton to form isobutene gas. The chloride ion (if HCl was used) attacks the OH of the carbamate and causes $CO_2$ to be released from the molecule and the free amine to be produced. As the medium is acidic the amine will become protonated in solution with the formation of amide salt.

### 2.1.1. Methods Used for the Removal of the N-Boc Group

It is possible to remove the N-Boc group using high temperature water in batch to form the free amine without any undesired salts. Water eliminates the need for unfavourable solvents, and using higher temperatures, acts as a dual acid base catalyst.[96] The main products being the desired amine, $CO_2$, and *tert*-butanol. The large volume of water present causes this carbocation to react with the surrounding water to create *tert*-butanol instead. Thermolytic removal of the N-Boc group was shown the be effective in 1985 by Rawal and Cava who, serendipitously, discovered it whilst acquiring melting points of protected compounds they had synthesised.[97]

Thermolytic removal was investigated using a continuous flow process adapted for a telescoped multistage synthesis. This process made use of HPLC pumps, high temperature reactor (Phoenix flow reactor by ThalesNano) and a sample switching device enabling sampling of the reaction stream.[98] A coupling reaction between 4-fluorobenzoyl chloride and 1-Boc-6-aminoindoline followed by the deprotection at 300 °C and 100 bar allowing a second reaction to be telescoped with 4-chlorophenyl chloroformate was carried out (Scheme 15). The deprotection was performed in MeCN only without the need for extra reagents giving >99% conversion removing the need for large excesses of harmful reagents.
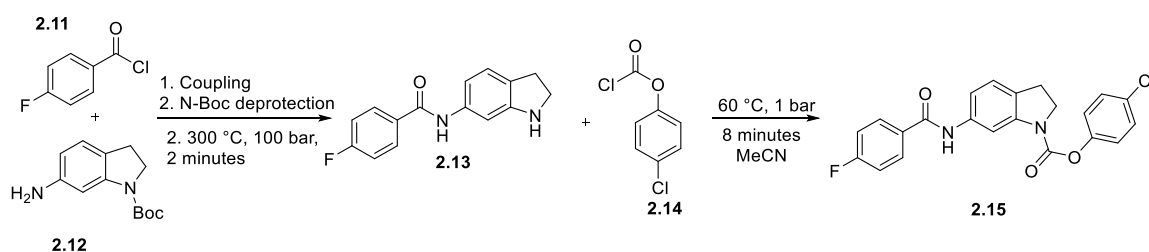


*Scheme 15 Reaction conditions for the high temperature thermolytic N-Boc removal carried out in a multistage continuous flow process at 300 °C, 100 bar using the Phoenix reactor system.*[98]

N-Boc deprotection may be carried out using fluorinated alcohols in a microwave reactor, which expedites the reaction compared to batch. More specifically; 2,2,2-trifluoroethanol (TFE) and hexafluoroisopropanol (HFIP) were used as solvents during deprotection of N-Boc protected amines as well as being used in the selective N-Boc deprotection of a doubly Boc-protected species. In this case the first protecting group was removed in TFE at 150 °C over the course of 2 hours (80%) whilst the second group was removed in HFIP at 150 °C over a further 2 hours (81%) (Scheme 16).[99]
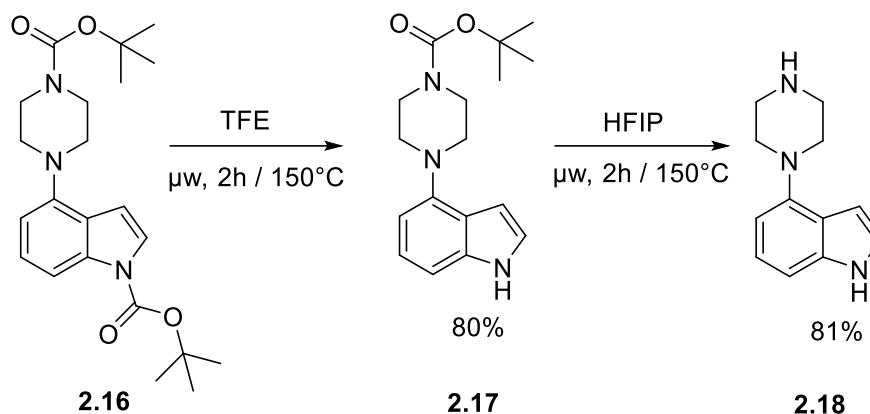
*Scheme 16 Selective N-Boc removal using TFE and HFIP under microwave conditions providing the desired product in 81% yield.*[99]

Removal can also be achieved by use of halogen containing reagents acting catalytically or by acting as a nucleophile. Tetra-*n*-butyl ammonium fluoride is a highly versatile compound commonly used for many fluoride-assisted reactions as a base in certain base-catalysed reactions. It has shown to work effectively in THF under mild conditions for many different carbamates, including the N-Boc group (Scheme 17).[100]
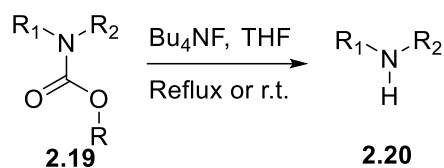


*Scheme 17 The use of TBAF to remove the N-Boc protecting group from amines under mild conditions.*[100]

Solvent free N-Boc deprotection was developed using iodine as catalyst to deprotect a diverse array of N-Boc protected amines. Isolated yields were as high as 98% after a 30-minute reaction for this process with no solvent. Comparatively, reactions carried out in different solvents (DCM, MeCN, benzene and toluene) varied from 5 h to 12 h with the more polar solvents having a shorter reaction time and greater yield. Mechanochemical reactions involved grinding reagent together with solid iodine for the duration of the reaction time before being washed and the deprotected product extracted (Scheme 18).[101] This mechanochemical method produced the desired product in 98% yield within 30 minutes. Activation of the carbonyl group by iodine led to cleavage of several bonds affording $CO_2$, isobutene and the desired amine.

**2.21**                                    **2.22**
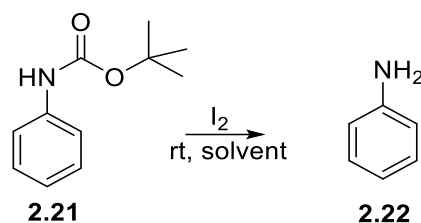
*Scheme 18 A solvent free N-Boc removal method using elemental iodine as catalyst.*[101]

Work by Ravindranath outlined the use of solid supported $NaHSO_4$ and HY-Zeolite to act as heterogeneous catalysts for the selective N-Boc deprotection of amines. A selection of different N-Boc protected amines were refluxed with either HY-zeolite or $NaHSO_4 \cdot SiO_2$ in dichloromethane until the reaction had proceeded to its endpoint. The shortest reaction time, 1.5 hours, provided 92% isolated yield for the deprotection of N-Boc aniline (Scheme 19).[102]



**2.23**                                    **2.24**
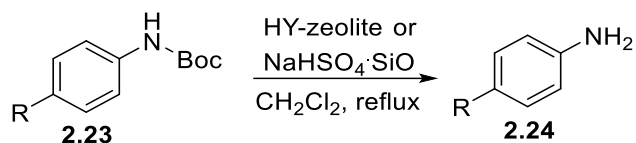
*Scheme 19 N-Boc deprotection of an amine using HY-zeolite or solid supported sodium hydrogen sulphate.*[102]

Removal of the N-Boc protecting group has been shown to be effective and selective whilst using $Fe^{III}$ salts. Many different protected amines were compiled and reacted on small scale using a batch method with stoichiometric amounts of iron salts. Again, reaction times were greatly reduced from traditional acid removal techniques and compared to other metal[III] salts with the shortest reaction time being 15 minutes for $FeCl_3$ giving >99% yield and 24 hours for $FeBr_2$ (Scheme 20).[103]



**2.25**                                    **2.26**
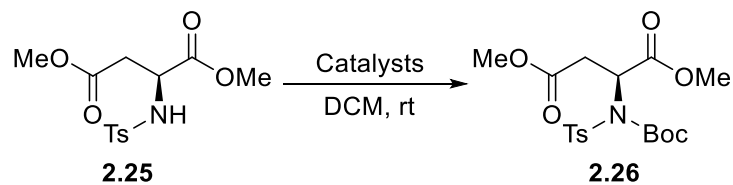
*Scheme 20 Amine formation via cleavage of the N-Boc protecting group using $Fe^{III}$ salts as catalyst.*[103]

*Table 5 Comparison table between 6 different methods of deprotection for the N-Boc group from thermolytic and use of fluorinated reagents to catalytic iodine, zeolite and iron (III) salts.*

| Deprotection Method | Reaction time | Yield/ % | Conversion/ % | Benefits | Drawbacks |
|---|---|---|---|---|---|
| **Thermolytic[98]** | 2 mins | - | >99 | • Short reaction times<br>• no harmful reagents<br>• reaction telescope possible<br>• Easier reaction control<br>• Can be optimised | • Requires specialist equipment<br>• Hazardous temperature and pressure |
| **Microwave Conditions Using HFIP & TFE[99]** | 1st: 2 hours<br>2nd: 2 hours | - | 1st: 80<br>2nd: 81 | • Thermolytic microwave conditions<br>• Selective deprotection | • Uses potentially carcinogenic reagent<br>• Uses highly corrosive reagent |
| **Fluorinated Reagents, Bu$_4$NF[100]** | 1-36 hours | 80-98 | - | • Mild conditions<br>• Selective deprotection | • Potentially carcinogenic solvent<br>• Corrosive materials |
| **Mechanochemical Iodine[101]** | 30 mins | 98 | - | • No solvent<br>• Mechanochemical | • Iodine is a very harmful reagent |
| **Zeolite Catalysis[102]** | 1.5 hours | 92 | - | • Selective N-Boc removal of two species | • DCM is a non-ideal solvent<br>• Specialised zeolite |
| **Iron Salt Catalysis[103]** | 1st: 15 minutes<br>2nd: 25 hours | >99 | - | • Mild conditions<br>• Cheap catalyst | • DCM is a non-ideal solvent<br>• Very specific Fe salt<br>• Large loading |

A summary comparison of each of these methods of N-Boc deprotection is shown in Table 5. Seemingly the best is the thermolytic method, which may also be used for scale-up. The other methods summarised have comparatively long reaction times or use unideal reagents.[100,102,103] Problems may occur with the thermolytic method when using thermally unstable molecules. Using continuous flow enables system pressurisation allowing significantly higher temperatures and the reduction of residence times whilst maintaining yield. Use of water as reagent negates the use of base for acid quench and thus reduces solvent and reagent use. The lack of acid creates the desired amine product in its free amine form, dissolved in solution. Therefore, reaction telescoping can occur without crystallisation, separation and purification.

## 2.2. Summary

- N-Boc protection reactions are quick, easy and can be important when performing long, multistep processes for reactions to occur throughout the molecule.
- Removal of the N-Boc group, using traditional means, often have long reaction times and use large amounts of acid and base. Non-environmentally friendly solvents are typically used which are troublesome and costly to dispose of.
- There are alternative ways in which the N-Boc group may be removed to form the desired amine.
- The most industry applicable: thermolytic removal or high temperature water deprotection.
- Both reduce reaction time, require no quench, crystallisation, separation or purification and use more environmentally friendly solvents.

## 2.3. Aims

This work aims to develop a continuous flow process effective at removing the N-Boc group from an API. Key points for the process must include; minimal solvent, use of green solvents, alternatives to acidic reagents and a timescale of 30 minutes or less.

Large scale, API syntheses would benefit from use of green solvents as costs relating to disposal of non-environmentally friendly solvents are significantly reduced. Green chemistry can be assured by following the "12 principles of green chemistry" several of which can be accounted for throughout this research project:

- Prevent waste: Telescoping the process means solvent is reused and not wasted.

- Less hazardous synthesis: Use of water and minimal green organic solvent negates the need to use substances which are toxic or hazardous to the environment.

- Safer Solvents and Auxiliaries: No separation means no separation agent therefore improving process "greenness" and reducing energy consumption and toxicity.

- Design for energy efficiency: Smaller channel diameters (for flow reactors) means more efficient heating of the process, reducing energy requirements. Optimisation algorithms are much more efficient and reduce the working time of the reactor.

- Real-time monitoring and pollution prevention: On-line analysis reduces exposure to unnecessary chemicals as well as signs the process needs to be changed before problems arise.

- Safer chemistry for accident prevention: Flow chemistry reduces working volume of reaction, minimising exposure risk to chemicals.

Furthermore, a robust and versatile process to be widely used in industry, allowing scale-up, whilst maintaining high conversions will be created. Mass spectrometry is ideal as the starting material does not contain a chromophore but does contain many ionisable groups and starting material/product mass difference is distant enough to not invoke complications doing optimisation.

This work preceded and, later, enabled work to be completed at AstraZeneca to scale-up the reaction to the pilot plant for synthesis of large quantities to be used for clinical trials. Due to pharmaceutical time-scales, it was necessary for this reaction to have a maximum residence time of 30-minutes. Any higher and the process would not produce the amount of material required within a reasonable period.[104,105]

### 2.4. The N-Boc Deprotection of an Active Pharmaceutical Ingredient

#### 2.4.1. Process Requirements

For a feasible and improved process, on lab scale and large scale, it is essential to create a robust, green, minimally hazardous method. The alternate proposed route uses minimal green solvent (MeOH, IPA or MeTHF) and water at elevated temperature and pressure (Scheme 21) but below that used by Bogdan[98]. The molecule of interest is an active pharmaceutical ingredient from AstraZeneca, AZD5634. Elevated temperatures, high equivalents of $H_2O$ and the $CO_2$ evolved create a more acidic environment for deprotection essentially mimicking the acid conditions used in the traditional route. Using this alternative route would be highly

beneficial industry as the process is much greener, reduces energy consumption and allows subsequent stages to be telescoped.
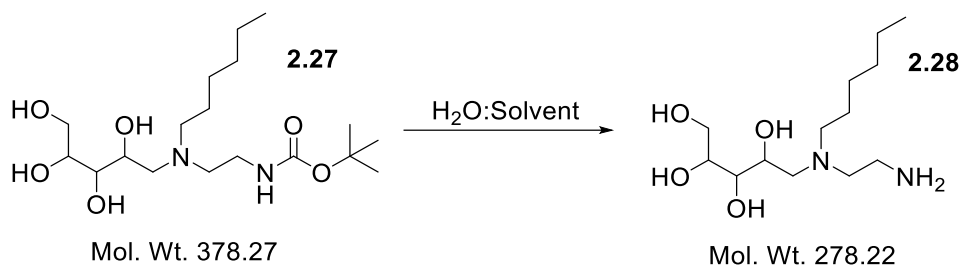


*Scheme 21 N-Boc deprotection of AZD5634 using H₂O as a catalyst. Solvent = MeOH, IPA or MeTHF.*

### 2.4.2. Initial Batch Reaction Scoping

An initial study, using a microwave reactor (CEM Discover® SP), was completed to determine feasible compound and process conditions as well as characterise the starting material and desired product. MS was employed to understand each species and how the reaction responded to different variables of interest for optimisation; reaction time, temperature and solvent composition (Table 6). The solvent needed to be able to easily dissolve the starting material, up to 10% w/v, as this concentration is typical of a large scale or commercial process. MeOH, IPA and MeTHF were explored but the latter only during solubility studies.

AZD5634 initially appears to be a glass like solid but is an incredibly viscous, colourless to pale-yellow liquid with limited solubility of water hence the use of minimal organic solvent. Reaction conversion was lowest when all conditions were at their lowest boundary *i.e.* a temperature of 130 °C, residence time of 5 minutes and a water content of 20%. As the reaction conditions were changed – higher water content (90%) and higher temperature (150 °C) – the conversion also increased, to almost 65%, suggesting that at least one of these factors was important. The final two entries in the table, 6 & 7, were from a sample which was initially reacted for 5 minutes, analysed, and reacted again for a further 5 minutes (entry 6) and re-analysed (entry 7). This provides evidence that the reaction time plays a role in increasing the conversion of the starting material. This allowed a preliminary flow process to be constructed to verify the finding of the small-scale batch findings. Higher conversions were observed when using longer reaction times, higher temperatures and a greater volumes of water (40 minutes, 140 °C, 90% H₂O, 10% MeOH).

All three solvents (MeOH, IPA & MeTHF) similarly showed optimal conversions at higher temperatures and longer reaction times when 1% w/v, 150 °C for 20 minutes were used. A greater focus could be made on progressing the reaction to conditions more likely to be used

on a larger scale process, in particular, higher concentration. However, even though MeTHF also performed highly it is not completely miscible with $H_2O$, creating an unideal biphasic system.

*Table 6 Summary of microwave experiments completed as preliminary work; including solvent system and temperature conditions for each sample. The reaction time for each was set at 5 minutes.*

| Sample | Solvent System (MeOH:H$_2$O) | Temperature/ °C | Conversion/ % |
|--------|------------------------------|-----------------|---------------|
| 1 | 80:20 | 130 | 0.0 |
| 2 | 80:20, 0.1 mL HCl | 130 | 49.4 |
| 3 | 50:50 | 130 | 12.5 |
| 4 | 50:50 | 140 | 6.4 |
| 5 | 50:50, 10 mg Amberlyst | 140 | 23.0 |
| 6 | 10:90 | 150 | 42.0 |
| 7 | 10:90 | 150 | 64.9 |

The range for residence times was set between 10 and 40 minutes ensuring the reaction had ample time to proceed to completion at 140 °C. This led to preparation of an experimental design, summarised in Table 7. The program MODDE was used to determine the experimental conditions before being put into an edited version of the SNOBFIT algorithm. As it was an initial DoE, the output was decided to be the normalised ratio between starting material, 379 Da ([**4**+H]$^+$) and product 279 Da ([**5**+H]$^+$) giving a good indication of reaction conversion. Results for the experimental design proved to reinforce those from initial batch and flow reactions.

### 2.4.3. Initial Experimental Design

Modde, a statistical design of experiments computer program, provides set reactions to be carried out based on starting variables, along with their ranges *e.g.* temperature, 100 °C to 140 °C before defining what outputs are to be modelled *e.g.* conversion. Once defined, the type of DoE can then be selected, in this case "screen", where a full factorial design was used for 2-factors; residence time and temperature (Table 7). Addition of water was kept constant at 90% of the total volume given previous reactions suggested higher conversions. DoE simplicity was ensured by removing "water %" as a variable due to a limited number of pumps. The reaction system was kept under 100 psi of pressure to ensure all solvents remained liquid.

*Table 7 Experimental design conditions and run order as given by the program Modde for the N-Boc deprotection of AZD5634.*

| Run Order | Temperature/ °C | Residence Time/ min | Flow Rate/ mL min$^{-1}$ |
|:---:|:---:|:---:|:---:|
| 1 | 140 | 40 | 0.075 |
| 2 | 140 | 10 | 0.3 |
| 3 | 140 | 10 | 0.3 |
| 4 | 130 | 25 | 0.12 |
| 5 | 120 | 40 | 0.075 |
| 6 | 140 | 40 | 0.075 |
| 7 | 120 | 25 | 0.12 |
| 8 | 130 | 25 | 0.12 |
| 9 | 130 | 25 | 0.12 |
| 10 | 130 | 40 | 0.075 |
| 11 | 130 | 10 | 0.3 |
| 12 | 120 | 10 | 0.3 |

DoE provides statistical data showing data reproducibility (Figure 29), regression ($R^2 = 0.931$), model validity ($Q^2 = 0.825$)[i] robustness of the process, from all reactions carried out. This data is used to create a contour plot visualising each factor in the experimental space with minima/maxima and optima outlined (Figure 30). The initial DoE was completed using MeOH, as this solvent could dissolve the starting material most efficiently, at 1% w/v as not to over saturate the system during the initial stages and minimise starting material usage due to limited supply.

---

[i] The $R^2$ term denotes the fit of the model to each point which has a maximum value of 1 and should be as close to this as possible. The $Q^2$ term outlines the ability of the model to predict a point from the given data set and should be as close to the $R^2$ value as possible.
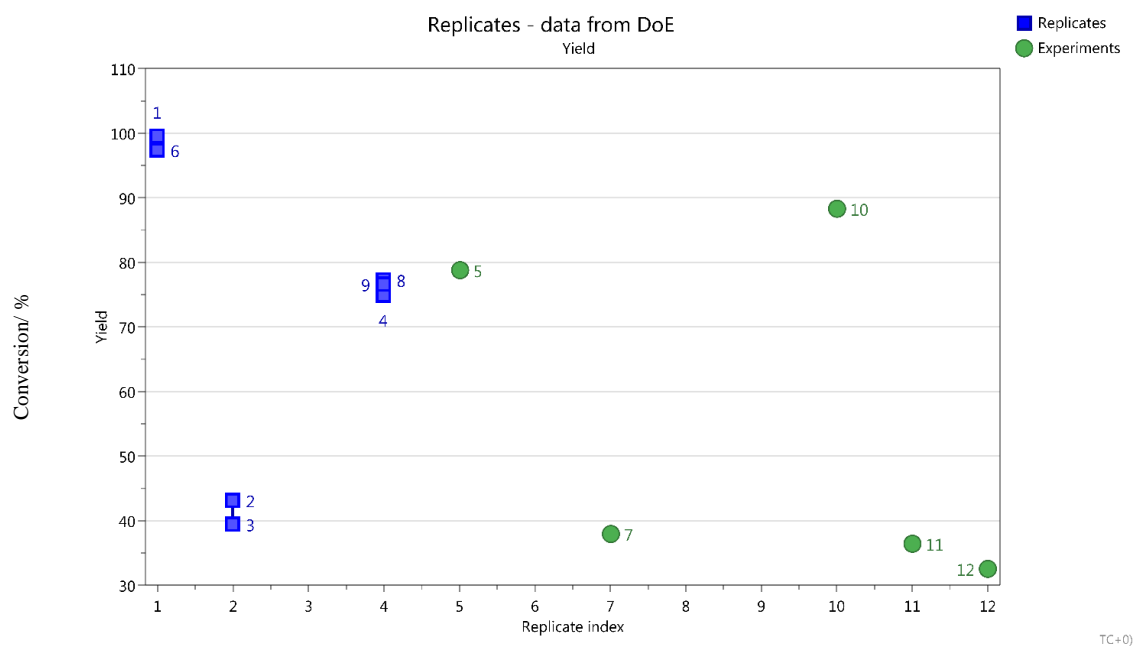
Figure 29 Replicates plot showing the variation from replication in the N-Boc deprotection of AZD5634. The blue square points, ideally, should be completely overlapped and all points should be spread across the experimental area.
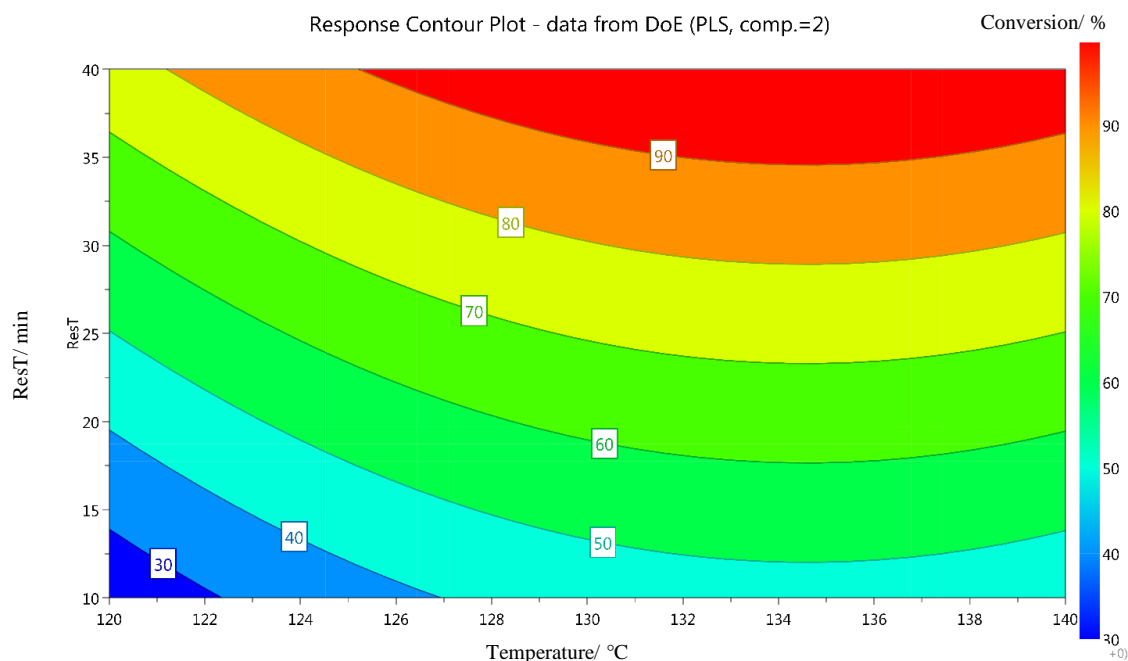


Figure 30 Response contour plot outlining the effect of residence time and temperature for the N-Boc deprotection of AZD5634. The optimum area; 125-140 °C and 34-40 minutes. Curvature can be seen in the plot due to a square term in the reaction.

The same series of reactions were carried out in IPA to make sure the solvent was compatible with the system and the reagent was completely miscible. The initial concentration, 1 % w/v, proved to work equally as well as it did in the previous solvent, giving a very similar statistical

profile with high $R^2$ (0.929) and $Q^2$ (0.888) values (Figure 30). Dissolution of the starting material, in IPA, was more challenging than dissolution in MeOH even when at the lowest concentration of 1 % w/v. The solvent and reagent were left in warm water and sonicated to ensure complete dissolution before the bulk solvent ($H_2O$) was added to the starting mixture.

### 2.4.4. Reaction Concentration Study for AZ Scale-up

The previous study saw starting material used at a concentration of 1 % w/v to ensure material conservation during initial runs. Many industrial processes use high working concentrations to maximise process efficiency to increase the throughput of material. Following these successful runs, starting material concentration was increased from 1 % w/v to 10 % w/v; the preferred concentration for the reaction at AZ. Preliminary batch reactions were carried out to gauge an understanding of the new concentration before continuing into flow. The microwave batch reactor was again employed and 5 mL of 10 % w/v in IPA was tested at 150 °C. Formation of $CO_2$ caused phase separation, which wasn't problematic at 1 % w/v. As the concentration was increased by a factor of ten the volume of $CO_2$ also increased by a factor of ten creating a large excess in pressure, calculated to be 10.27 bar, which causes the glass vial to fail. The reaction was repeated but with a total volume of 1 mL, a five times reduction in volume, so the excess $CO_2$ would only cause 2.06 bar of pressure (twice that observed at 1 % w/v). At 10 % w/v conversion were still high, >70%, so the reaction was transferred to flow. The higher concentration of 10 % w/v in flow was problematic again due to the large excess of $CO_2$ being produced. The previous flow reaction was carried out using a BPR of 100 psi, which when used again at the higher concentration, causing biphasic separation *i.e.* the solution segments were separated by gaseous segments. However, not entirely problematic, the excess gas increased reactor pressure therefore reducing residence times and lowering conversions (Figure 31). As reaction residence time needs controlling for reliable results a higher pressure BPR (250 psi) was used to keep the evolved $CO_2$ in solution keeping residence times at the required length. Several batches of the starting material were used for the work carried out. The first and third, as previously described, a highly viscous liquid and the second, a brown, crystalline powder. Problems occurred when using the second batch of material with the formed products, during and after the reaction. At 10 % w/v a thick, brown oil which deposited inside the reactor tubing and sample loop (Figure 32) upon cooling. This was not ideal for on-line analysis *via* MS as the residual coating caused the MS background to become dominated by a mixture of starting material and product. This batch was used until a higher quality sample was provided, where

the reactions were repeated. However, even with the purer sample small amounts of viscous substance was formed meaning its identity is from a side reaction or by-product of the reaction.
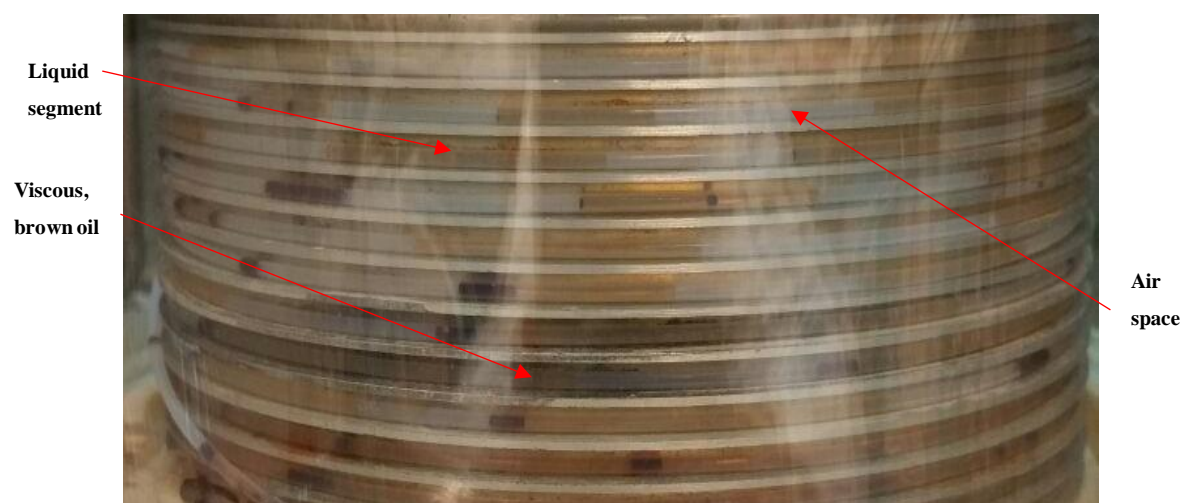


*Figure 31 Image showing the biphasic system after increased starting material concentration to 10 % w/v. The large excess of $CO_2$ can be seen between liquid phases. The image also shows the formation of the thick, brown oil on the reactor.*



*Figure 32 Image showing the formed thick, brown oil coating the inner reaction tube prior to the sample loop.*

The oil was suspected to be an oxidation side-product which was hoped to be batch specific. Reaction concentration was reduced to 5 % w/v proving beneficial in terms of $CO_2$ and oil production allowing for better and more reliable data from on-line MS.
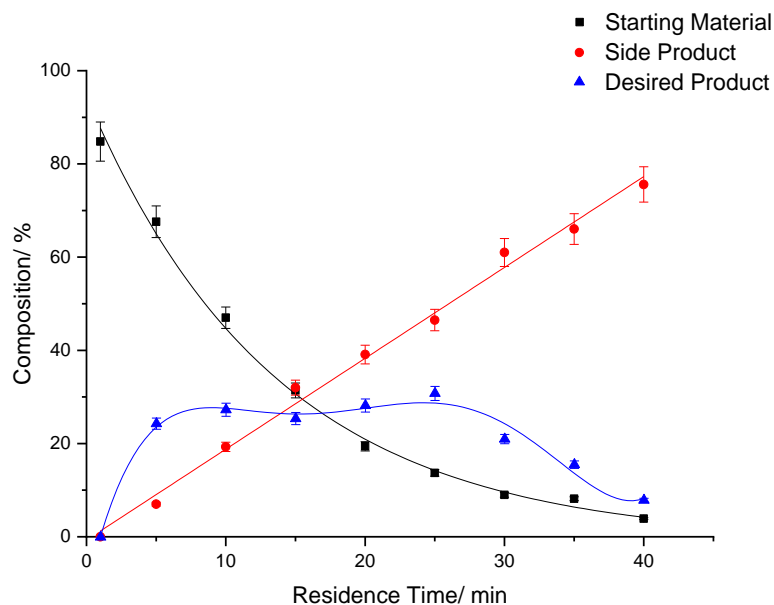
*Figure 33 Graph showing the reaction profile for the N-Boc deprotection of AZD5634 at 150 °C over residences time from 1-40 mins. The starting material can be seen to decrease in amount whilst the product has an initial increase. The starting material then reaches a steady state plateau as a by-product is produced before being completely degraded where the product level can be seen to decrease.*

A time profile study was carried out using batch two of AZD5634 to gauge an understanding of the process and to form a kinetic profile of the results from off-line LCMS. The time profile graph (Figure 33) clearly shows an increase in conversion upon and increase in residence time. However, initially the desired product forms but plateaus from 10 minutes until 25 minutes before decreasing when using longer residence times. The species seen to dominate the reaction at longer residence times has an *m/z* at 236 Da corresponding to the cleavage of the tertiary amine (Figure 36). This profile also provides evidence of another potential impurity when the reaction was performed at longer residence times and higher temperatures, although the species was only observed in small amounts and only with the batch with which this study was conducted with.

The AZ batch used may have a higher concentration of species which may facilitate reactions under the conditions used. $S_N1$ reactions will be much more prevalent under the used conditions as the medium becomes more acidic. As water is a major component in this reaction it is expected that more $H^+$ ions are produced at higher temperature. However, this does not occur. The pH of water indeed decreases with temperature decrease but this is balanced by the falling pOH which will be equal to the pH. Water at 100 °C has a pH of 6.14 but also a pOH of 6.14;

equal numbers of $H^+$ and $OH^-$ ions meaning neutrality is maintained. Acidity increase is likely due to the evolution of $CO_2$ in the reaction. Stoichiometic amounts of $CO_2$ are produced following N-Boc cleavage and, as the reaction is kept under pressure, this $CO_2$ will stay dissolved in solution causing dissociation to form carbonate ions, $HCO_3^-$, reducing the pH of solution (Scheme 22).
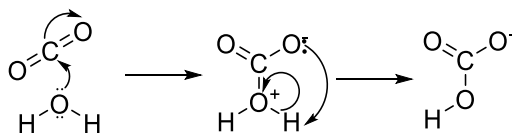


*Scheme 22 Mechanism for the formation of $HCO_3^-$.*

The work completed in this section used temperatures between 120 °C and 150 °C as limited by the polar bear reactor which has a maximum usable temperature of 150 °C. In order to achieve higher temperature and look more into reducing the total residence time of the reaction, a new high temperature reactor (Figure 34) was designed and built within the department.

### 2.4.5. High Temperature N-Boc Deprotection
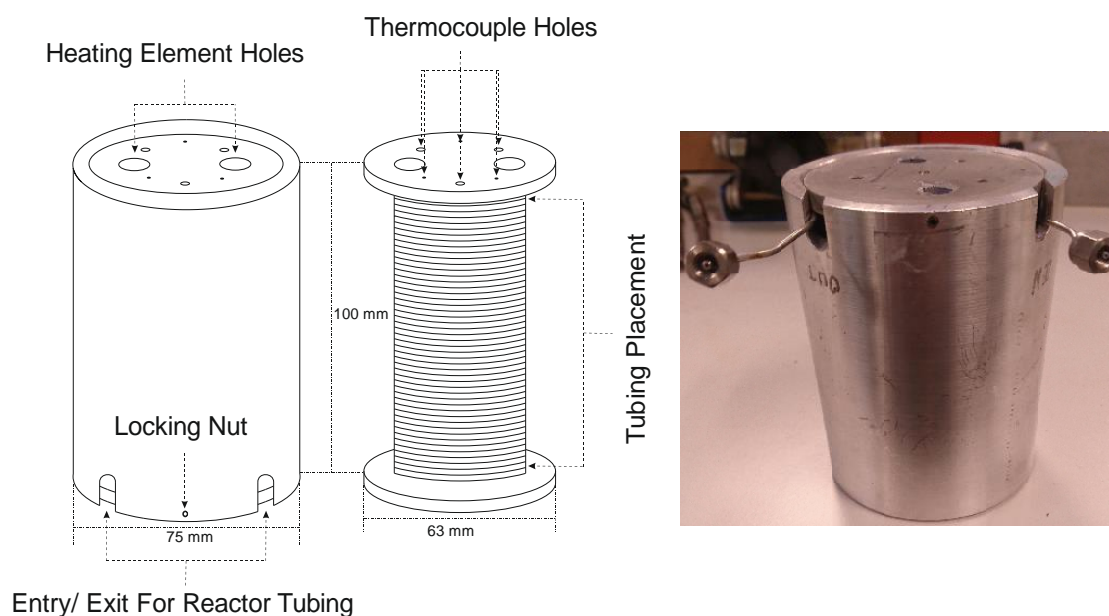


*Figure 34 Photograph (right) and 3D-schematic (left) of the high temperature reactor (cotton reel reactor) used for the N-Boc deprotection of AZD5634. Temperatures ranged from 130 °C to 200 °C to achieve the highest possible conversions possible during experimental designs.*

The newly designed reactor can reach temperatures above 150 °C for the deprotection of AZD5634. 180 °C was used for residence times of 5, 10 and 20 mins to gauge the duration needed for 95% - 100% conversion. On-line MS and off-line LCMS were both used to monitor the reaction. Previous work showed that higher temperatures are more beneficial and give high conversions (>90%). With a residence time of 20 minutes the deprotection achieved full conversion according to LCMS results (Figure 35, bottom chromatogram) which showed a single, large peak in the chromatogram with an *m/z* value of 279.5 Da. With a residence time of 5 minutes at the same temperature (180 °C) a conversion of >90% to the desired product was achieved (Figure 35, top chromatogram). There is evidence of other side-products forming in the reaction, due to the acidity of the system, which have been verified by the team at AZ and through accurate mass LCMS. One of the main side-products to for is the *tert*-butyl form of AZD5634 (Figure 36) which has an *m/z* value of 335 Da.
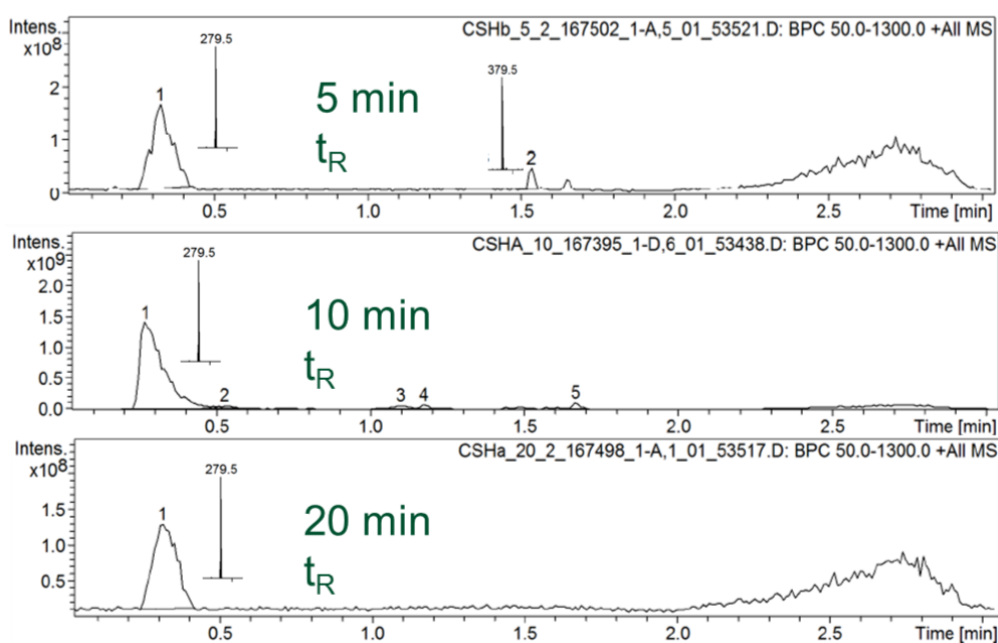


*Figure 35 LCMS chromatograms showing the N-Boc deprotection of AZD5634 to its amine form, the desired product can be seen at ~0.32 mins (m/z: 279.5, peak 1), starting material at ~1.54 mins (m/z: 379.5). 100% conversion has been achieved at 180 °C in 20 mins and >90% conversion for both 5 and 10 mins.*

$C_{11}H_{26}NO_4^+$
Mol. Wt. 236.33
**2.29**

$C_{17}H_{39}N_2O_4^+$
Mol. Wt. 335.51
**2.30**
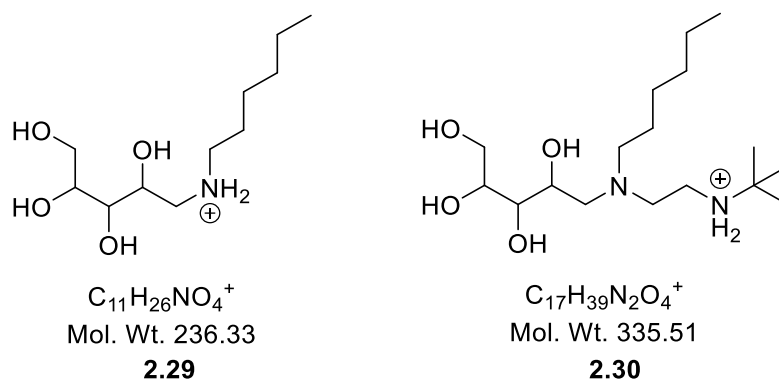
*Figure 36 Molecules **2.29** and **2.30** are both side-products generated during the N-Boc deprotection of AZD5634. **2.29** was mainly seen during the reaction profile study using material batch 2 whilst **2.30** has been observed by accurate mass, mass spectrometry.*

Calibration of reaction components, both starting material and desired product, were used and manually sampled *via* MS also helping to observe the if reactions stay within the instruments linear dynamic range at high process concentrations. Starting material calibration was completed from 10% w/v to 100% w/v (

Figure 37) whilst product calibration was carried out from 50% w/v to 100% w/v (Figure 38).
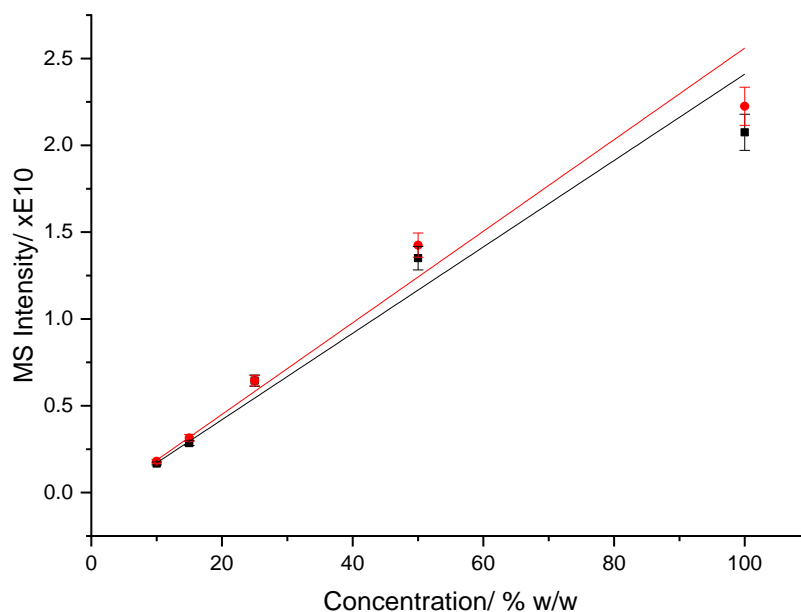


Figure 37 Calibration graph of AZD5634 starting material using online mass spectrometry. Two different calibrations were completed enabling a comparison with one another.
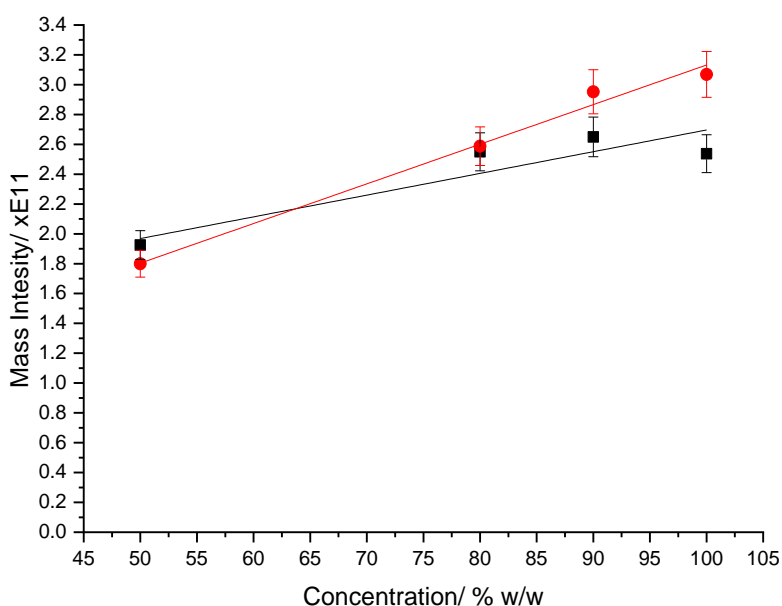
*Figure 38 Calibration graph of AZD5634 product, synthesised during flow reactions, using online mass spectrometry. Two different calibrations were completed enabling a comparison with one another.*

Following the previous studies, a three variable DoE was setup and completed by the team at AZ; temperature, residence time and % water added using IPA as solvent. This study showed that there was a much greater preference for higher temperature on this system compared to residence time (which had less effect on the reaction) and a greater preference for a larger percentage of water in the system.

A full overview of the three variable model is shown in Figure 39, containing a replicates plot, summary of fit, coefficients plot and a residuals plot. The model has an $R^2$ of 0.962 and a $Q^2$ of 0.799 suggesting low variability and good predictability of the model. The high validity shows statistically that the model does not contain any outliers or problems within the observed or predicted data. Reproducibility is high (0.94) and can be observed through the replicates as each blue square is close to being stacked one on top of the other. The residuals show good linearity meaning that they are normally distributed suggesting a good estimate of the statistical error in the model. Figure 40 outlines the observed *vs* predicted plot and how the predicted data deviates from the observed data. Each point is localised around the central diagonal line with even spacing above and below. This good spread dictates good model predictability with a normal distribution *i.e.* all points being within 0.5 units of the central line.

Both Figure 41 and Figure 42 show the response surface plots for the design model. The best conditions to use can be found at the most extreme conditions; high temperature, long residence times and high % water, matching what previous models suggest. The coefficients plots also shows that water ratio and temperature play the largest role in conversion followed by the residence time (Figure 39 – bottom left plot).
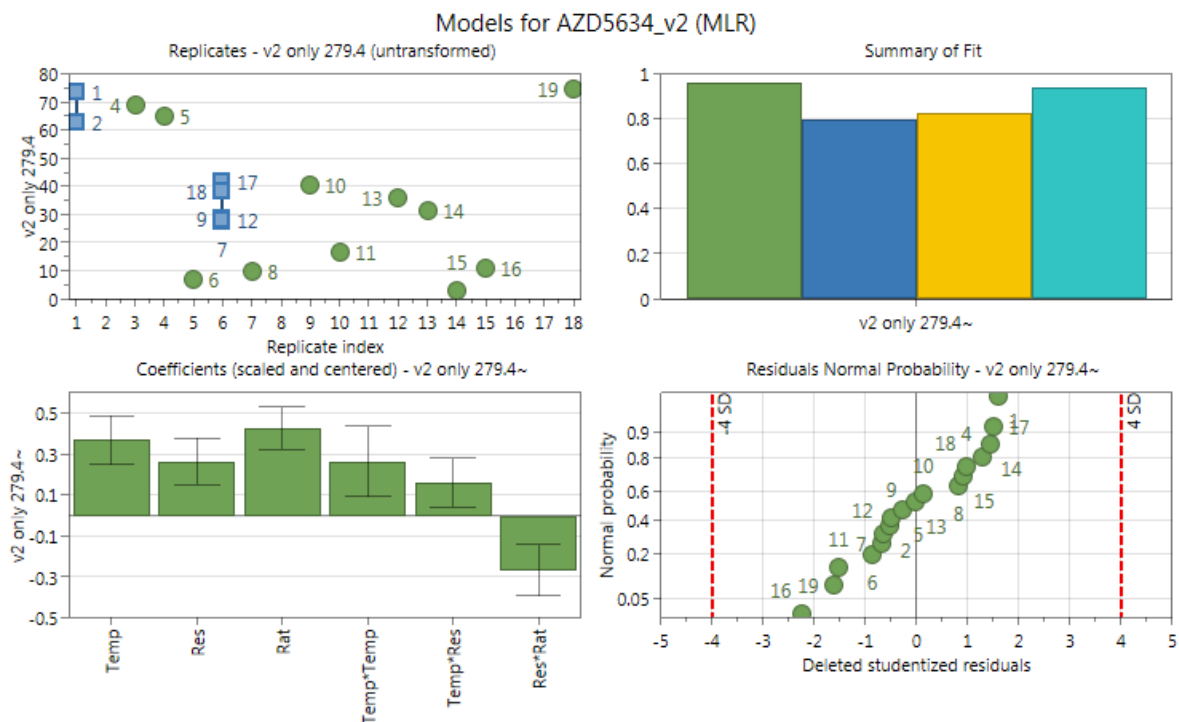


*Figure 39 An overview plot of the 3 variable model performed by the AZ team. From top to bottom and left to right: replicates plot, fit of the model ($R^2$, $Q^2$, validity and reproducibility), coefficients plot and residuals plot.*
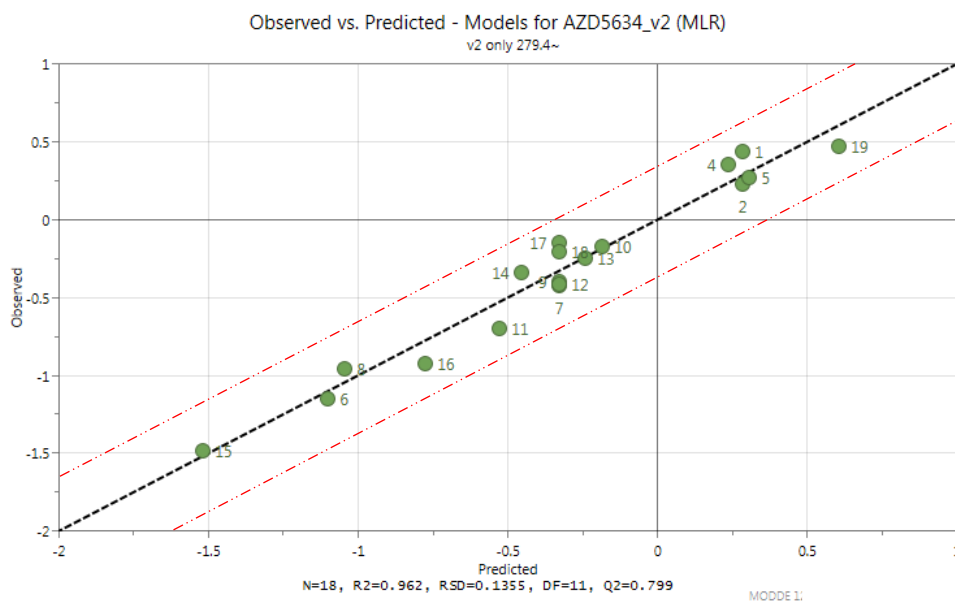
*Figure 40 Observed vs predicted plot for the 3 variable model performed. The points can be seen to be localised around the regression line, above and below within statistical error (red dotted line).*
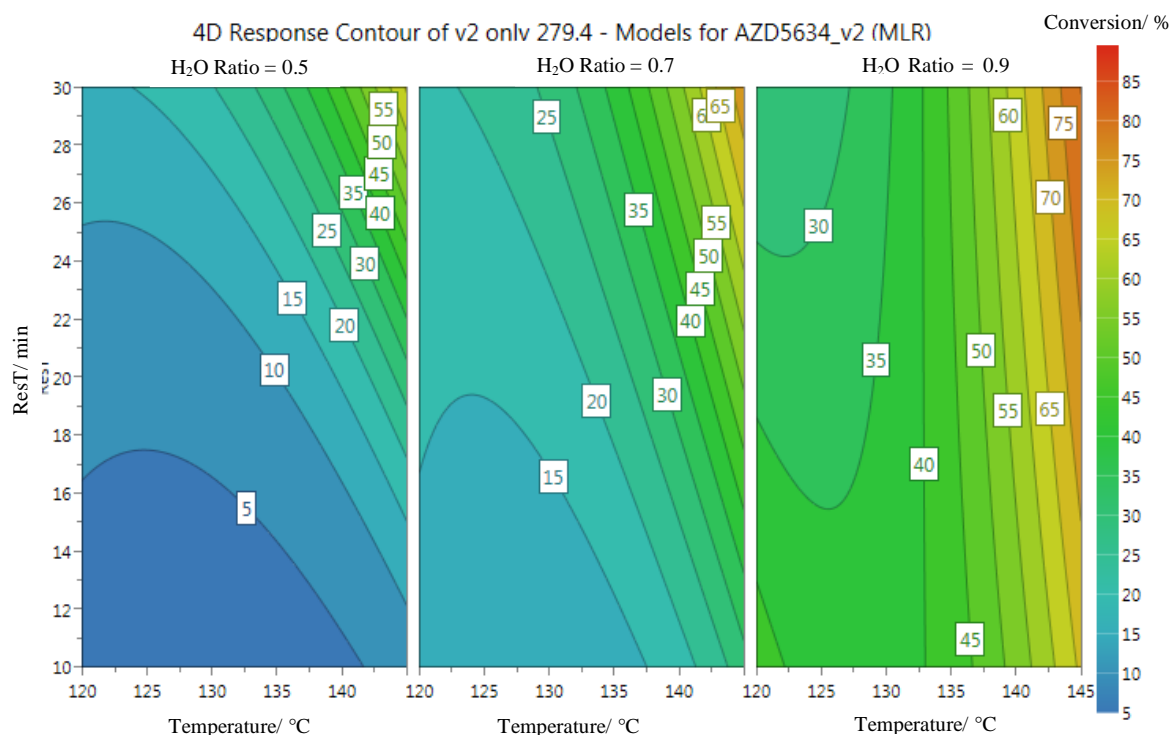


*Figure 41 Contour plot for the 3 variable DoE carried out. The best conditions can be seen at the far top right of the plot at high temperature, long residence time and high water concentration.*
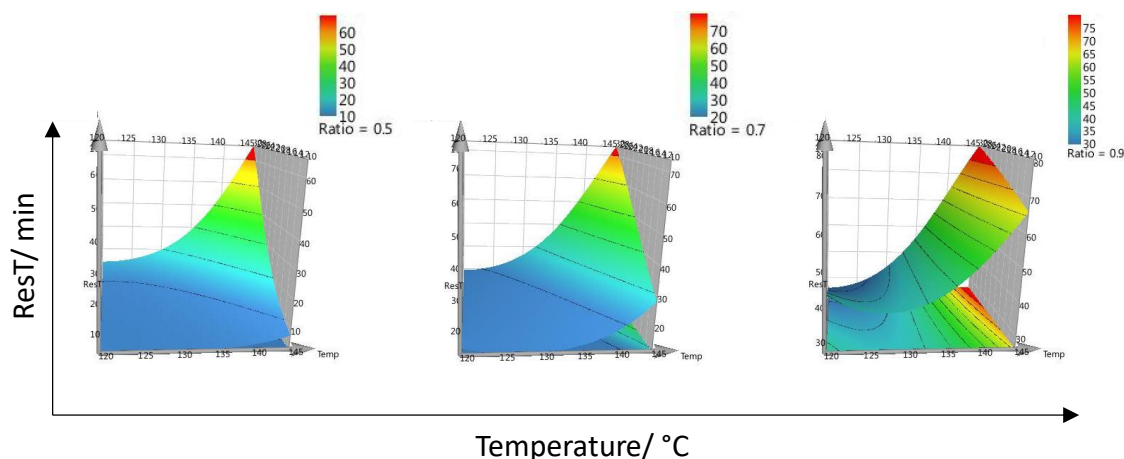
*Figure 42 Surface plots for the 3 variable DoE undertaken at different ratios of water. Left: 50%, middle: 70% and right: 90%.*

The above work was completed whilst knowledge of the subsequent reaction of AZD534 would need to be carried out in IPA as to avoid a solvent swap stage in the process (as would be needed if MeOH had been used). However, the subsequent stage was changed to improve reagent throughput (and therefore productivity), enabling use of MeOH. It was also shown that both starting material and product are more soluble in MeOH than IPA, agreeing with previous observations, where *ca.* twice as much starting material is dissolved. 20 reactions for the DoE were performed (in MeOH) within the experimental design space dictated by previous work (Table 8). This experimental design closely matched those previously ran when using IPA as the process solvent.

Reaction 0 (test) is needed to ensure analytical, thermal and flow equilibrium, acting as a "dud" and can be discounted. The first reaction failed to inject so both 0 (test) and first repeat have been discounted from the model. The algorithm monitors reaction conditions deeming when the reactor is at steady state and automatically takes a sample. This, however, caused the subsequent experiment to fail as there were no significant changes in the process conditions and so no sample was taken.[ii] DoE data suggests relationships between temperature and

---

[ii] The optimisation program used at AZ was a newer and more updated version than the program originally used in Leeds. This program had many improvements, some being; inclusion of dead-time conditions whilst sample analysis was performed to reduce waste of material, an easier to use GUI with more simplistic information input, the ability to add/change reactions in the current run list, text output to tell when a sample had been taken.

residence time for highest conversions. 78% conversion was seen at 180 °C whilst 42% conversion was seen at 160 °C, (Table 9). Within blocks of similar temperatures there was no direct relationship between volume of water, although the amount of water changed the residence time of the reaction which did affect the conversion. Within the temperature block of 170 °C, highest conversion was shown to be 75% with a corresponding residence time of 8.6 mins and water addition at 0.21 mL min$^{-1}$ (2.63 eq.). A similar reaction at 170 °C provided a conversion of 71% with a residence time of 8 mins and water flow rate of 0.23 mL min$^{-1}$ (2.88 eq.). Similar patterns can be seen in the temperature blocks of 177 °C and 162 °C.

*Table 8 Table outlining an experimental design conducted for the N-Boc deprotection of AZD5634 in MeOH and H$_2$O. *This reaction did not inject due to the test reaction beforehand having the same conditions. A discrepancy in the code caused injection failure as there was not a significant change in the conditions.*

| Order | Pump Flow rates/ mL min$^{-1}$ | | | Temperature/ °C | Ratio (water)/ eq | Residence Time/ min |
|---|---|---|---|---|---|---|
| | Water | SM & IS | Dilution | | | |
| 0 (test) | 0.22 | 0.08 | 0.60 | 170.00 | 2.75 | 8.33 |
| 1* | 0.22 | 0.08 | 0.60 | 170.00 | 2.75 | 8.33 |
| 2 | 0.22 | 0.08 | 0.60 | 180.01 | 2.75 | 8.33 |
| 3 | 0.34 | 0.08 | 0.84 | 177.40 | 4.25 | 5.95 |
| 4 | 0.16 | 0.08 | 0.48 | 177.40 | 2.00 | 10.42 |
| 5 | 0.17 | 0.08 | 0.50 | 177.40 | 2.13 | 10.00 |
| 6 | 0.36 | 0.08 | 0.88 | 177.40 | 4.50 | 5.68 |
| 7 | 0.45 | 0.08 | 1.06 | 170.00 | 5.63 | 4.72 |
| 8 | 0.15 | 0.08 | 0.46 | 170.00 | 1.88 | 10.87 |
| 9 | 0.22 | 0.08 | 0.60 | 170.00 | 2.75 | 8.33 |
| 10 | 0.21 | 0.08 | 0.58 | 170.00 | 2.63 | 8.62 |
| 11 | 0.23 | 0.08 | 0.62 | 170.00 | 2.88 | 8.06 |
| 12 | 0.22 | 0.08 | 0.60 | 170.00 | 2.75 | 8.33 |
| 13 | 0.15 | 0.08 | 0.46 | 170.00 | 1.88 | 10.87 |
| 14 | 0.17 | 0.08 | 0.50 | 162.60 | 2.13 | 10.00 |
| 15 | 0.34 | 0.08 | 0.84 | 162.60 | 4.25 | 5.95 |
| 16 | 0.16 | 0.08 | 0.48 | 162.60 | 2.00 | 10.42 |
| 17 | 0.36 | 0.08 | 0.88 | 162.60 | 4.50 | 5.68 |
| 18 | 0.22 | 0.08 | 0.60 | 170.00 | 2.75 | 8.33 |
| 19 | 0.22 | 0.08 | 0.60 | 159.99 | 2.75 | 8.33 |
| 20 | 0.22 | 0.08 | 0.60 | 170.00 | 2.75 | 8.33 |

*Table 9 Table outlining the conversions for each of the reactions completed in the experimental design laid out in table 3. *1 injected sample (HPLC&MS) during initial dud run which can be discounted.*

| Sample | On-line HPLC Conversion/ % | On-line MS Conversion/ % | Off-line HPLC Conversion/ % |
|---|---|---|---|
| 1* | 44.88 | 58.82 | 57.51 |
| 2 | 78.33 | 76.58 | 77.65 |
| 3 | 76.91 | 74.56 | 75.47 |
| 4 | 63.56 | 57.58 | 70.91 |
| 5 | 67.34 | 62.96 | 67.61 |
| 6 | 67.10 | 65.08 | 63.62 |
| 7 | 54.88 | 50.00 | 69.36 |
| 8 | 69.15 | 67.01 | *DoE left to run over night, no off-line samples taken |
| 9 | 71.64 | 71.82 | |
| 10 | 75.90 | 74.07 | |
| 11 | 71.64 | 68.49 | |
| 12 | 70.08 | 69.62 | |
| 13 | 70.97 | 67.63 | |
| 14 | 54.56 | 51.54 | |
| 15 | 41.47 | 41.67 | |
| 16 | 47.48 | 43.33 | |
| 17 | 39.39 | 38.51 | |
| 18 | 66.57 | 63.95 | |
| 19 | 41.95 | 41.75 | |
| 20 | 61.90 | 60.25 | |

Initially, both off-line and on-line comparison samples were taken. However, it was difficult to determine the time point for off-line sampling. If taken too soon there is a risk steady state has not been reached, too late, the next set of conditions may cause a slight change in the conversion and therefore a greater level of variation. A full set of off-line samples could not be taken as the DoE was left to run over night. Dead-time[iii] conditions (Jasco pumps – 0.02 mL min$^{-1}$) are assumed immediately after an on-line sample is taken, giving time for next conditions to equilibrate, to conserve material. The HPLC and MS together allow a more accurate comparison to be made between results, particularly for product concentration and starting material conversion. AZ created a HPLC method (through analytical services) to separate and detect the starting material to allow conversions to be calculated. MS conversion was calculated directly from mass spectra using the calibration plots created previously. These

---

[iii] Dead-time conditions are those at which occur when the system is not running an experiment as to reduce reagent use and save potentially expensive materials.

plots allowed the concentration of each reagent to be determined which could then be related to the conversion of starting material.

Reaction for the DoE were completed using a steady state set point of 1.5 reactor volumes before a sample was taken. Calculated conversions are lower than previously seen using IPA, suggesting that the steady state set point of 1.5 may be too low, stopping the reaction from reaching its real conversion. The steady state point was set to 2 reactor volumes, high enough to reach peak conversion but low enough to still conserve material.

Results suggest that the water content is not largely important for high conversions, effects may be due to its change in reaction residence time. The higher temperatures used seem to be the main contributing factor in affecting the reaction conversion *i.e.* reactions were above thermolysis temperature. Further evidence from the repeat process, using a set point of 2 reactor volumes for steady state, also shows that the higher temperature gives the higher conversion. The general increase in conversion and the close proximity of the repeat points in the steady state shows that the reaction is at steady state for the DoE run. As seen in previous work, temperature has the greatest effect on the conversion to the desired product, whilst residence time also has an effect, lower importance. Interestingly, the water content doesn't seem to play a large role in the process and on conversions but still needs to be kept in excess to the organic solvent. A recent paper looking into the mechanism of thermolysis of N-Boc deprotections suggests that the initial loss of $^t$Bu-group is kinetically favoured by high temperatures with the second stage, loss of $CO_2$, much slower. It is shown that thermolysis of the bond begins at a temperatures above of 200 °C but is highly dependent on reagent structure, dictated by electronics. Below the limit of thermolysis,[106] water% in the reaction may have a more significant role in the N-Boc deprotection mechanism shown by previous DoE results.

It was necessary to test if each reaction carried out was reaching steady state. Steady state[iv] is important in a continuous flow process as it ensures the true conversion and output of desired product are reached. Reactions were run using the same volumes and flow rates of SM/IS and $H_2O$ through the reactor at the most extreme temperature (180 °C); the highest observed conversion. Each reaction was sampled 3 times in order to get a profile of the reaction progressing up to its highest point. Due to previous problems with the program not sampling

---

[iv] Steady state is a term used in continuous flow processing which dictates that the reaction system is at a state were all conditions are constant, a change will only occur if the reaction conditions are changed.

when required, the temperature of each was varied by 0.5 °C whilst the dead-time temperature was set to 150 °C. This ensured enough change for the reactor to cool slightly during sampling and the have a new temperature for the program to set. The steady state of the reactor was tested by observing the conversions at 1 reactor volume and 2 reactor volumes. 1.5 was not looked at as sufficient data has been collected from the DoE ran.

The data from the steady state study showed lower conversion with a steady state reactor volume of one, getting only 40% conversion. At the upper limit of 2 reactor volumes the overall conversion was 88%. Graphs of each show the extent towards steady state by giving a profile; a reactor at steady state rises to a plateau, the maximum conversion for the used conditions (Figure 43). It can be seen from profiles and total conversions that the ideal steady state to be used in subsequent DoE's would be two reactor volumes. Above this and there will be no change in conversion and will waste material.
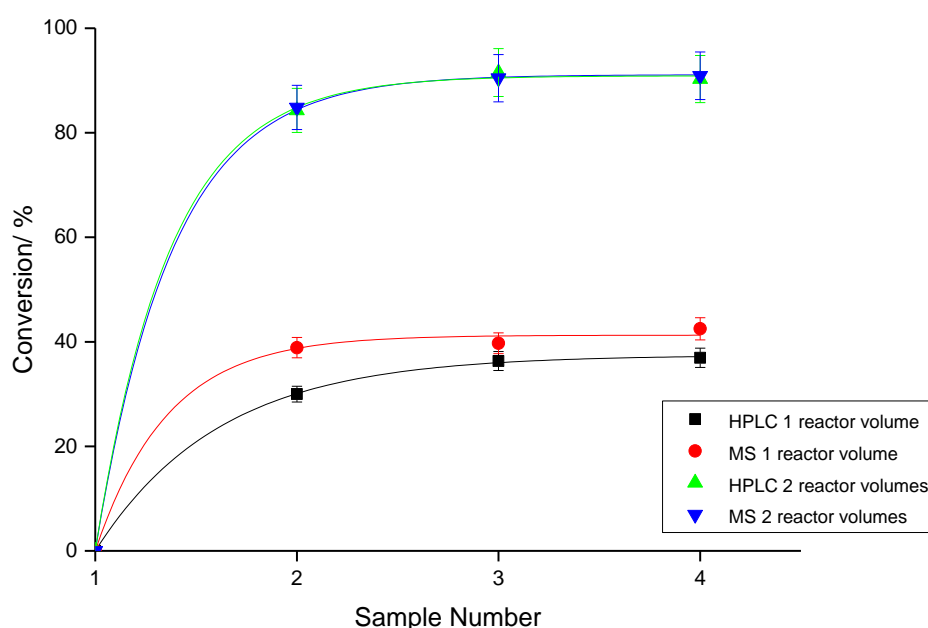


*Figure 43 Graph to show the difference between 1 and 2 reactor volumes for both HPLC and MS showing steady state during the flow N-Boc deprotection reaction of AZD5634.*

The DoE was repeated but using a steady state volume of 2 to ensure conversions have reached a constant plateau (Figure 5) for all experiments. A test reaction was, again, placed at the start of the run to ensure complete purge of the reactor system, but this time the test condition was set 0.5 °C higher than the actual run in order for the system to sample when required. The results from this DoE again show a similar data set and relationship between factors as the first.

The most extreme temperature 180 °C, allowed a conversion of 82% to be reached almost 5% more than the previous attempt. The majority of experiments on this second attempt were higher with the newer steady state set volume. Several samples failed to take shown by empty HPLC chromatographs and no file for mass spectrometry which could mean that the sample was out of alignment causing no sample to be injected. From the acquired data, the same trends can be seen between temperature blocks, similar conversions where conditions are the same and slight increases in conversion when residence time is longer. The similarities in conversions provides evidence that the flow technique gives reproducible data although the data set will have to be looked at in much greater detail from a statistical point of view. The second DoE was also a good opportunity to have a look further into the quantitative nature of the MS. On-line samples were collected using LC-MS and conversions calculated for both LC and MS separately. A look further into the kinetics of the system (by the AZ team) will show how important the water content is, particularly at higher temperatures, and whether it plays a major role in the reaction mechanism. Water may play a major role at lower temperatures if it is acting as a dual acid-base catalyst but maybe less important at higher temperatures if the reaction mechanism proceeds in a thermolytic fashion.

### 2.5. AZD5634 Reaction Analysis

The above DoE made use of AZD5634 product previously synthesised during other flow reactions and experimental designs, separated and isolated from the original mixture. The main outcome for this was to fully characterise the compound to look at side-products from this and previous reactions. All accurate mass spectra and data can be found in Table 10. Entries **2.34** and **2.38** are both AZD5634 starting material and product respectively. The high temperature of the MS ionisation source can provide significant thermal energy, for ~$10^{-4}$ s,[104] which may cause initial thermal degradation of the starting material, in this case, loss of the *tert*-butyl group. Mass spectrometry of a pure sample of starting material shows this loss and a fragment ion can be observed at *m/z* 323, entry **2.37**. This loss of the terminal *tert*-butyl group can either form a *tert*-carbocation, and subsequently isobutene, or $^t$BuOH by attack of the carbocation with water as it is highly abundant in the system. The *tert*-carbocation allows reactions to occur in the mixture with nucleophilic atoms in the molecule most commonly seen with the nitrogen atoms and forms species shown in entries **2.32** and **2.35**. The high temperatures may induce further fragmentation in the reaction system which may cause the deprotected product to
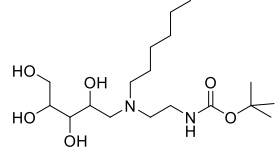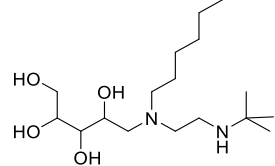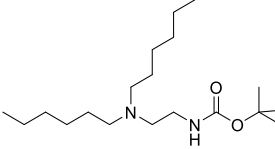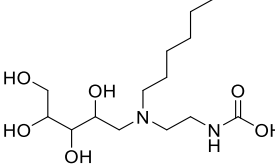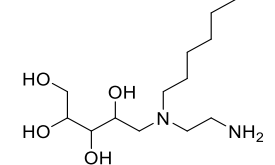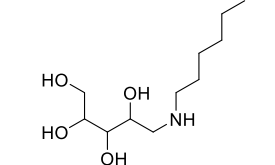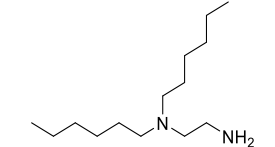
decompose and lose the terminal amine group to form the species seen in entry **2.39**. Dimerization of the product is seen to occur with the starting material, to produce entry **2.31**. The large evolution of $CO_2$ gas formed during the reaction facilitates $S_N1$ reactions with nucleophiles generated in the process. This could potentially be reduced by using gas permeable tubing to ensure all generated $CO_2$ is released from the system, reducing the acidity of the reaction. This has not been experimentally tested but may be worth revisiting in the future to see how it may affect conversions, yields and the formation of side-products.

Work completed by the team at AZ looked at the reaction at varying temperatures (100, 115, 120 & 126 °C) monitored *via* NMR spectroscopy. NMR spectroscopy allows protons to be monitored throughout the reaction and determined whether or not to decrease at that position or whether the proton appears due to a newly formed group in the molecule. The terminus of the hexyl chain was used as an internal molecular standard as this position (0.9 ppm, 3H) is not involved in the reaction and should stay constant in the starting material and product. A resonance assigned to the desired product was observed with an increasing peak intensity over time (2.82 ppm, 2H) and a resonance assigned to the starting material (3.2 ppm, 2H). Assuming that the terminal methyl-group is a relative constant in the molecule, a relative concentration plot of the product can be obtained giving an approximate percent yield (Figure 44 and Figure 45).

*Table 10 An outline of all structurally determined species, via accurate mass spectrometry, from the AZD5634 thermal N-Boc deprotection in continuous flow.*

| Entry | Formula Weight | [M+H]$^+$ | Molecular Formula | Structure |
|---|---|---|---|---|
| **2.31** | 582.78 | 583 | $C_{27}H_{58}N_4O_9$ |  |
| **2.32** | 434.62 | 435 | $C_{22}H_{46}N_2O_6$ |  |
| **2.33** | 428.48 | 429 | $C_{17}H_{36}N_2O_{10}$ |  |

| | | | | |
|---|---|---|---|---|
| **2.34** | 378.51 | 379 | $C_{18}H_{38}N_2O_6$ | |
| **2.35** | 334.50 | 335 | $C_{17}H_{38}N_2O_4$ | |
| **2.36** | 328.54 | 329 | $C_{19}H_{40}N_2O_2$ | |
| **2.37** | 322.40 | 323 | $C_{14}H_{30}N_2O_6$ | |
| **2.38** | 278.39 | 279 | $C_{13}H_{30}N_2O_4$ | |
| **2.39** | 235.32 | 236 | $C_{11}H_{25}NO_4$ | |
| **2.40** | 228.42 | 229 | $C_{14}H_{32}N_2$ | |
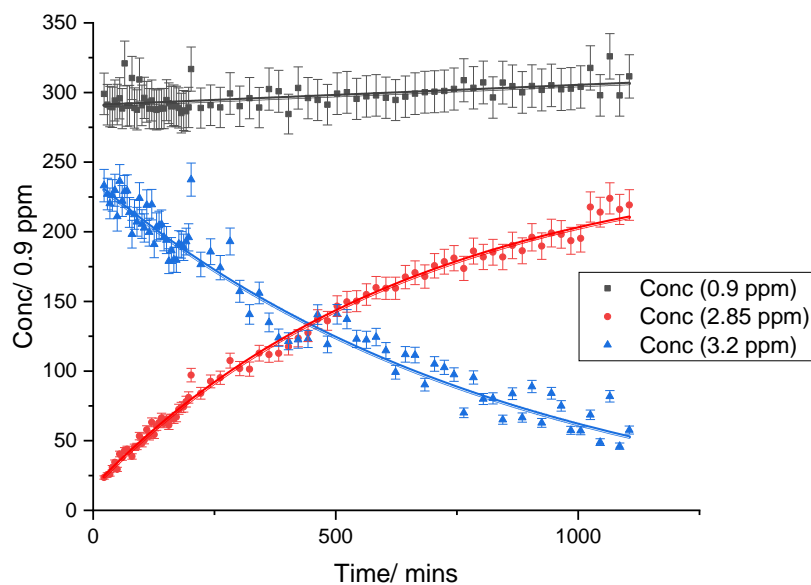
*Figure 44 Graph showing the integral of the product hydrogen environment forming (orange, 2.82 ppm, 2H) and the decrease in the starting material hydrogen environment (grey, 3.2 ppm, 2H) against time. A constant hydrogen environment for the hexyl chain in both starting material and product (blue, 0.9 ppm, 3H) was used as an internal standard.*
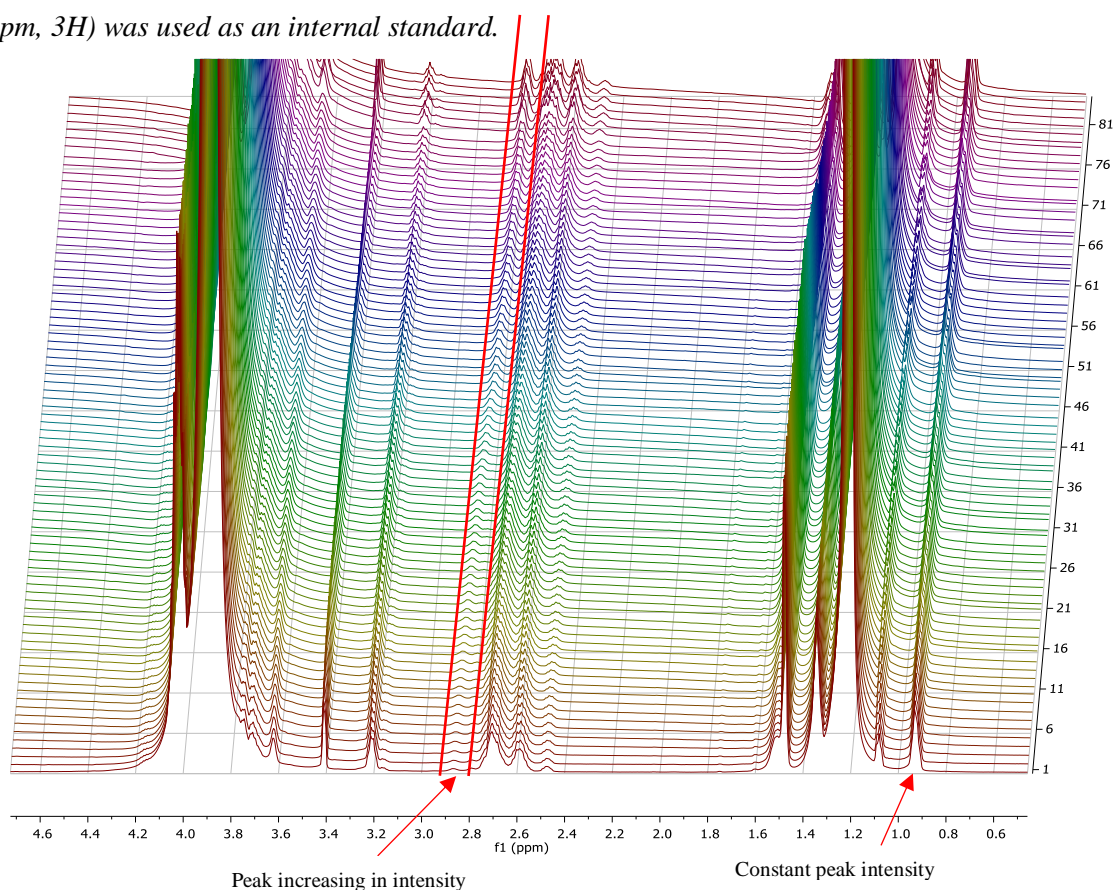


*Figure 45 Layered NMR spectrum showing the constant peak intensity for the terminal methyl group on the hexyl chain for all components (0.9 ppm) and the increasing intensity for the resonance assigned to the desired product (2.82 ppm).*

From this data observations of the 3H methyl resonance at 0.9 ppm increased over the first three hours until a steady state was reached. This may be attributed to a slow dissolution rate of the starting material (the sample began with a white material deposited at the bottom of the tube). If it is assumed that the 3H methyl is a relative constant it can be seen that there is a gradual increase of the relative concentration of the product to *ca.* 80% from the methyl resonance. This resonance at 3.2 ppm shows a steady decline but plateaus at around 16 hours. Monitoring was stopped after 19 hours and allowed to cool where ^{1}H NMR (Figure 46) and ^{13}C NMR data were acquired at *ca.* 18 °C. This data shows improvement over the higher temperature with respect to line and resolution of multiplicity so the methyl at 0.88 ppm can be attributed to 3H.



doc160810061/400
E16-010005
AZD5634 Montoring thermal deBoc at 120oC (Tset 384.6K)
Pressure tube
BBFO/In tube
MNunn/10Aug16

Post reaction data at Tset=289.5K
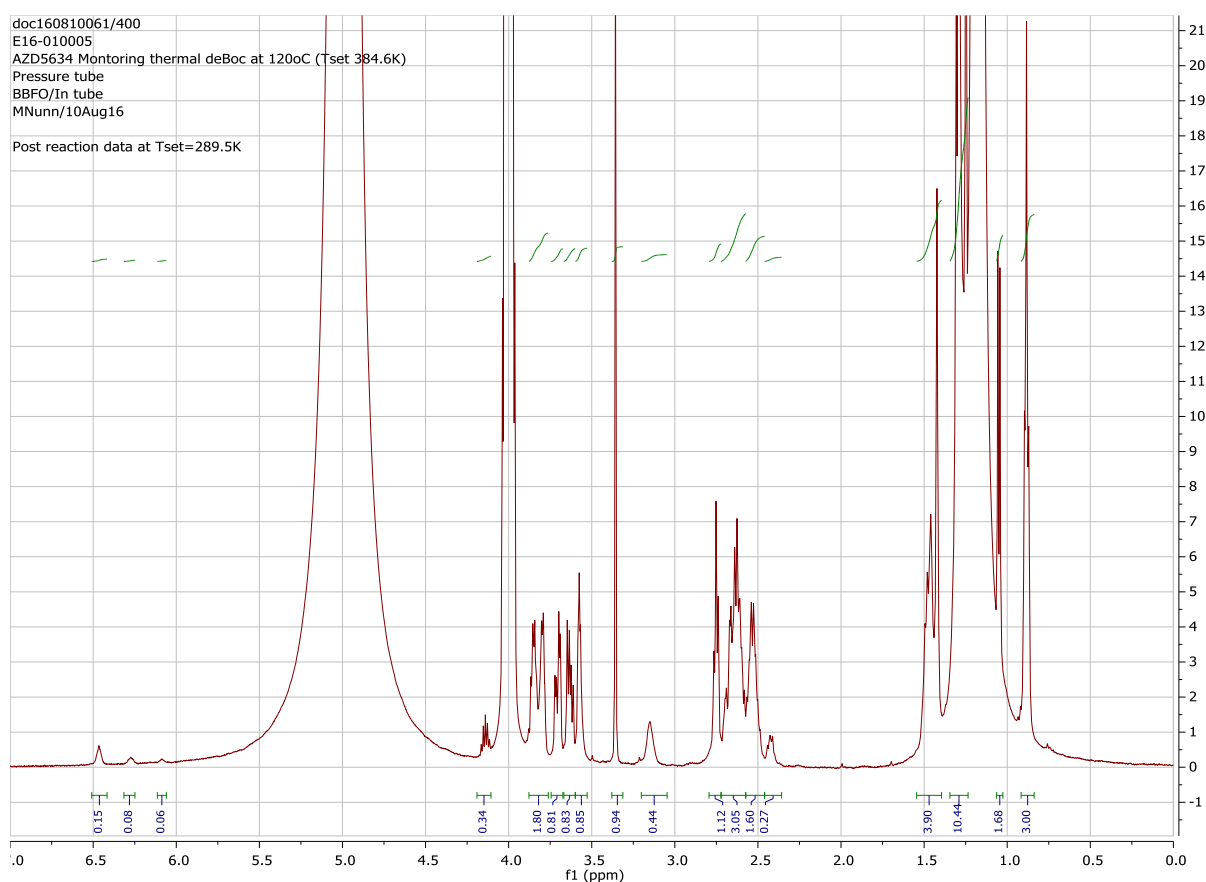
*Figure 46 AZD5634 Monitoring of thermal N-Boc deprotection at 120 °C. ^{1}H NMR spectra shows the reaction post reaction at ca. 25 °C.*

Most importantly, the septet at 4.2 ppm and the doublet at 1.04 ppm suggest that a new isopropyl containing species was formed. The could be credited to the reaction of isopropanol with the liberated *t*Bu-cation fromed during the deprotection (*i.e. tert*-butyl isopropyl ether). A The singlet found at 1.43 ppm can be assigned to *t*Bu-group of the *tert*-butyl isopropyl ether or

*tert*-butyl alcohol. Remaining starting material can be observed from the carbonyl resonance in the $^{13}$C NMR at 158 ppm. It is possible to monitor this thermal N-Boc deprotection *via* $^{1}$H NMR but there will be greater error associated with conversions and yields due to poor peak shape and poor baseline resolution

## 2.6. Summary and Future Work

This reaction studied was a simple reaction and success has been shown in both batch results and flow results. Several deprotection reactions of the N-Boc group have been shown to occur in batch with $H_2O$ at 100 °C with high yields but having long reaction times.[100,102,103] This work demonstrates that elevated temperatures and pressures can be used to promote reactivity in an organic transformation to improve overall conversions. It has also been shown that the most important effect on N-Boc deprotection is temperature as the process is still highly effective at 180 °C with short residence times. A method of N-Boc deprotection has previously been stated using temperatures of 300 °C in MeCN *via* a thermolytic reaction mechanism.[98] The study for $H_2O$ content did show that the higher the amount of water in the reaction the better the final conversion to the product this was seen much better at the lower temperatures, however, the role of water may become less major as the temperature rises if the final mechanism is thermolytic. Another group has shown that $H_2O$ can act as a dual acid-base catalyst under temperature and pressure and so can remove the N-Boc group. It may be beneficial to explore at what temperature the reaction become purely thermolytic at when the role of $H_2O$ stops in the mechanism.

The use of on-line MS has been ideal for this reaction as the structure of the starting material is highly complex, creating complicated NMR spectra and without any chromophores to be monitored *via* LC-UV. The MS can detect the starting material, product, by-products and impurities created in the reaction. The calibration has also helped to indicate the conversion much more accurately then at the beginning of the work when the peak ratio between starting material and product were used.

As the bulk of optimisation of conditions have been completed for the reaction, favouring high temperatures and greater volumes of $H_2O$, a closer look into the E-factor of the system could be a valuable next step. As industry takes great care to reduce costs and improve the environmental impact each process takes, being able to put a value on the reaction as a whole could be highly advantageous in the long run. The solvent system has already been switched

to something greener, MeOH to IPA, but could still be made to be completed in an even greener solvent, MeTHF.

All work completed for this project at the University of Leeds and at AstraZeneca played a part in key process development for the reaction to be done in continuous flow on the large scale. Initial work helped develop the process and showcase how effective continuous flow would be for this type of reaction. Subsequently, this process was successfully scaled up to the large-scale lab (LSL) to generate larger amounts of AZD5634 starting material to be used for further process development. The successful scale-up of this process also enabled the initial stages to be developed for the pilot plant for pre-commercial running to synthesis enough material for clinical trials. Currently, this molecule is still in development, but the use of continuous flow chemistry has made this process much easier for industry to adopt and apply on a much larger scale.

It would be an interesting study to look both into the kinetics of thermolysis of deprotection but also water mediated deprotection. Being able to determine separate ranges of conditions for many molecules may prove useful in developing reaction model but also in training predictive algorithms. This work could be completed in flow but may be easy to form a library using high-throughput methods before testing and optimising each compound in continuous flow using DoE and algorithmic optimisation similarly to the study detailed.

Other N-Boc protected amines will also be tested to see how applicable it is across different substrates containing different functional groups. Boc-Trp-OH and Boc-Arg(Pbf)-OH (Figure 47) have been chosen due to different structures, more complex structures and can be monitored by mass spectrometry and as they were cheap and available N-Boc protected substrates.



**2.41**                    **2.42**
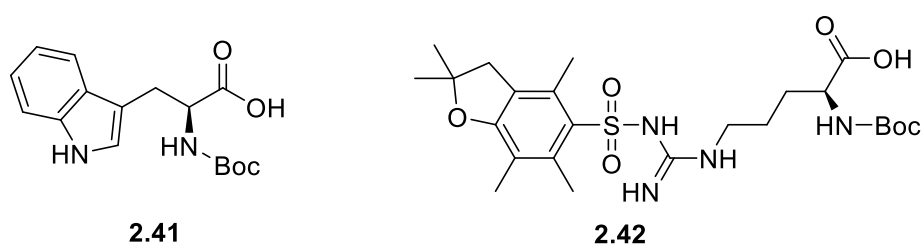
*Figure 47 Chosen N-Boc protected molecules which will be used to continue N-Boc deprotection studies.*

### 2.7. Experimental

### 2.7.1. General Experimental Methodology and Experimental Setup

Unless otherwise stated, all reagents, starting materials and solvents (all used as HPLC grade) were purchased from Sigma Aldrich, Fischer Chemical, Acros Organics, Alfa Aesar, Merck, VWR International or Fluorochem and were used without further purification. Each new starting material was analysed using $^1$H NMR spectroscopy to check purity and to check characterisation. NMR spectra were obtained using Bruker Advance 500 at 500 MHz or Bruker DPX 300 at 300 MHz for $^1$H NMR. NMR samples were recorded in CDCl$_3$. Starting material samples were also analysed using mass spectrometry to gain an understanding of molecular ion peaks, adducts and fragments. Mass spectra were obtained using the Advion Compact Mass Spectrometer (CMS) either in APCI+ or ESI+ mode using capillary PEEK tubing (OD 1/16", ID 0.005", max. pressure 550 bar). Further detail of ionisation conditions is shown in each individual experimental section. HPLC analysis was completed on an Agilent HP-1100 instrument (ethernet controlled) fitted with a Sigma Ascentis Express C18 column (5 cm x 6.6 mm x 2.7 µm particle size). Both mass spectrometric and HPLC analysis were take using a Vici Valco 4-port sample injection valve, (CI4W.06 manual valve, DCI4W.06 rotor/DCI4W.5 rotor, medium torque EUDA actuator, 0.06 µL/0.5 µL injection volume). All HPLC solvent gradients are described in the relevant sections. UV-vis analysis was performed using a PCRmax Lambda Spectrophotometer, reference solvent MeCN. In-line UV-Vis measurements were made using a an OceanOptics USB2000+ detector with industrial grade fibre optics cables. Microwave reactions were completed in a CEM Discover SP Microwave Synthesizer with autosampler. Samples we made in a 10 mL vial with PTFE seal and silicone rubber lid with escape gas exit. The microwave was kept in constant power mode using 200 W of microwave energy to heat the sample to the required temperature. The CEM microwave keeps the reaction under pressure to achieve temperatures above those of the solvents used. Once at the required temperature the sample would be left to reaction for the required time before being cooled and ejected by the autosampler and analysed. Accurate mass mass spectrometry was carried out on a Waters Acquity UHPLC coupled to a Waters Synapt G2Si ion mobility quadrupole time of flight mass spectrometer which uses and swaps between both APCI+ and ESI+ modes

repeatedly LCMS analysis was completed using both a Bruker Daltonics HCTultra LCMS with ESI ionisation and a Bruker AmaZon Speed LCMS with ESI ionisation.

### 2.7.2. Self-Optimising Flow Reactor Set-up
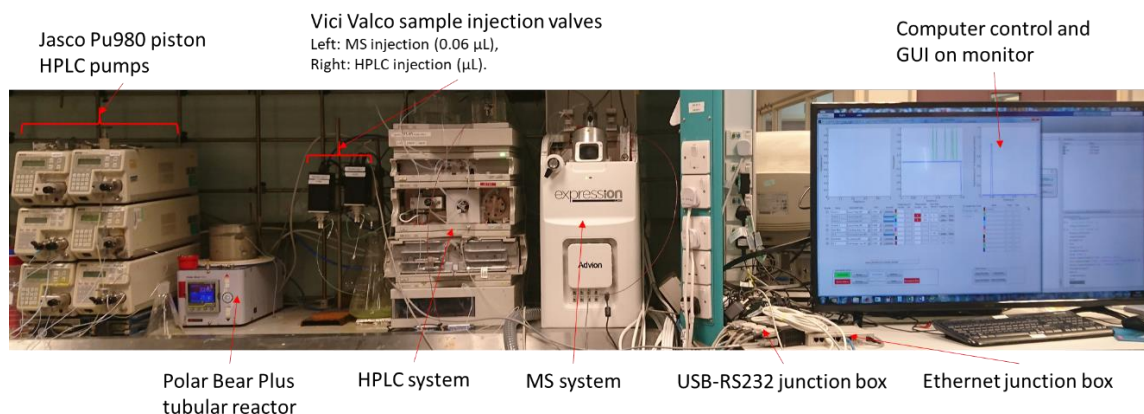


*Figure 48 Annotated photograph of the SOFR setup used for work throughout this thesis at the University of Leeds. It is comprised of HPLC piston pumps, tubular reactor, sample valves, HPLC, MS and control computer.*

Flow chemistry reactions were completed using Jasco PU980 dual piston HPLC pumps (RS232 controlled, Figure 49). To calibrate each pump set to 1 mL min$^{-1}$ and proceed to pump $H_2O$ for 10 minutes (after fully purging) and collect the solvent in a pre-weighed measuring cylinder. If the collected volume is 10 mL (=10.0 g for water) the pumps are running as expected. If a different volume is collected *e.g.* 10.6 mL, divide by 10 and convert to a percentage (1.06 %) and use this value to reduce the pump speed so exactly 1 mL min$^{-1}$ is pumped (0.94 mL min$^{-1}$). It is useful to note the date of calibration for future reference. This procedure can be repeated for various flow rates to check accuracy. If flow rates are greater than ± 5% the pump heads require maintenance.
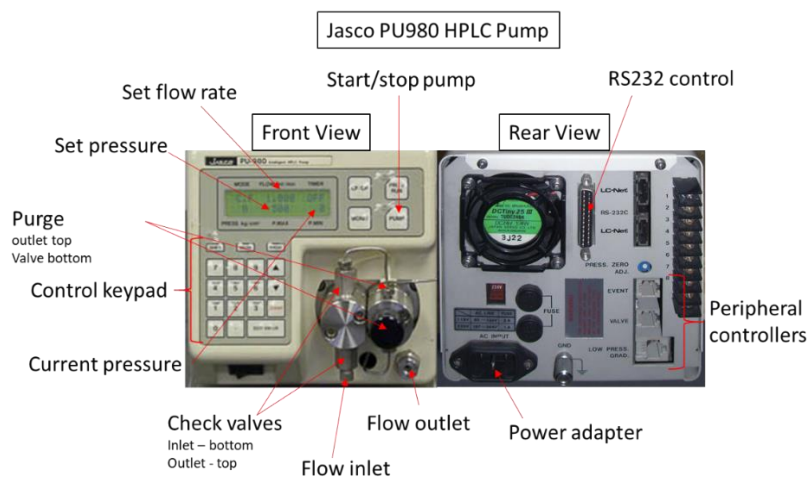
*Figure 49 Jasco PU980 pump diagram outlining key components for use in a self-optimising flow reactor system.*

Tubing used for reactions was either PTFE from Polyflon (OD 1/16", ID 1/32", max. pressure 35 bar), stainless steel from Sigma Aldrich (OD 1/16", ID 0.03") or Hastelloy from Thames Restek (OD 1/16", ID 0.04"). These were connected to pumps or reactors using Swagelok stainless steel 1/16" unions and/or tee-pieces.

For flow reactions themselves a Polar Bear Plus from Cambridge Reactor Design (ethernet controlled) was employed as well as an in house designed and built 'Cotton Reel' tubular reactor (RS232 controlled, Figure 50). Both reactors were made use of a Eurotherm temperature controller (Nanodac and 3200 respectively).
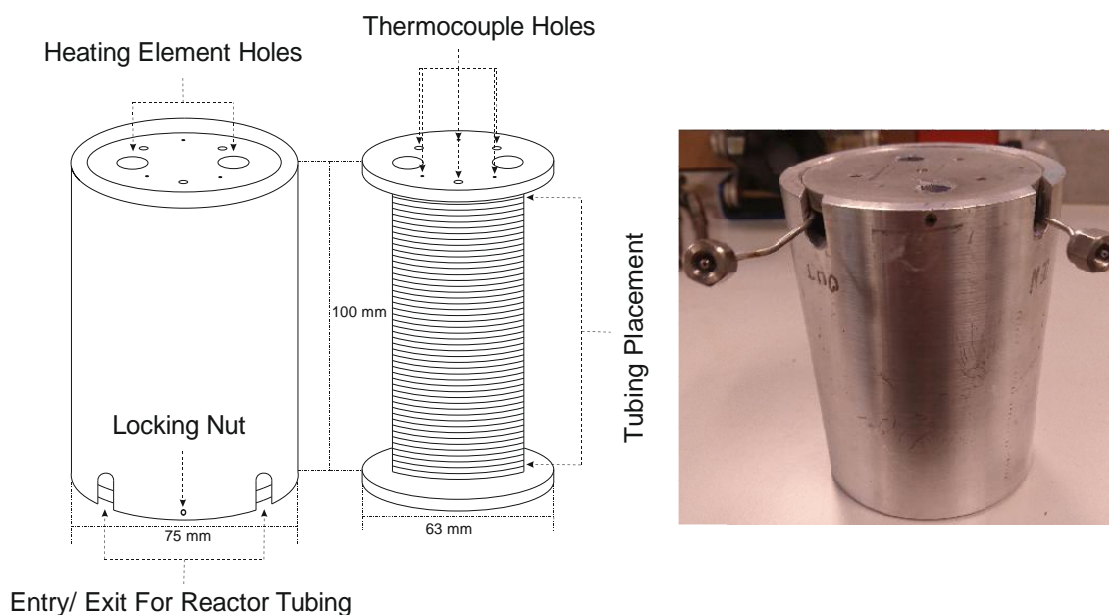
*Figure 50 Image and 3D drawing of the high temperature reactor (cotton reel reactor) used for the N-Boc deprotection of AZD5634. Temperatures ranged from 130 °C to 200 °C to achieve the highest possible conversions possible during experimental designs.*

The continuous flow stream connected to a Vici Valco 4-port sample injection valve to sample the reaction using on-line MS and HPLC. The internal volume of valve for HPLC was 0.5 uL and 0.06 uL for MS analysis. Fittings used were Jasco ferrules, stainless steel tubing for port adaptors and 1/16" Swagelok stainless steel unions. A contact closure cable is used which connects the sample actuator to the HPLC. Once the actuator has injected the sample to the analysis stream, contact closure starts the HPLC method for analysis. A pin connection diagram is shown in Figure 52. The 3 pin – RS232 computer control cable is supplied when purchasing the Vici Valco sample valve. This connects to a single 9 pin port at the rear of the computer and the correct socket (A or B) on the rear of the sample valve. It is possible to connect a cable to both sockets A and B where two computers can be used to control the actuator. The actuator can be controlled manually using the provided output display as well as computer control. Several useful commands can be used when computer controlling the actuator:

- DT[0000] – Delay time between actuator movements from e.g. DT[3000] gives a delay time of 3 seconds.
- TT – trigger the sample valve using the DT time delay.
- LRN – the valve finds and relearns positions A and B before stopping in the home position (A).
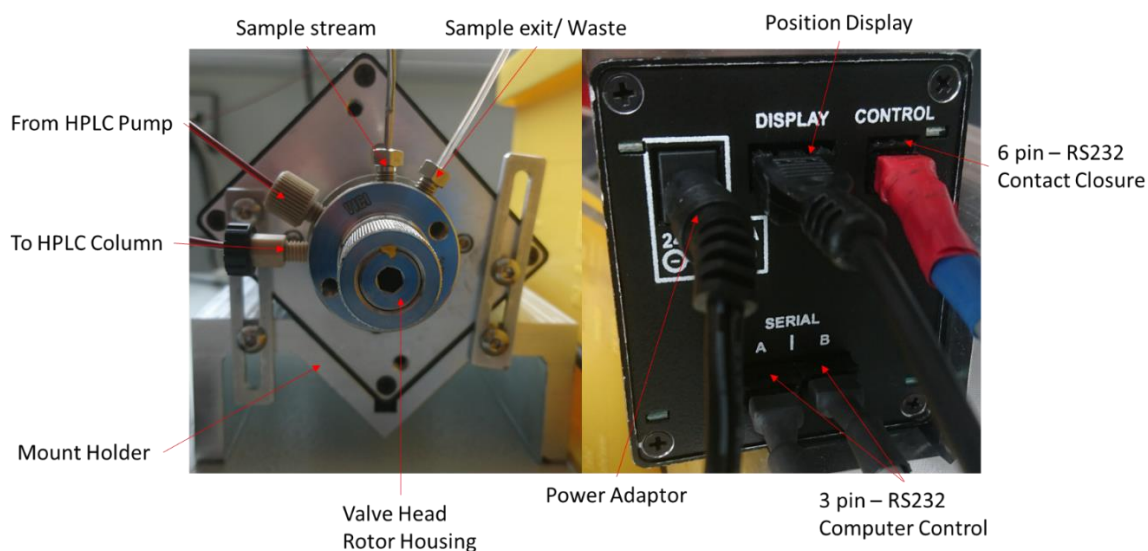
*Figure 51 Vici Valco sample valve front (left) and back (right) with annotated components.*
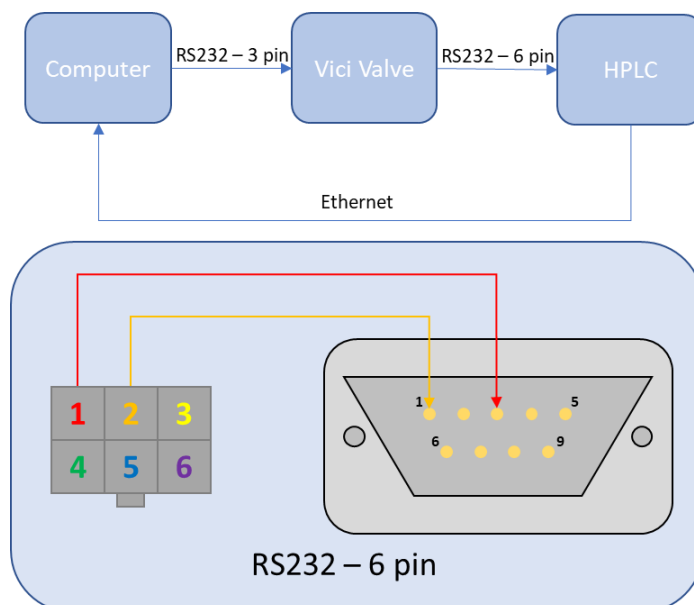


*Figure 52 Cable connection configuration and pin connection diagram for the Vici sample actuator/HPLC contact closure cable.*

All peripheral equipment, unless stated otherwise, was controlled using RS232 and wired to a junction box (EasySync USB2-H-1008-M, Figure 53). The junction box allows up to 8 separate RS232 controlled peripherals to be connected and controlled *via* the computer controller and user interface.

*Figure 53 EasySync USB2-H-1008-M junction box when connects all RS232 controllers to the PC.*

The HPLC system comprised of a degasser, quaternary pump with 4-feeds (A-D), a column heater and a diode array detector (DAD). The mobile phase for HPLC was typically MeCN with 0.1% TFA and $H_2O$ with 0.1% TFA. The pump output feed, using capillary tubing, was connected to port-P on the sample valve and continued through port-C to the HPLC column. Port-S contains the flow stream from reactor and continues through to port-W for collection/waste. The HPLC system was configured to be controlled using an external injector to wait for a contact closure signal before the method to begin. The MS was connected similarly to the HPLC, but the column output was replaced for the HPLC feed. Where the HPLC was being used independently a Jasco PU980 pump was replaced for the HPLC pump.

A contact closure and control cable for the MS allowed the system to control the HPLC, start a method and begin a method on the MS. This played a big role when running LCMS of reaction components but also allowed MS analysis to begin alongside HPLC analysis for a representative reaction sample.

### 2.7.3. MATLAB and Coding

The user interface itself utilizes the program MATLAB. MATLAB (Matrix Laboratory) is a mathematical computer software package which uses its own proprietary programming language developed by MathWorks. The software allows many different basic mathematical functions to be calculated and the development of more complicated functions, programs and user interfaces. Added toolboxes for the software open many more opportunities to integrate new interfaces from other programming languages and

access much more in-depth computational mathematics for a wider range of scientific areas to develop algorithms for manipulating mathematical expressions and objects. MATLAB allows the creation of graphical user interfaces (GUIs) to make the opening and running of programs and algorithms quick, simple & user friendly and can incorporate user made functions with highly complex multifunctional methods for many different outcomes. The codes, GUIs and functions described are all ones written especially for use with the self-optimising flow reactor, written in-house or downloaded for free.
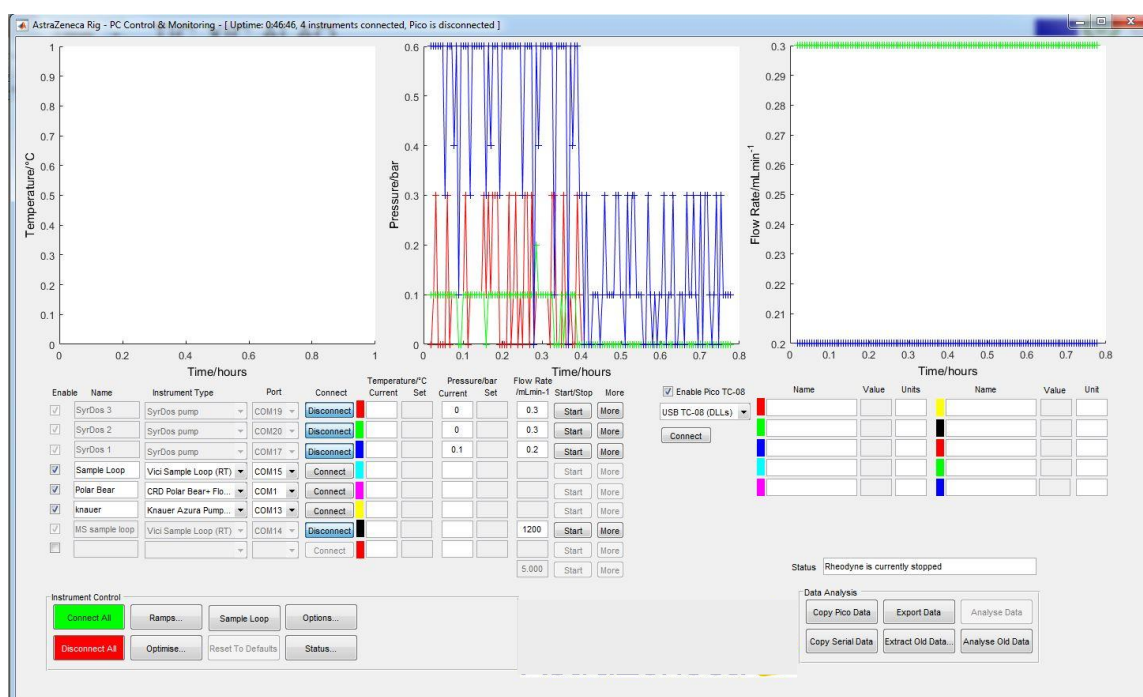


*Figure 54 The control interface used to connect and monitor each device involved in the reactor. The interface allows various optimisation sequences to be selected and ran.*

The CO2GUI is a GUI (Figure 54) which gives the user complete control of the system and allows equipment to be connected to the computer and MATLAB as well as monitor the conditions the devices are currently at. As each piece of equipment is directly connected to the computer the user can stop, start or change the conditions of each when necessary without leaving the computer. Live graphical plots output temperature of reactors as well as flow rates and pressures of pumps which can be useful for visually seeing if there are any problems which need to be addressed. It is also from here where the all other options for running set ramps, optimisations and options can be selected.

Each piece of equip can be connected to the software by either selecting 'connect all' or individually selecting and connecting each in turn. The same method can be applied when disconnecting equipment from the GUI. The status of each device can be further examined by selecting 'Status'. The 'Options' setting gives detail on which outputs on the GUI can be seen for each device. Current temperatures, pressures and flow rates can be toggled on or off. Other useful changeable options include: Data export format (default to Excel) and the refresh rate of each timer connected to the GUI (default 10/s). The functions 'Ramps' and 'Optimise' can be used either to run a set user defined program, utilize the optimisation algorithms or run a series of predefined experiments for a DoE approach. The ramps function allows each connected device to be controlled by a user set method which changes each conditions in set increments. This can be useful in the early stages of a reaction where an equation to add into Snobfit hasn't been found or to trial the reaction at many different conditions using an OFAT approach. The ramps themselves are selected and controlled by another interface (Figure 55) and can have several functionalities.
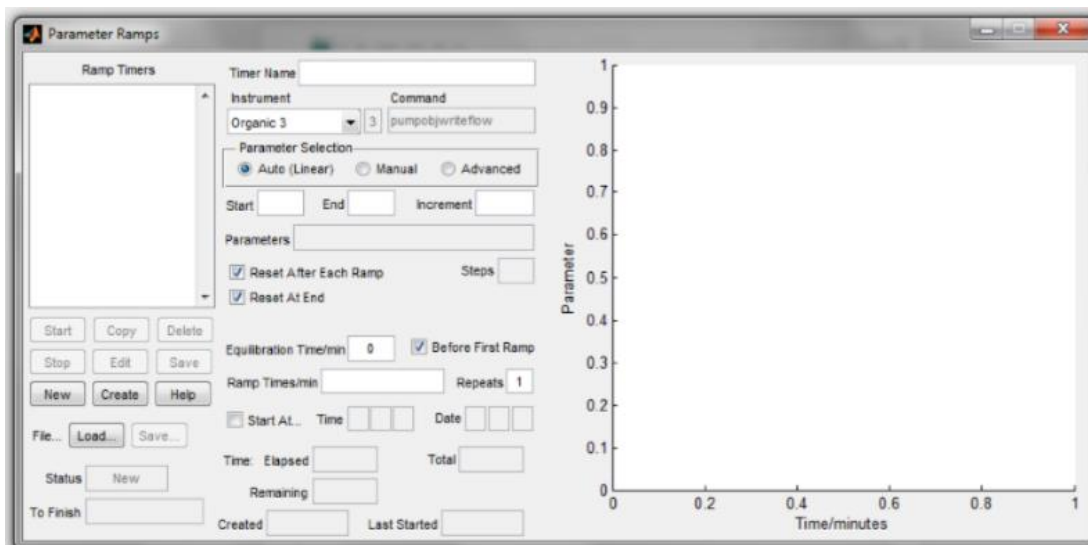


*Figure 55 Interface to create ramps using the optimisation equipment for each connected device.*

Once the 'New' button has been pressed the user inputs the name, the device, and selects whether they want to complete an auto, manual or set their own method ('Advanced'). The 'Advanced' section gives the user control of each added device and lets them set the parameter, *e.g.* flowrate for pumps, from the start to the end and the increment between each and the total time the set parameters are held for or changed over (Figure 56). The

ramps themselves are displayed in lines corresponding to the order they occur in and are shown graphically alongside. Once programmed into the interface the user can click the start button to begin the run, where the program then runs until the end of the ramp. The sample loop can also be set to trigger at the end of each programmed experiment, so samples are taken, and chromatographs and spectra are taken.
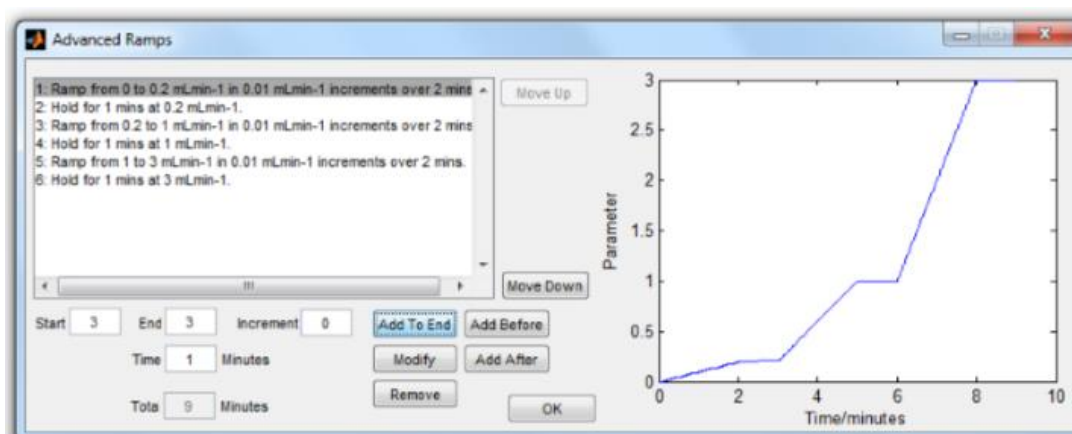


*Figure 56 Advanced ramps interface where the user creates their own method sequence to run on the optimisation equipment.*

Each ramp selected runs independently of another so care must be taken to ensure that the timings overlap in the correct places. All the ramps programmed can be viewed on the parameters screen overlaid from start to finish and can be edited, deleted, copied or saved. Saved ramps can be loaded back in if repeats are needed or the ramp needs to be restarted. In this case the example ramps can be seen to be 'A HPLC pump' and 'Heater' and are set to run three temperatures and three flow rates over the course of nine minutes (Figure 57).
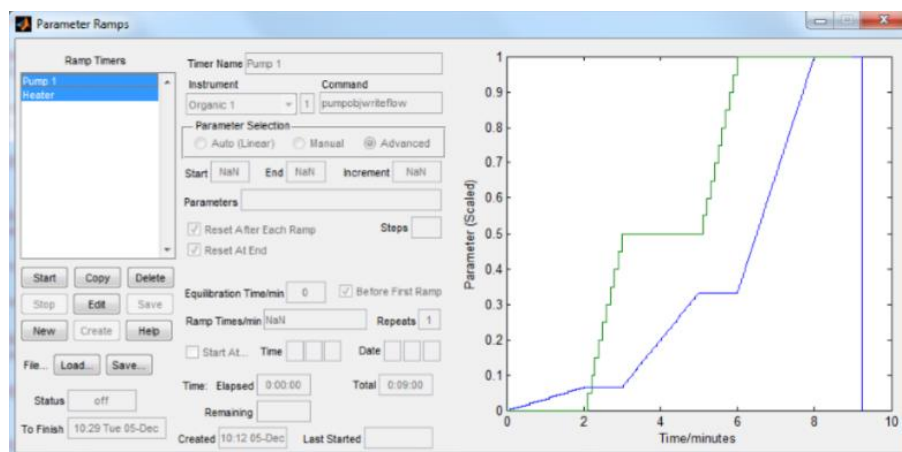
*Figure 57 Ramps interface with user added programs for 'A HPLC pump' and 'Heater' over the course of 9 minutes.*

The 'Optimise' setting provides the user with several options for the optimisation of their reaction by choosing another GUI to add in parameters and reaction details. From this 'Optimise Options' interface (Figure 58) the user has the choice of 9 different programs to choose from: First Version, Super-Mod, Snobfit-HPLC, Snobfit-MS, PSearch-Seg, Extraction, Pareto, IBM Optimisation and Extraction W/ Email. The first two options in the list are the simplest programs to use for optimisation and only run a hybrid simplex/modified simplex or the super-modified simplex. The next two options run the original Snobfit algorithm with an edit in each either for HPLC analysis and data extraction or MS analysis and data extraction. A pattern search algorithm and Pareto algorithm can be chosen as well as the Snobfit algorithm with an adaption to allow optimisation using pump ratios and a separate interface which provides email notification of successful runs or problems.
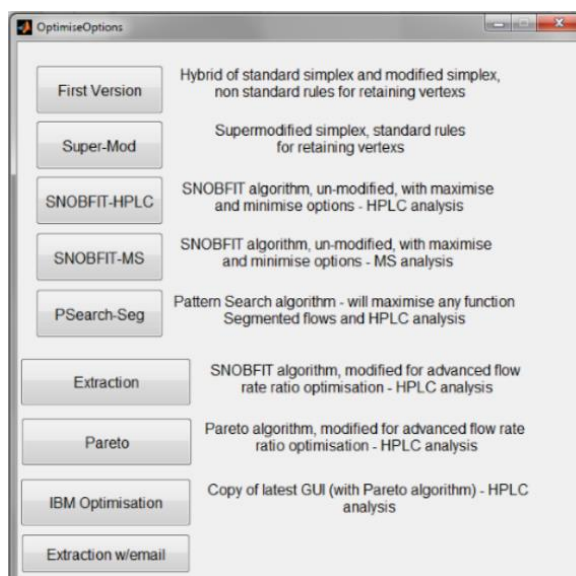
*Figure 58 Algorithm list from the control GUI given in Figure 54. A description of each algorithm is given after each button option.*
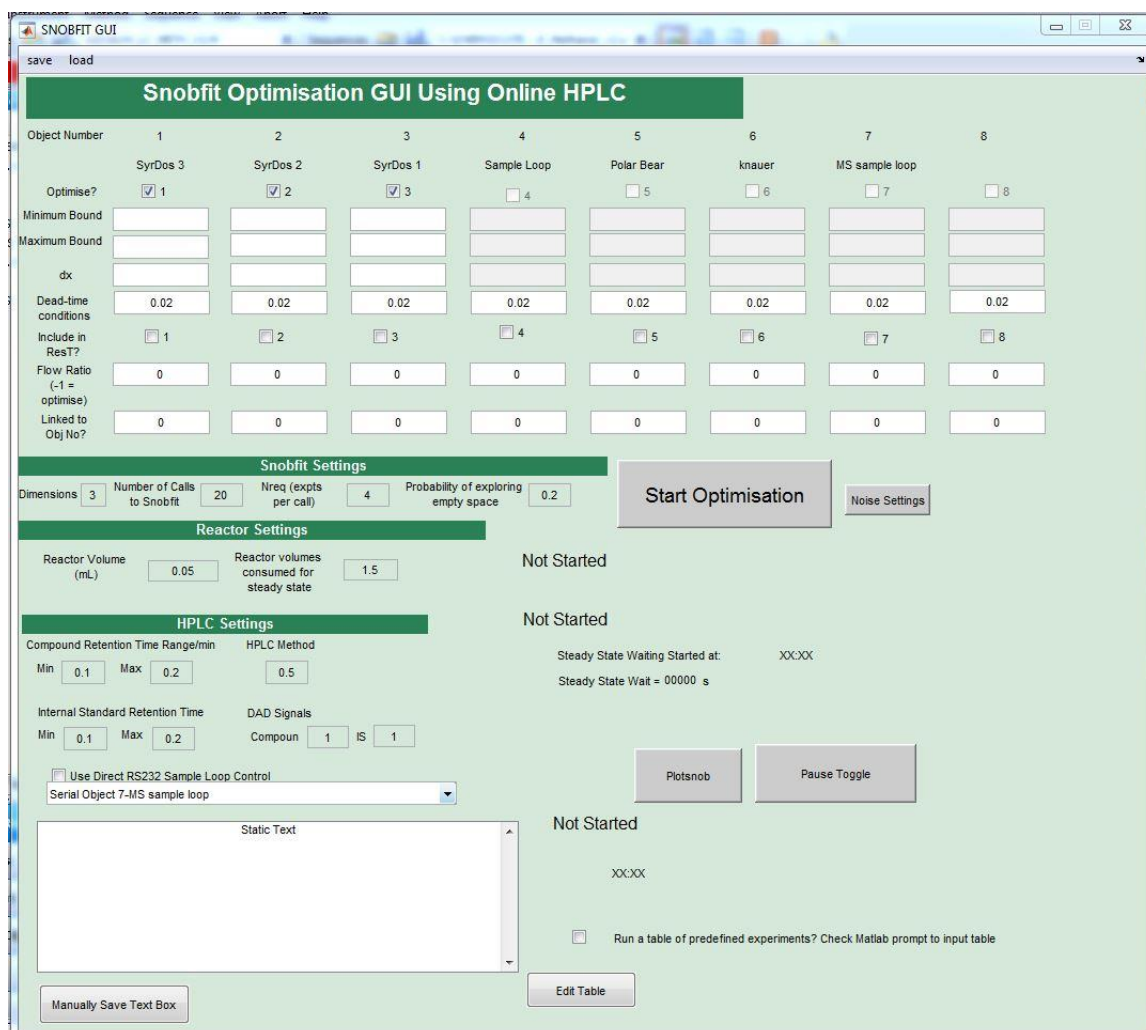


*Figure 59 Snobfit Optimisation GUI interface using HPLC to optimise the reaction.*

The optimisation GUI required user input which is used by the program to set the upper and lower bounds of conditions for optimisation for each peripheral *i.e.* pump, reactor or sample actuator. Conditions for pumps are flowrates, temperature for reactor. Each pump can be added into the residence time calculation or kept out if being used for dilution/quench of the reaction stream. Dead-time conditions are those which the equipment assumes when a sample is running to minimise material loss but keep the flow constant. The Snobfit setting can be changed here if further dimensions need to be examined. 'Number of calls to Snobfit' is the total number of times the code is cycled through in an optimisation with 'Nreq' being the number of experiments per call. Algorithm exploration can be changed if more exploration is needed to balance with Snobfit's exploitive capabilities. Other setting can also be changed here such as the reactor settings and the HPLC settings. If RS232 control of the sample loop is needed, then the box 'Use Direct RS232 Sample Loop Control' must be selected and the correct injector selected if multiple exist.

If running a design of experiments the experiment list is first created in Modde along with range of conditions. These can be added as a lowest to highest *e.g.* 30-150 °C or as a list of set conditions *e.g.* 30, 70, 110, 150 °C. Once the list is generated it can be copied and pasted into a MATLAB table file ready for the program to read. There is a selection box in the Snobfit GUI which sets the program to use a predefined list of conditions. This should be selected when performing a DoE which will allow the user to load the MATLAB conditions file. Subsequently the system will the cycle through the list and carry out each set of conditions, run HPLC analysis and save the data ready for extraction. Full copies of code used to perform DoE, self-optimisation and automated continuous flow reactions can be found in Appendix I.

### 2.7.4. Continuous Flow Setup
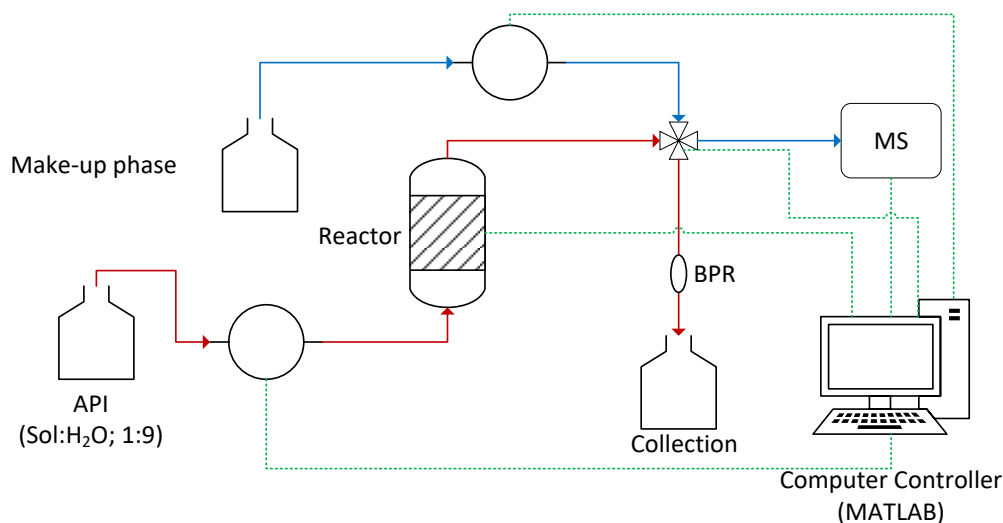


*Figure 60 Schematic of the first continuous flow setup used for the N-Boc deprotection of AZD5634 API using a single HPLC pump with the API in minimal solvent and $H_2O$.*

Continuous flow work was completed as per Figure 48 using set up as shown in the schematic in Figure 60 using a single HPLC pump (Jasco PU-980). A Jasco PU-980 pump was also used to pump the mobile phase for the mass spectrometer (Advion Expression CMS) system from a stock solution. The sample loop employed was a Vici Valco 4-port micro-volume sample injector with tubular reactor (Cambridge Reactor Design Polar Bear Plus Flow Synthesiser) was used.

When one pump is used the reagent stream flow directly into the reactor then the sample loop and waste. For a three-pump system the SM stream is mixed with the correct percentage of $H_2O$ before being diluted by solvent. The stream the flows through the reactor, sample loop and to waste. The sample loop switches a portion of the reaction to the analysis stream where it is taken to the MS.

The first experimental setup used for AZD5634 only employed a single pump to feed the reagent API in a solution of solvent and $H_2O$ which was determined to be the best environment for the reaction to occur in.
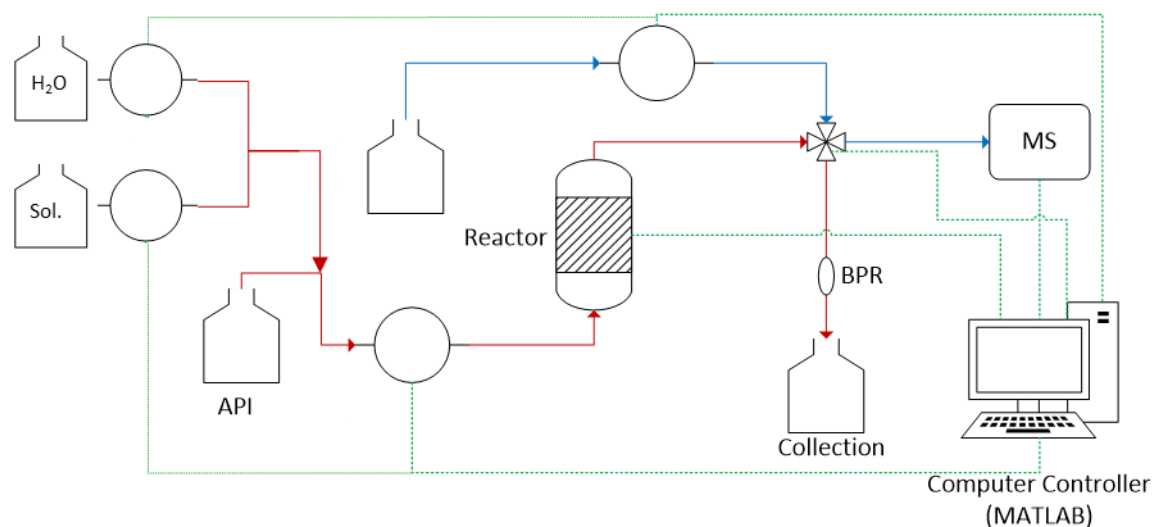
*Figure 61 Schematic of the second continuous flow setup used for the N-Boc deprotection of AZD5634 API using a three HPLC pumps with the API in solvent, solvent make up and $H_2O$.*

The expanded DoE to include $H_2O$ percentage as a variable used a three-pump system where the SM was mixed with the correct amount of $H_2O$ before being diluted to the correct concentration by solvent (Figure 61).

For both versions of DoE, the reagent stream was passed through the reactor before flowing to the sample loop and waste. After injection by the sample loop the parallel flow stream flowed the reaction slug into the on-line MS for analysis.

### 2.7.5.   Microwave Studies

All microwave experiments were performed in batch in either 10 mL or 35 mL vials, sealed with a PTFE and silicon rubber lid using a CEM Discover SP Microwave Synthesizer with autosampler. The microwave was kept in constant power mode using 200 W of microwave energy to heat the sample to the required temperature. The CEM microwave keeps the reaction under pressure to achieve temperatures above those of the solvents used. Once at the required temperature the sample would be left to reaction for the required time before being cooled and ejected by the autosampler and analysed (**2.7.1**).

#### *2.7.5.1. N-Boc Deprotection in IPA and Batch*

A solution of compound AZD5634 was made by dissolving the compound (0.05 g, 0.13 mmol, 1 % w/w) in IPA (0.5 mL), sonication was used to help dissolution, and then

adding $H_2O$ (4.5 mL). Samples were prepared in a total of 5 mL of solvent. The reaction was left for 15 mins at 150 °C (**2.7.1**). Mass spectra were obtained by manual injection with the instrument in APCI+ mode, and mobile phase of MeOH with 1 % TFA. Mass Spectra: 279 *m/z*, 323 *m/z* and 379 *m/z* (279 *m/z* required, found 279 *m/z*).

### *2.7.5.2. Batch N-Boc Deprotection in IPA at 10 % w/w*

A 10 % w/v solution of AZD5634 was prepared in a flask using the starting compound (0.5 g, 0.13 mmol)in solvent (5 mL); IPA and $H_2O$ in a ratio of 10:90. The reaction was left for 20 mins at 150 °C. Mass spectra were obtained by manual injection with the instrument in APCI+ mode, and mobile phase of MeOH. Mass Spectra: 279 *m/z*, 323 *m/z* and 379 *m/z* (279 *m/z* required, found 279 *m/z*).

### *2.7.5.3. Batch N-Boc Deprotection in IPA at 10 % w/w*

A 10 % w/v solution of AZD5634 was prepared in a flask (0.1 g, 0.025 mmol) the starting compound in solvent (1 mL); IPA and $H_2O$ in a ratio of 10:90. The reaction was left for 20 mins at 120, 130, 140 & 150 °C. Mass spectra were obtained by manual injection with the instrument in APCI+ mode, and mobile phase of MeOH. Mass Spectra: 279 *m/z*, 323 *m/z* and 379 *m/z* (279 *m/z* required, found 279 *m/z*).

### *2.7.5.4. N-Boc Deprotection in MeTHF in Batch*

A solution of compound AZD5634 was made by dissolving the compound (0.05 g, 0.13 mmol, 1 % w/w) in MeTHF (0.5 mL), sonication was used to help dissolution, and then adding $H_2O$ (4.5 mL). Samples were prepared in a total of 5 mL of solvent. The reaction was left for 20 mins at 150 °C. Mass spectra were obtained by manual injection with the instrument in APCI+ mode, and mobile phase of MeOH with 1 % TFA. Mass Spectra: 279 *m/z*, 323 *m/z* and 379 *m/z* (279 *m/z* required, found 279 *m/z*).

### 2.7.6.  Mass Spectra Method

Mass spectra were gathered in APCI+ mode from *m/z* 25 to 550 using a scan time of 875 ms and an acquisition time of 40 seconds. Other settings of the MS are summarized below. The mobile phase used was 100% MeCN which was pumped at 0.3 mL min$^{-1}$.

Polarity: Positive,

Capillary Temperature: 250 °C

Capillary Voltage: 180 V

Source Voltage Span: 30 V

Source Voltage offset: 20 V

Source Gas Temperature: 350 °C

APCI Current: 5 µA

### 2.7.7. Flow Feasibility Studies

#### 2.7.7.1. Flow N-Boc Deprotection

A 1 % w/v solution of AZD5634 was prepared in a flask using the starting compound (1 g, 2.6 mmol) in solvent (100 mL); MeOH and $H_2O$ in a ratio of 10:90. Reactor metrics were manually programmed into the flow system as shown in Table 11. Mass spectra were obtained with the instrument in ESI+ mode, and mobile phase of MeOH with 1 % TFA, and a sample volume of 0.06 µL. Mass spectra: 279 *m/z*, 323 *m/z* and 379 *m/z*

*Table 11 Reactor metrics for the deprotection of AZD5634 in continuous flow.*

| Run | Temperature/ °C | Sample Time/ s | Flow Rate/ mL min$^{-1}$ | Residence Time/ min |
|-----|------|-----|-------|------|
| 1 | 130 | 90 | 0.300 | 10 |
| 2 | 130 | 90 | 0.150 | 20 |
| 3 | 140 | 180 | 0.150 | 20 |
| 4 | 140 | 180 | 0.075 | 40 |

#### 2.7.7.2. N-Boc Deprotection in IPA and Flow

A 1 % w/v solution of AZD5634 was prepared in a flask the starting compound (1 g, 2.6 mmol) and solvent (100 mL); IPA and $H_2O$ in a ratio of 10:90. Reactor metrics were manually programmed into the flow system as shown in Table 12. Mass spectra were obtained with the instrument in APCI+ mode, and mobile phase of MeOH and a sample volume of 0.06 µL. Mass Spectra: 279 *m/z*, 323 *m/z* and 379 *m/z* (279 *m/z* required, found 279 *m/z*).

*Table 12 Table outlining the reactor metrics used for the initial N-Boc deprotection of AzD5634 using IPA as solvent.*

| Run | Temperature/ °C | Residence time/ mins |
|:---:|:---:|:---:|
| **1** | 120 | 10 |
| **2** | 120 | 40 |
| **3** | 140 | 10 |
| **4** | 140 | 40 |
| **5** | 130 | 25 |

### 2.7.7.3. Flow N-Boc Deprotection in IPA at 10 % w/w

A 10 % w/v solution of AZD5634 was prepared in a flask using the starting compound (5 g, 13.2 mmol) and solvent (10 mL); IPA and $H_2O$ in a ratio of 10:90. The 3 mL reactor was heated to 150 °C and the starting solution flowed through at 0.075 mL min$^{-1}$ for a residence time of 40 mins. A think brown oil was produced on the reactor separated by gas bubbles. Initial pressure of the system was 75 psi, however, a new BPR (250 psi) was added. Mass spectra were obtained with the instrument in APCI+ mode, and mobile phase of MeOH, and a sample volume of 0.06 μL. Mass Spectra: 279 *m/z*, 323 *m/z* and 379 *m/z* (279 *m/z* required, found 279 *m/z*).

### 2.7.7.4. Flow N-Boc Deprotection in IPA at 5 % w/w

A 5 % w/v solution of AZD5634 was prepared in a flask using the starting compound (1 g, 2.6 mmol) and solvent (20 mL); IPA and $H_2O$ in a ratio of 10:90. The 3 mL reactor was heated to 150 °C and the starting solution flowed through at 0.075 mL min$^{-1}$ for a residence time of 40 mins. Mass spectra were obtained by manual injection with the instrument in APCI+ mode, and mobile phase of MeOH. Mass Spectra: 279 *m/z*, 323 *m/z* and 379 *m/z* (279 *m/z* required, found 279 *m/z*).

### 2.7.7.5. N-Boc Deprotection in IPA Time Profile

A 1 % w/v solution of AZD5634 was prepared in a flask using the starting compound (0.25 g, 0.066 mmol) and solvent (25 mL); IPA and $H_2O$ in a ratio of 10:90. Reactor metrics were manually programmed into the flow system as shown in Table 13. For each reaction, the reactor was allowed to get to steady state, 1.5x the volume of the reactor (or

1.5x the residence time). Mass spectra were obtained with the instrument in APCI+ mode, and mobile phase of MeOH, and a sample volume of 0.06 µL. Mass Spectra: 279 *m/z*, 323 *m/z* and 379 *m/z* (279 *m/z* required, found 279 *m/z*).

*Table 13 Reactor metrics for the time profile of N-Boc deprotection of AZD5634 in IPA at 1 % w/w.*

| Flow Rate/ mL min$^{-1}$ | T$_{res}$/ mins | Temperature/ °C | SS Time/ mins |
|---|---|---|---|
| **3.000** | 1 | 150 | 1.5 |
| **0.600** | 5 | 150 | 7.5 |
| **0.300** | 10 | 150 | 15 |
| **0.150** | 20 | 150 | 30 |
| **0.100** | 30 | 150 | 45 |
| **0.075** | 40 | 150 | 60 |

### 2.7.7.6. N-Boc Deprotection in IPA at 180 °C

A 5 % w/v solution of AZD5634 was prepared in a flask using the starting compound (2 g, 5.3 mmol) and solvent (40 mL); IPA and $H_2O$ in a ratio of 10:90. The 3 mL reactor was heated to 180 °C and the starting solution flowed through at 0.6, 0.3 & 0.15 mL min$^{-1}$ for a residence time of 5, 10 & 20 mins. Mass spectra were obtained with the instrument in APCI+ mode, and mobile phase of MeOH, and a sample volume of 0.06 µL. Mass Spectra: 279 *m/z*, 323 *m/z* and 379 *m/z* (279 *m/z* required, found 279 *m/z*)

### 2.7.7.7. Steady State Studies in MeOH

The same solution as prepared in 0 was again used in the work to determine at what reactor volume reached steady state. The solution was first allowed to pump for a total of one reactor volume a total of three times with a sample taken for each. The conditions and flow rates (Table 14) set out were those which gave the highest conversion in the previous DoE, 0. This methodology was repeated for two reactor volumes and both sets

of conversions, for HPLC and MS, were plotted to create a steady state curve (Figure 62). Mass Spectra: 279 *m/z*, 323 *m/z* and 379 *m/z* (279 *m/z* required, found 279 *m/z*).

*Table 14 Table of conditions for the steady state validation of the experimental design in 5.17.*

| DoE Reaction | Temperature/ °C | API flow/ mL min⁻¹ | Water flow/ mL min⁻¹ | Solvent flow/ mL min⁻¹ | Steady state vol |
|---|---|---|---|---|---|
| **N10** | 180 | 0.08 | 0.22 | 0.45 | 1 |
| | 180 | 0.08 | 0.22 | 0.45 | 2 |
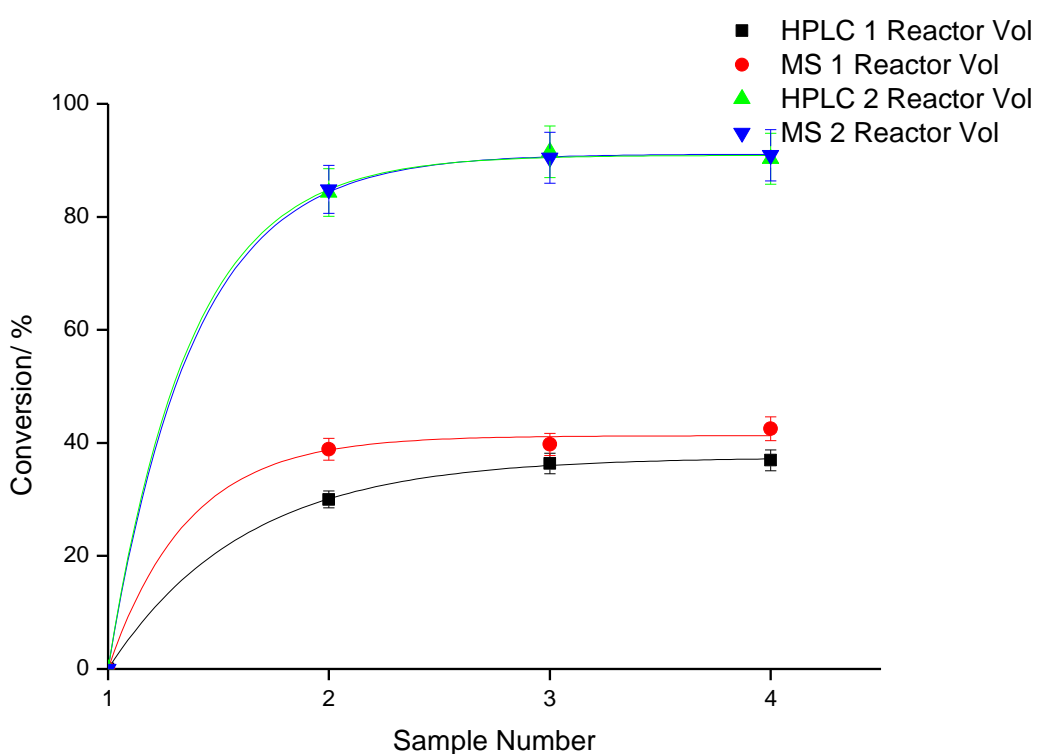


*Figure 62 Graph showing the steady state of the system in the experimental design in 5.17 for both the HPLC data and MS data at 1 and 2 reactor volumes.*

### 2.7.8. DoE Studies

#### 2.7.8.1. DoE N-Boc Deprotection in MeOH

A solution was prepared using the same method used to create the solution as on page 2.7.7.2 but with double the volume used *i.e.* 250 mL. A DoE list of experiments was compiled in Modde using residence time and temperature as variables of interest (Table 15) which were then loaded into the optimising program MATLAB. Two experimental designs were completed using both ionisation (ESI+ and APCI+) sources in the mass

spectrometer. The system set up reflected the schematic shown in and used only a single HPLC pump to house the API reagent in minimal solvent (enough to fully dissolve) and 90% $H_2O$. Both ionisation modes were used as a comparison to determine which performed better for future quantitative studies.

*Table 15 An outline of the experimental design conditions used in the deprotection of AZD5634 in continuous flow.*

| Run Order | Flow Rate/ mL min$^{-1}$ | Residence Time/ min | Temperature/ °C |
|:---:|:---:|:---:|:---:|
| 1 | 0.075 | 40 | 140 |
| 2 | 0.300 | 10 | 140 |
| 3 | 0.300 | 10 | 140 |
| 4 | 0.120 | 25 | 130 |
| 5 | 0.075 | 40 | 120 |
| 6 | 0.075 | 40 | 140 |
| 7 | 0.120 | 25 | 120 |
| 8 | 0.120 | 25 | 130 |
| 9 | 0.120 | 25 | 130 |
| 10 | 0.075 | 40 | 130 |
| 11 | 0.300 | 10 | 130 |
| 12 | 0.300 | 10 | 120 |

### 2.7.8.2.N-Boc Deprotection in IPA DoE

A solution was prepared as with the method from 2.7.7.2 but made up to a volume twice of 250 mL A DoE list of experiments was compiled in Modde (Table 16) and used with the optimising program in MATLAB. The experimental design was completed using the APCI+ ionisation source in the mass spectrometer. Mass Spectra: 279 *m/z*, 323 *m/z* and 379 *m/z* (279 *m/z* required, found 279 *m/z*).

*Table 16 An outline of the experimental design conditions used in the deprotection of AXD5634 in continuous flow in IPA.*

| Run order | Flow Rate/ mL min⁻¹ | Residence time/ min | Temperature/ °C |
|---|---|---|---|
| 1 | 0.075 | 40 | 120 |
| 2 | 0.300 | 10 | 140 |
| 3 | 0.120 | 25 | 120 |
| 4 | 0.075 | 40 | 130 |
| 5 | 0.120 | 25 | 130 |
| 6 | 0.120 | 25 | 130 |
| 7 | 0.300 | 10 | 120 |
| 8 | 0.120 | 25 | 140 |
| 9 | 0.075 | 40 | 140 |
| 10 | 0.300 | 10 | 130 |
| 11 | 0.300 | 10 | 140 |
| 12 | 0.075 | 40 | 140 |

### 2.7.8.3. Experimental Design in MeOH

A solution of AZD5634 (12 g, 200 mg/mL) was prepared by slowly allowing the starting material to dissolve in the solvent (MeOH, 60 mL) before using sonication until full dissolution. A DoE list of experiments was compiled in Modde (Table 17) and used with the optimising program in MATLAB. The experimental design was completed using LCMS with the APCI+ ionisation source in the mass spectrometer. The same solution was used again after the first attempt at the DoE did not get the data required. The DoE repeat used the same table of conditions but made use of two reactor volumes to reach steady state rather than one. Mass Spectra: 279 *m/z*, 323 *m/z* and 379 *m/z* (279 *m/z* required, found 279 *m/z*).

*Table 17 Table outlining an experimental design conducted for the N-Boc deprotection of AZD5634 in MeOH and H$_2$O. *This reaction did not inject due to the test reaction beforehand having the same conditions. A discrepancy in the code caused injection failure as there was not a significant change in the conditions.*

| Order | Reaction | Water/ mL min$^{-1}$ | SM:IS/ mL min$^{-1}$ | Dilution (water)/ mL min$^{-1}$ | Temp/°C | Water/ eq | ResT/ min |
|---|---|---|---|---|---|---|---|
| 0 (test) | N15 C | 0.22 | 0.08 | 0.60 | 170.00 | 2.75 | 8.33 |
| 1* | N15 C | 0.22 | 0.08 | 0.60 | 170.00 | 2.75 | 8.33 |
| 2 | N10 | 0.22 | 0.08 | 0.60 | 180.01 | 2.75 | 8.33 |
| 3 | N2 | 0.34 | 0.08 | 0.84 | 177.40 | 4.25 | 5.95 |
| 4 | N4 | 0.16 | 0.08 | 0.48 | 177.40 | 2.00 | 10.42 |
| 5 | N8 | 0.17 | 0.08 | 0.50 | 177.40 | 2.13 | 10.00 |
| 6 | N6 | 0.36 | 0.08 | 0.88 | 177.40 | 4.50 | 5.68 |
| 7 | N11 | 0.45 | 0.08 | 1.06 | 170.00 | 5.63 | 4.72 |
| 8 | N12 | 0.15 | 0.08 | 0.46 | 170.00 | 1.88 | 10.87 |
| 9 | N16 C | 0.22 | 0.08 | 0.60 | 170.00 | 2.75 | 8.33 |
| 10 | N13 | 0.21 | 0.08 | 0.58 | 170.00 | 2.63 | 8.62 |
| 11 | N14 | 0.23 | 0.08 | 0.62 | 170.00 | 2.88 | 8.06 |
| 12 | N17 | 0.22 | 0.08 | 0.60 | 170.00 | 2.75 | 8.33 |
| 13 | N12 | 0.15 | 0.08 | 0.46 | 170.00 | 1.88 | 10.87 |
| 14 | N7 | 0.17 | 0.08 | 0.50 | 162.60 | 2.13 | 10.00 |
| 15 | N1 | 0.34 | 0.08 | 0.84 | 162.60 | 4.25 | 5.95 |
| 16 | N3 | 0.16 | 0.08 | 0.48 | 162.60 | 2.00 | 10.42 |
| 17 | N5 | 0.36 | 0.08 | 0.88 | 162.60 | 4.50 | 5.68 |
| 18 | N17 | 0.22 | 0.08 | 0.60 | 170.00 | 2.75 | 8.33 |
| 19 | N9 | 0.22 | 0.08 | 0.60 | 159.99 | 2.75 | 8.33 |
| 20 | N17 C | 0.22 | 0.08 | 0.60 | 170.00 | 2.75 | 8.33 |

### 2.7.9. Calibration

#### *2.7.9.1.Calibration of Starting Material*

Samples of AZD5634 starting material were prepared in the concentrations (% w/w) of 5, 10, 15, 50 & 100 and manually fed into the MS where 3 samples of each solution were taken. The total abundance for each peak mass (379) were averaged and plotted; Figure 63. Mass spectra were obtained with the instrument in APCI+ mode, and mobile phase of MeCN and a sample volume of 0.06 µL. Mass Spectra: 323 *m/z* and 379 *m/z* (379 *m/z* required, found 379 *m/z*).



| Model | Asymptotic1 | | |
|---|---|---|---|
| Equation | y = a-b*c^x | | |
| Reduced Chi-Sqr | 5.03762 | 2.28297 | |
| Adj. R-Square | 0.9756 | 0.989 | |
| | | Value | Standard Error |
| 1. Average MS 379.7 | a | 3.69231 | 1.63742 |
| | b | 3.87248 | 1.58556 |
| | c | 0.99086 | 0.00531 |
| 2. Average MD 379.7 | a | 4.03445 | 1.27133 |
| | b | 4.21482 | 1.23409 |
| | c | 0.99124 | 0.00359 |
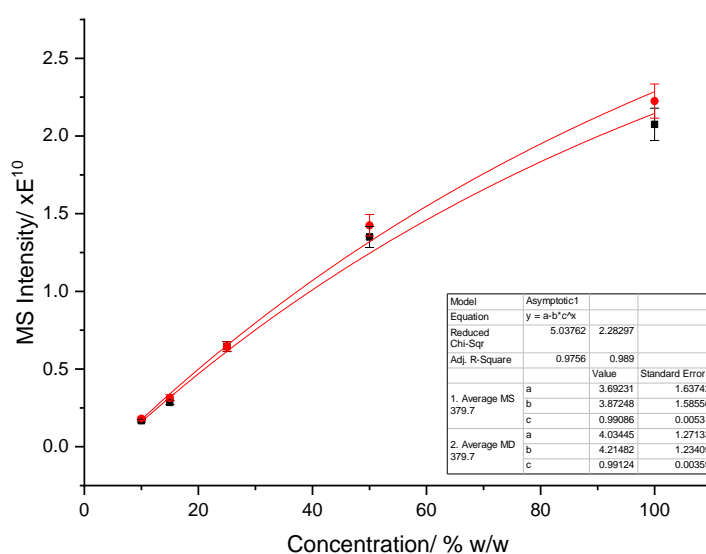
*Figure 63 Calibration graph of AZD5634 starting material; MS intensity (xE$^{10}$) vs concentration (% w/w).*

#### *2.7.9.2. Calibration of Product*

Samples of AZD5634 starting material were prepared in the concentrations of 50%, 80%, 90% & 100% and manually fed into the MS where 3 samples of each solution were taken. The total abundance for each peak mass (279) were averaged and plotted; Figure 64. Mass spectra were obtained with the instrument in APCI+ mode, and mobile phase of MeCN and a sample volume of 0.06 µL. Mass Spectra: 279 *m/z* (279 *m/z* required, found 279 *m/z*).
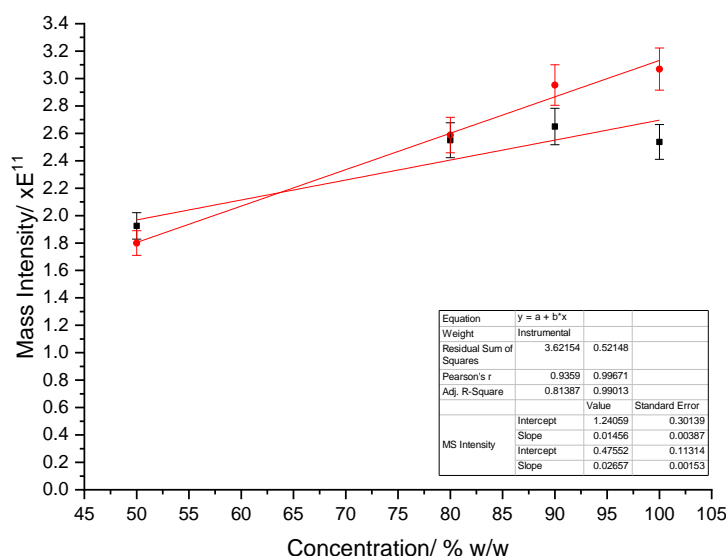
*Figure 64 Calibration graph of AZD5634 product; MS intensity ($xe^{11}$) vs concentration (% w/w).*

### 2.7.10. NMR Reaction Monitoring of AZD5634 by AZ

A mixture of the starting material was prepared in 5:1 IPA/$H_2O$ in a Young's pressure 5 mm NMR tube. The NMR probe was heated to the required temperature and allowed to thermally equilibrate for 5 minutes and the sample was shimmed repeatedly in order to achieve the best possible line shape. The spectrometer was set to acquire a 1H spectrum [ns = 32; aq = 3.17 s; p1 = 5.5 µs; d1 = 1.5 s] and a solvent supressed spectrum [ns = 32; aq = 0.82 s; p1 = 5.5 µs; d1 = 3.0 s] with two suppression points to supress the methyl and OH resonances for IPA; every 5 minutes for 3 hours then every 20 minutes for 24 hours.

### 2.7.11. Kinetics Studies in IPA

The kinetics study for the N-Boc deprotection of AZD5634 used the last remaining volume of the solution, which was diluted from the original 200 mg mL$^{-1}$ to 100 mg mL$^{-1}$. The solution was pumped through the reactor using the conditions given in Table 18. As the solution was already diluted, a dilution pump was not used before the HPLC-MS. The first two temperatures used were 185 °C and 165 °C at the concentration stated. For the third and final temperature, 175 °C, the solution was diluted again from 100 mg mL$^{-1}$ to 50 mg mL$^{-1}$ and pumped through the reactor using the conditions in Table 18. Mass Spectra: 279 *m/z*, 323 *m/z* and 379 *m/z* (279 *m/z* required, found 279 *m/z*).

*Table 18: Conditions for the kinetic study carried out for the N-Boc deprotection of AZD5634.*

| Concentration/ mg mL$^{-1}$ | Temperature/ °C | Starting Material Flow/ mL min$^{-1}$ | Water Flow/ mL min$^{-1}$ | Residence Time/ mins |
|---|---|---|---|---|
| 10 | 185 | 0.59 | 2.97 | 0.983146067 |
| 10 | 185 | 0.09 | 0.44 | 6.603773585 |
| 10 | 185 | 0.88 | 4.38 | 0.66539924 |
| 10 | 185 | 0.14 | 0.68 | 4.268292683 |
| 10 | 185 | 0.3 | 1.50 | 1.944444444 |
| 10 | 165 | 0.08 | 0.38 | 7.608695652 |
| 10 | 165 | 0.22 | 1.10 | 2.651515152 |
| 10 | 165 | 0.13 | 0.63 | 4.605263158 |
| 10 | 165 | 0.04 | 0.2 | 14.58333333 |
| 10 | 165 | 0.5 | 2.50 | 1.166666667 |
| 20 | 175 | 0.07 | 0.34 | 8.536585366 |
| 20 | 175 | 0.10 | 0.48 | 6.034482759 |
| 20 | 175 | 0.74 | 3.72 | 0.784753363 |
| 20 | 175 | 0.31 | 1.56 | 1.871657754 |
| 20 | 175 | 0.16 | 0.81 | 3.608247423 |

Chapter 3

**Monitoring, Modelling and Optimising an N-Boc Deprotection Reaction Using a Hybrid DoE Kinetic Model**

## 3. Hybrid model optimisation: Deprotection of an N-Boc Protected Amine in Continuous Flow Monitored by On-line Mass Spectrometry

Chapter 3 will build on previous work completed on the N-Boc deprotection of an amine using high temperatures and water. However, a model compound has been used instead of an industrially relevant species to show the technique is versatile. The work presented outlines the origins of the project as well as detailing problems encountered throughout the project with solutions to each. The chapter will begin with a general introduction to discuss key concepts for the work and uses of the final molecule after N-Boc removal. Work carried out was fully automated, using several PATs and employing a unique experimental design tailored for this process in order to provide several levels of detailed information; statistical, optimal and kinetic.

### 3.1. Introduction

The CPACT[107] (Centre for Process Analytics and Control Technology) feasibility study arose from a collaboration between AstraZeneca and the CPACT centre at the University of Strathclyde. The premise of the project was to optimise a chemical system using a structured design of experiments approach and gain multiple data sets from various analytical techniques. Using this approach, ideally, an optimum operating region in continuous flow can be found as well as enough data to form kinetic profiles ready for scale-up of the chemical system. The first reaction to be studied comprised a turbo Grignard reaction which was unsuccessful due to the degradation of the turbo Grignard reagent when small amounts of water were in the system, from atmospheric moisture. This moisture usually localised at the union joints, tee-pieces or where there was a joint in the flow stream. Degradation arises from the water protonating the Grignard reagent and forming an insoluble magnesium hydroxide species which cannot be reverted back to the active reagent. Due to this repeated problem a new system was found based on previous work completed on N-Boc deprotection reactions (Chapter 2). The previous work shows how the N-Boc protecting group can be cleaved off primary amines in continuous flow using only high temperatures and water as reagent acquiring high conversions. This project was eventually scaled up to large scale (10s of grams) in the

LSL at AstraZeneca for further trials of the API. The process itself is currently being scaled up to the pilot plant for a much greater output of the desired drug AZD5634.

The good forecast for this work provided enough evidence that another study could be carried out on behalf of the CPACT group for their feasibility study, using a new model compound. The chosen compound would need to be large enough and thermally stable to be trialled in continuous flow using mass spectrometry (MS) as a real time analytical technique. After numerous attempts using several compounds (Figure 65) one showed promise in both being detected by HPLC and MS. The said compound (Boc-Arg(Pbf)-OH, **3.4**) was easily accessible to buy from several suppliers thermally stable (except for the N-Boc group) and provided a mass spectrum with a similar mass pattern to the one seen during the work using AZD5634. As several analytical techniques would be used for the design, it was necessary for the compound to be detected multiple ways *i.e.* ionisable groups would be needed and a chromophore would be necessary. Compounds **3.1**, **3.2** and **3.3** were all trialled to be used for the CPACT feasibility study. However, each of these showed much lower thermal stability levels when attempting the thermally deprotect each compound and lack of sufficient ionisation within the MS. It is useful to know that thermal deprotection of compounds is highly dependent on the compound itself. Some compounds may be more useful to deprotect using lower temperatures, water as additive or another method previously outlined in Chapter 2. Due to better thermal stability and easy detection by HPLC and MS compound **3.4** was chosen to proceed with this study.
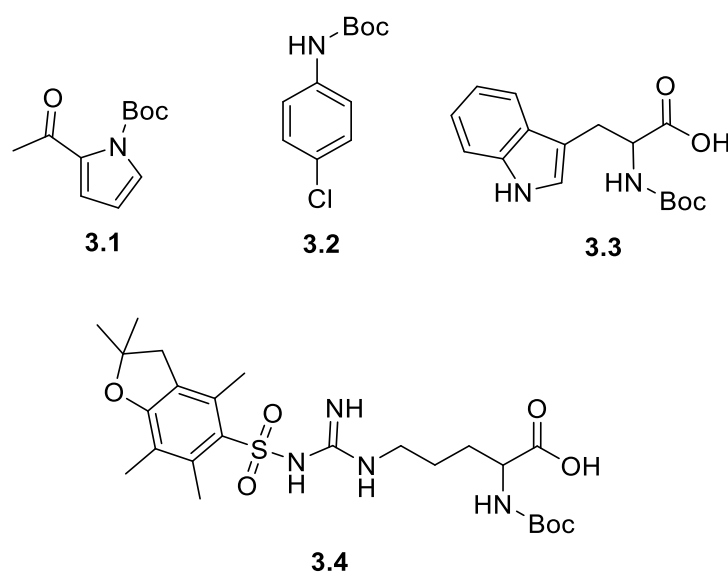


*Figure 65 N-Boc protected compounds trialled to be used for the CPACT feasibility study.*

### 3.1.1. Uses of Boc-Arg(Pbf)-OH and Arginine

Compound **3.4** has been used in the development of nanotechnology-based drug delivery systems where nanometre micelles are formed using methoxy poly(ethylene glycol)-*block*-poly($\varepsilon$-caprolactone) (PEG-PCL).[108] PEG-PCL gives the micelle greater structural integrity enabling prolonged circulation during transportation to the lesion. Upon modification of the PEG-PCL micelle using arginine, a higher payload of drug can be transported into the cancer cells giving the anticancer drug a higher efficacy and overcoming problems seen with overdose of certain anticancer drugs in order to get the required effect from one which may be slowly released.[108]

L-arginine has several uses in the body and can be seen as a prodrug in these circumstances as L-arginine is broken down to nitric oxide (NO) by NO synthases. The breakdown of L-arginine is highlighted below (Scheme 23).[109,110]
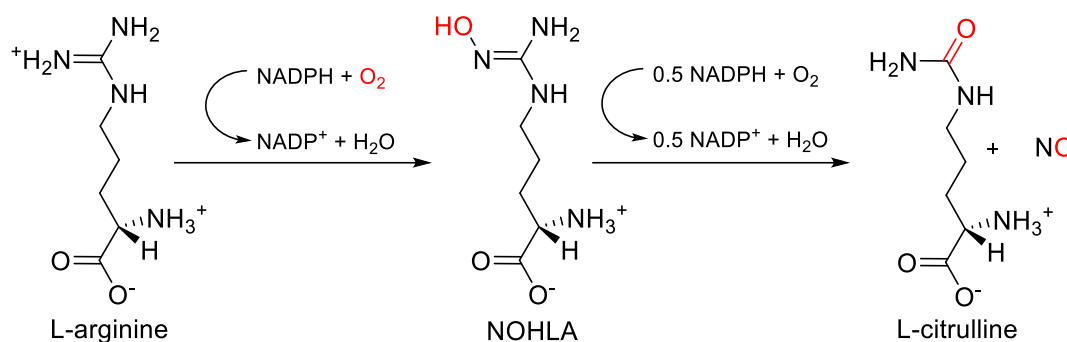


*Scheme 23 Catalysis of L-arginine by NO synthases, in a five-electron oxidation procedure to produce, N$^\omega$-hydroxy-L-arginine (NOHLA) as an intermediate before L-citrulline and nitric oxide (NO).*

Nitric oxide synthases form NO in a five-electron oxidation process of L-arginine. The oxidation initially occurs through two consecutive mono-oxygenation reactions which produced the intermediate species N$^\omega$-hydroxy-L-arginine (NOHLA). The NOHLA then reacts again with another molecule of $O_2$ to produce L-citrulline and NO. Once NO is formed in the body it can be used in the brain acting as a neurotransmitter, immune system for mediating the defence of host cells and cardiovascular system acting as a vasodilator.[111]

This N-Boc deprotection reaction follows the same principle as for AZD5634; using minimal organic solvent, water as reagent and high temperatures (Scheme 24). The high temperatures were achieved with use of a BPR which ensures the entire system is under

constant pressure in order to access a greater range of conditions (typically above solvent boiling points) which are not normally used in laboratory batch chemistry.
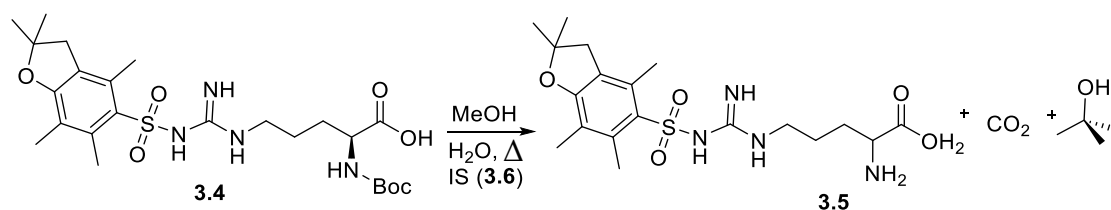


*Scheme 24 Reaction of Boc-Arg(Pbf)-OH using heat and water to remove the N-Boc protecting group in continuous flow.*

The adjoining protecting group 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulphonyl, (Pbf), is one specifically employed to protect the guanidine group of arginine. It is highly TFA sensitive, still stable whilst arginine participates in reactions and can be removed easier than its counterpart 2,2,5,7,8-pentamethylchroman-6-sulfonyl (PMC). It was hoped that the method of N-Boc deprotection would be selective enough to remove only the N-Boc group and not the Pbf group. However, as the Pbf group is also acid sensitive it would allow the model to consider a second potential process and provide a slightly more complex system to model.

### 3.1.2. Optimisation Models

As mentioned in the introduction (Chapter 1) DoE is a useful optimisation technique to model a reaction surface; gather information about an optimum and other statistical data for the model. Most industrial processes require a full data profile of the reaction usually containing different forms of analysis, best conditions, kinetics, impurities and statistical data, to name a few, for health and safety. Most of this data is gathered stage by stage individually which can add a lot of time to the manufacturing process and extra costs.
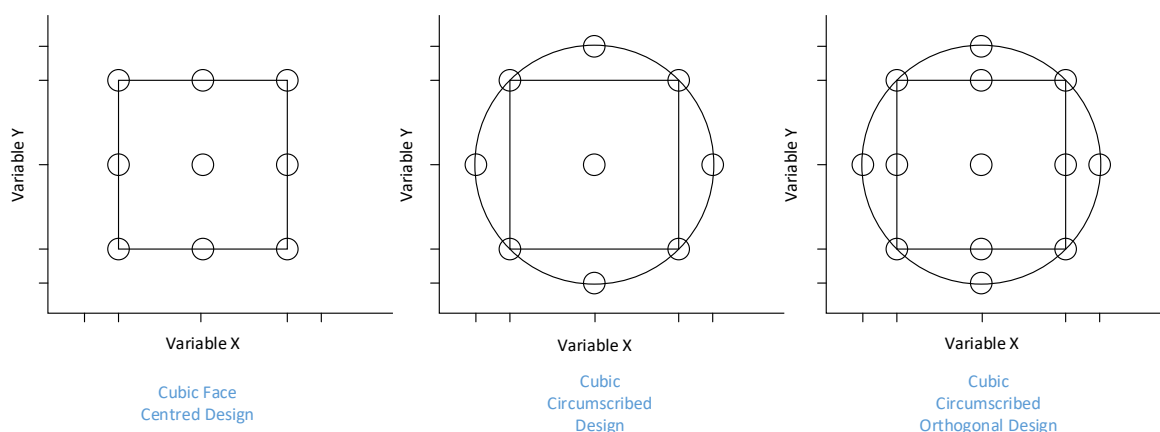
*Figure 66 DoE optimisation models for 2 variables, X & Y. Left; shows a cubic face centred model (FCC), middle; a cubic circumscribed model (CCC) and right; a cubic circumscribed orthogonal design (CCO). Both FCC and CCC span over 3 levels whilst the CCO spans over 5 levels.*

Hybrid DoE models can be used to incorporate multiple sets of data into one model condensing the time taken to gather the data set. Hybridisation allows a process to gather large amounts of information from multiple sources, in this case, combining DoE, response surface methodology and kinetics together using data from HPLC, MS and IR. The DoE model used for this work was a cubic circumscribed orthogonal design (CCO) a combination of cubic face centred (CCF) and cubic circumscribed (CCC). The main differences between the three are; CCF has centrum points, all corner points, and a point in the centre of each face; CCC has centrum points, corner points and circumscribed face points an equal distance away from the centrum as each corner is. The CCO design contains the centrum points, face centred points and circumscribed points essentially creating a cube within a sphere. This means each variable will span over 5 points rather than the typical 3 seen in the other designs (Figure 66).



| Hillside | Peak | Rising Ridge | Saddle |
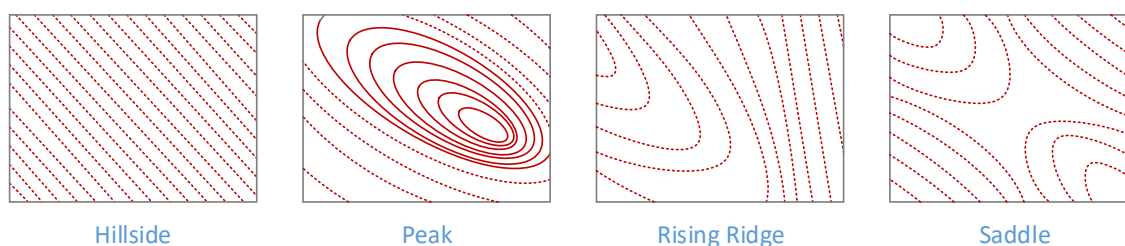
*Figure 67 The four types of response surface outputs given from a response surface methodology design. From the left; "hillside", "peak", "rising ridge" and "saddle".*

A response surface methodology (RSM) design is a type of optimisation design which allows the gathering of a useful surface plot of the desired design space. The experiment is designed to allow estimation of interactions between variables and quadratic effects to provide an overall visual of the space under investigation. RSM designs will find improved or optimal process settings, troubleshoot process problems finding any weak points in the method, and test the robustness of the method against other external, non-controllable influences. Subsequent experimental data addition into the DoE program will allow a surface to be created which makes use of the experimental data and data predicted by the model. In doing so, a surface plot can be created which outlines the area of optimality for the process as given by the model. A set of experimental conditions,

which all should provide an optimum point, are also created ready for the robustness testing. Several types of response surface can be given (Figure 67); A "Saddle" (may also be a "Steep Ridge") type shows a trough area between two peaks which could be an optimum if the sought response requires the lowest value possible *e.g.* cost. A "Hillside" surface may typically show the optimum at either extreme or maybe an area of the surface which is far from the optimum from the given experimental conditions. A "Rising Ridge" surface may entail that the optimum is at the uppermost plateau or focuses on an area of the experimental area leading to an optimum. Finally, A "Peak" (may also be a "trough") type surface show that the optimum is located at the top of the plateau. These four types give a good visual representation of the experimental area.

### 3.2. Summary

- A collaborative study with AZ and CPACT has been performed to develop a large data set using multiple analytical techniques and a hybrid DoE and kinetic motif.
- Thermal N-Boc deprotection of compounds largely depends on the thermal stability of the compound.
- N-Boc protection is applicable to many compounds and is often essential to pharmaceutical development.
- The second protecting group on the amino acid uses similar deprotection conditions as N-Boc so facilitates side-product formation.
- Several DoE models can be used to optimisation reactions, CCF, CCC & CCO. This method uses a CCO design for 5 levels for each variable.
- Several response surface methodologies can be used within DoE to show optimum areas of operation.

### 3.3. Aims

The main purpose of this project was to monitor the N-Boc deprotection of a model compound using a variety of process analytical techniques (HPLC-UV, MS and IR) to build a large database on the reaction. This database would include; analytical, statistical and kinetic information where it would help to find optimum operating conditions and kinetic profiles of the process.

The DoE approach adopted for this process was tailored to put together a central composite design along with ranges of temperatures with set residence times to encompass kinetic experiments.

The model compound chosen would have to be readily available to buy, cheaply and one which would be able to be detected by HPLC-UV, MS and IR. It must also be stable under more harsh environmental conditions during reaction and have a higher thermal stability for MS analysis. Being more thermally stable would allow this process to be effectively monitored by on-line MS and to show further how the MS system can be used to quantitate reactions similarly to HPLC.

As with the previous process completed for AZD5634, this process must be green and environmentally friendly. Green solvents were employed, minimal MeOH and $H_2O$ as well as ensuring a time scale of 30 mins or less for the reaction residence time. This low residence time helps to reduce total time taken if the process to be scaled-up as well as helping to increase the total throughput of reaction in the lowest time possible. Again, using $H_2O$ rather than strong acids reduces the large volumes of base (NaOH) to neutralise and work-up the whole system. Making use of water as a reagent will counter this and potentially remove the necessity of work up, permitting the process to be telescoped. In doing this the free amine will be formed in solution, in a safer and greener manner compared to its batch counterpart (Chapter 2).

## 3.4. The N-Boc Deprotection of a Protected Amine using a Hybrid Model Design of Experiments

Initial work essential for the development of this project, saw the generation of a HPLC method which would be sufficient to separate each component in the reaction with good resolution. Other methods for the on-line analytical instruments would also have to be developed prior to DoE procedure. Unfortunately, tests conducted by the team at AZ using the ReactIR 45m, failed to get a usable method as the chosen molecule (**3.4**) was difficult to characterise with differing volumes of water. The complexity of the species; many NH groups and OH groups as well as the large excess of water in the system inhibited the ReactIR to monitoring specific groups either formed or destroyed in each reaction. Due to this, only HPLC-UV and MS would be used for the reaction system during optimisation *via* DoE.

The method was carried out in LCMS mode to ensure each peak generated in the chromatogram was fully determined. The HPLC chromatogram for reactions shows internal standard (IS) at 2.6 minutes (*m/z* 150), desired product (DP) at 3.7 minutes (*m/z* 427), main (SP) at 4.0 minutes and the starting material (SM) at 5.1 minutes (*m/z* 527) (Figure 68). Characterisation of each peak was completed using the optimisation system by feeding in each reagent through the sample actuator to manually sample, using the developed HPLC method and coupled MS to determine the mass of each species. This technique allowed verification of the IS, DP and SM but not the side products. As discussed previously, the main SP was thought to be the second protecting group on the arginine amino acid. Purification of the waste from all reactions was hoped to get enough of each component to be able to analyse and characterise completely using MS and NMR. Previous work used a UHPLC and accurate MS to determine the structures, but this instrument was not available to use for this project.

Generation of this data was allowed by carrying out the reaction in small scale batch using the microwave reactor. The microwave reactor is used as it is a closer representative of a continuous flow reaction under pressure at high temperature in comparison to the reaction carried out using a typical batch methodology. The reaction in the microwave was carried out at (180 °C for 15 minutes) The starting material was prepared in MeOH and $H_2O$ (80:20) so some product in the reaction could be generated so a full HPLC separation method could be determined.

During the previous study with AZD5634 there were problems encountered with the solubility of the starting material in solution and with the product at certain concentrations. System parameters for this compound would, ideally, match those used for the previous N-Boc study, in particular, a high water content to help facilitate the reaction. To limit the problems of solubility with this system a 1 % (w/v) solution was prepared in MeOH. $H_2O$ was systematically added to a series of the solution from 5 % to 100 % to permit a visual of the Boc-Arg(Pbf)-OH precipitating at its limit of water solubility which was found to be 60 % $H_2O$ compared to MeOH. To limit the chance of the compound precipitating out of solution the maximum percentage of water added to the reaction was restricted to 55% of the total flow rate of starting material.

Once this information was known the reaction was transferred to continuous flow in order to see if it behaved the same as in the microwave reactor and as the previously carried out AZD5634. Several reactions were attempted to get a variety of data points

which show variable conversions of the starting material to the product. The conditions carried out also helped to play a role in determining the correct lower and upper boundaries for creating an experimental design or for self-optimisation of the chemical system.

Again, parameters were selected to run preliminary experiments in continuous flow to check that the set-up was correct; 150, 175 and 200 °C each with 10, 20 and 50% water, each with a 20-minute residence time, giving a total of 9 experiments. These 9 different experiments acted as a screen for the process to ensure the chosen factors were important for the process.

Many attempts were made before the final successful DoE was completed. These repeated attempts were mainly due to hardware issues, code issues and problems with sampling and analysis for prolonged runs at high pressure.

Initial runs saw experiments being sampled when the system reached steady state, but issues were observed upon closer inspection of subsequent HPLC data. Low signals of each peak in the HPLC were seen where the internal standard was barely noticeable above the baseline and the starting material had a total HPLC area of *c.a.* units. This low signal could be put down to two things; either the starting material pump was failing to pump the required flow rate, or the sample loop volume was too small. The easiest fix was changing the sample rotor in the sample loop to a larger volume; 0.06 μL to 0.5 μL, which should give a signal *ca.* 8 times greater than originally seen if this was the only problem for the low signal. The second fix was to check each pump was pumping as required for the DoE model.

Problems seemed to arise from the HPLC pumps when needing them to pump the reaction constituents at low flow rates under higher pressure. It was expected that there would not be a problem with this because HPLC pumps are designed to pump at a range of flow rates, potentially below 0.05 mL min$^{-1}$, at pressure above 30 bar to give good separation and resolution of compounds in a HPLC column. However, the pumps used are old and have been used for many different chemical applications other than pumping pure solvents (as typical for HPLC). This means that the efficiency of each pump may have dropped significantly for the quoted lowest flow rate, 0.001 mL min$^{-1}$. As the lowest flowrate in the DoE 0.0118 mL min$^{-1}$ actions had to be taken to ensure each pump was working effectively. When working correctly each pump builds up a pressure, when pumping against a BPR, usually a little higher than the pressure of the BPR. If not

working correctly the pressure may not build or may slowly decrease over time causing the flow stream to stop flowing. To test each was working, a tap was introduced at the outlet of the pump where, when closed, the pump could pressurise at the low flow rate until it exceeded the set maximum pressure of the pump. If this occurred, then each pump was deemed to be pumping correctly and could be used for the DoE at the desired flow rates.

The third issue arose from use of MS as an on-line analytical technique. The mobile phase used to transfer the reaction sample to the MS was a 50:50 mixture of MeCN and $H_2O$, which is usually effective but, in this case, caused precipitation of the reaction constituents inside the sample loop and MS inlet tube. This problem was easily rectified by changing the solvent from a 50:50 mixture of MeCN and $H_2O$ to 100% MeCN. Tests were conducted first to guarantee that the internal standard, starting material and any products did not precipitate.

The final problem encountered during the first runs was errors caused by the MATLAB coding; 'timer errors'. As the optimisation algorithm is initiated a program timer is set for each section so the code can progress through each stage of the algorithm correctly. These timers are necessary for keeping the algorithm on track but can become misaligned if information is not gathered from each device appropriately. Call-back functions in the algorithm collect information from each device continuously and then match them to the set conditions. If, for whatever reason, the conditions don't match then the set conditions typically override the conditions which are displaying.

There are numerous solutions to this problem which have been developed over the course of the DoE as 'timer errors' were a recurring problem.

This series of code allows all active timers to be loaded, assigned and then restarted by typing each line into the MATLAB command window.

```
A=timerfindall;
A(6);
Start(a(6));
```

The next code is alternate to the one above but works by stopping all timers and restarting them all simultaneously, again by typing each line into the command window of MATLAB.

```
stop(timerfindall('Tag','snobfit_tim'))
set(timerfindall('Tag','snobfit_tim'),'StartDelay',10)
start(timerfindall('Tag','snobfit_tim'))
```

Developed after the final, successful DoE run a code was written which replaces any '0' or 'NaN' values with '1'. This rectifies the program timer where an error is caused by a failed call-back. This section of code was placed within the Snobfit code after a call-back as a check and counter if there are empty cells present.

```
ttt = allowedDeviation>0;
if ttt==0
   allowedDeviation = 1;
else
end
ttt2 = allowedNoise>0;
if ttt2==0
   allowedNoise = 1;
else
end
```

HPLC spectra gathered throughout the experimental design show all necessary species from the reaction. This is due to both starting material and product containing a chromophore, which is UV active, as well as the internal standard. HPLC was also useful to detect the formation of side products during reactions. Figure 68 shows a HPLC chromatogram at conditions 175 °C, 37.5% $H_2O$, 3% w/v over 19 mins. The internal standard can be observed as the first peak to elute in the system followed by the product, side product and finally a very small amount of starting material. Peak identities if IS, DP and SM were verified by HPLC analysis of pure samples using the same method and LCMS analysis. The side product at retention time 4.0 minutes is that of the main side product formed in the reaction, thought to be the second protecting group, from the starting material, cleaved from the amino acid.
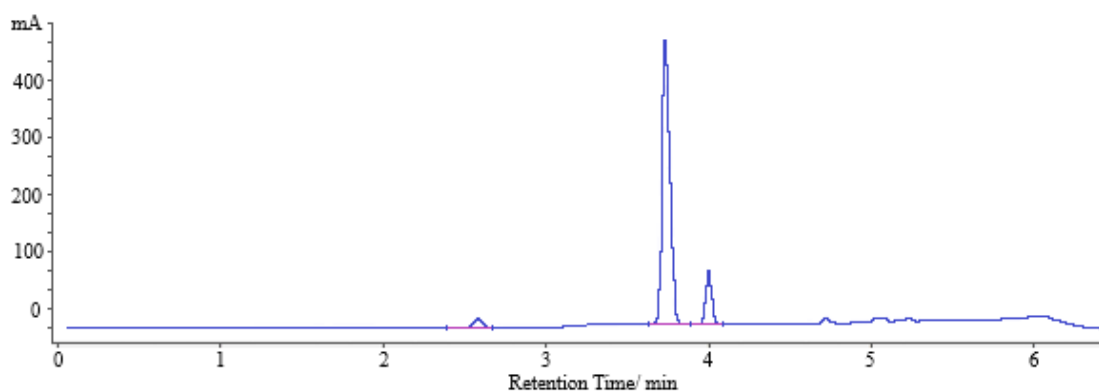
*Figure 68 HPLC chromatogram at conditions 175°C, 37.5 % water, 3% w/v and 19 minutes. The first peak is the IS, second is the DP, third SP. SM appears at 5.17 minutes retention time. Conversion = 92.6%, yield = 68.8%.*

The second HPLC, Figure 69, of note is one were the preceding reaction was carried out at a high temperature over a long residence time (more extreme conditions). In this HPLC the IS can still be seen along with DP, SP and some SM with other SPs. The HPLC show that at high temperatures, large water content, high concentration and a long residence time the reaction proceeds to a larger majority of both DP and SP. Each HPLC was taken at the end of every reaction when it had reached steady state. A total of sixty-nine HPLCs were taken over the course of the DoEs twenty-eight hour run time.
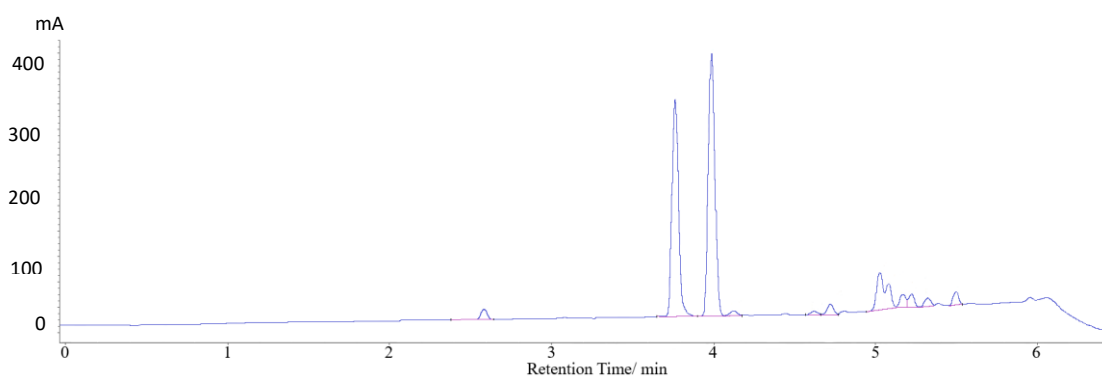


*Figure 69 HPLC chromatogram at conditions 193°C, 50.433 % water, 4.478% w/v and 19 minutes. The first peak is the IS, second is the DP, third SP. SM appears at 5.17 minutes retention time. Conversion = 92.8%, yield = 36.7%.*

Comparatively, the MS spectra associated (closest time stamp) with each of the chromatograms above shows a similar profile for each set of reaction conditions. Each MS was taken as a collective sample of the reaction stream *i.e.* components were not separated, so all species appear together. The mass spectra shown were acquired over 60

seconds with an interval of 15 seconds (a MS taken every 75 seconds) where 1592 spectra were taken throughout the DoE for 28 hours.
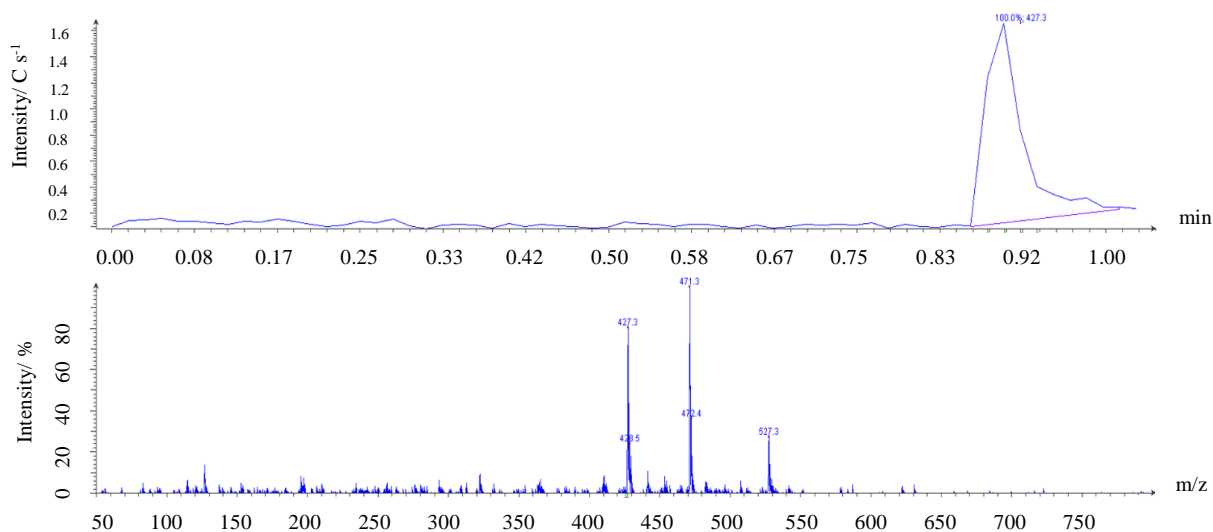


*Figure 70 MS total ion chromatogram (top) and mass spectrum for the single peak at conditions 175°C, 37.5% water, 3% w/v and 19 minutes. DP can be seen at m/z 427, and SM at m/z 527. The peak at m/z 471 is a fragment peak from SM, loss of $^t$Bu group from the N-Boc functional group.*

The total ion chromatogram (TIC) and mass spectrum for reaction conditions 175°C, 37.5% water, 3% w/v and 19 minutes shows peaks for the DP and SM along with SM fragment (loss of $^t$Bu group from N-Boc) similarly to AZD534.

*Table 19 Full set of conditions for the experimental design on the CPACT feasibility study. The DoE has 4 conditions; temperature, water %, concentration and residence time with the first 3 variables at 5 levels and the final variable at 4 level.*

| Expt Name | Run order | Temp/ °C | Water/ % | Conc./ % w/v | Residence time/ Mins |
|---|---|---|---|---|---|
| N1 | 49 | 156.524 | 24.567 | 1.52194 | 1 |
| N2 | 60 | 156.524 | 24.567 | 1.52194 | 7 |
| N3 | 84 | 156.524 | 24.567 | 1.52194 | 13 |
| N4 | 58 | 156.524 | 24.567 | 1.52194 | 19 |
| N5 | 36 | 193.476 | 24.567 | 1.52194 | 1 |
| N6 | 11 | 193.476 | 24.567 | 1.52194 | 7 |
| N7 | 63 | 193.476 | 24.567 | 1.52194 | 13 |
| N8 | 77 | 193.476 | 24.567 | 1.52194 | 19 |
| N9 | 71 | 156.524 | 50.433 | 1.52194 | 1 |
| N10 | 7 | 156.524 | 50.433 | 1.52194 | 7 |
| N11 | 25 | 156.524 | 50.433 | 1.52194 | 13 |
| N12 | 6 | 156.524 | 50.433 | 1.52194 | 19 |
| N13 | 18 | 193.476 | 50.433 | 1.52194 | 1 |

| | | | | | |
|---|---|---|---|---|---|
| **N14** | 73 | 193.476 | 50.433 | 1.52194 | 7 |
| **N15** | 65 | 193.476 | 50.433 | 1.52194 | 13 |
| **N16** | 5 | 193.476 | 50.433 | 1.52194 | 19 |
| **N17** | 78 | 156.524 | 24.567 | 4.47806 | 1 |
| **N18** | 67 | 156.524 | 24.567 | 4.47806 | 7 |
| **N19** | 81 | 156.524 | 24.567 | 4.47806 | 13 |
| **N20** | 19 | 156.524 | 24.567 | 4.47806 | 19 |
| **N21** | 24 | 193.476 | 24.567 | 4.47806 | 1 |
| **N22** | 28 | 193.476 | 24.567 | 4.47806 | 7 |
| **N23** | 55 | 193.476 | 24.567 | 4.47806 | 13 |
| **N24** | 42 | 193.476 | 24.567 | 4.47806 | 19 |
| **N25** | 54 | 156.524 | 50.433 | 4.47806 | 1 |
| **N26** | 38 | 156.524 | 50.433 | 4.47806 | 7 |
| **N27** | 32 | 156.524 | 50.433 | 4.47806 | 13 |
| **N28** | 75 | 156.524 | 50.433 | 4.47806 | 19 |
| **N29** | 23 | 193.476 | 50.433 | 4.47806 | 1 |
| **N30** | 46 | 193.476 | 50.433 | 4.47806 | 7 |
| **N31** | 21 | 193.476 | 50.433 | 4.47806 | 13 |
| **N32** | 13 | 193.476 | 50.433 | 4.47806 | 19 |
| **N33** | 22 | 150 | 37.5 | 3 | 1 |
| **N34** | 20 | 150 | 37.5 | 3 | 7 |
| **N35** | 62 | 150 | 37.5 | 3 | 13 |
| **N36** | 44 | 150 | 37.5 | 3 | 19 |
| **N37** | 33 | 200 | 37.5 | 3 | 1 |
| **N38** | 45 | 200 | 37.5 | 3 | 7 |
| **N39** | 80 | 200 | 37.5 | 3 | 13 |
| **N40** | 27 | 200 | 37.5 | 3 | 19 |
| **N41** | 1 | 175 | 20 | 3 | 1 |
| **N42** | 64 | 175 | 20 | 3 | 7 |
| **N43** | 57 | 175 | 20 | 3 | 13 |
| **N44** | 74 | 175 | 20 | 3 | 19 |
| **N45** | 31 | 175 | 55 | 3 | 1 |
| **N46** | 16 | 175 | 55 | 3 | 7 |
| **N47** | 68 | 175 | 55 | 3 | 13 |
| **N48** | 69 | 175 | 55 | 3 | 19 |
| **N49** | 37 | 175 | 37.5 | 1 | 1 |
| **N50** | 2 | 175 | 37.5 | 1 | 7 |
| **N51** | 30 | 175 | 37.5 | 1 | 13 |
| **N52** | 82 | 175 | 37.5 | 1 | 19 |
| **N53** | 47 | 175 | 37.5 | 5 | 1 |
| **N54** | 83 | 175 | 37.5 | 5 | 7 |
| **N55** | 51 | 175 | 37.5 | 5 | 13 |
| **N56** | 17 | 175 | 37.5 | 5 | 19 |
| **N57** | 61 | 175 | 37.5 | 3 | 1 |
| **N58** | 48 | 175 | 37.5 | 3 | 7 |
| **N59** | 14 | 175 | 37.5 | 3 | 13 |
| **N60** | 35 | 175 | 37.5 | 3 | 19 |
| **N61** | 9 | 175 | 37.5 | 3 | 1 |
| **N62** | 72 | 175 | 37.5 | 3 | 7 |
| **N63** | 76 | 175 | 37.5 | 3 | 13 |
| **N64** | 8 | 175 | 37.5 | 3 | 19 |
| **N65** | 4 | 175 | 37.5 | 3 | 1 |

| N66 | 66 | 175 | 37.5 | 3 | 7 |
| N67 | 70 | 175 | 37.5 | 3 | 13 |
| N68 | 53 | 175 | 37.5 | 3 | 19 |

The full table for the DoE is shown in (*Table* 19). Originally each factor (temperature, water %, and concentration and residence time) had a total of 5 levels giving 85 experiments. However, when calculating the flow rate of each pump, flow rates were lower than thought comfortable for the pumps even after checks (see above). Due to this it was decided that all 25-minute residence times would be taken out of the model to make this a 4-level factor rather than the intended 5-level factor. This reduction was due all pumps not being able to pump reliably at the necessary flow rates (below 0.02 mL min$^{-1}$)

Three pumps were used for each reagent; water, MeOH dilution and starting material with internal standard. The starting material was diluted with MeOH before being mixed together with the required water % to bring it down to the correct concentration and so the total amount of water was exact.

The first model gathered (Figure 71); HPLC conversion, shows a good spread of data points with higher conversions when conditions are more extreme. This matches results gained from the previous case study, AZD5634, where maximum conversions were seen at higher temperatures, longer residence times and a greater volume of water added to the reaction.

The model outline has an $R^2$ value of 0.83 and a $Q^2$ value of 0.72. The value of 83% indicates that the model can explain a high level of the variability in the response data. The higher $R^2$ the more variability in the model can be explained. $Q^2$ should not exceed $R^2$ shows how good the model is at predicting responses. The higher the value then the better predictive ability of the model.

$R^2$ does not necessarily show whether the model is a good or a bad fit. To be able to see this the residuals plot should be examined. The residual of an observed value is the difference between its observed and estimated value. The residuals plot outlines the response *vs* the normal probability of the distribution. A slight "S" shape can be seen which is typical of an underlying trend which is not descried in the model, such as a cubic term, but the effect is relatively minor. This is shown in Figure 71 where all the low conversion points can be seen below the line and all the high conversion points above the line.
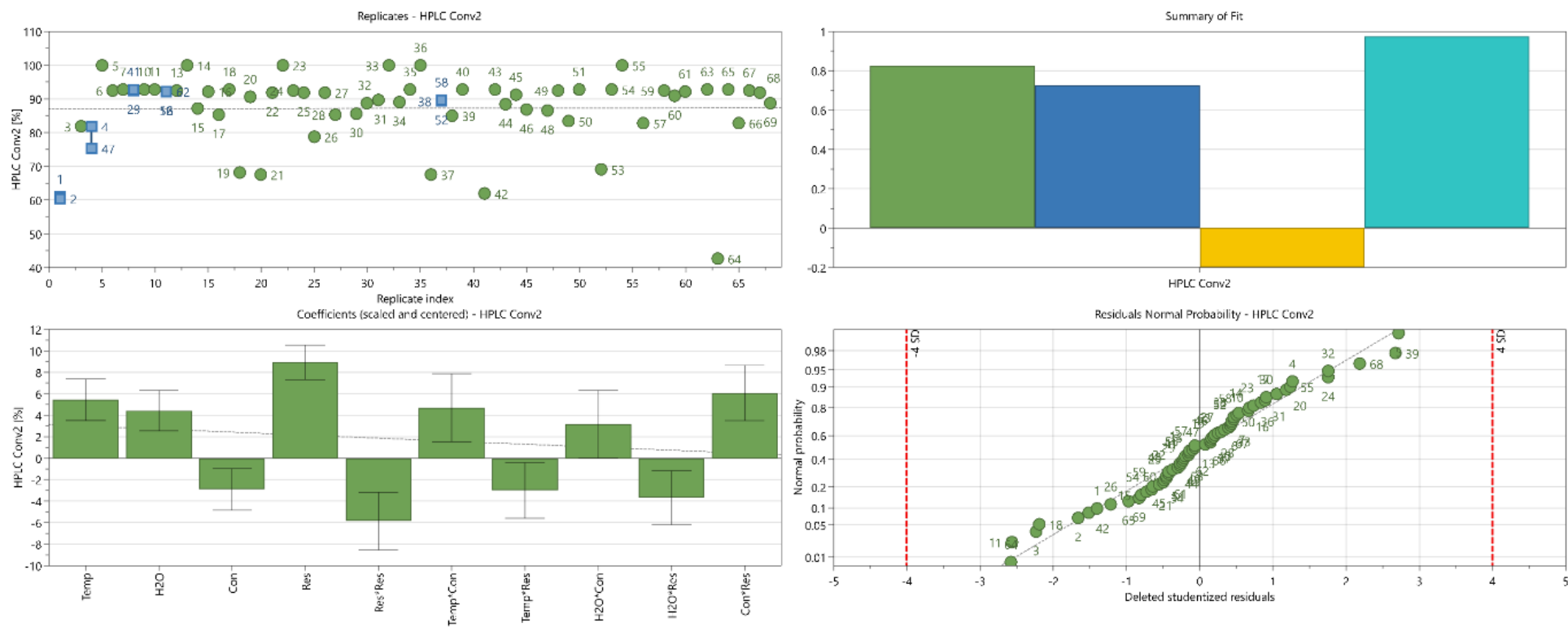
*Figure 71 CPACT feasibility model summary of HPLC data. From top to bottom & left to right; replicates plot, summary of fit, coefficients plot and residuals plot.*

The model itself considers 67 of the 69 experiments as one file was not saved after HPLC acquisition. The other data point was over the set 4 standard deviation lines for outliers. The file in question was a replicate point from the model so can be used without detrimental effects to the model as longs as the rest of the replicate points do not vary much. Entries N57, N61 & N65 correspond to the first set of replicates (1 missing, $S^2 =$ 43.06); N58, N62 & N66 the second set ($S^2 = 0.051$); N59, N63 & N67 the third set ($S^2 = 0.019$) and N60, N64 & N68 the final set ($S^2 = 0.019$). The high variation of the first set of replicates is because one being missing from the data set and the other two having more spread between them. However, neither of these can be taken out of the model because it is uncertain which value is the true value for these conditions. The final, missing point would corroborate which value is correct if it had been available for the model. Furthermore, when attempting to exclude either of these two points from the model there is no significant change in $R^2$ or $Q^2$ suggesting that the three other replicates play a greater role in giving the model the fit which it exhibits.

The model validity, yellow bar, is quite low for this model which could be due to the program attempting to fit its polynomial to points which have very small error bars. As the reproducibility is very high, the error associated in the model will be much smaller *i.e.* small error bars. Trying to fit a model to all point within small error bars can be a challenge which may show a model of low validity even if it is not.

The model shows a good spread of data where the model can predict the surface well from the experimental data (Figure 72). The observed *vs* predicted plot outlines how the predicted data deviates from the experimental information added into Modde. This plot should show all data points spread evenly above and below the central line for a good model where more are localised towards the line and less at the edges as per a Gaussian fit. The actual plot shows this very well with no point being a large distance from the line and more points can be seen localised around the line in comparison to the sparser edges. The coefficients plot shows which variables in the model are significant terms and which are not. The bottom left plot in Figure 71 shows only significant terms as depicted with error bar not over the 0 line. The larger the bar the more effect it had on the response, in this case residence time has the greatest positive effect on the conversion *i.e.* the longer the residence time the higher the conversion.
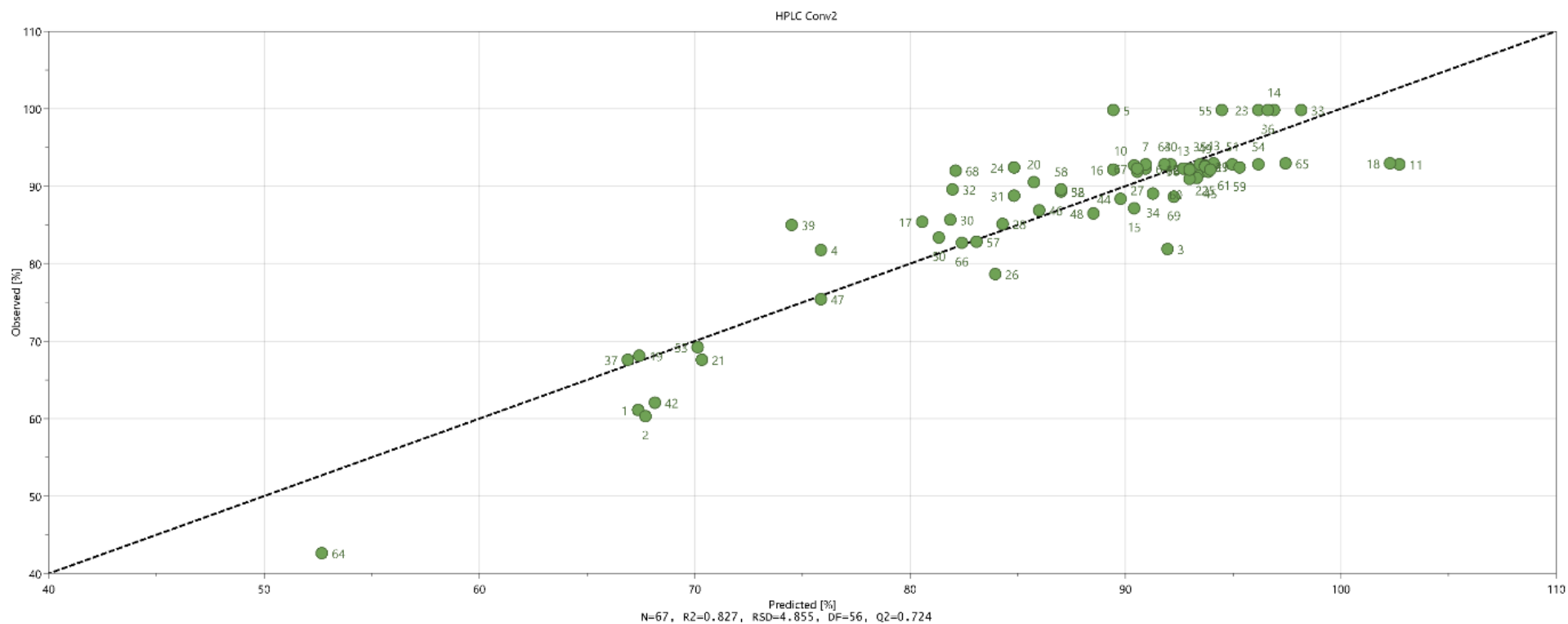
*Figure 72 Observed vs predicted plot for the CPACT feasibility study. The points can be seen to be localised around the regression line, above and below within statistical error.*

This is closely followed by the temperature, water % and concentration which have positive effects on the response. For the concentration bar which can be seen below the line, this indicates that a lower concentration in beneficial for a higher conversion. The model contains a squared term (Res*Res) which shows curvature in the model with respect to residence time, which can also be seen in the surface response and contour plots (Figure 73,Figure 74,Figure 75 andFigure 76 respectively), the higher the residence time the more curved the response surface.

There are also several interaction terms; the main two being between temperature and concentration (significant by a small margin) and between concentration and residence time. At lower temperatures the level of conversion does not significantly affect the conversion. It is only at high temperatures where the higher concentration gives a greater conversion, as you would expect. This is also the same with the second interaction between residence time and concentration. Higher conversions can be seen only when longer residence times are used with a high starting material concentration. With more starting material to react a greater proportion will be converted to products. This relationship may be due to how the acidity of the system changes at the higher concentrations. As there is more starting material containing the N-Boc group, once the reaction begins to occur, more $CO_2$ will be produced and dissolved in the system. As the levels of $CO_2$ rise the acidity of the system will continue to decrease facilitating more N-Boc removal and other $S_N1$ reactions. This theory could be shown in the kinetics plots of the rate of reaction increases for higher concentrations.[112]

The contour plot, Figure 73, indicates the best operating region to get the highest conversion of starting material. In this case the best region localised around the most extreme conditions; highest temperatures, residence times, water % and concentration. It is also possible to get a high conversion (*ca.* 90%) by carrying out the reaction at low temperature, high residence time but with low concentration and water %. This region can be seen in the top left corner of the bottom left plot. These conditions may be more favourable if other side reactions begin to compete with the DP. The contour plot can be represented by a surface plot to further increase understanding and give a more visual representation of the surface in three dimensions.
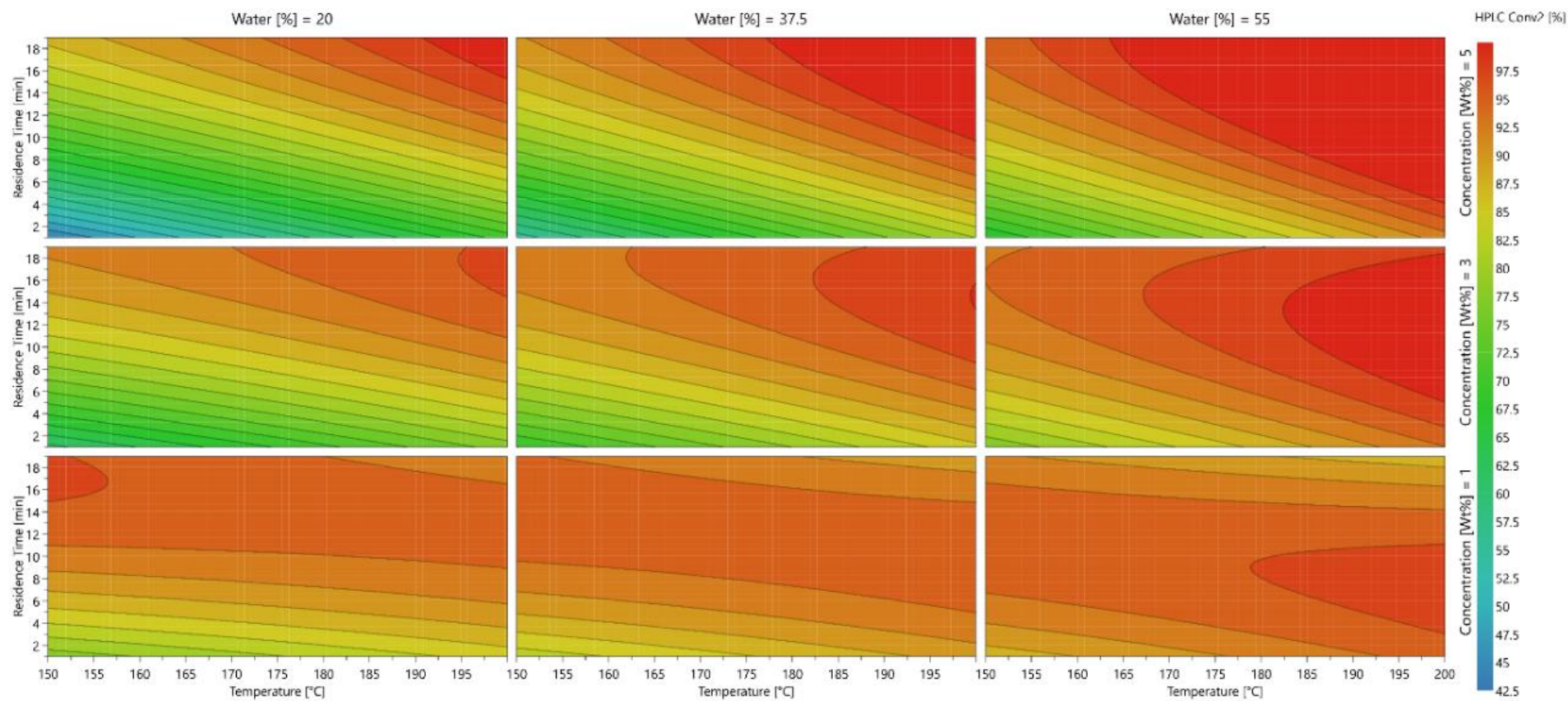
*Figure 73 4D response contour plot for the CPACT feasibility study conversion. The red area shows conditions which can get 95% or more conversion of starting material.*

The first two plots in each row shows a typical horse saddle shape with a large curvature in the z-axis (residence time, Res*Res). As the conditions are made more extreme the plots change to a rising ridge and form a peak. Even with a model of conversion and a found optimum area, this does not mean that the yield of DP is high. The ideal outcome of a reaction is to have full conversion of SM and 100% yield to the DP with no side products forming. The reaction and DoE carried out allows models of the conversion of reaction and the yield of DP formed with best conditions for each. The next model outlines, in the same way as the model for conversion, fit of experimental data, predicted data and a contour plot to visually represent the yield profile.

From HPLC data the yields can be seen to increase as both the temperature and residence time increase. However, during the reaction, specifically at higher temperatures over longer residence times, there is a greater prevalence for the reaction to form other side-products causing the desired product yield to decrease. This causes a change in the model to find the most optimum conditions for the formation of the desired product. Due to there being a greater propensity of the reaction to over react at higher temperatures the region of optimisation is reduced to conditions that are less harsh *i.e.* lower temperatures. This is shown in the contour plot of the DoE model where the optimum region, shown in red, is found localised around the central region (Figure 73).

Based on this optimum region and contour plot, the program used (Modde) can run a test simulation to output a list of potential conditions which should get the stated output conversion and yield. It can also be used to perform a robustness test on the list to provide a set of conditions which it thinks should get to the highest conversion and highest yield reproducibly over many tried.

Table 20 shows the list, determined by Modde, should reach the required outputs and the final result in the table, given by an *R* are the conditions which have been robustness tested.

All the surface plots for the CPACT N-Boc deprotection reaction can be seen in Figure 74, Figure 75 and Figure 76. These show how the reaction surface changes when each factor changes. It can be seen that at the lower extreme conditions a clear plateau is reached with a defined optimum region. However, it can also be seen that as the conditions become more extreme the defined plateau disappears and become much more linear in shape where the optimum region is at the very corner of the surface plot.
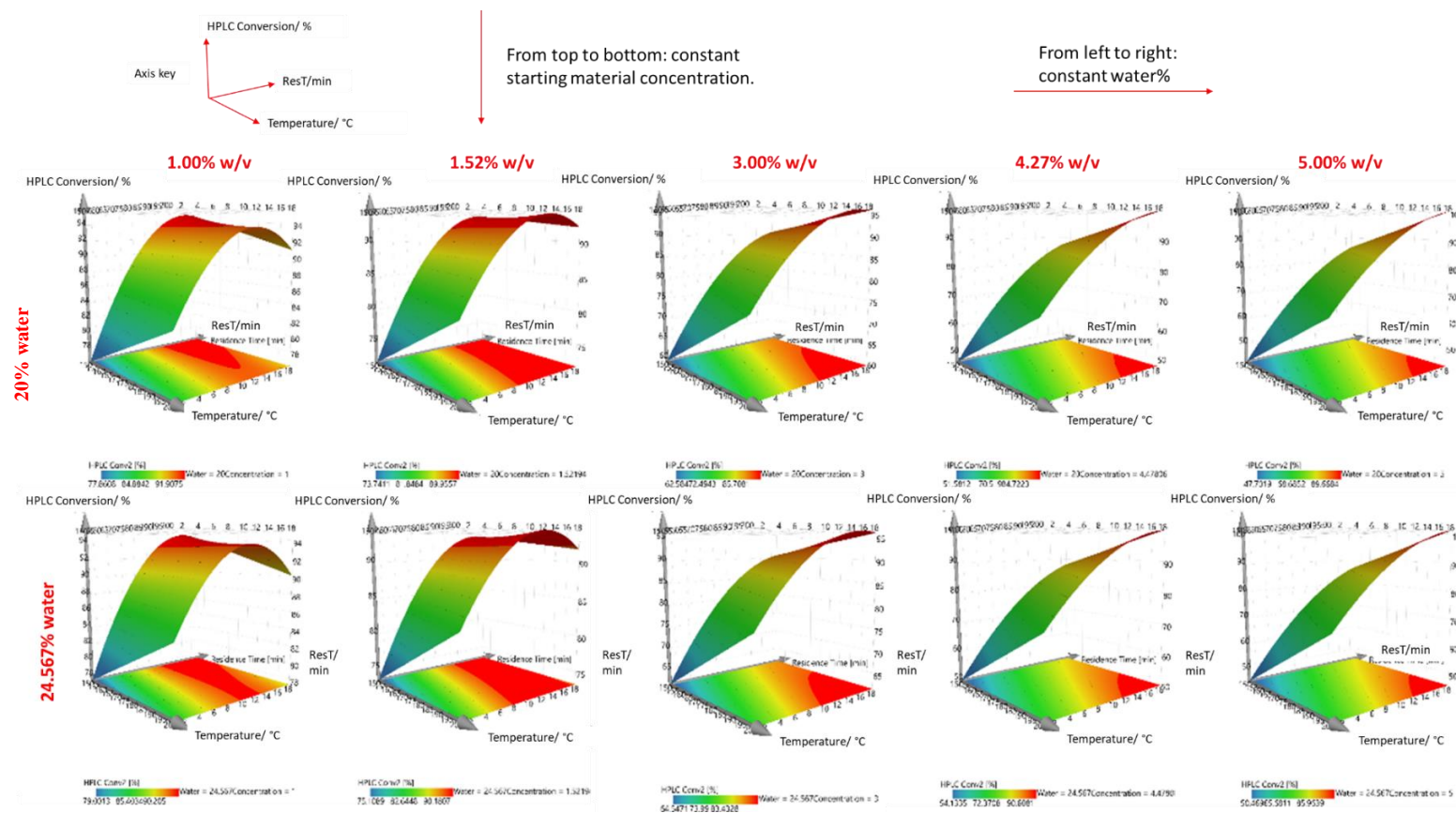
*Figure 74 Surface response plots for the CPACT feasibility study for total conversion of the SM. The y-axis shows conversion, x-axis shows temperature and z-axis residence time. Top row: constant H₂O at 20% and changing concentration (1, 1.52, 3, 4.27 and 5% w/v). Bottom row: constant H₂O at 24.567% and changing concentration (1, 1.52, 3, 4.27 and 5% w/v) for each column entry.*
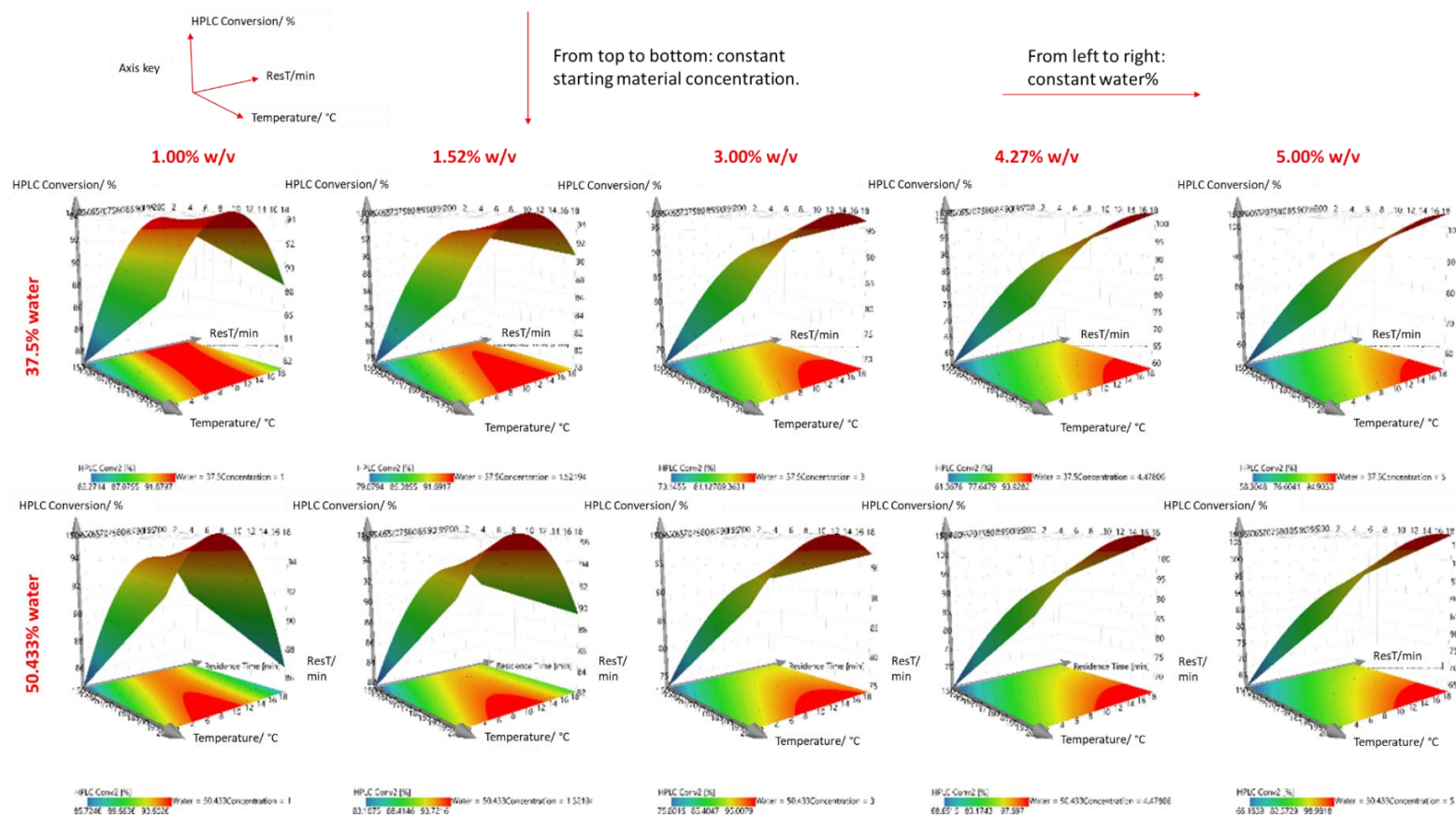
*Figure 75 Surface response plots for the CPACT feasibility study for total conversion of the SM. The y-axis shows conversion, x-axis shows temperature and z-axis residence time. Top row: constant H₂O at 37.5% and changing concentration (1, 1.52, 3, 4.27 and 5% w/v). Bottom row: constant H₂O at 50.433% and changing concentration (1, 1.52, 3, 4.27 and 5% w/v) for each column entry.*
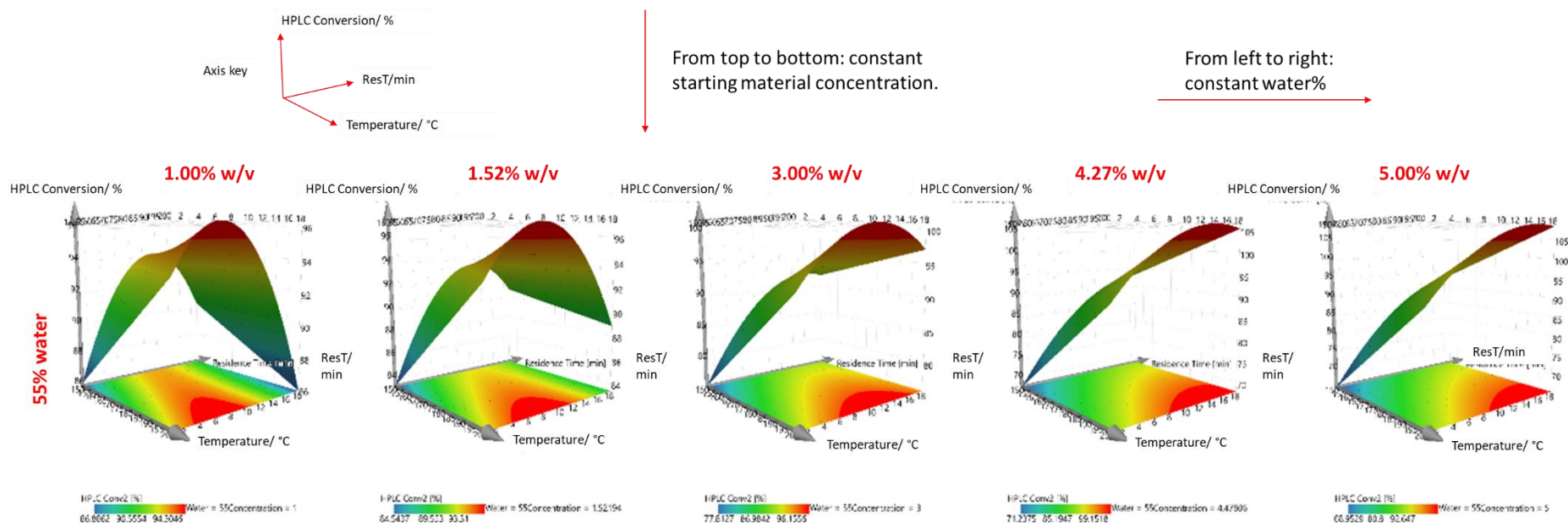
*Figure 76 Surface response plots for the CPACT feasibility study for total conversion of the SM. The y-axis shows conversion, x-axis shows temperature and z-axis residence time. Across the row shows constant H₂O at 55% and changing concentration (1, 1.52, 3, 4.27 and 5% w/v) for each column entry.*

*Table 20 Outline of all the conditions which the program Modde suggest meet the requirements to get the stated output yield and conversions for the N-Boc deprotection reaction studied. The final entry in the table are robustness tested conditions; 170 °C, 41 % water, 3.4 % w/w and 11.8 minutes.*

| Temperature / °C | Water / % | Concentration / % w/v | Residence Time/ min | HPLC Yield | HPLC Conv | Iterations |
|---|---|---|---|---|---|---|
| **161.42** | 54.897 | 4.99841 | 18.3276 | 92.325 | 96.746 | 309 |
| **160.479** | 54.968 | 5 | 18.9685 | 92.434 | 96.615 | 301 |
| **159.151** | 54.542 | 4.97093 | 18.9491 | 92.306 | 96.089 | 280 |
| **159.911** | 54.999 | 4.98061 | 18.992 | 92.413 | 96.429 | 219 |
| **159.981** | 54.795 | 4.99283 | 18.9954 | 92.4 | 96.426 | 210 |
| **159.099** | 54.91 | 4.99574 | 18.7905 | 92.375 | 96.167 | 203 |
| **159.974** | 54.857 | 4.99954 | 18.9207 | 92.409 | 96.437 | 185 |
| **158.967** | 54.996 | 4.99964 | 18.9918 | 92.425 | 96.191 | 183 |
| **160.107** | 55 | 4.9289 | 19 | 92.338 | 96.395 | 173 |
| **160.02** | 55 | 5 | 18.7911 | 92.413 | 96.461 | 166 |
| **161.388** | 54.888 | 4.98695 | 18.9579 | 92.391 | 96.832 | 159 |
| **163.785** | 54.965 | 4.87866 | 18.8686 | 92.151 | 97.301 | 158 |

| | | | | | | |
|---:|---:|---:|---:|---:|---:|---:|
| **159.634** | 54.859 | 4.99875 | 18.9621 | 92.41 | 96.34 5 | 151 |
| **161.585** | 54.997 | 5 | 19 | 92.42 6 | 96.94 3 | 141 |
| **160** | 54.893 | 4.99748 | 19 | 92.42 3 | 96.46 2 | 137 |
| **161.316** | 55 | 4.9986 | 18.6509 | 92.38 8 | 96.80 8 | 115 |
| **161.524** | 53.933 | 4.99248 | 18.9999 | 92.24 7 | 96.67 6 | 107 |
| **161.514** | 54.903 | 4.9856 | 18.7825 | 92.36 9 | 96.84 3 | 96 |
| **160** | 55 | 4.99732 | 19 | 92.43 9 | 96.48 5 | 94 |
| **162.194** | 54.976 | 4.97673 | 18.8629 | 92.35 9 | 97.05 | 84 |
| **170** | 41 | 3.4 | 11.8 | 82.09 4 | 91.57 | *R* |

### 3.4.1. Mass Spectrometry Data

The first thing to notice when looking at the spectra gained from the MS (Figure 70) throughout the DoE is the low level of internal standard (IS) peak at *m/z* 150 [**3.6**+H$^+$]$^+$. The small IS peak suggests that analysis of the reaction system is encountering suppression in favour of one other component in the reaction possible the SM and DP ions. This could be due to the much higher concentrations used and where both SM and DP have multiple ionisation points which may cause preferential ionisation. SM & DP peaks are observed in a similar pattern to those observed in the previous AZD5634 section. The SM gives a peak for [**3.4**+H$^+$]$^+$ at *m/z* 527 and [**3.4**-$^t$Bu]$^+$ at *m/z* 471 (Figure 70) whilst the DP can be seen at *m/z* 427 for its [**3.5**+H]$^+$. MS spectra also allows the observation of SPs formation so can help to characterise. However, it is also difficult to quantitate SM, DP and SP without IS present. An initial ratio between SM, DP & SP can be used to estimate the conversion of SM to all other components and compared to HPLC analysis.

Spectra were gathered in bulk, as previously stated, every 75 seconds meaning it should be possible to get an understanding of the reactor steady state as well as the conversion and yield of reaction at steady state. From one reaction to the next there should be a change in both or individual outcomes whether conversion or yield. However, each spectra shows an increase or decrease over the time the reactor takes to reach steady state. Ideally, between each reaction there should be a flush procedure to save starting material and so each reaction and steady state can be measured more accurately. However, this would generate much more solvent waste and could potentially add on extra time to the total optimisation.

The model from MS data, where the ratio between SM, DP and other SPs (total contributions) was used, is much more prone to error than the HPLC conversion counterpart. Without an internal standard to help with quantitation a reliable model could not be found. The model generated makes use of the ratio between SM and DP *m/z* values from each spectrum. All spectra time signatures were extracted and compared with HPLC time signatures to locate those MS which were take close to or at steady state. The ensured that each spectra showed maximum conversions and yields for each reaction in the list.

### 3.4.2. Kinetic profiles of reactions

From the DoE there were 17 sets of kinetic conditions carried out in an automated fashion (Table 21). Each of the 5 temperatures (150, 156, 175, 193 and 200 °C) was used to run reactions over each of the 4 residence times (1, 7, 13 and 19 minutes) at varying levels of water addition and concentration. A program developed within the group to fit the results to a kinetic model and provide equilibrium constants and activation energies was used. Both conversion of SM and the concentration of each species created in the reaction were needed to develop the kinetic model. Concentrations of knowns and unknowns were calculated in comparison to the IS and kinetic model fit to the HPLC data. Using the known reaction schematic of starting materials to products a general relationship and kinetic equation could be generated for the reaction. This relationship was added to the program along with conditions and calculated concentrations. The program will then analyse all the data and conditions and output kinetic variables such as equilibrium constants and activation energies. However, from the input data and conditions, no values could be determined due to "insufficient condition spread". Residence time is essential to generate a large spread of data from low conversion to high conversion. As there were only four residence times the program could not find a sufficient relationship for each set of conditions. Typical kinetic studies make use of ten or more data points relating to the total reaction time and therefore conversion at that time point. Originally five residence times were set to be used but the longest, 25 minutes, had to be taken out due to flow rates being too low to pump. If this had been included or another residence time, lower than one minute, had been used then it may have been sufficient for the kinetics study to be completed. It would be useful to determine how many data points would be sufficient for kinetic parameters to be estimated. The program developed within the group has only been tested on simulated data sets so may not be powerful enough to give sufficient data which so few data points or with real experimental error within reaction data. Kinetic parameters may also not be estimated as certain combinations of reactions have been missed creating gaps in the kinetic experimental surface. Another DoE and kinetic hybrid model maybe able to be designed but at the risk of creating many more experiments rather than less experiments with maximised data.

*Table 21 Kinetic experiments carried out within the hybrid DoE model for the CPACT study N-Boc deprotection of Boc-Arg(Pbf)-OH.*

| Set Number | Temperature/ °C | Water % | Conc. Wt.% | Residence/ time/ min |
|---|---|---|---|---|
| 1 | 150 | 37.5 | 3 | 1, 7, 13, 19 |
| 2 | 156.524 | 24.567 | 1.522 | 1, 7, 13, 19 |
| 3 | 156.524 | 50.433 | 1.522 | 1, 7, 13, 19 |
| 4 | 156.524 | 24.567 | 4.478 | 1, 7, 13, 19 |
| 5 | 156.524 | 50.433 | 4.478 | 1, 7, 13, 19 |
| 6 | 175 | 20 | 3 | 1, 7, 13, 19 |
| 7 | 175 | 55 | 3 | 1, 7, 13, 19 |
| 8 | 175 | 37.5 | 1 | 1, 7, 13, 19 |
| 9 | 175 | 37.5 | 5 | 1, 7, 13, 19 |
| 10 | 175 | 37.5 | 3 | 1, 7, 13, 19 |
| 11 | 175 | 37.5 | 3 | 1, 7, 13, 19 |
| 12 | 175 | 37.5 | 3 | 1, 7, 13, 19 |
| 13 | 193.476 | 24.567 | 1.522 | 1, 7, 13, 19 |
| 14 | 193.476 | 50.433 | 1.522 | 1, 7, 13, 19 |
| 15 | 193.476 | 24.567 | 4.478 | 1, 7, 13, 19 |
| 16 | 193.476 | 50.433 | 4.478 | 1, 7, 13, 19 |
| 17 | 200 | 37.5 | 3 | 1, 7, 13, 19 |

### 3.5. Summary and Future Work

The designed process, continuous flow hybrid DoE and kinetic model works but is generally limited by the equipment *i.e.* pump flow rates and algorithm programming. It is essential to have a robust program which works continually when performing a large number of reactions as any error can set experimentation to a halt. The desired flow rates themselves should have been well within the range of the useable pump flow rates but, due to pump age and efficiency, flow rates for a residence time of 25-minutes were not feasible. The only way around this would be to buy new pumps working at 100% capacity or fix and refurbish the existing pumps. However, this is also unfeasible as brand-new pumps cost a substantial amount. Coding errors are generally easier to fix in the short run, as shown on page 118. These short-term fixes allowed the full DoE to be carried out

with only four timer-errors over the DoE time period. A more permanent fix, more so than ensuring no zeros are present in the matrix, is being sought and should be fine-tuned with help from Mathworks and AZ.

The chosen molecule, **3.4**, gave sufficient data to generate a good DoE model for conversions and yields both using HPLC and MS. Even though molecule **3.4** was more thermally unstable than the previous AZD5634, side product formation improved the data set as this could be considered with the potential to be modelled alone. **3.4** was also sufficient to be monitored by both HPLC-UV and MS, although a further third technique would have been desirable. Rather than IR, bench-top NMR could have been a valuable asset which providing more data for this work. **3.4** contains many unique functional groups but also has SM and DP which differ. The change of N-Boc group to $NH_2$ group could have been the focus of NMR monitoring. The fact that NMR is also an in-line analytical technique may also have added another layer of depth into the optimisation and data potentially providing evidence of intermediate formation or mechanistic information.

However, the thermal instability of **3.4** did make quantifying MS results much more difficult. APCI has been shown to be more quantitative than other ionisation techniques[113] but in this case the high temperature of the ionisation source led to uncontrollable breakdown of the compound skewing conversions and yields. Lowering the ionisation temperature and using a smaller sample volume may have helped improve quantitation. Lowering the temperature will ensure less thermal degradation of the compound but will also affect the amount of compound which reached the gas state and therefore ionised. Lowering sample volume may have brought the concentration much further down the linear dynamic range of the instrument allowing easier quantitation.

The second protecting group on starting compound, Pbf, has a similar deprotection procedure to N-Boc also preferring strong acid facilitated removal. This is likely to have started a competing deprotection reaction and given the SP although only small amounts were observed during many of the DoE conditions with more being produced at higher temperatures and longer residence times. This suggests that the group is slightly more thermally stable than the N-Boc group potentially allowing a selective procedure to be developed in the future.

### 3.6. Experimental

The general setup for flow work performed is described in section 2.7.1 where details on pumps, reactors, analysis and programming can be found. Specific details for each reactions setup are provided in this section for the work with Boc-Arg(Pbf)-OH the DoE and kinetic hybrid model and work carried out prior to this run. Other details outlining reagents and solvents with purities and grades can be also found in this section along with details of how analyses of samples were performed.

### 3.6.1. Miscibility of Boc-Arg(Pbf)-OH in MeOH and $H_2O$

A 1% w/v stock solution of Boc-Arg(Pbf)-OH (1.00 g, 1.90 mmol) was dissolved in MeOH (100 mL). Aliquots of this solution were then added to sample tubes before water was added to each in turn (Table 22). Observations: white precipitate 60% water and over.

*Table 22 Summary of solubility conditions of Boc-Arg(Pbf)-OH in MeOH with portions of water added.*

| MeOH % | Water % | Observations |
|--------|---------|--------------|
| **90** | 10 | Full dissolution of compound |
| **80** | 20 | Full dissolution of compound |
| **70** | 30 | Full dissolution of compound |
| **60** | 40 | Full dissolution of compound |
| **50** | 50 | Low level of precipitate then full dissolution |
| **40** | 60 | White precipitate, no dissolution |
| **30** | 70 | White precipitate no dissolution |
| **20** | 80 | White precipitate no dissolution |
| **10** | 90 | White precipitate no dissolution |

### 3.6.2. Microwave Batch Reactions of Boc-Arg(Pbf)-OH

A sample of Boc-Arg(Pbf)-OH (0.05 g, 0.095 mmol) was dissolved in MeOH (1 mL) with N,N-dimethyl benzamide (10 mol%) to act as internal standard. Water was added to the solution to create a ratio of 60:40 (MeOH to Water). This solution was then reacted under microwave conditions at 180 °C for 15 minutes. A 1 mL aliquot was taken and then injected manually into the HPLC and MS, used for flow reactions, to determine the

compositions prior to method development. Mass Spectra: 150.2 *m/z*, 427 *m/z*, 471 *m/z* and 379 *m/z* (279 *m/z* required, found 279 *m/z*). HPLC: 2.61 min (IS), 3.76 min (DP), 4.04 min (SP), 5.18 min (SM).

### 3.6.3. HPLC and MS Method Development

A sample of Boc-Arg(Pbf)-OH (0.05 g, 1.90 mmol) was dissolved in MeOH (1 mL) with N,N-dimethyl benzamide (10 mol%) to act as internal standard. This solution was connected to a HPLC pump whilst deionised water was connected to second HPLC pump. The pumps were set to flow at 0.1 mL min$^{-1}$ each using a 3 mL heated tubular reactor. The solution was pumped through at room temperature *c.a.* 22°C, before a sample was taken *via* HPLC and MS. This process was repeated at 150 °C to generate product to develop a HPLC method. Mass spectra: 150.1 *m/z*, 427.2 *m/z*, 471.2 *m/z*, 527.2 *m/z* (required; 150.1 *m/z*, 427.2 *m/z*, 471.2 *m/z*, 527.2 *m/z*, found; 427.2 *m/z*, 471.2 *m/z* and 527.2 *m/z*). HPLC: 2.61 min (IS), 3.76 min (DP), 4.04 min (SP), 5.18 min (SM).

*Table 23 HPLC method and conditions developed to separate components generated for the N-Boc deprotection of Boc-Arg(Pbf)-OH during the CPACT feasibility study.*

| Time/ min | % B | %A | Flow Rate/ mL min$^{-1}$ |
|-----------|-----|-----|----------------------------|
| **0** | 5 | 95 | 1 |
| **1** | 50 | 50 | 1 |
| **5** | 95 | 5 | 1 |
| **6** | 95 | 5 | 1 |
| **6.5** | 5 | 95 | 1 |

### 3.6.4. Mass Spectra Method

Mass spectra were gathered in APCI+ mode from *m/z* 100 to 800 using a scan time of 875 ms and an acquisition time of 60 seconds. Other settings of the MS are summarized below. The mobile phase used 100% MeCN and was pumped at 0.3 mL min$^{-1}$.

Polarity: Positive,

Capillary Temperature: 250 °C

Capillary Voltage: 180 V

Source Voltage Span: 30 V

Source Voltage offset: 20 V

Source Gas Temperature: 350 °C

APCI Current: 5 µA

### 3.6.5. Calibration of Boc-Arg(Pbf)-OH

A stock solution of Boc-Arg(Pbf)-OH (1.00 g, 1.90 mmol) was prepared in MeOH (10 mL) along with N,N-dimethyl benzamide (10 % w/v) as internal standard. This solution was then diluted to give standards with concentration of 0.007 M, 0.015 M, 0.038 M, 0.075 M, 0.113 M and 0.15 M for SM and 0.0015 M, 0.003 M, 0.0075 M, 0.015 M, 0.023 M and 0.03 M for IS. Each was, in turn, injected into the HPLC and MS a total of three times each and an average taken before being plotted as a calibration graph (Figure 77Figure 78). Mass spectra: Mass spectra: 427.2 *m/z*, 471.2 *m/z*, 527.2 *m/z* (required; 150.1 *m/z*, 427.2 *m/z*, 471.2 *m/z*, 527.2 *m/z*, found; 427.2 *m/z*, 471.2 *m/z* and 527.2 *m/z*). HPLC: 2.61 min (IS), 3.76 min (DP), 4.04 min (SP), 5.18 min (SM).
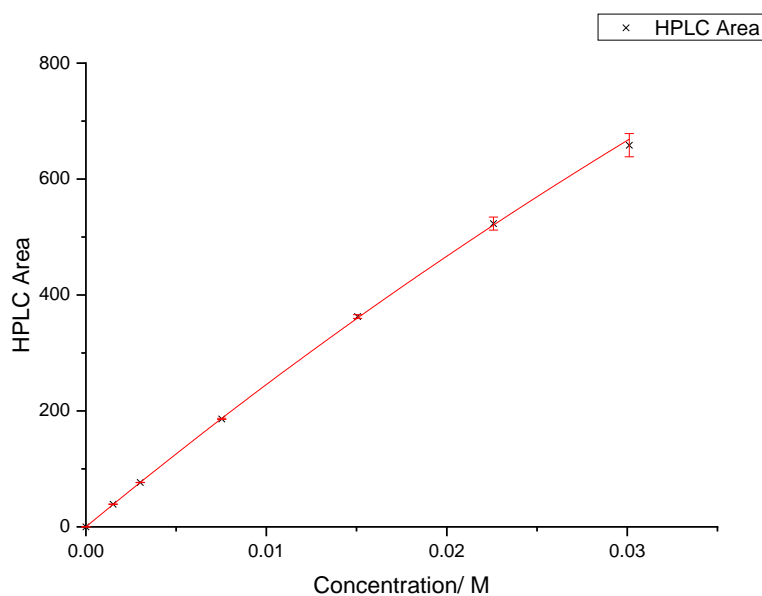


*Figure 77 Internal standard (N,N-dimethylbenzamide) calibration graph through HPLC data for the CPACT N-Boc deprotection study.*
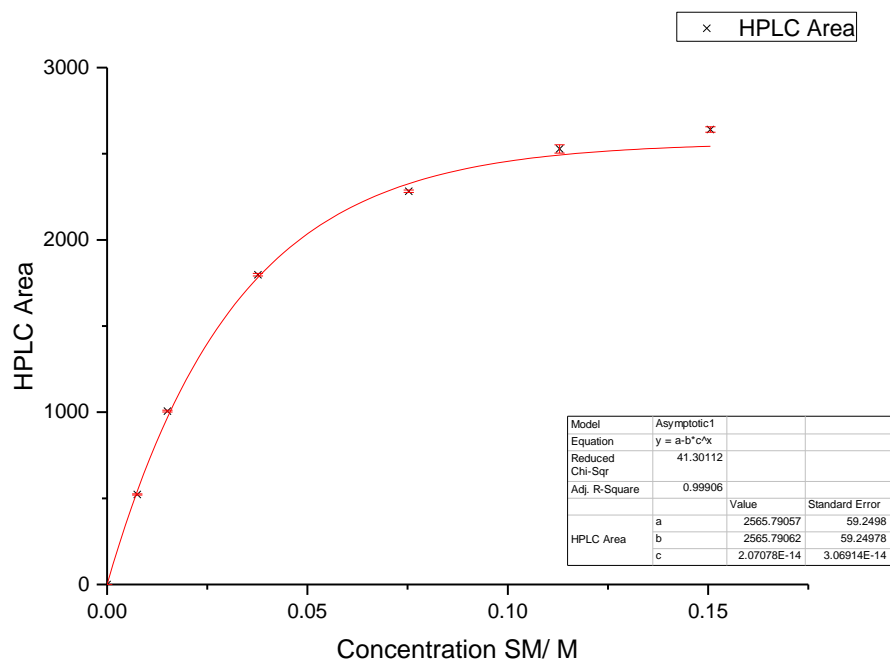
*Figure 78 Starting material (Boc-Arg(Pbf)-OH) calibration graph through HPLC data for the CPACT N-Boc deprotection study.*

### 3.6.6. N-Boc Deprotection DoE Optimisation of Boc-Arg(Pbf)-OH
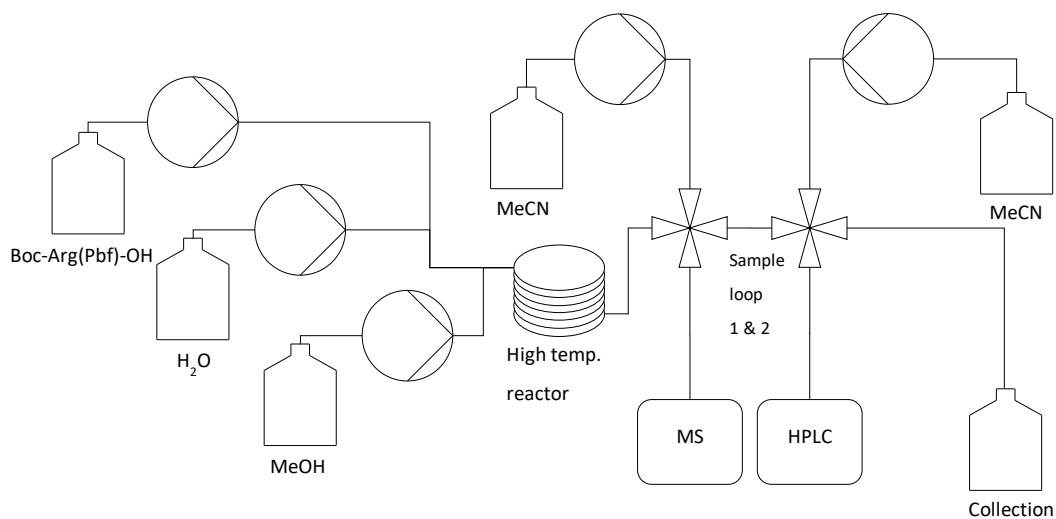


*Figure 79 Reactor setup used for the hybrid modelling optimisation of Boc-Arg(Pbf)-OH monitored using on-line HPLC and on-line MS.*

A summary of the performed DoE is shown in Table 19. The starting compound, Boc-Arg(Pbf)-OH (12.67 g, 24.1 mol) was dissolved in MeOH (160 mL) along with N,N-dimethyl benzamide (10 mol%), connected and primed through a HPLC pump. Deionised water (500 mL) was connected and primed through a HPLC pump whilst HPLC grade MeOH was used and primed (500 mL) through a HPLC pump. The DoE conditions used (*Table* 19) were generated using Modde (by AstraZeneca) through a CCO design to ensure each factor had 5 levels and were randomly spaced. The HPLC was setup using the designed method in section 3.6.3 whilst the MS was setup with a one-minute acquisition time to sample every 75 seconds. A fourth pump contained HPLC grade MeCN for the mobile phase of the MS. The DoE conditions were added to the DoE algorithm in MATLAB before the reaction series was started. A schematic of the reactor setup is shown in Figure 79. The SM mixes with $H_2O$ to the required water% before this is then diluted to the correct concentration for each experiment in the list. All three pumps dictated the residence time through the reactor and flowed through to the collection flask through two separate sample loops; one for the MS and one for the HPLC.

Mass spectra: 427 *m/z*, 471 *m/z*, 527 *m/z* (required; 427 *m/z*, 471 *m/z*, 527 *m/z*; found; 427 *m/z*, 471 *m/z*, 527 *m/z*). HPLC: 2.61 min (IS), 3.76 min (DP), 4.04 min (SP), 5.18 min (SM)

Chapter 4

**A Full Optimisation Platform for the $S_N$Ar reaction used During Active Pharmaceutical Ingredient Synthesis.**

### 4. A Continuous Flow Process for the S$_N$Ar Reaction Between Ethyl-4-fluorobenzoate and *cis*-2,6-dimethylpiperazine.

Chapter 4 will discuss the exploration and optimisation of an S$_N$Ar reaction using DoE, self-optimisation and kinetic approaches. The reaction itself, between ethyl-4-fluorobenzoate and *cis*-2,6- dimethylpiperazine, is a reaction of interest as a step in a multistage synthesis in the pharmaceutical industry. Both compounds are readily available from multiple sources. The chapter will begin with a general introduction to discuss key concepts for the work and uses of the final molecule after reaction. Work carried out was fully automated, using HPLC and MS as analysis, and employing a basic experimental design to get statistical data from the reaction before self-optimisation and kinetic experiments were carried out to gain as much useful information about the reaction as possible.

### 4.1. Introduction

The formation of carbon-carbon or carbon-heteroatom bonds are typically achieved by either cross-coupling reactions, using transition metal catalysts *e.g.* [CpRu(*p*-cymene)]PF$_6$, or by nucleophilic aromatic substitution reactions (S$_N$Ar).[114] Often the easiest and the best approach is to use the S$_N$Ar reaction as there is no need for catalytic extraction post-reaction. However, S$_N$Ar reactions are not without their own set of problems. Reagents used for S$_N$Ar reactions often contain electron deficient groups and multiple reactive sites (commonly two, *para vs ortho*). It is these multiple groups which cause reactions to have problems with regio-selectivity.[115] If the initial starting material contains only one reactive site then it may also be problematic when using a nucleophile which has multiple sites to react, again creating regio-selectivity problems.

There have been many studies into the *ortho* and *para* reactivity with some studies showing high regio-selectivities of >95% for *ortho*-substituted compounds using non-polar solvents.[114] In general, parameters which can affect regio-selectivities are: substrate and leaving group, solvent polarity, nucleophiles, temperature, reaction time *etc*. Regio-selectivity has been explored using flow reactors where isomer and disubstituted species were controlled but had limitations with homogenous conditions and high temperatures.

$S_NAr$ reactions have two pathways which they can proceed through, either radical or ionic, the latter being most common.[115] The ionic pathway consists of the nucleophile attacking the electrophilic starting material to form a transition species, a resonance stabilized Meisenheimer complex, before the leaving group is displaced by oxidative elimination from the *ortho* or *para* position. The reaction rate can be increased by use of a more electron withdrawing species (*e.g.* $NO_2$) or a more nucleophilic species.

Reizman and Jensen[116] developed a fully automated system which calculates kinetic parameters, fits the results to a model using computer algorithms and determines an optimum value for the reaction. The reaction chosen was an $S_NAr$ reaction between 2,4-dichloropyrimidine and morpholine, using ethanol as solvent, to produce the desired product 2-substituted aminopyrimidine and two side products; 4-substituted aminopyrimidine and 2,4-disubstituted aminopyrimidine (Scheme 25). This reaction has pharmaceutical interest as the products formed are used in APIs which exhibit a kinase inhibitory effect in cancerous cells.
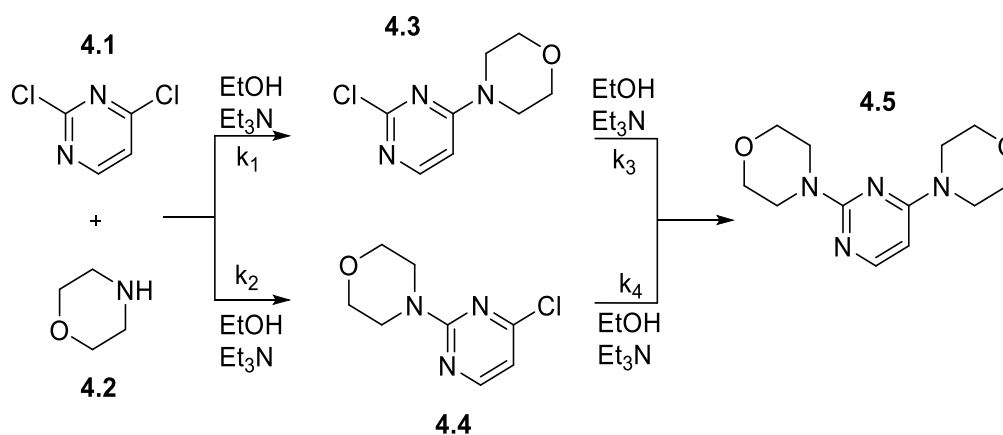


*Scheme 25 Multistep reaction pathway for the $S_NAr$ reaction between 2,4-dichloropyrimidine to 4,4'-(2,4-pyrimidinediyl)-bis-morpholine.*[116]

A factorial experimental design was used to run reactions using on-line HPLC to generate a feedback loop, helping the algorithm to calculate best-fit kinetic parameters and sensitivity coefficients. This part of completing the cycle relies on an initial user-defined model to get the best agreement with data from the DoE. The algorithm could progress to an optimum point (*via* minimisation of sensitivity coefficients) using a D-optimal program which selects new conditions for an experiment based on prior experimental data. However, this sort of optimisation method does not have a function to stop the

optimisation from happening meaning it will continue to iterate until the user terminates the design, Figure 80.
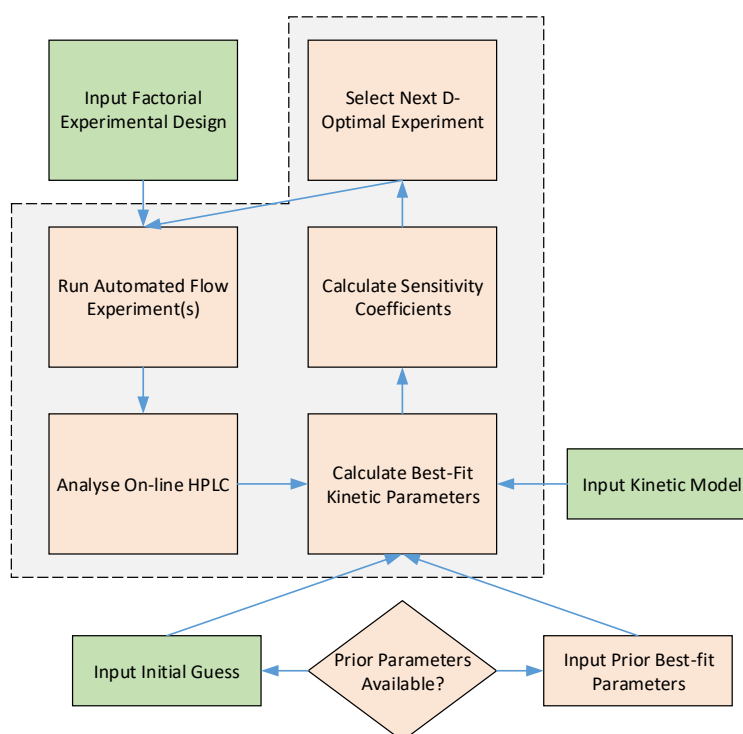


*Figure 80 Logic flow diagram for the automated kinetic parameter estimation in continuous flow.* The DoE was made of three different stages, each with a designed, specific outcome. The first stage, initial set of experiments, were used to determine 8 kinetic parameters based on rate equations for each step in the reaction scheme and determine the rate laws governing each species' generation and consumption. The next stage parametrised each step of the reaction individually before the final stage calculated the 8 kinetic parameters concurrently using deduced values from previous experiments. The four factors of exploration were those which were decided to be essential to the reaction; residence time ($t_{res}$), temperature (T), initial concentration ($C_{i0}$) and the equivalents of morpholine used during each reaction. As it was necessary for them to include a user defined model they included assumptions and general kinetic theory in the automated program; in this case, it was assumed that the reaction followed a second-order kinetic system and that the entire system followed ideal plug flow kinetics.

The initial factorial DoE provided them an output of 12 automated experiments where the data was plotted and fitted to a model. The algorithm which they used, D-optimal, allowed them to ensure that the uncertainties in parameter estimates were made as small as possible and was added to the reaction program after the initial set of experiments

were performed. Like with any other DoE, data was collected at each individual point in the surface; where each estimate in the model was recalculated considering these upper and lower bounds. This process was continued for each of the reactions previously performed giving 24 in total as to compare the model for each, plotted against residence time and the program manually terminated. There is good agreement with both of their final models, however, the experimental data underestimates the yields of both the 2-substituted and 4-substituted species whilst overestimating the conversion of the starting material, particularly at long residence times. This could be quite problematic if attempting to run the reactions for an expensive compound where it is essential to get both an accurate conversion and yield. Due to this they decided to isolate each reaction and find reaction kinetics at each stage using the same approach described.

From these individual modelling experiments it was shown that parameter estimation can occur using DoE to get a large amount of kinetic data which can be used to optimise 8 parameters simultaneously. However, this method may only be of use for small scale reactions of cheap molecules as the initial work to optimise the 8 parameters was unsuccessful. Several of these 8 parameters contained up to 20% uncertainty which is nonideal when attempting to scale-up reactions. Isolating each reaction stage allowed each sensitivity parameter to be minimised significantly reducing the uncertainty and improved results for optimal yield for the formation of the 4-substitued species.

Hone *et al.*[117] used a similar methodology to explore the creation of kinetic motifs using an automated flow system but rather than optimising the kinetic parameters outright, focussed more on under and over reaction products. This approach extended the range of overall reaction conditions and provide a much richer data set to help increase the confidence of the final kinetic model. The reaction chosen this time was the $S_NAr$ between 2,4-difluoronitrobenzene and pyrrolidine. Multiple reaction products were monitored; *ortho*-substituted (desired product), *para*-substituted and *bis*-substituted, using on-line HPLC (Scheme 26).
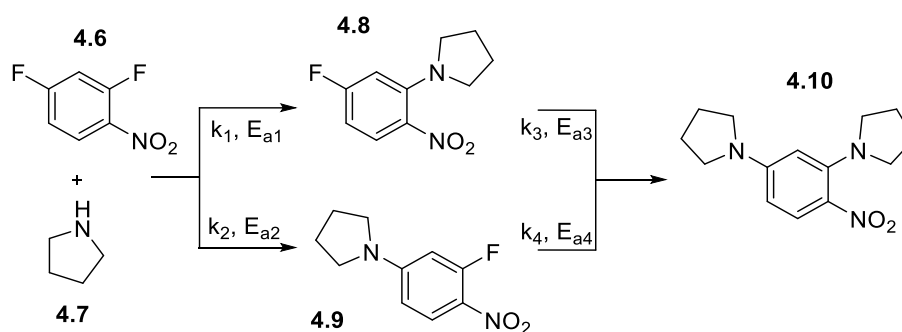
*Scheme 26 Four step reaction pathway of an $S_NAr$ reaction of 2,4-difluoronitrobenzene with pyrrolidine, in ethanol and using $NEt_3$ as base.*[117]

The use of gradient flow ramps made it easier to explore the full reaction profile in one prolonged run. The gradient was setup similarly to a HPLC gradient only with maximum flow rates to begin along with a varied ratio of starting materials. To get accurate results for the first sample the reactor was allowed to reach steady state before the flow rates were decreased by a set amount (0.836 mL $s^{-2}$) allowing the residence time to be altered whilst keeping constant process concentrations. More conditions (Figure 81) were explored from mildest (1.5 mol eq., 0.5 min, 30 °C) to harshest (7 mol eq., 2 min, 120 °C) based on work by Hessel *et al.*[5] All gradients were preprogramed into and controlled by MATLAB.
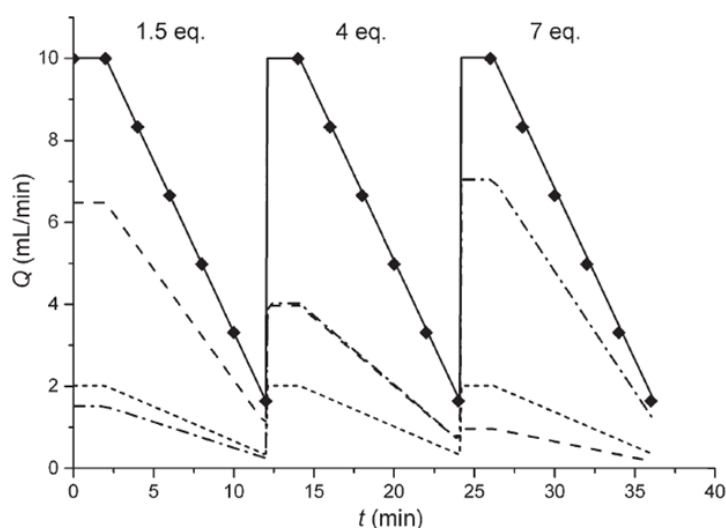


*Figure 81 Changes in volumetric flow rate over time, where $Q_{P1}$, $Q_{P2}$, $Q_{P3}$ and $Q_{total}$ were for; ···· a HPLC pump (Ar in EtOH), --- pump 2 (EtOH), ---· pump 3 (pyrrolidine in EtOH), — total volumetric flow rate and ♦ HPLC injection respectively. The linear flow ramps correspond to pyrrolidine to 2,4-difluoronitrobenzene molar ratios: (i) 1.5:1, (ii) 4:1 and (iii) 7:1 using $Q_{total}$ from 10 to 1.5 mL $min^{-1}$.*[117]

These reactions were repeated whilst changing the temperature (30, 60, 90, 120 °C) giving a total of 12 reaction profiles from 72 data points over the course of 3 hours. From this data it was found that the reaction was second order and that rate constants could be determined using the Levenburg-Marquardt algorithm[118] under isothermal conditions (90 °C). Calculated kinetic parameters closely matched the experimental data ($R^2$ = 0.9995) and calculate rate constants for the *ortho*-substituted product to be *ca.* 20 times larger than the rate constant for the *para*-substituted product and activation energies being 33.3 $\pm$ 0.3 kJ mol$^{-1}$ and 35.3 $\pm$ 0.5 kJ mol$^{-1}$ respectively. These results suggest that temperature has an influence on the rate of reaction but does not directly affect the positional selectivity in the molecule.

The work by Hone *et al.*[117] shows that kinetic modelling can be automated and that it is possible to acquire large data sets from many reaction profiles in a short period of time. The model gained from the data fit 4 reactions into one process using 8 fitting parameters with a total of 4% uncertainty. Their latter experiments show that it is possible to run kinetics without waiting for the reactor to reach steady state, reducing the amount of material used. It may also help to generate models far earlier in process development to help gauge the sensitivity of the product and which process parameter settings are most important.

The reaction used for this work was between ethyl-4-fluorobenzoate and *cis*-2,6-dimethylpiperazine (Scheme 27). The desired product of the $S_NAr$ reaction is of interest to AstraZeneca as an essential API for the final drug molecule AZD4547.
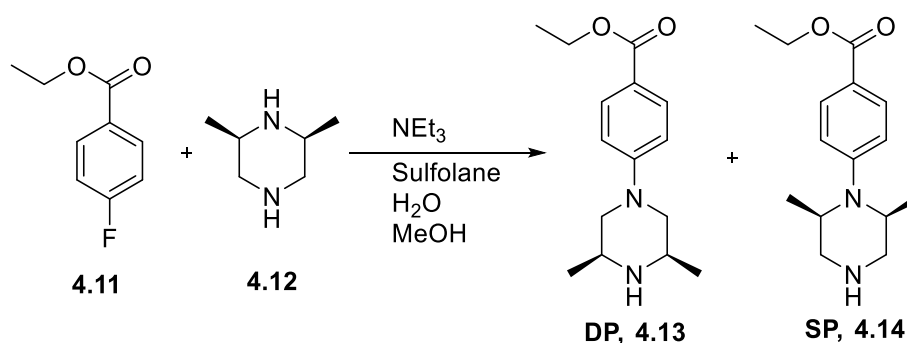


*Scheme 27 $S_NAr$ reaction scheme between cis-2,6-dimethylpiperazine and ethyl 4-fluorobenzoate.*

Currently, this $S_NAr$ reaction is carried out as a batch process using the solvents sulfolane and $H_2O$ at 120 °C. Sulfolane is a good solvent to use when the reaction must be carried out at higher temperatures as it has a much higher boiling point (285 °C) than

conventional solvents used in industry *e.g.* MeOH, THF or IPA. However, the high boiling point of the solvent can cause problems during separation. This reaction has been initially transferred into flow by an AZ team in Bangalore where they used DMSO and $H_2O$ as solvents, at 130 °C. Unfortunately, they were unable to get a good yield from the reaction, 14%, although no side-products were generated. DMSO is not a good solvent to use for this reaction as it has the propensity to go off at elevated temperatures and DMSO has the unique ability to transfer compounds dissolved in it through the skin. The batch process mentioned above also is very problematic for this reaction. The current batch procedure employed by AZ takes between 2-4 days at 120 °C to react, and does not reach completion, gaining a low 14% yield. At the elevated temperatures used the piperazine reagent sublimes and condenses as a solid on the top of the reactor or any cold surface reducing the chance of the reaction occurring.

The starting material compound itself, ethyl-4-fluorobenzoate, contains an ester which is a mildly electron withdrawing group meaning it will slightly activate both *ortho* and *para* positions on the benzene ring (Scheme 28). The *para* group consists of a fluorine which is highly electronegative, drawing electrons towards itself, making it a much poorer leaving group than other halides. The reagent used, *cis*-2,6-dimethylpiperazine, contains two methyl-groups both adjacent to a nitrogen atom. Methyl groups are electron donating groups which give the nitrogen atom a higher density of electrons *i.e.* it is more nucleophilic, in comparison to the second nitrogen in the molecule. Due to this, the increased nucleophilicity of one nitrogen will make it more reactive which will inevitably decrease the rate of reaction to the desired product. It was hoped that completing this reaction in flow and with self-optimisation capability, that the reaction could occur much more quickly and where conditions could be found to carry it out in a robust fashion.
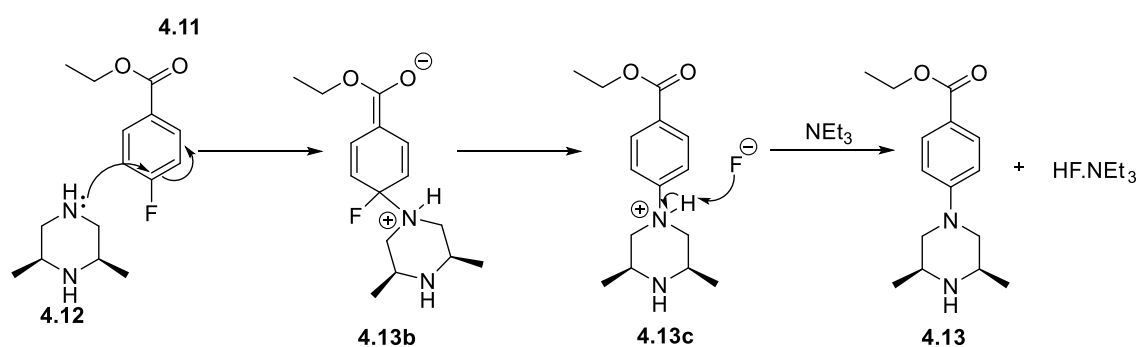
*Scheme 28 Outline of the $S_NAr$ reaction mechanism to get the desired product and by-product hydrofluoric acid, HF.*

A by-product of the reaction is hydrofluoric acid, HF, forming due to the nature of the $S_NAr$ reaction and *cis-2,6*-dimethylpiperazine (Scheme 28). This would not normally be a problem as adding a base to the reaction allows the base to react with the HF to neutralise it and make it safer. However, more problems can arise with an acid like this in continuous flow, particularly, at high temperatures and under pressure. At high temperature the HF formed during reaction will most likely exist in its free acid form and not in a stable inert form with the base. Due to this, precautions have to be made to ensure the reaction and reactor setup is non-hazardous. At elevated temperature use of a stainless-steel tubular reactor would normally be used, but in this case a Hastelloy reactor was employed (1/16" OD, 0.04" ID). Hastelloy is a superalloy comprised of high percentages of nickel, chromium and molybdenum containing low levels of iron (~3%) which gives it a much better nature for dealing with high temperature, corrosive systems. The downside to using a corrosive resistant tubular reactor is the much higher expense in comparison to stainless steel; £70 per metre compared to just £5 per metre respectively.

Apart from the desired product, this reaction has the potential to produce a side-product *i.e.* the regio-isomer from the attack of the piperazine at the most hindered position. It is necessary to have a robust analytical method which can separate and distinguish between the two different isomers formed in this reaction (Figure 82), HPLC and MS. As both starting material and product contain a chromophore, functional group dense in carbon-carbon double bonds with resonance, it is possible to separate and detect all species. MS can be used either as a detector to determine the mass of each HPLC peak or as a technique to determine all constituents in the sample at once and help determine steady state.
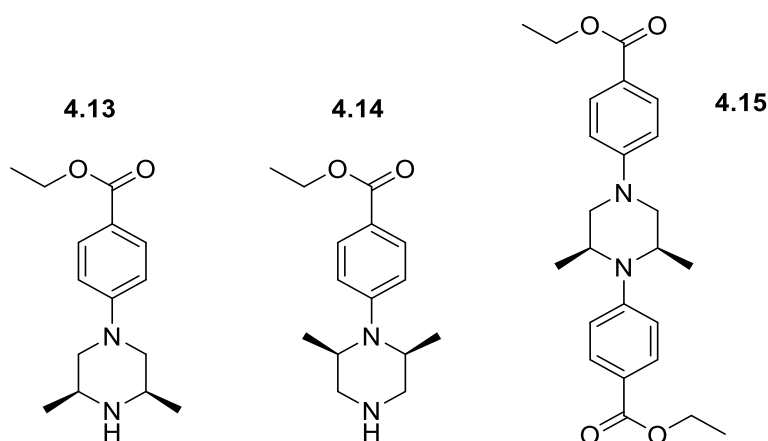
*Figure 82 Desired product (**4.13**), regio-isomer side product (**4.14**) and di-substituted side product species (**4.15**) formed in the $S_NAr$ reaction carried out on continuous flow.*

The DoE design was created in the program Modde to look at four different variables to screen and find the most important. Temperature, residence time, % $H_2O$ and equivalents of piperazine were initially chosen for the reaction. Other factors which may have been considered are the choice of solvent, volume and type of base used, and overall concentration of the reaction. Being able to find the most important factors in a process is essential to simplify future work when self-optimising as the less factors that affect the process the quicker the self-optimisation should be. The actual design chosen was a 2-level full factorial approach which accounts for model interactions of factors. This type of design is a balanced design where all combinations of factor levels are covered, and the main effects and all interactions are accounted for and not confounded. The total number of experiments to complete this is 19; $4^2$ combinations plus 3 centre point replicates.

The algorithm used, SNOBFIT, has been described in section (section **1.4.4.3.**) and has its uses in automation chemistry due to its ease of use and its balanced nature in the exploration and exploitation of the experimental system.

Use of this algorithm requires the development of several other functions within the MATLAB program. A basic response function which contains the desired calculation and functions to monitor and access the HPLC data which helps the algorithm decide where to experiment next. This algorithm has been fully explored and used within the group to explore the reaction surface efficiently and find the optimum value fairly quickly. This means it can be seen as the quickest and most data rich algorithm making it ideal to use for this reaction system. Typically, an optimisation algorithm is used when there are many variables to consider and an optimum point is difficult to determine.

Initially the reaction would look at 4 different variables to optimise; temperature, residence time, reaction concentration and substrate equivalents. Concentration of the starting solution had to be disregarded as a controlled variable as some of the flow rates necessary were well below the capability of the pumps.

The experimental approach used for this research closely mimics that used within the industrial sector. Beginning with a reaction screen completed on small scale before progressing to a feasibility study in the chosen reaction vessel. Then a DoE is performed on the system to find out which are the most important factors in the process as well as an optimum region for operation. This usually entails a full DoE study and a kinetic study in industry, but the research has made use of an optimisation algorithm to find the optimum area of operation and an automated approach to get kinetic data.

### 4.2. Summary

- The $S_N Ar$ reaction is a simple and well-established reaction.
- It can be completed in several ways to improve selectivity between different substitution products.
- $S_N Ar$ reactions can be carried out in continuous flow and automated.
- Automation can help to perform DoE studies or larger kinetic studies for greater certainty of results.
- Algorithms can be employed to take previous experimental data, new kinetic data and fit a robust model.
- It is possible to get kinetic data from reactions which are not at steady state and get accurate kinetic parameters which agree with experimental data.
- This may make it possible to perform kinetic studies *in silico* with high accuracy.

### 4.3. Aims

The aims of this work were to ultimately try and improve upon the batch synthesis of this reaction by carrying it out in continuous flow, monitored by on-line analytics. Currently, the reaction takes up to 4 days in batch without getting to completion, meaning flow may benefit the transformation. Current flow work has only managed to give a low yield of 14% so use of the self-optimisation kit may help improve yields.

It was hoped to carry out a DoE on the system to gain reaction and process understanding and to allow the reaction to be carried out in a robust fashion on a larger scale. As well

as DoE, self-optimisation was needed to compare to the DoE results allowing verification of the optimum operating region. It was also necessary to ensure that the on-line analytics used could separate and detect enough species in the reaction for qualitative and quantitative data. It was interesting to see how MS could, if at all, be used to quantify this reaction by sampling of the entire reaction stream, unseparated, as MS cannot distinguish between regio-isomers.

As mentioned above, HF if formed in the reaction as a by-product so health and safety requirements must be considered to ensure the flow process can be used on scale without problems.

It will be necessary to look at how the temperature, residence time, piperazine equivalents and amounts of water affect the reaction. The amount of base used may also be a factor of interest as the use of base will dilute the reaction and potential slow the rate of desired product formation. However, lack of base will also make the process much more unsafe with HF formation. Alternate methods of HF quench will have to be used if use of base becomes a factor to test. These have been covered in a recent paper by A. J. Blacker where the HF quench was completed subsequent to the reaction using calcium salts (carbonate, sulphate and propionate) in water.

### 4.4. $S_N$Ar results and discussion

Before work could be started using the required AZ substrates, matching ones were found in order to gauge the system and help setup analysis methods. Ethyl-4-fluorobenzoate (reaction nucleophile) is cheap and readily available to buy (£42.80 for 25 g) so could be used straight away for the reaction. 2,6-*cis*-Dimethylpiperazine is also readily available but had a longer lead time at the time of investigation. An initial alternative, piperazine (Scheme 29), had to be used for process exploration for MS and HPLC method development as well as for verification of previous work carried out by AZ in Bangalore using a microwave reactor.
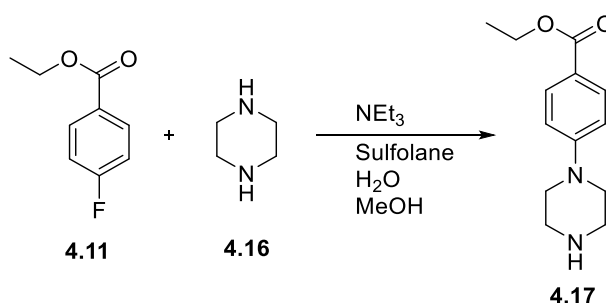
*Scheme 29 Alternative nucleophile, piperazine, used initially to verify AZ work and develop HPLC and MS methods prior to 2,6-cisdiemthyl piperazine use.*

Microwave reactions allowed conditions to be more forcing as the pressure enforced on the vessel allows solvents to be used above their boiling points. It was necessary to synthesise some product to develop a HPLC method which would separate all species in the reaction. Initial conditions used mimicked the conditions used by the AZ Bangalore team, 130 °C for 30 minutes using sulfolane as solvent. Unfortunately, no conversion was seen so the conditions were changed in order for a reaction to occur. The next set of conditions changed the amount of water relative to the SM and varied temperature (180 °C and 220 °C) over the course of a 20-minute reaction time (Table 24).

*Table 24 Reaction conditions used during the microwave batch study of ethyl-4-fluorobenzene and piperazine to form the desired product with the corresponding HPLC conversion (see section 4.6.1).*

| Reaction Number | SM:H2O | Temperature/ °C | HPLC Conversion to Product/ % |
|---|---|---|---|
| **1** | 70:30 | 180 | 12.0 |
| | | 220 | 17.7 |
| **2** | 60:40 | 180 | 23.9 |
| | | 220 | 30.6 |
| **3** | 50:50 | 180 | 38.5 |
| | | 220 | 44.9 |
| **4** | 40:60 | 180 | 77.7 |
| | | 220 | 82.8 |

Clearly from the data it can be seen that the amount of water in the reaction has a pronounced effect on the conversion of the reaction with temperature making a much smaller difference. At 220 °C when 60 % water was in the reaction there was a total and highest conversion of almost 83% with the lowest conversion at 180 °C and 30% water. With this in mind, the DoE performed will use 4 factors; the water percentage, residence time of reaction, temperature and the equivalents of piperazine used.

MS analysis of each batch reaction was completed in conjunction with the HPLC to separate and then analyse the mass of each peak in the reaction. The initial method developed was short and could separate the reaction components where 4 peaks were observed; piperazine, trimethylamine, product and internal standard (Figure 83).
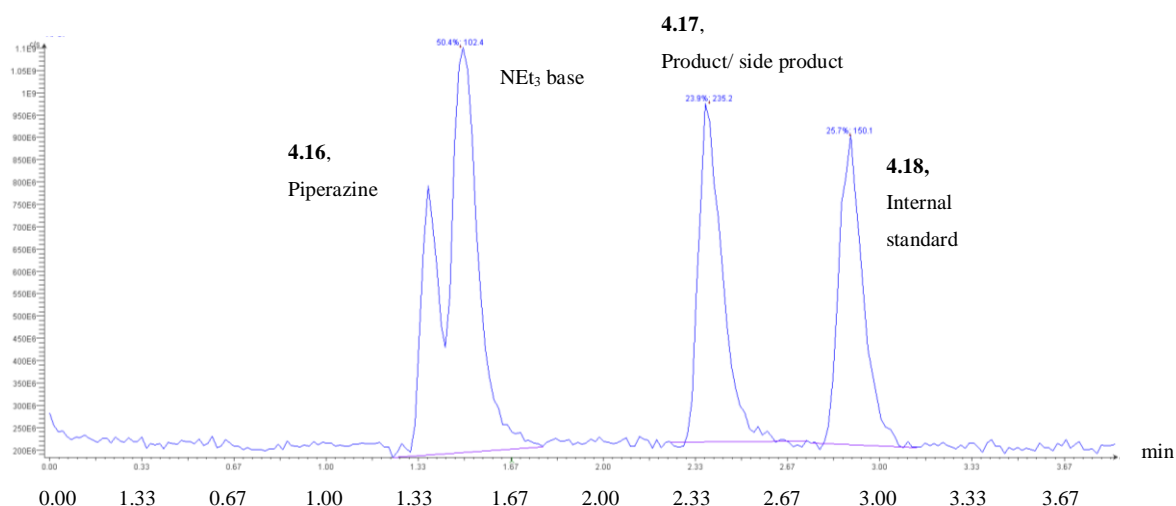


*Figure 83 MS total ion chromatogram showing 4 peaks (provided by HPLC retention times) relating to piperazine, NEt₃, desired product and chosen internal standard (the SM was not observed in the MS) see Figure 84 for isolated masses of each peak.*

Each peak from the TIC relates to a specific species in the reaction proved from HPLC separation giving four different retention times. The first at 1.35 minutes is from the piperazine nucleophile and gives a peak at 87 Da and 128 Da for its [**4.16**+H$^+$]$^+$ and [**4.16**+H$^+$+MeCN]$^+$ respectively. Similarly, both peaks at 2.37 and 2.87 minutes give a *m/z* values of 235 Da and 276 Da for the S$_N$Ar product [**4.17**+H$^+$]$^+$ and [**4.17**+H$^+$+MeCN]$^+$ and 150 Da and 191 Da for the IS [**4.18**+H$^+$]$^+$ and [**4.18**+H$^+$+MeCN]$^+$ respectively. The base used, NEt₃ can also be seen in the MS at 102 Da and 143 Da again for its [Et₃N+H$^+$]$^+$ and [Et₃N +H$^+$+MeCN]$^+$ (Figure 84).

As well as using the MS for qualitative analysis of the reaction pure samples of each component, apart from product, were prepared and injected into the HPLC using the

same method. These chromatograms show components of the reaction which cannot be picked up *via* the MS *i.e.* the starting material, as well as those which do show and also contain a chromophore. The starting material is only detected *via* HPLC, internal standard and product by both HPLC and MS whilst piperazine and base can only be seen in the MS. Together both these techniques can be used to monitor the reaction gaining knowledge of different components to create a more powerful process. MS struggles to differentiate between regio-isomers and so HPLC will be able to separate and quantitate these aspects and MS other species. This was particularly useful when carrying out the DoE in continuous flow as both HPLC and MS were used as the monitoring techniques.
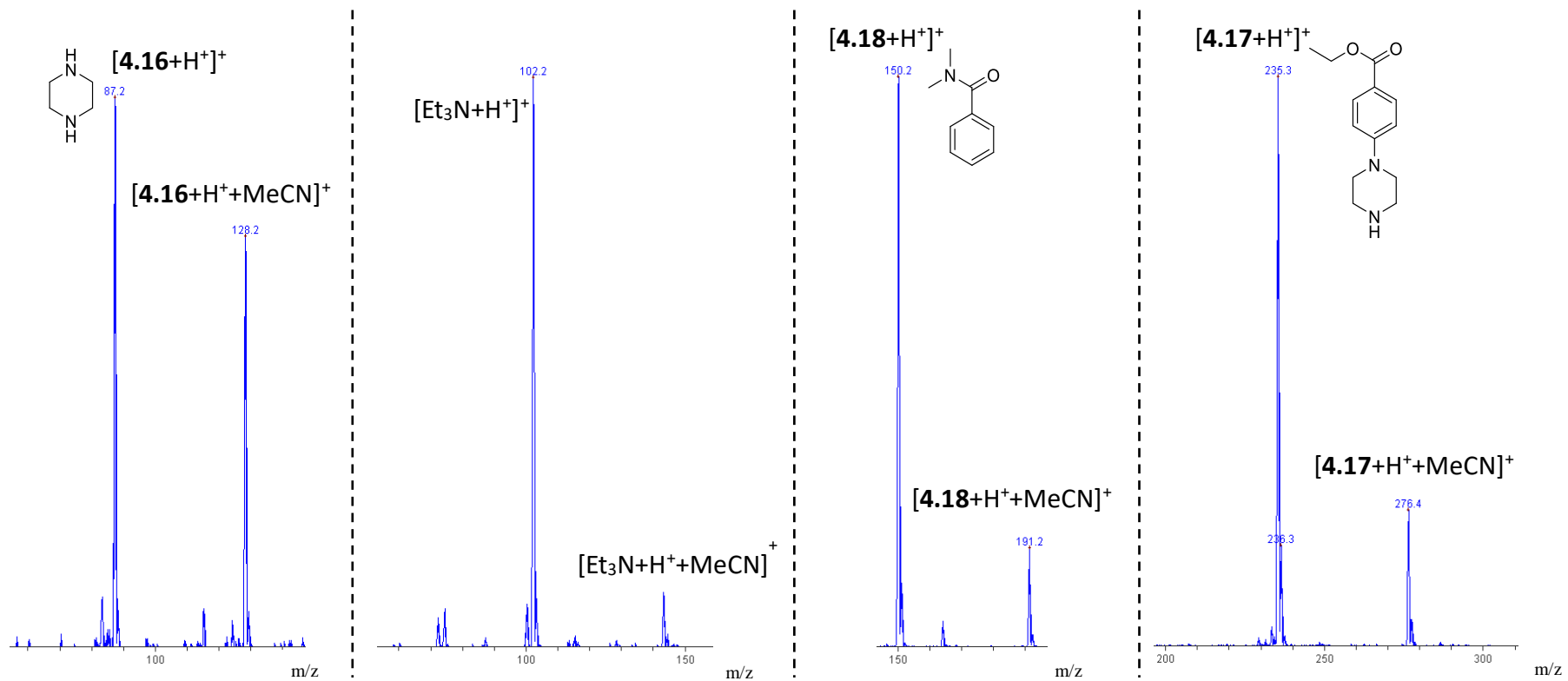
*Figure 84 Isolated MS spectra from the TIC in Figure 83. The first outlines piperazine, second the base in the reaction, third the internal standard and final spectrum from the desired $S_NAr$ product. Each spectrum shows the molecules $[M+H^+]^+$ peak and $[M+H^++MeCN]^+$ adduct peak.*

The DoE list was prepared using Modde and a 2-level full factorial design for screening using the upper and lower boundaries found in Table 25. The maximum boundary used for temperature was increased slightly to push conversions higher particularly at the most extreme conditions of the DoE. The DoE was carried out using 2,6-*cis*-dimethylpiperazine in sulfolane using the MS method previously developed. The HPLC method had to be adapted for use with this compound and its likelihood of developing a regio-isomer in the reaction. The final two DoE experiments did not run due to pump over pressure causing the program to fail. However, there is still enough data to put together a model which was gained *via* HPLC and MS (Figure 82).

Sampling of the reaction was completed using two different sample loops (section **4.6.4**, Figure 93); the first for analysis *via* on-line MS and the second *via* on-line HPLC. HPLC is already used as a quantitative tool but has typically long method times and cannot show which species is which. Most often, MS is used as a detector for HPLC to provide qualitative data on each of the HPLC peaks. This work has tried to show that using on-line MS can be both qualitative and quantitative to maximise the data output from one analytical technique which can be completed rapidly.

The starting material and desired product (as well as side products) are all easily detected by on-line HPLC due the occurrence of a chromophore. The final HPLC method saw species elute at 1.6 (IS), 2.3 (DP), 2.7 (SP), 2.9 (SP), 4.6 (SP), 5.5 (SM) minutes, with good resolution and a steady base line, over five different wavelengths; 210, 230, 254, 270 and 280 nm.

These molecules, however, are not as easy to ionise for the MS to detect. APCI was used as the ionisation mode as the reaction system involves more non-polar species. APCI positive mode relies on atoms in the analyte having greater electron density which can accept a proton, in order to be ionised. Both the IS and *cis*-2,6-dimethylpiperazine were detected and showed peaks at *m/z* 150 and 115 respectively.

*Table 25 DoE boundaries used for the S$_N$Ar reaction to study AZD4547.*

|  | **Temperature/ °C** | **tR/ mins** | **% Water** | **Eq Piperazine** |
|---|---|---|---|---|
| **Lower** | 130 | 5 | 40 | 1 |
| **Centre** | 180 | 17.5 | 65 | 1.75 |
| **Upper** | 230 | 30 | 90 | 2.5 |

Using HPLC data an initial model was put together for reaction conversion using 15 of 19 data points. Two data points were not gathered whilst the other two were seen to be outliers by the model. The data from the DoE showed that the % water added to the reaction had the greatest effect on SM conversion followed by residence time and equivalents of nucleophile added. Temperature of the reaction had a lower effect on the conversion than the other three factors. The model gives and $R^2$ value of 0.974 and a $Q^2$ value of 0.921 with high reproducibility. The residuals plot outlines almost a perfect straight line suggesting good fit of the model and predictability from the observed results (Figure 85).

The fit of the model can also be shown using the observed *vs* predicted plot which plots the input data from the experimental DoE against the statistical predicted data to compare the two data sets. From this plot the given data matches really well with the predicted data where the majority of data points are found on the regression line with a few being found above or below (Figure 86).

From the reaction prediction profile, a contour plot can be generated to show the reactivity surface and show areas to operate for maximum conversion output. In this case the model thinks that the best area to operate the reaction is when all 4 factors are at their highest points (Figure 87) to give a maximum conversion of ~67%.

A model for the yield of reaction was also developed using Modde. This one made use of 17 of 19 experiments where two were not gathered due to pump over pressure. The data showed a wide range of results with very good reproducibility. An $R^2$ value of 0.605 and $Q^2$ of 0.35 show a reasonable model fit and predictability with the residuals plot showing good linearity and reasonable data fit. Along with these, the model shows that temperature and residence time affect the yield of reaction but % water and equivalents of nucleophile do not (Figure 88).

The observed *vs* predicted plot outlines more variance in the data but at a constant level. This variance in the model could be due to how the model predicts the important variables and how they affect the output. In this case, the model and disregarded two main factors (base equivalents and SM concentration) as affecting the output which may cause variance in the predicted data. Each point or set of points is above and below the regression line *ca.* equal distance again reflecting the residuals plot of the extremes being both under and overestimates for yield. Formation of a side-product may contribute to the prediction variance as the model has not considered how side-product formation

affects the yield of desired product. The factors percentage water and nucleophile equivalents were both important in terms of conversion of the starting material. The fact that they have now been disregarded for yield may also suggest that side-product formation plays a large role in the reaction. HPLC data provides evidence for the formation of a competing side product most likely the regio-isomer in the $S_NAr$ reaction. Designing a self-optimisation which could maximise the HPLC peak of desired product whilst also minimising the HPLC peak of the regio-isomer would be highly beneficial to the reaction.

Optimum conditions can still be estimated through the generation of a contour plot like the plot established for conversion (Figure 89). This allows the best conditions to be found as well as an estimated outcome for those conditions. The plot shows that at 198.5 °C and 29.9 minutes a total yield of 38.5% should be achievable.

The solvent choice for the DoE ran was a sulfolane:$H_2O$ mixture. Use of water in the system seems to be important from the DoE as it increases overall polarity of the reaction system which could stabilise the Meisenheimer complex. There is the potential of swapping the water to another polar solvent, MeOH, as this solvent dissolves all components of the reaction but also maintains polarity of the system. Sulfolane is better to use than DMSO as it is more stable at higher temps (DMSO degrades at higher and prolonged temperatures) but use of sulfolane in flow could be problematic as it is viscous. Using another solvent *i.e.* with a low boiling point, may be beneficial as it will be easier to remove a lower boiling point solvent and separate products rather than using various liquid-liquid extractions with sulfolane, increasing the waste of the process.
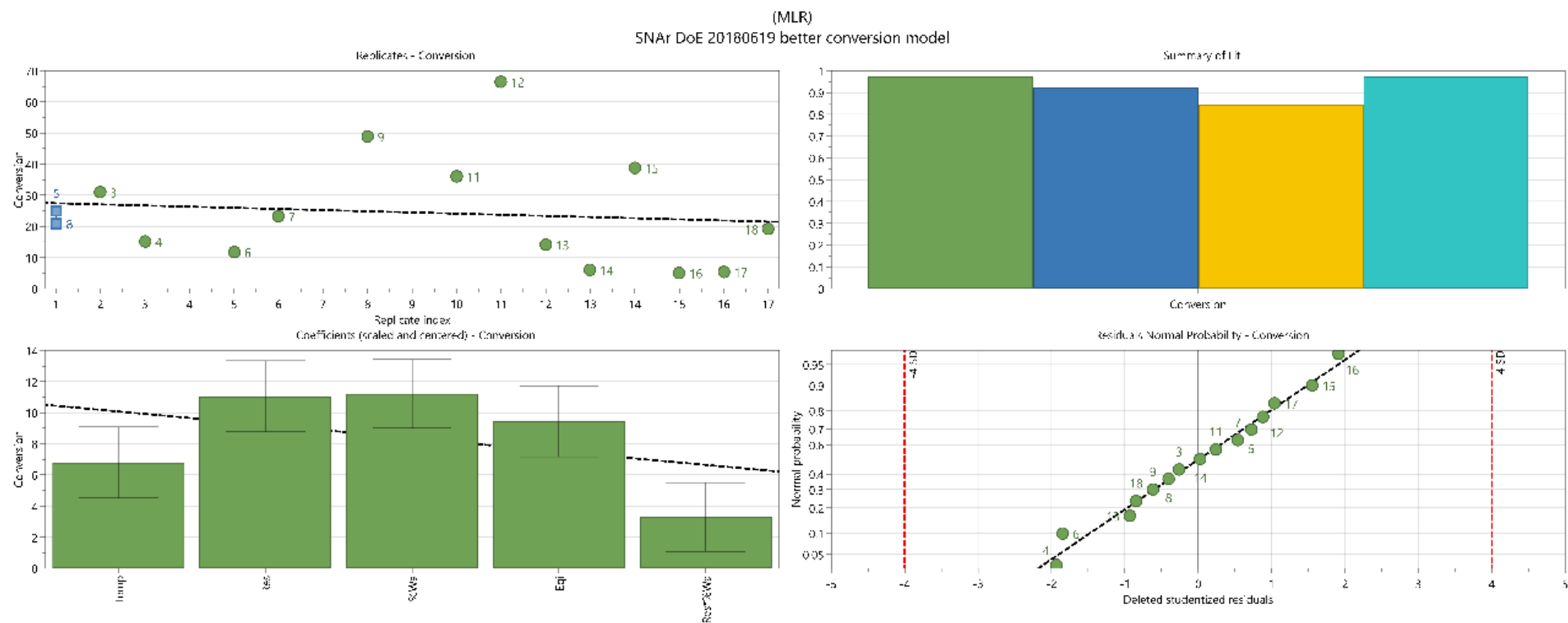
*Figure 85 HPLC conversion model summary for the DoE carried out for the S$_N$Ar reaction of ethyl-4-fluorobenzoate and 2,6-cis-dimethylpiperazine.*
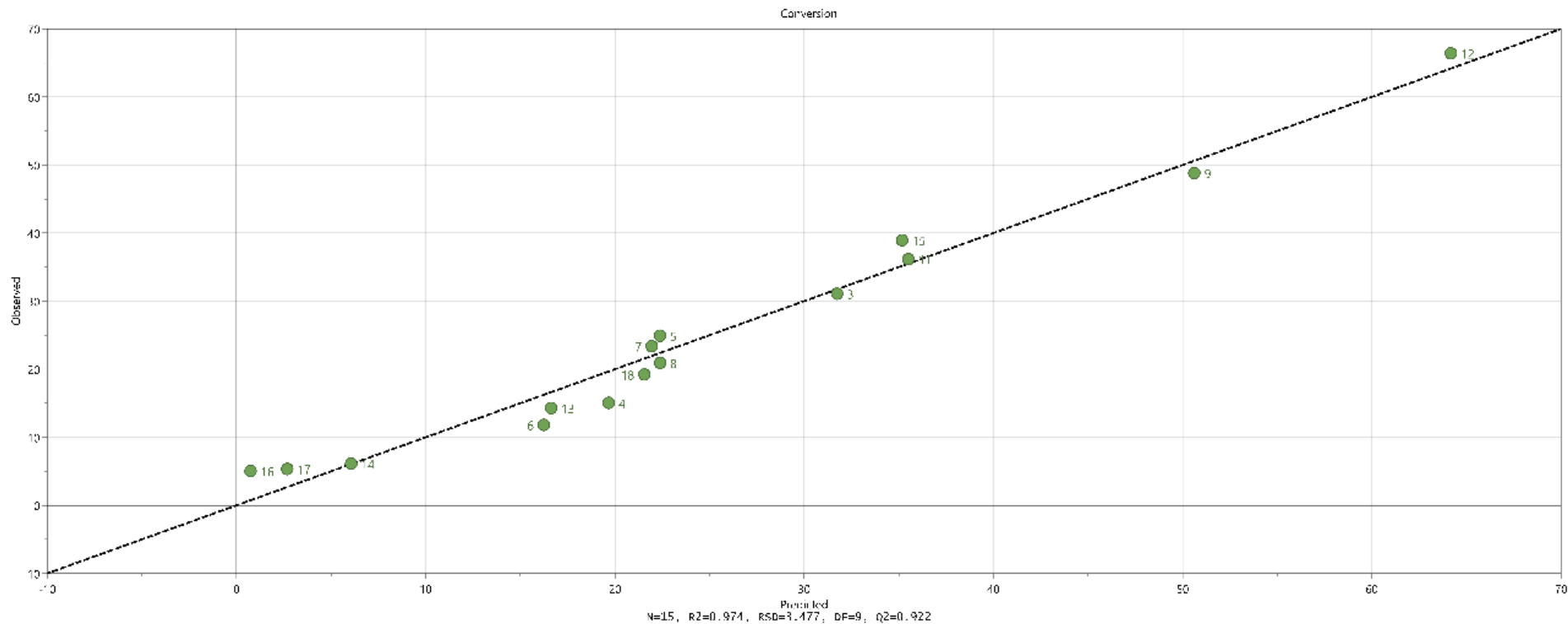
*Figure 86 S$_N$Ar DoE observed vs predicted plot outlining the given data against statistical predicted data.*
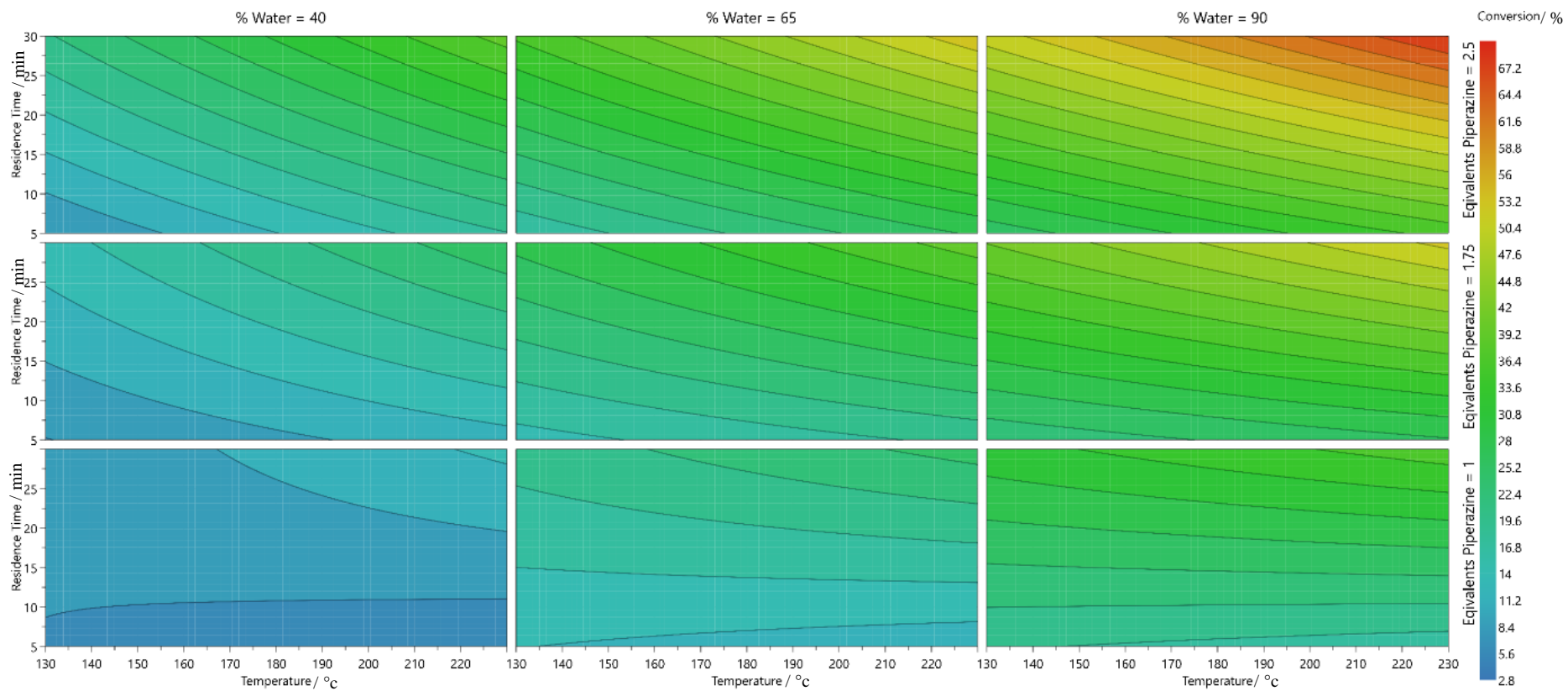
*Figure 87 S<sub>N</sub>Ar DoE contour plot outlining the best operating conditions using all 4 factors. The best area can be seen at the top right corner of the plot when all conditions are at their extreme points.*
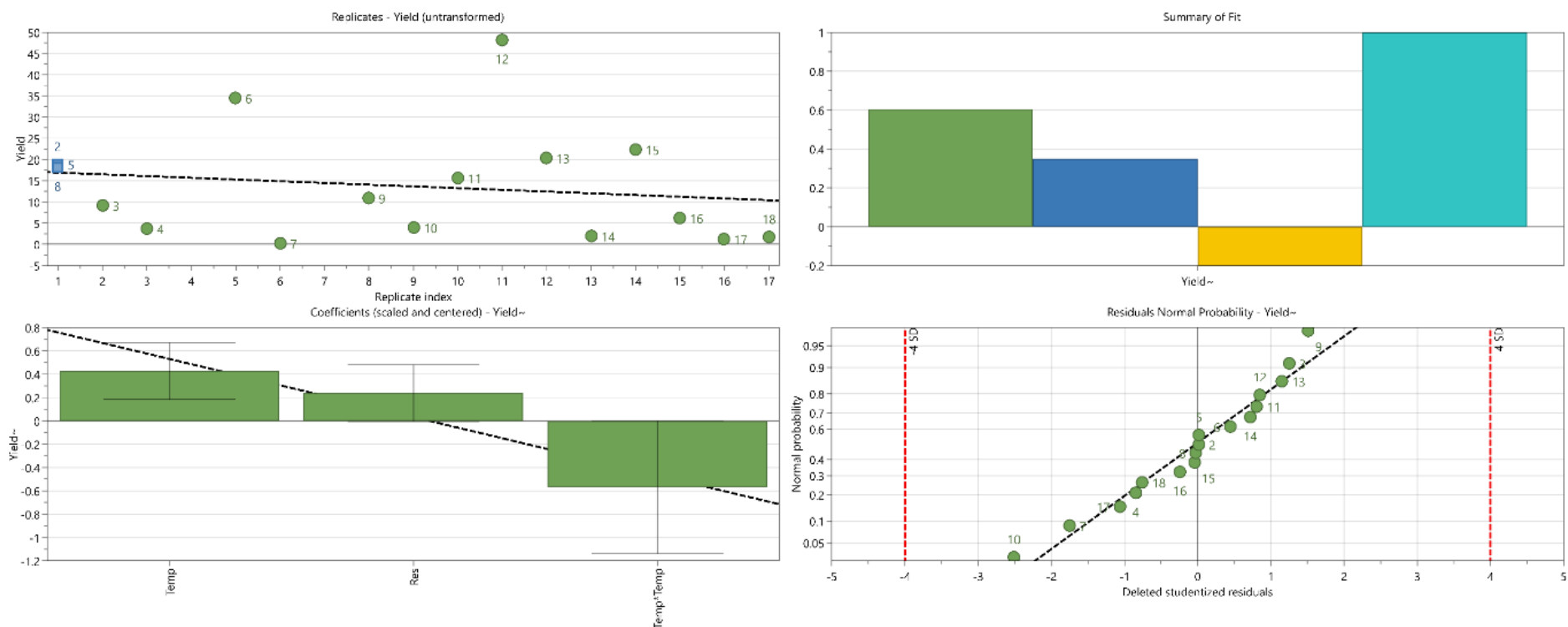
*Figure 88 DoE summary of fit plot for HPLC yield of the $S_NAr$ reaction studied.*
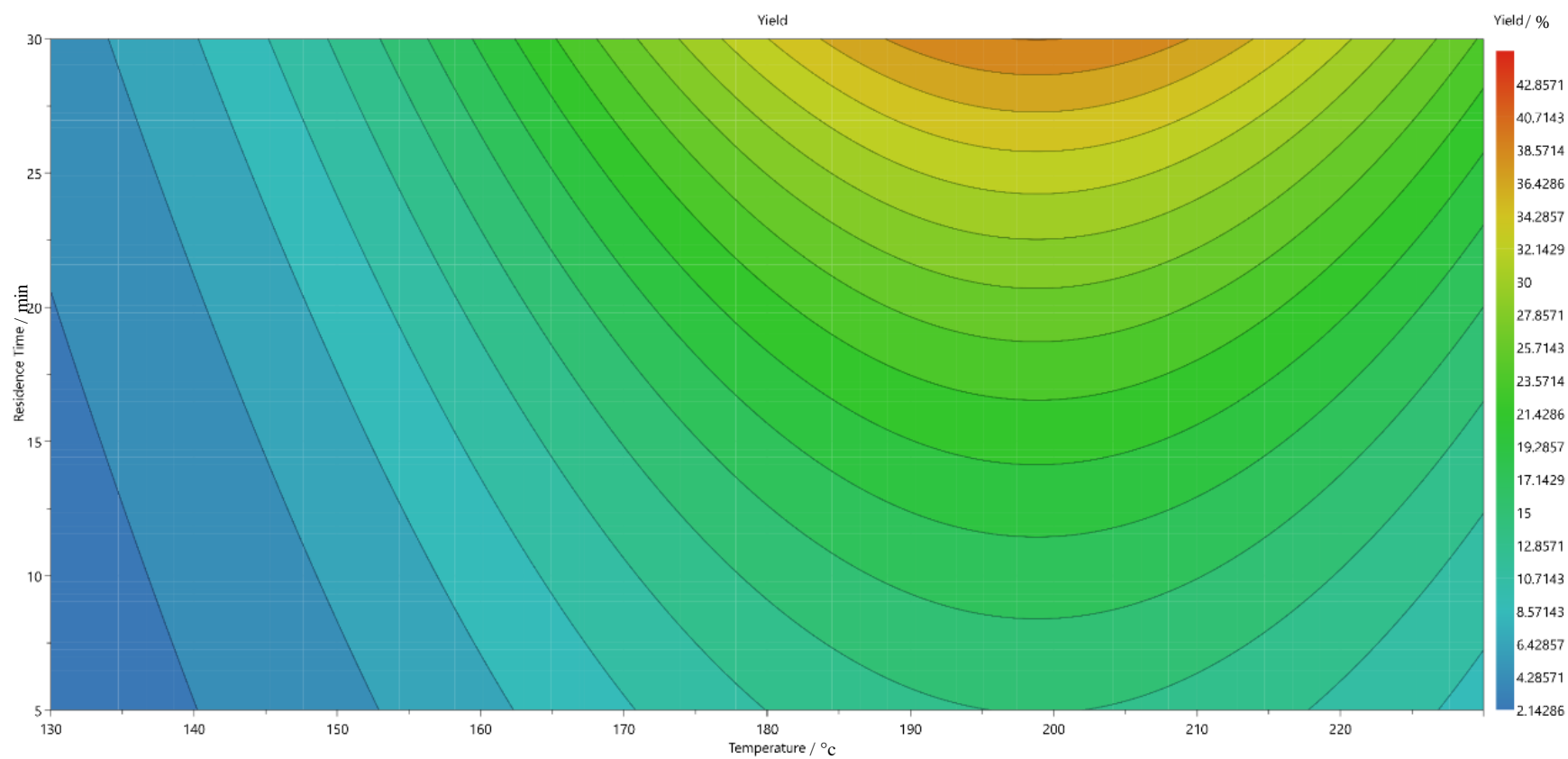
*Figure 89 Contour plot for yield of $S_NAr$ reaction calculated during the DoE.*

For further development of the system, an automated, continuous flow calibration was completed using the same starting solution of reagents that were used for subsequent self-optimisation and kinetic reactions. The flow rates were changed for each of the 3 pumps involved to adapt the concentration of starting material, equivalents of *cis*-2,6-dimethylpiperazine and dilution. Table 26 summarises the conditions used to get the concentrations necessary for the calibration. Each set of conditions were repeated, and an average taken for the final data before being plotted; HPLC area *vs* concentration. The IS was doped into the starting solution alongside the SM at 10 mol%. Equivalents of piperazine were set as a 1:1 ratio of the SM and was included to make dilution to the required concentration easier. The program used for the calibration was also the same as the one previously described so it was essential to include it for the run to occur.

*Table 26 Conditions used in the automated calibration of ethyl-4-fluorobenzoate, cis-2,6-dimethylpiperazine and N,N-dimethyl benzamide.*

| Concentration SM/ M | Flow rate 1/ mL min$^{-1}$ | Flow rate 2/ mL min$^{-1}$ | Flow rate 3/ mL min$^{-1}$ | Temperature/ °C |
|---|---|---|---|---|
| **0.05** | 0.43 | 0.43 | 2.14 | 25 |
| **0.075** | 0.64 | 0.64 | 1.71 | 25 |
| **0.1** | 0.86 | 0.88 | 1.29 | 25 |
| **0.125** | 1.07 | 1.07 | 0.86 | 25 |
| **0.15** | 1.29 | 1.29 | 0.43 | 25 |

Calibration of reaction components was monitored using on-line HPLC and on-line mass spectrometry using the methods described (section **4.6.5.1**, Figure 94, Figure 95 and Figure 96) where plots for both IS and SM show very good linearity and estimate the concentration of reaction accurately. Unfortunately, as the SM could not be properly detected *via* MS this method of analysis was not as useful as HPLC.

The initial self-optimisation used the below response function (equation **K**), a minimisation function, to calculate a response for the algorithm to optimise but had several problems.

$$R = (areaA)/(areaB)*(areaISO)^5 \qquad \qquad (equ.\ \textbf{K})$$

---

[5] areaA = desired product, areaB = side product and areaISO = internal standard.

HPLC retention times were unstable, undergoing drift which caused the HPLC extraction file to miss integrated peaks giving a zero value. Once or twice, files were not saved meaning there was no data for the file to extract and thus were calculated as zero due to having "empty" values. This led to algorithm confusion meaning manual input of peaks and response values as well as a "dud" file had to be included for progression.

The second response function (equation **L**) accounted for the amount of SM used and the peak area of the DP with 0.001 added to each to ensure a value entered at all times into the algorithm.

$$R=((areaA+0.001)/((areaA+0.001)+(areaB+0.001)+(areaC+0.001))*(areaISO+0.001)^6$$

$$(equ.\ \boldsymbol{L})$$

The algorithm proceeded for 55 experiments without any "empty" values or missing files. However, upon analysing the data the program seemed to have gotten in a cycle and stuck in a loop. The 4D plot of conditions and response shows localised reaction runs on one face in a rough quadrangle shape. There is a large amount of design space which has not been explored which is problematic when using machine learning optimisation algorithms. Due to this, the run was stopped so recalculation could be done so the optimisation could be restarted.

The next self-optimisation involved a final change of solvent due to the sulfolane availability and problems with maintaining consistent flow with the high viscosity of the solvent. Both starting materials were not miscible in $H_2O$, so another main solvent had to be used. MeOH was chosen as it fully dissolved both SM, reagents and IS and is easy to remove. It is important to consider the solvent used in a continuous flow reaction as it is important the reaction still occurs efficiently but reduces the problems when pumping or passing the stream through the smallest parts of the reactor *i.e.* sample loop or BPR. Another problem with using sulfolane is its high boiling point, 285°C. A boiling point this high makes it very difficult to remove and costly particularly if use in a large-scale commercial process.

Initial observations of $S_NAr$ conditions shows the SM:Pip ratio is set to a constant 1:1 ratio after 12 experiments. An early lucky hit may have caused the program to search for a set of conditions which give the pest output signal when using 1 equivalent of nucleophile. This meant that from experiment 12 onwards the program was looking for

---

[6] areaA = desired product, areaB = side product 1, areaC = starting material and areaISO = internal standard where 0.001 was added to each to ensure values were supplied to the algorithm.

conditions building from this experiment and would repeat it and try conditions leading away from this. However, it seems like the program may have gotten stuck in a loop trying conditions which were not as good as this first set and so cycled repeats without any useful optimum progress.

The first instance of a high conversion was found after 15 experiments, the program then decided to explore other condition whilst fine tuning the supposed optimum. This supposed optimum was revisited several times with the latter experiments cycling and repeating suggesting the program was at an end and the best conversion and yielding experiment had been found (Figure 90). The SNOBFIT optimisation preceded a pre-determined kinetics experimental list for a more complete profile of the $S_NAr$ reaction.

The kinetics run was designed within the same parameter space in order to get a complete data set and see if the optimum point *via* DoE and self-optimisation were genuine. The kinetic study was also used to see if kinetic parameters could be estimated to enrich the data set.
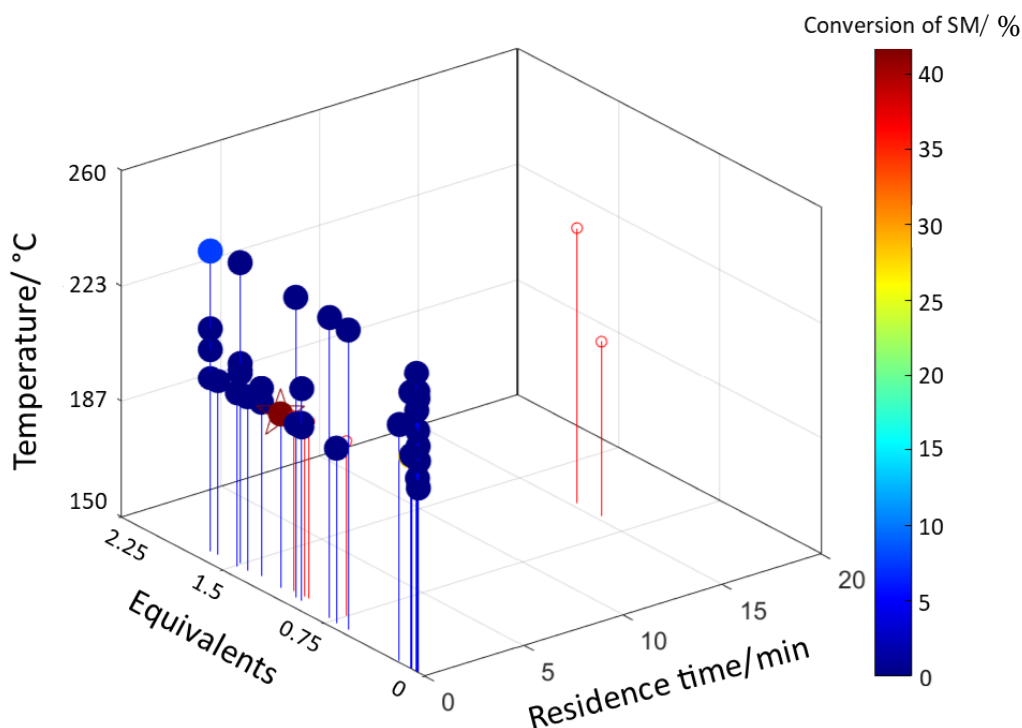


*Figure 90 SNOBFIT plot from the self-optimisation run of the $S_NAr$ reaction of AZD4547. The starred point in red is the optimum point with conditions; 201 °C, 1 eq. and 1 minute.*

Kinetic simulation of the gathered data could only estimate several K values (equilibrium constants) for the data as conversions were low and not enough residence times were used for each temperature.

For both sets of equivalents of nucleophile the low K value suggests that there is more reactant in the solution than products adhering to the low observed conversions and yield. The kinetic simulator used (developed within the group) failed to assign activation energies for the reaction as the amount of data was insufficient for it to get a good profile. These results also indicate that there is a higher tendency of the reaction to form the desired product over the side-product with most being around twice that in relation to the desired product. The highest conversion observed from the HPLC data is shown when at 1.5 equivalents, 225 °C and 20 minutes residence time. Contrasting this, the highest K value from those calculated, is observed at 225 °C and 1 equivalent (Table 27).

*Table 27 Outline of equilibrium constants gained from kinetic reactions during the $S_NAr$ study.*

| Equivalents | Temperature/ °C | K Value for Product/ x10^-6 | K Value for Side-Products/ x10^-6 |
|---|---|---|---|
| | 175 | 4.880 | 1.508 |
| 1 | 200 | 5.060 | 3.919 |
| | 225 | 53.00 | 7.480 |
| | 175 | 4.000 | 1.100 |
| 1.5 | 200 | 44.70 | 25.40 |
| | 225 | 4.300 | 2.350 |

Plotting each temperature series (for each equivalent) on the same graph of conversion *vs* residence time or sample number, there is a general increase in conversion as the temperature is increased. A maximum point is reached at a temperature of 225 °C before conversion levels decrease and more SM remains in solution (Figure 91). The higher temperatures may begin to facilitate side reactions and begin to degrade reaction components. It has been shown that triethylamine degrades, into methane and nitrogen, at high temperatures and low pressures.[119] Piperazine may follow a similar pattern and degrade stopping the reaction and limiting conversion. Another reason could be the high temperatures boiling the solvent, but this is unlikely as the reaction was kept under 1000 psi of pressure using a BPR.
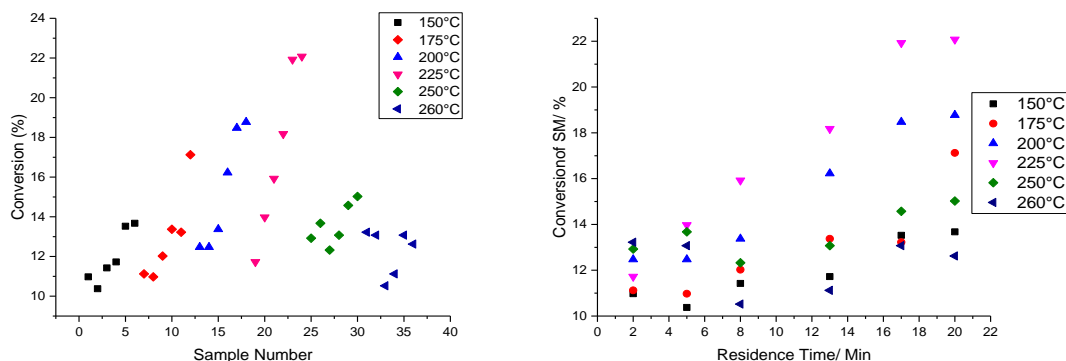
*Figure 91 Graph comparison when plotting conversion vs sample number and conversion vs residence time. Both show the general increase of conversion as the temperature increases up until 225 °C.*

Yield calculations using data from the kinetics experiments assumes that the desired product has a response factor the same as the starting material, so the calibration of starting material can be used to calculate concentration. The calculations then measure the yield of the sample taken (50 μL) based on the maximum amount of SM in the reaction at the start. Calculations still show low yields of desired product across all temperatures which do not allow kinetics to be calculated but give maximum of 3.8% alongside the highest conversions (225 °C, 20 mins, 1.5 eq.) and an average yield of 1.7%.

## 4.5. Chapter 4 Conclusion

Previous work carried out on the compound related to AZD4547 by teams at AZ showed the difficulties in getting a good conversion and yield of the desired product. Using DMSO and water, both high boiling point polar solvents, 14% yield of the desired product was gained when carrying the reaction out 130 °C for an unknown amount of time. Given that there was limited reaction information provided with certain scale-up problems noted, the reaction was adapted into flow well and had few issues. Development of a HPLC method proved simple to resolve both regio-isomers and starting material reproducibly for a continuous flow process. MS method development proved more difficult as only certain species were detected; in particular base, IS and nucleophile. It was possible to detect the starting material but with much more variation in amount, so quantitation proved difficult. Similarly, desired product was difficult to

detect but even if detection could be achieved reliably, it is impossible to separate the desired product and regio-isomer as both have the same mass making quantitation impossible. HPLC results were much easier to quantitate for the DoE, self-optimisation and kinetic study.

A good, reliable model was achieved for the DoE with high reproducibility and good predictability. This model showed that it was possible to achieve ~67% conversion and that % water, had the highest effect on the output closely followed by equivalents, residence time and then temperature. The DoE also showed that a yield of 38% can be achieved although this model shows that %water and equivalents do not affect this output. Both of these show significant improvements on the current conversion and yields gained by AZ, stated above.

Self-optimisation of the system showed results that were unlike the DoE for its optimum set; ~200 °C, 1 eq. and 1 minute residence time. This may be due to the algorithm getting stuck in a loop after selecting well performing conditions causing a repetition. However, the SNOBFIT algorithm does not normally get stuck in loops so it may also suggest that the algorithm had malfunctioned to get stuck in a loop. However, even at these conditions a conversion of 39% was obtained with 22% yield of the desired product.

With regards to the kinetic experiments, the simulator used was unable to use much of the data for a model and useful kinetic parameters. Some K values were obtained but show what much of the kinetic data shows, low conversion and yields outlining that conditions are far away from optimum points. The best conditions (225 °C, 20 minutes and 1.5 eq.) show a conversion of only 21% and ~4% product yield. These are quite different to what the self-optimisation plot shows but align well with the predicted results for the DoE plots in particular the conversion plot (~19% conversion).

### 4.6. S$_N$Ar Reaction Experimental

The general setup for flow work performed is described in section 2.7.1 where details on pumps, reactors, analysis and programming can be found. Specific details for each reactions setup are provided in this section for the work with Boc-Arg(Pbf)-OH the DoE and kinetic hybrid model and work carried out prior to this run. Other details outlining reagents and solvents with purities and grades can be also found in this section along with details of how analyses of samples were performed.

#### 4.6.1. Microwave Batch Reactions

Ethyl-4-fluorobenzoate (0.03 g, 0.18 mM) was dissolved in sulfolane (2.5 mL) before piperazine (0.015 g, 1 eq.) and N,N-dimethyl benzamide (2.7 mg, 10 mol%) were also added to the solution. Triethylamine (37 μL, 2 eq.) was added to the previous solution before water was added (2.5 mL) to make a 50:50 sulfolane/water mix. Other solutions like this were prepared but with varying levels of water and sulfolane (Table 28). Each solution was then reacted using a microwave reactor (as in **2.7.1**) either at 180 or 220 °C for 20 minutes (Table 28). Mass Spectra:

HPLC: 1.60 (IS), 2.26 min (DP), 2.73 min (SP1), 2.92 (SP2), 4.63 min (SP3), 5.46 (SM).

*Table 28 Experimental conditions used for the batch reactions carried out in a microwave reactor.*

| Composition/ S:H$_2$O | Temperature/ °C | Time/ mins |
|:---:|:---|:---|
| **70:30** | 180 | 20 |
| | 220 | 20 |
| **60:40** | 180 | 20 |
| | 220 | 20 |
| **50:50** | 180 | 20 |
| | 220 | 20 |
| **40:60** | 180 | 20 |
| | 220 | 20 |

### 4.6.2. Mass Spectrometry Method

Mass spectra were gathered in APCI+ mode from *m/z* 50 to 800 using a scan time of 938 ms with a total acquisition time of 60 seconds. Other settings of the MS are summarized below.

Polarity: Positive

Capillary Temperature: 300 °C

Capillary Voltage: 100 V

Source Voltage Span: 0 V

Source Voltage offset: 20 V

Source Gas Temperature: 350 °C

APCI Current: 5 µA

### 4.6.3. HPLC Method Development

#### 4.6.3.1. HPLC Method: Microwave Batch Experiments

An injection of each batch reaction was made into the HPLC instrument where each peak was monitored to allow sufficient separation and resolution of each species. HPLC: 1.60 (IS), 2.26 min (DP), 2.73 min (SP1), 2.92 (SP2), 4.63 min (SP3), 5.46 (SM).

*Table 29 HPLC conditions used for separations of the reaction mixture of microwave batch experiments when using piperazine as nucleophile.*

| Time/ mins | % B | Flow Rate/ mL min$^{-1}$ | Max Pressure/ Bar |
|---|---|---|---|
| **0** | 30 | 1 | 450 |
| **0.5** | 30 | 1 | 450 |
| **3** | 60 | 1 | 450 |
| **3.5** | 80 | 1 | 450 |
| **4.2** | 30 | 1 | 450 |

#### 4.6.3.2. HPLC Method: DoE, Self-Optimisation and Kinetics

A solution of ethyl-4-fluorobenzoate (0.03 g, 0.18 mM), 2,6-*cis*-dimethylpiperazine (0.015 g, 0.0013 mmol,1 eq.), N,N-dimethyl benzamide (2.7 mg, 10 mol%) and triethylamine (37 µL, 2 eq.) was dissolved in sulfolane (3 mL) before water (2 mL) was added to create a 60:40 solvent mixture. This solution was then reacted for 20 minutes at 220 °C under microwave conditions. An aliquot of the reacted mixture was injected into

the HPLC for separation. A pure sample of the desired product (0.03 g) was prepared in MeOH (1 mL) and injected into the HPLC to determine which peak was the desired product peak. The method in section 4.6.3.1 was adapted for the separation of the desired product regio-isomer. The rates of change for each solvent were kept constant but the overall method length was lengthened to 6.5 minutes (Table 30). HPLC: 1.60 (IS), 2.26 min (DP), 2.73 min (SP1), 2.92 (SP2), 4.63 min (SP3), 5.46 (SM).

*Table 30 HPLC conditions used for separations of the reaction mixture of microwave batch experiments when using 2,6-cis-dimethylpiperazine as nucleophile.*

| Time/ mins | %B | Flow Rate/ mL min$^{-1}$ | Max Pressure/ Bar |
|---|---|---|---|
| 0 | 30 | 1 | 450 |
| 0.5 | 40 | 1 | 450 |
| 3 | 60 | 1 | 450 |
| 6 | 90 | 1 | 450 |
| 6.5 | 30 | 1 | 450 |

### 4.6.4. Automated DoE reactions
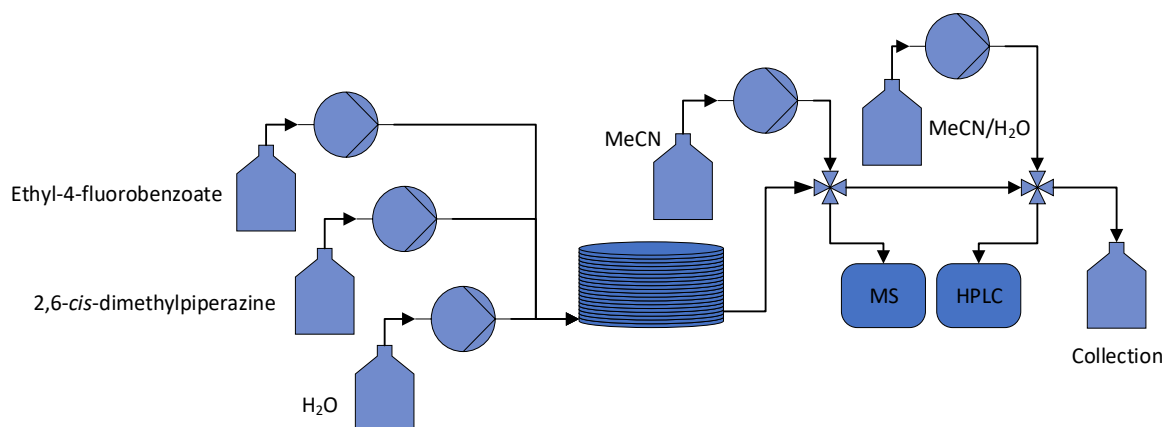


*Figure 92 Automated reactor schematic used during the S$_N$Ar DoE.*

Ethyl-4-fluorobenzoate (1.60 g, 9.5 mmol) was dissolved in sulfolane (50 mL) before N,N-dimethyl benzamide (0.14 g, 10 mol%) and trimethylamine (1.93 mL, 2 eq.) were added to the solution and then connected to a HPLC pump. 2,6-*cis*-Dimethylpiperazine (1.09 g, 9.5 mmol) was dissolved in MeOH (50 mL) and connected to a HPLC pump. Pure MeOH was added to a HPLC pump to change the concentration or reaction and ensure all residence times were met (Figure 92). Mass Spectra: 115.0 *m/z*, 129.3 *m/z*,

143.3 *m/z*, 150.2 *m/z*, 156.3 *m/z*, 170.3 *m/z* (required; 115.2 *m/z*, 150.1 *m/z*, 169.2 *m/z*, 263.4 *m/z* found; 115.0 *m/z*, 150.2 *m/z*). HPLC: 1.60 (IS), 2.26 min (DP), 2.73 min (SP1), 2.92 (SP2), 4.63 min (SP3), 5.46 (SM).

### 4.6.5. Fully Automated S$_N$Ar Reaction and Optimisation Study
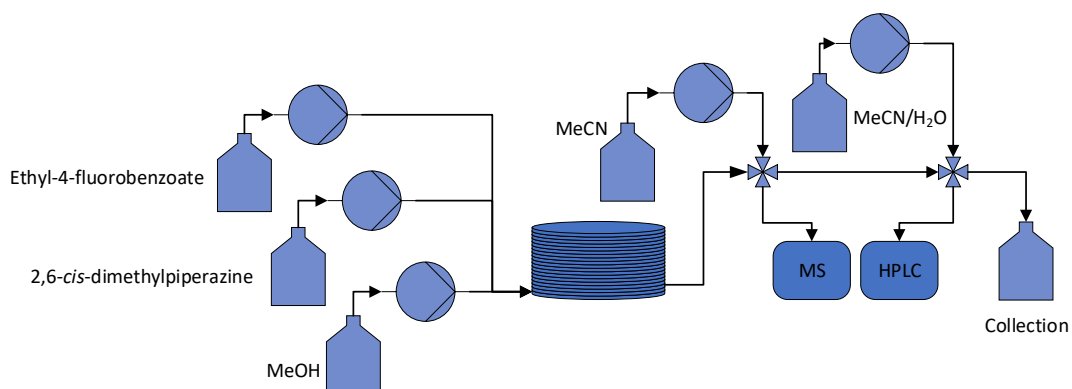


*Figure 93 Automated reactor schematic used during the S$_N$Ar auto-calibration, self-optimisation and kinetic study.*

Ethyl-4-fluorobenzoate (31.0 g, 0.18 mol) was dissolved in sulfolane (250 mL) before N,N-dimethyl benzamide (2.75 g, 10 mol%) and trimethylamine (19.3 mL, 0.19 mol) were added to the solution and then connected to a HPLC pump. 2,6-*cis-D*imethylpiperazine (20 g, 0.18 mol) was dissolved in MeOH (250 mL) and connected to pump 2. Pure MeOH was added to a HPLC pump to change the concentration or reaction and ensure all residence times were met (Figure 93).

### *4.6.5.1. Auto-Calibration*

Hall HPLC pumps were connected to a reservoir feed and purged before being pumped at the set flow rates shown in Table 31. Five different concentrations were selected, and each was analysed by HPLC three times to create calibration curves from average results (Figure 94, Figure 95 and Figure 96). HPLC: 1.60 (IS), 2.26 min (DP), 2.73 min (SP1), 2.92 (SP2), 4.63 min (SP3), 5.46 (SM).

*Table 31 Automated calibration conditions for the continuous flow reactor.*

| Concentration/ M | SM Flow/ mL min⁻¹ | Pip Flow/ mL min⁻¹ | Solvent Flow/ mL min⁻¹ | Temperature/ °C |
|---|---|---|---|---|
| **0.05** | 0.429 | 0.429 | 2.143 | 25 |
| **0.075** | 0.643 | 0.643 | 1.714 | 25 |
| **0.1** | 0.857 | 0.857 | 1.286 | 25 |
| **0.125** | 1.071 | 1.071 | 0.857 | 25 |
| **0.15** | 1.286 | 1.286 | 0.429 | 25 |
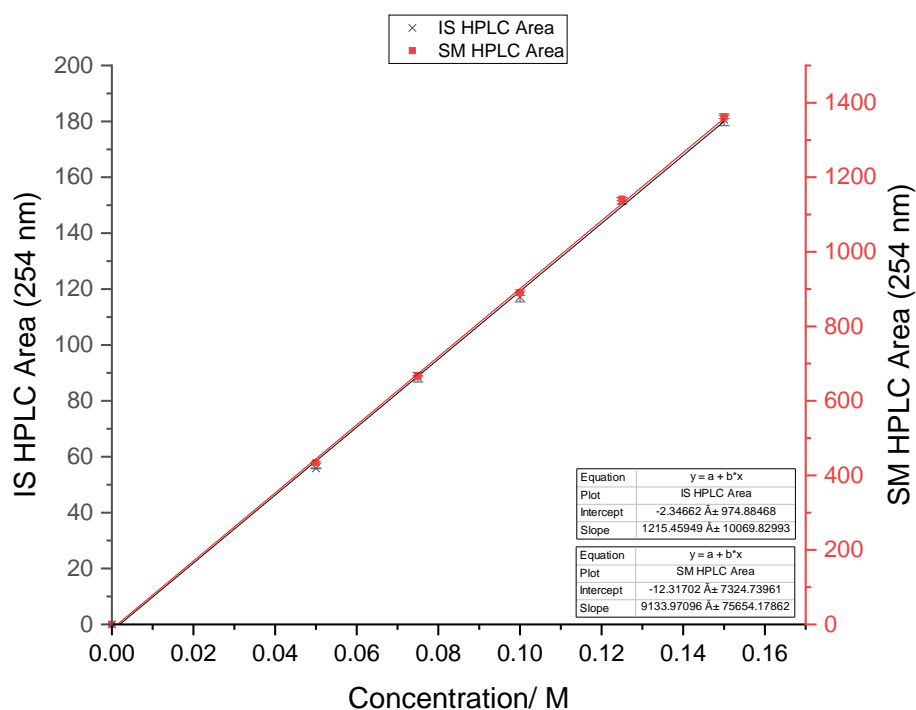


*Figure 94 Calibration graph of ethyl-4-fluorobenzene and internal standard N,N-dimethyl benzamide at 254 nm.*
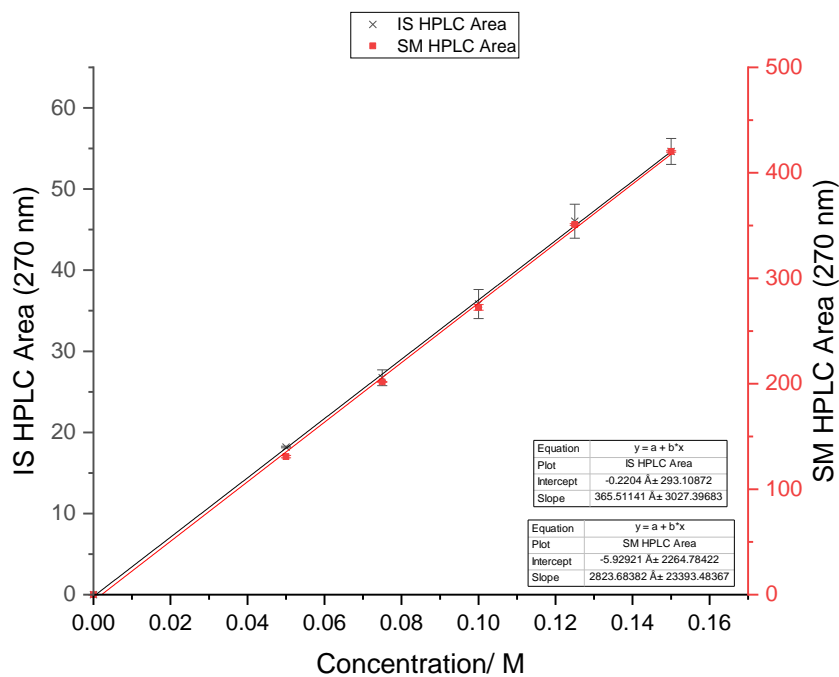
*Figure 95 Calibration graph of* ethyl-4-*fluorobenzene and internal standard N,N-dimethyl benzamide at 270 nm.*
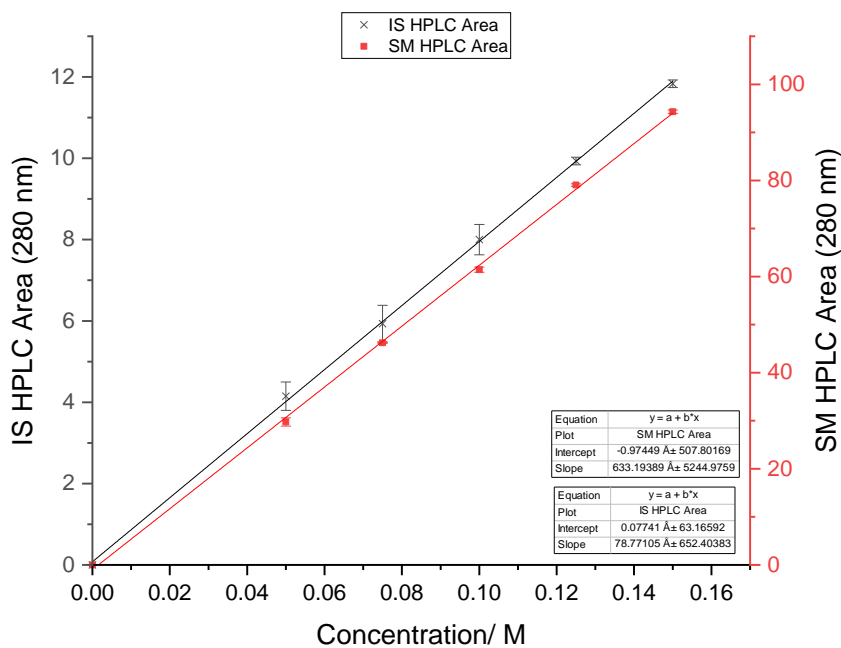


*Figure 96 Calibration graph of ethyl-4-fluorobenzene and internal standard N,N-dimethyl benzamide at 280 nm.*

### *4.6.5.2. Self-Optimisation*

After auto-calibration the self-optimisation was setup using the boundary conditions found in Table 32. Once these had been entered into the computer the optimisation was left to run until an optimum was found. Mass Spectra: 115.0 *m/z*, 150.0 *m/z*, 156.0 *m/z* (required; 115.2 *m/z*, 150.1 *m/z*, 169.2 *m/z*, 263.4 *m/z* found; 115.0 *m/z*, 150.0 *m/z* ). HPLC: 1.60 (IS), 2.26 min (DP), 2.73 min (SP1), 2.92 (SP2), 4.63 min (SP3), 5.46 (SM).

*Table 32 Self-optimisation boundary conditions used for the $S_NAr$ reaction.*

|  | Residence Time/ mins | Ratio | Concentration/ M | Temperature/ °C |
|---|---|---|---|---|
| **Minimum** | 2 | 1 | 0.05 | 150 |
| **Maximum** | 20 | 1.5 | 0.11 | 260 |

### *4.6.5.3. Automated Kinetics Reactions*

A list of 108 experiments was put together to look into the system in more detail. Concentration of the reaction was set to 0.1 M over a range of 6 different temperature, 6 difference residence time and 3 different equivalents of 2,6-*cis*-dimethylpiperazine to generate a total of 108 different experiments to put together kinetic profiles (Table 33). Mass Spectra: 60.1 *m/z*, 115.0 *m/z*, 123.9 *m/z*, 149.8 *m/z*, 155.9 *m/z*, 181.8 *m/z* (required; 115.2 *m/z*, 150.1 *m/z*, 169.2 *m/z*, 263.4 *m/z* found; 115.0 *m/z*, 149.8 *m/z*). HPLC: 1.60 (IS), 2.26 min (DP), 2.73 min (SP1), 2.92 (SP2), 4.63 min (SP3), 5.46 (SM).

*Table 33 Kinetic experimental conditions used for the $S_NAr$ reaction. For each of the 3 different equivalents tested there are 6 temperatures and 6 residence times creating 36 different experiments for a total of 108.*

| Temperature/ °C | 150 | 175 | 200 | 225 | 250 | 260 |
|---|---|---|---|---|---|---|
| **Residence Time/ mins** | 2 | 5 | 8 | 13 | 17 | 20 |
| **Equivalents** | 1 | | 1.25 | | 1.5 | |

Chapter 5

**An Electrochemical Method for Synthesising Iron-NHC Complexes in Continuous Flow Monitored by On-line Mass Spectrometry; Towards a Telescoped Catalytic C-H Hydroxylation Reaction.**

# 5. Formation of Iron-NHC Complexes and Their Application in C-H Activation Reactions

Chapter 5 introduces carbene species, in particular NHCs, their synthesis and catalytic applications. Research has been conducted in forming iron N-heterocyclic carbene (Fe-NHC) complexes in continuous flow using a newly developed electrochemical flow (ECF) cell whilst being monitored by on-line analytical techniques, in particular on-line mass spectrometry. The development of this system will advance the field of electrochemically synthesised catalysts and their potential use in telescoped catalytic reactions.

## 5.1. Introduction

### 5.1.1. Carbenes and N-Heterocyclic Carbenes

Carbenes are divalent carbon species with the carbenic carbon atom containing two non-bonding electrons and only six valence electrons, hence it does not obey the octet rule. The non-bonding electrons can typically exist in two different forms. The first, having anti-parallel spins, creates a singlet carbene with both electrons existing in the same $sp^2$ orbital (named $\sigma$). The second type, having parallel spins, creates a triplet carbene where the electrons exist in different orbitals, one in the $sp^2$ ($\sigma$) orbital and one in a p orbital ($p_\pi$). The atoms adjacent to the carbenic carbon determine the relative energies of the $\sigma$ and $p_\pi$ orbitals, hence determines the spin state of the carbene.

N-Heterocyclic carbenes (NHCs) are contained within a heterocycle with nitrogen atoms adjacent to the carbenic carbon. The electron withdrawing nature of the adjacent nitrogen atoms stabilises the carbene $\sigma$ orbital, whilst the mesomeric effect of the nitrogen lone pairs increases the energy of the $p_\pi$ orbital, thus NHCs exist in the singlet spins state (Figure 97).[120]

**A**    Singlet Carbene        **B**    Triplet Carbene

π-Electron donating R-Groups   σ-Electron withdrawing R-groups     π-Electron withdrawing R-Groups   σ-Electron donating R-groups
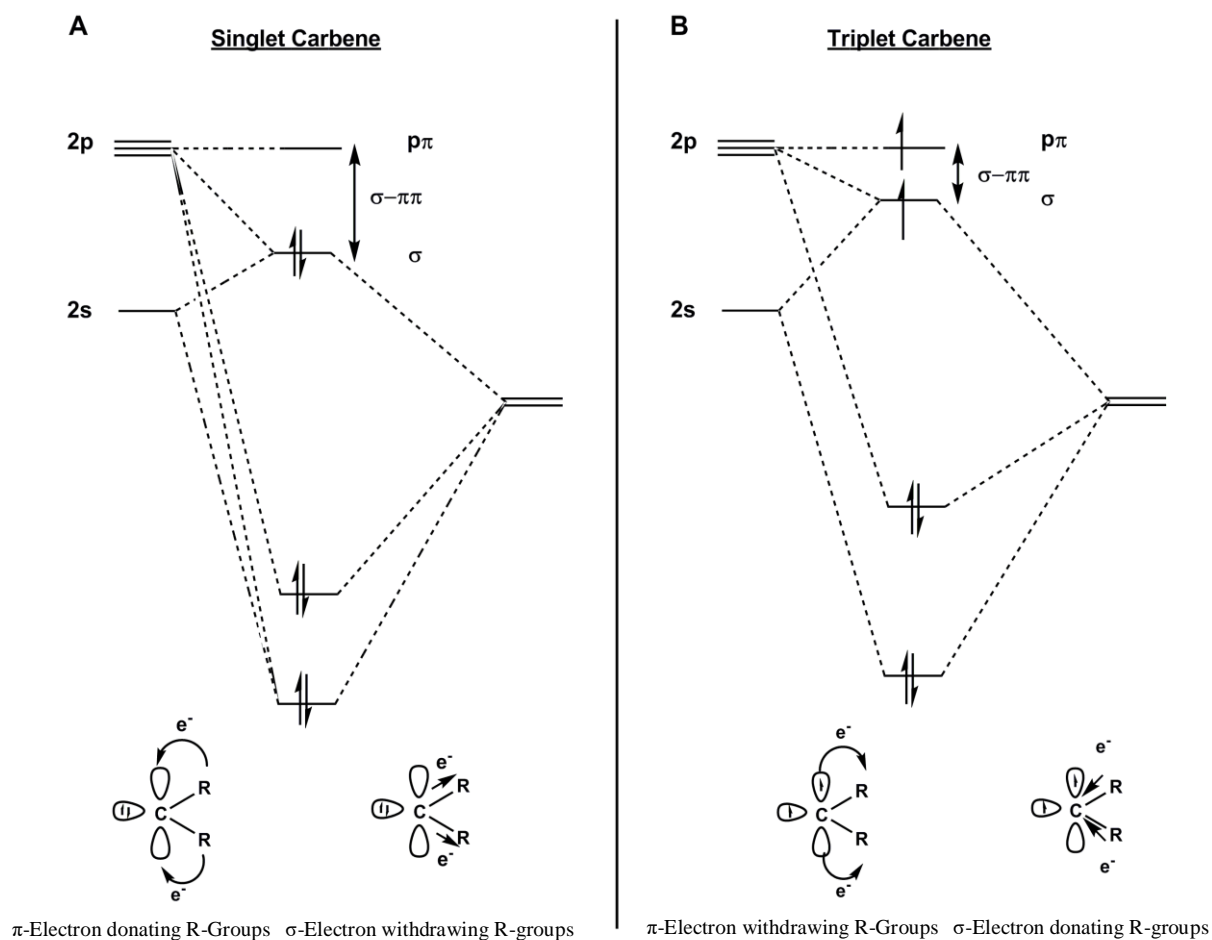
*Figure 97 Molecular orbital diagrams, molecular species to show the lone pairs for both singlet carbene and triplet carbene.*

The first NHC to be isolated in the absence of a metal centre was reported in 1991 by Arduengo, in which an N,N'-diadamantyl-imidazol-2-ylidene was described as the first 'bottleable' NHC (Scheme 30).[121] Since then the field of NHCs has flourished, with these compounds being developed for catalytic, biomedical and materials applications.[120,122–124] NHCs commonly comprise two bulky substituents on either side of the carbene which improve the kinetic stability of the NHC.
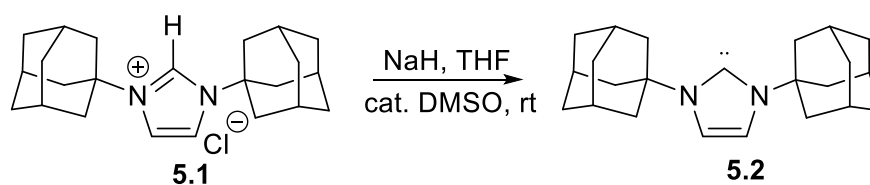


*Scheme 30 Synthesis of N,N'-diadamantyl-imidazol-2-ylidene.*[121]

NHCs are electron rich, neutral σ-donor ligands and have greater electron donating ability than other similar ligands *e.g.* phosphines. 5-Membered NHCs can easily be tuned

by changing the R-groups on both N-atoms (steric properties) and the ligand backbone (electronic properties).[120]

Stabilisation can also occur from partial-aromaticity of the NHC. A double bond backbone, lone pairs on each nitrogen atom and the carbene electrons provide 10 pi-electrons proving Hückel's rule ($4n + 2 = 10 \rightarrow n = 2$) as true for NHCs with an unsaturated backbone. However, a double bond and therefore aromaticity is not essential for NHC stabilisation as there are many examples of stable NHCs containing a saturated backbone.[125, 126]

### 5.1.2. Transition Metal-NHC Complexes

NHCs form very strong bonds with most transition metals, with metal-NHC complexes generally being more thermally and oxidatively stable than their metal-phosphine counterparts. This makes them ideal for use in catalytic applications as it can prevent the catalyst from degrading and increase activity. In addition, typical metal-phosphine complexes show a cone-shaped arrangement of their substituents whereas NHCs generally adopt an umbrella type shape with the nitrogen R-groups perpendicular to the carbene and facing into the metal, thus have greater steric influence at the metal.

Metal-NHCs are synthesised in a variety of ways, usually through base deprotonation of an imidazolium salt, with the metal coordination often carried out *in situ* (Scheme 31).[127]



5.3          5.4
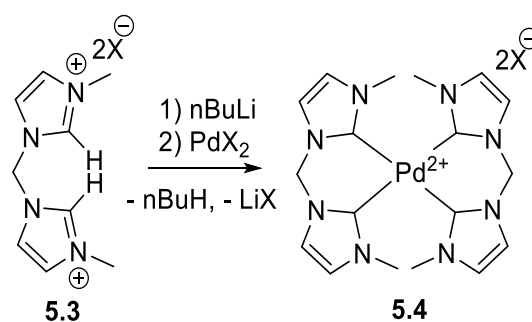
*Scheme 31 Metal NHC formation via deprotonation with an external base.*[127]

A deprotonation reaction can also occur by using a metal precursor containing a basic ligand. This method of synthesis is convenient as many metal precursors contain an acetate, hydride or acetylacetonate ligand (Scheme 32).[127]
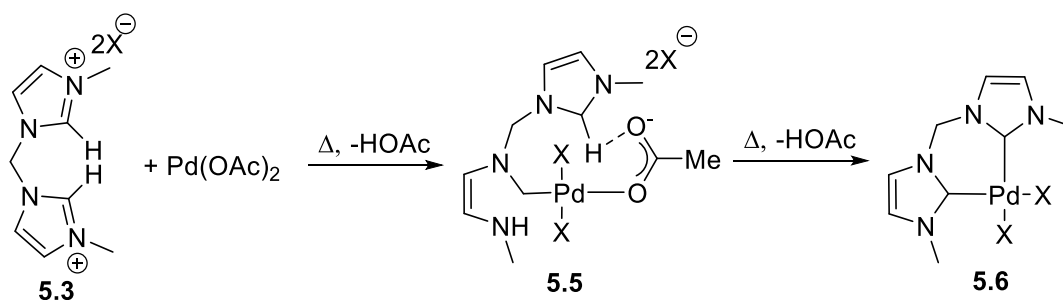
*Scheme 32 Metal-NHC formation via deprotonation with a basic metal precursor.*[127]

Metal-NHC formation can be achieved using transmetallation *via* a silver-NHC complex. A silver-NHC is generated using a basic silver precursor (*e.g.* $Ag_2O$) with transmetallation occurring due to the relatively weak Ag-NHC bond (Scheme 33). This method negates the need for strong bases and strict inert conditions.[127]
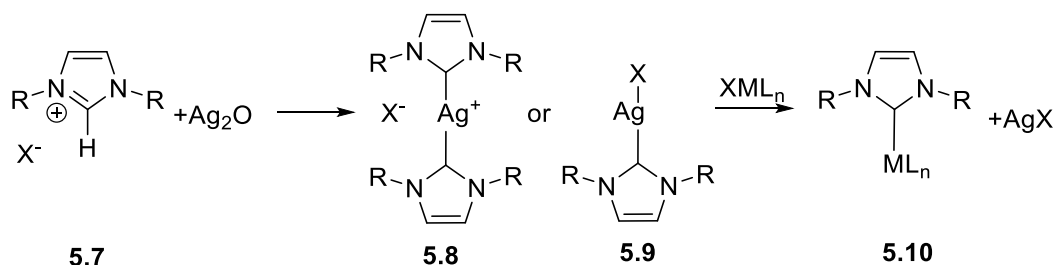


*Scheme 33 Metal NHC formation via transmetallation from a silver-NHC.*[127]

Metal-NHC complexes may also be synthesised using electrochemical means either in batch or in flow. In this case it is electrons acting as the reactive species to form the complexes with only $H_2$ as the by-product (Scheme 34).[128]
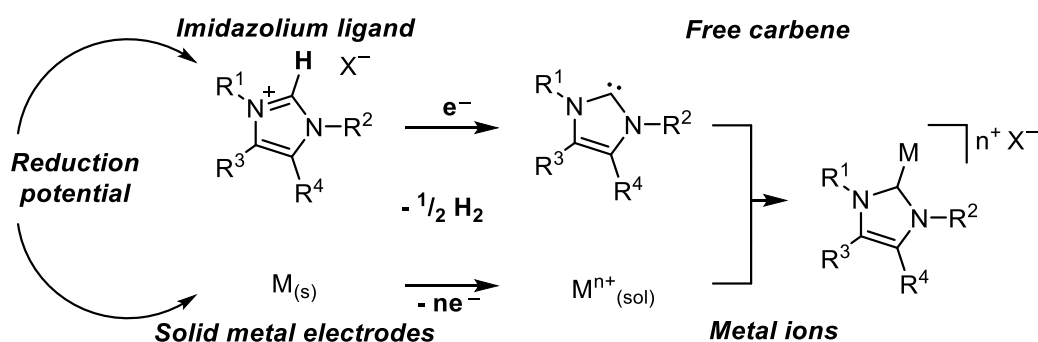


*Scheme 34 Electrochemical reaction performed in continuous flow to generate Fe-NHC complexes, monitored by on-line mass spectrometry.*

## 5.2. Applications of Metal-NHC Complexes

Metal-NHCs have been used in many catalytic processes including; metathesis, hydrogenation,[129] Sonagashira,[130] Kumada,[131] Stille[132] and Suzuki[133] couplings and many more (Table 34). Metal-NHCs have the main advantage when used for catalysis chiefly due to their structural versatility of the bound NHC; they can be made to be chiral, easily functionalised, immobilised on solid surfaces, chelating and made more polar for water solubility or for ionic liquids.[134]

*Table 34 Summary of varying metals which can be bound to NHCs to partake in a variety of chemical reactions.*

| Metal | Reactions | Catalyst |
|---|---|---|
| **Ruthenium**[135] | Olefin metathesis | Grubbs $2^{nd}$ generation |
| **Palladium**[136] | Suzuki-Miyaura | [(NHC)Pd(allyl)Cl] |
| | Allylic alkylation | [(NHC)Pd(allyl)]+ |
| | Borylation of aryl chlorides | - |
| | Direct arylation | [(NHC)Pd(PPh$_3$)I$_2$] |
| | Heck reaction | [(NHC)Pd(Phos)Cl] |
| | Hiyama raction | [(NHC)Pd(3-ClPy)Cl$_2$] |
| | Kumada | [(NHC)Pd(Py)Cl$_2$] |
| | Negishi | [(NHC)Pd(3-ClPy)Cl$_2$] |
| | Sonogashira | [(NHC)$_2$PdBr$_2$] |
| | Stille | - |
| | Buchwald-Harwig | [(NHC)Pd(p-allyl)Cl] |
| **Iridium**[137] | Oppenauer-Type oxidation | Ir$^{II}$(Cp)(OTf)$_2$(NHC) |
| **Osmium**[135] | B-H & O-H activation | [(NHC)OsCl(p-cymene)][OTf] |
| **Copper**[138] | Allylic oxidation | [(NHC)CuCl] |
| **Nickel**[139] | Carbene ransfer reactions | Raney Nickel |
| | Suzuki | [(NHC)Ni] |
| | Kumada | [(NHC)(PPh$_3$)NiCl] |
| **Iron**[135] | C-H activation hydroxylation | Fe(mep)/(PDP) |

### 5.2.1. Iron-NHC Complexes and Their Application

Fe-NHCs have been developed for use in catalytic process because of the high earth abundance of iron, low cost, stable prices and low biotoxicity. Iron is not typically associated with catalysis as metals like palladium or iridium (precious metals) dominate the scene. However, using iron in catalysis has brought some benefits to reactions. Iron-NHCs in particular can be used for polymerisation reactions and cyclisation reactions. It was reported by Louie and Grubbs[140] that $[(NHC)_2FeCl_2]$ is very efficient during the atom transfer radical polymerisation (ATRP) of styrene and methyl methacrylate (Scheme 35).
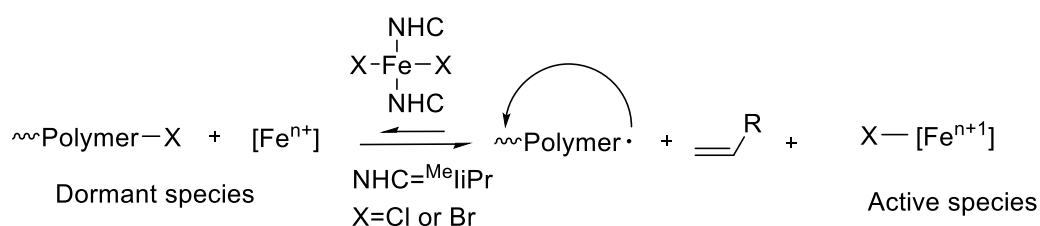


*Scheme 35 ATRP of vinyl monomers using an Fe-NHC as catalyst.*[140]

It has been shown that an intermolecular ring expansion can be successfully completed using an Fe-NHC as catalyst.[141] Styrene oxide was used alongside other dienes, acrylates, enynes or styrenes using a mixture of $FeCl_2$, phosphine, NHC (I*mes*) and a reducing agent and $Et_3N$ as base (Scheme 36). It is not stated which ligand, NHC or phosphine, bonds to the iron only that a mixture gives the greatest reactivity (iron dichloride, $PPh_3$ and I*mes*).
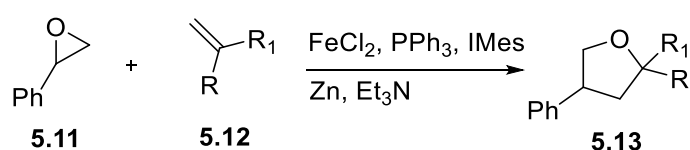


*Scheme 36 Ring expansion catalysis of styrene oxide using an Fe-NHC complex.*[141]

Not only are iron species and Fe-NHC complexes good for polymerisation and ring expansion but also show promise as catalysts for C-H activation reactions. C-H activation and functionalisation reactions are the simplest way to reduce the total number of steps in a synthesis as C-H bond are, more often than not, substituted multiple times before the desired functional group. The main problem is C-H selectivity. New catalysts are being developed which selectively target certain C-H bonds in a molecule, in particular iron catalysts.[142]

Chen and White[143] looked at designing iron catalysts with the intent of using them in C-H activation reactions and the hydroxylation of alkenes and tertiary carbon species. The developed iron catalyst is not a metal NHC complex but shows what can be done with iron as a metal for the development of catalytic species. The designed catalyst consists of a metal centre with 'mep' [N,N'-dimethyl-N,N'-*bis*(2-pyridylmethyl)-ethane, 1,2-diamine] a *tetra*-dentate ligand, two acetonitrile ligands and two antimony hexafluoride counter ions (Figure 98).
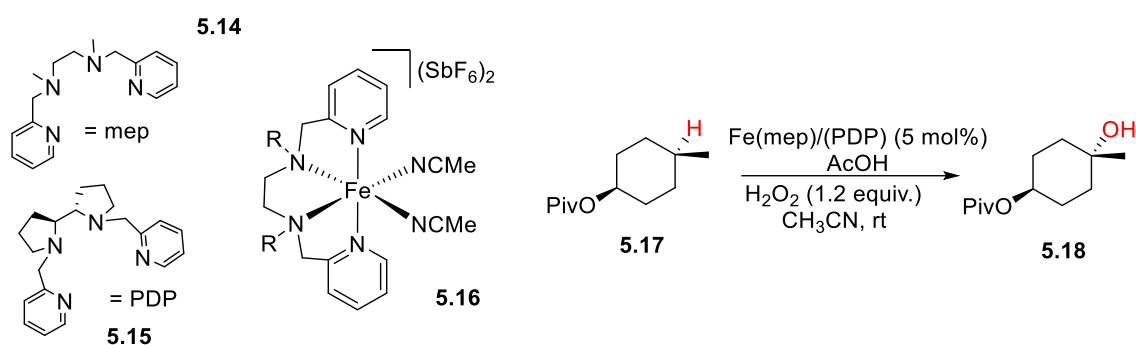


*Figure 98 Iron catalyst developed by the White group for C-H activation reactions where hydroxylation of 4-methylcyclohexane pivalate was explored.*[143]

The initial reaction of interest was the hydroxylation of 4-methylcyclohexane pivalate where the oxidative reaction targets the hydrogen on the *tertiary*-carbon (Figure 98) with good conversion and selectivity to the desired product. The best catalyst, Fe(S,S-PDP), was employed for a variety of other alkenes to show the reactivity of each alkene species and determine which carbon centre the catalyst oxidised predominantly. Low conversions but acceptable selectivity were observed with pivalate and Fe(mep) due to the flexibility of the 'mep' ligand backbone. Fe(S,S-PDP) showed greater rigidity and yields were improved to 51% by adapting the reaction conditions.

Positional reactivity and selectivity of the catalyst was tested using a series of 1°, 2° and 3° alkanes where greatest reactivities were shown from 3° > 2° >> 1°. Within the 3° group the greatest reactivity was observed when an electron withdrawing group was not present, with remote C-H bonds being more reactive than proximal C-H bonds (Scheme 37).
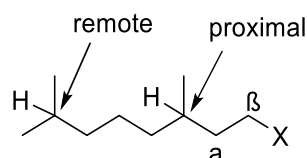
*Scheme 37 Dihydrocitronellol used to show the positional selectivity of the C-H activation reaction using Fe(2,2-PDP).*

### 5.2.2. Electrochemical Reactors and Fe-NHC Complexes

Chapman *et al.*[128] conducted work in exploring the generation of Cu-NHCs for catalysis. This work was completed using a new method of synthesis; an electrochemical continuous flow (ECF) cell. The initial cell made use of traditional batch techniques where the ligand precursor is charged into an appropriate vessel (3-necked round bottomed flask) and metal electrodes placed into the solution to complete the reaction. However, there are practical issues associated with this procedure:

1. A large over potential of the reaction system is required due to poor mass-transfer between electrodes.

2. The Faradaic efficiency of the cell is typically low (Q > 2).

3. Only small amounts of complex can be synthesised at any one time.
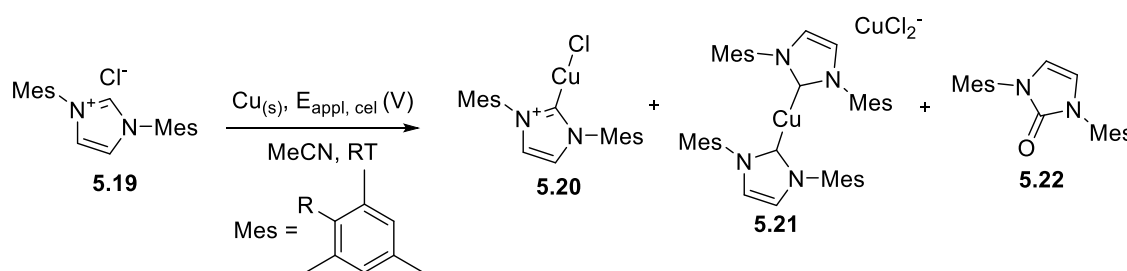


*Scheme 38 Electrochemical reduction/complexation of [ImesH]Cl in flow.*[128]

These drawbacks were overcome by the development of an ECF reactor for synthesis of Cu$^I$-NHC complexes (Scheme 38). Two prototypes were developed to determine which reactor performed best for the reaction shown in

Scheme 38; a linear reactor with PTFE gasket between two copper electrode plates with a separation gap of 2.5 mm giving a reactor volume of 1.9 cm$^3$ (

Figure 99). Initial runs showed limited formation of product **5.20** and larger amount of side product (**5.21**), put down to insufficient mixing of the NHC ligand precursor (**5.19**) and the generated [Cu(NCMe)$_4$]$^+$ in solution. Adaption of the cell to increase formation

of **5.20** saw glass beads inserted into the flow channel to increase mixing, decreasing the reactor volume to 1.05 cm$^3$. In doing so, the formation of **5.21** was completely suppressed increasing the yield of desired product (**5.20**, 36%, with conversions of 92% upon recirculating the reaction mixture through the cell). These ECF reactors have a much greater faradaic efficiency (Q < 2) than the batch reactions allowing much greater efficiency of charge to be transferred to facilitate the electrochemical reaction.
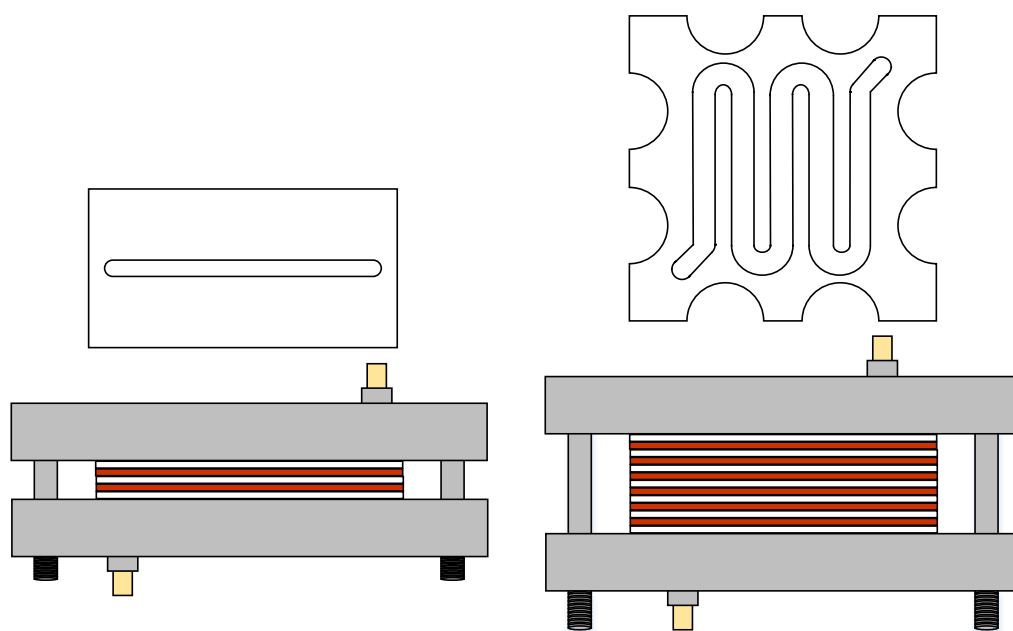


*Figure 99 Prototype electrochemical cells used in the formation of Cu-NHC complexes by Chapman et al. The left image shows the linear flow cell and the right the stacked flow cell. Each one consists of PTFE gaskets containing the flow pathway and Cu metal plates to create an anode and cathode.*[128]

The second cell type developed aimed at increasing the surface area of the metal plates to improve the surface area and efficiency of the cell. Copper plates separated by PTFE gaskets were assembled into a stacked arrangement (

Figure 99). It was necessary to alternate the applied voltage between plates creating an anode and a cathode to facilitate the electrochemical reaction. The modularity of this reactor allows its metrics to be changed accordingly and optimised; channel length, width, electrode separation, interfacial area and reactor volume may all be varied. This second reactor allows much smaller diffusion distances for ions, minimises the cells

resistivity and maximises the current density of the cell. In comparison to the first cell, 1.9 cm$^3$ volume (SA of 7.6 cm$^2$), the second cell used electrodes with a SA of 40 cm$^2$ giving a total SA of 240 cm$^2$ (for six plates). Immediate improvements were observed with regards to the output selectivity of the reaction. Using 1.94 V and 0.67 mL min$^{-1}$ the desired complex, **5.20**, was obtained in a yield of 94% whilst the imidazolinone, **5.22**, yield was reduced to just 3%. The improved yield was completed in a single pass with residence time of 360 seconds. A direct comparison between the two versions of the cell is shown in Table 35.

*Table 35 Comparison between the first and second generation reactors used for the electrochemical synthesis of Cu(IMes)Cl (5.19). $^a$ Registered at steady state. $^b$ In recirculation mode. $^c$ In single-pass mode.*

|  | **Linear Flow Cell** | **Stacked Flow Cell** |
|---|---|---|
| **Single Electrode Surface Area/ cm$^2$** | 7.6 | 40 |
| **Electrode Distance/ mm** | 2.5 | 1.0 |
| **Applied Potential/ V** | 2.50 | 1.94 |
| **Observed Current$^a$/ mA** | 2.7 | 10.0 |
| **Current Density/ mA cm$^2$** | 0.36 | 0.25 |
| **Residence Time/ s** | 126 | 360 |
| **Total conversion time of 0.132 mmol *10* to *11*/ min** | 80.0 (92%)$^b$ | 29.9 (97%)$^c$ |

With the success of the stacked flow cell it was used to examine a variety of imidazolium precursors with different anionic ligands. All species synthesised were done so with yields ($^1$H NMR) between 91% and 95% The group has also extended the reactor to using iron plates for the preparation of Fe-NHC complexes.

The work by Chapman *et al.* shows good initial scope in the development of an ECF reactor for the generation of metal-NHC complexes. By incorporating on-line analysis the need to extract and prepare samples for off-line analysis will be removed. This will enable the possibility to dispense the complex directly into a catalytic reaction.

### 5.3. Summary

- There are several types of carbene species; Fischer, Schrock, Foiled and Persistent.

- Persistent carbenes are generally stable and isolatable making them relatively easy to use in reactions.

- Current methods of metal-NHC synthesis involve hazardous reagents and harsh conditions.

- Method improvements have been made to metal-NHC synthesis involving an ECF cell using metal plates and metal source.

- Using electrons in synthesis can be seen as the ultimate green methodology.

- Metal-NHC complexes have a wide range of uses within chemistry.

- Fe-NHC complexes have been shown to be active in C-H activation reactions.

- It is possible to telescope the electrochemical generation of metal-NHCs with a catalytic reaction.

### 5.4. Aims

The aims of this chapter are to further develop Fe-NHC complex formation in continuous flow using a newly developed electrochemical reactor whilst monitoring their formation by on-line analytical techniques (MS and UV-Vis). Modelling the system with DoE allowed the process to be optimised and a map of the surface of complex formation used to find the best operating conditions. The Fe complexes will be applied to catalytic C-H activation, also be conducted in continuous flow using on-line HPLC to monitor the reaction. Both these reactions will be telescoped together to test catalyst formation and catalytic reaction in tandem, using optimum operating conditions.

### 5.5. Formation of Iron N-Heterocyclic Complexes in Continuous Flow

### 5.5.1. Initial Batch Work

**5.26**

**5.27**
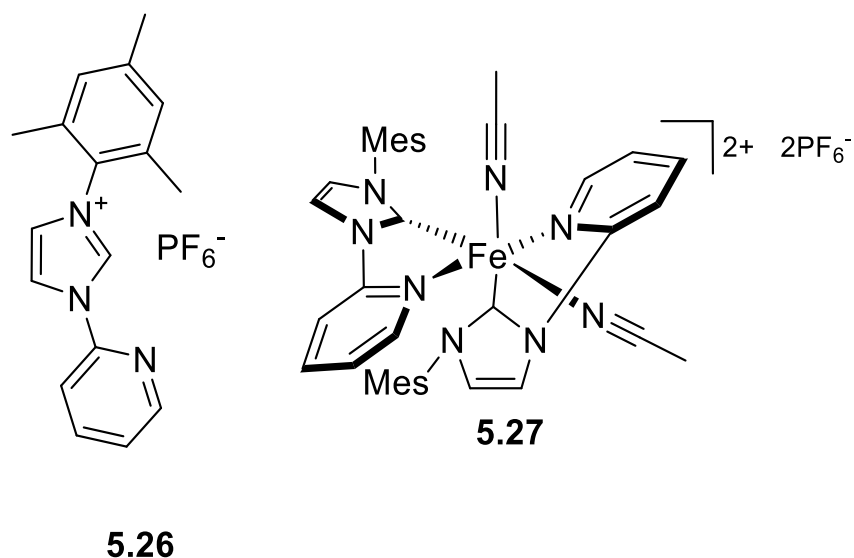
*Figure 100 Fe-NHC complex [Fe(MesPy)$_2$(MeCN)$_2$]2PF$_6$ and ligand precursor initially explored using the electrochemical batch cell.*

Fe-NHC complexes were previously synthesised within the group using electrochemical batch reactors. These samples were analysed using the bench-top Advion MS used for all flow work. The first complex to be analysed was [Fe(*Mes*Py)$_2$(MeCN)$_2$]$^{2+}$ 2[PF$_6$]$^{2-}$, **5.26**, and the NHC ligand precursor **5.27** (Figure 100). The ligand precursor gives a distinct peak at *m/z* 264 as its [**5.26**-PF$_6$]$^+$ ion and the iron complex gives peaks at *m/z* 332 and 695 corresponding to the [**5.27**-2PF$_6$]$^{2+}$ and [**5.27**-2PF$_6$+MeO]$^+$ (Figure 101). Other ligand precursors tested in the batch cell and *via* MS were 1,3-dipyridyl imidazolium hexafluorophosphate, 1,3-*bis*(2,6-diisopropylphenyl) imidazolium chloride, 1,3-dimesitylimidazolium hexafluorophosphate and 1,3-vinyl-(4-methoxypyridyl)-imidazolium bromide.
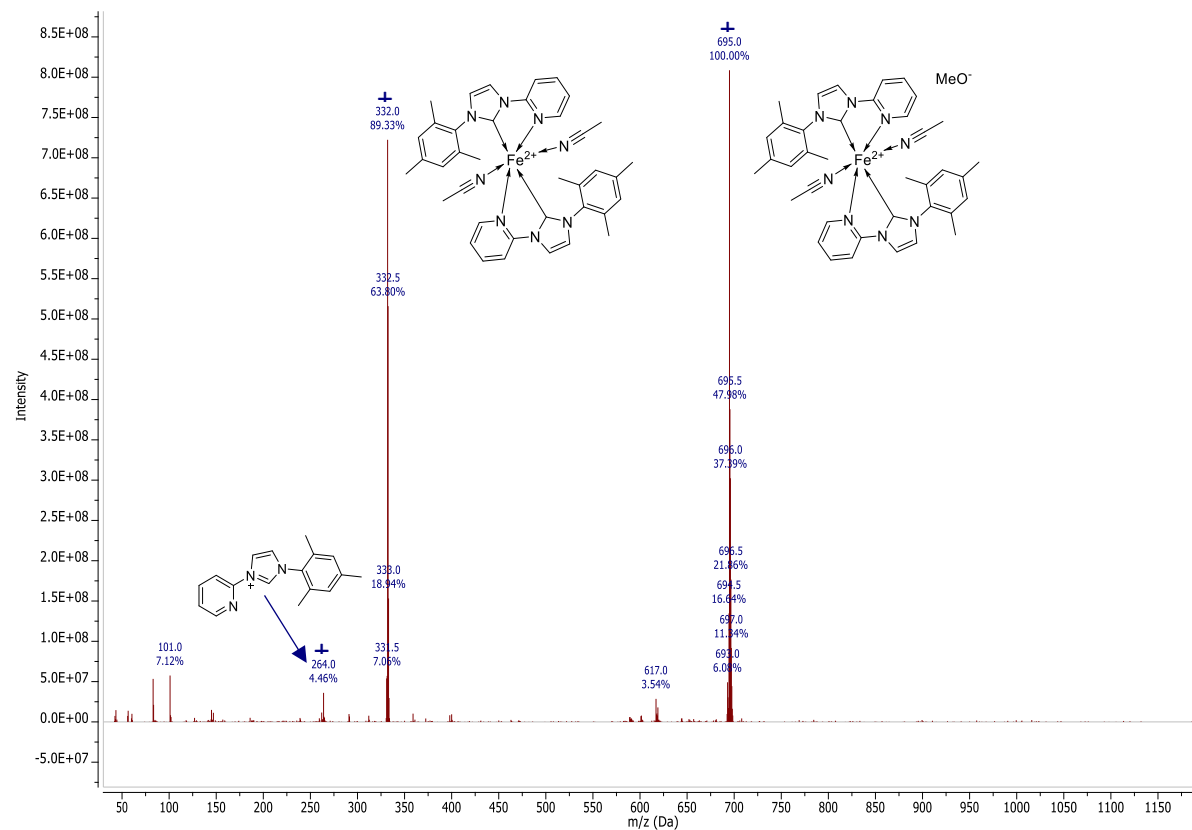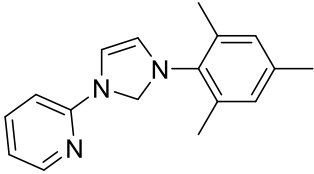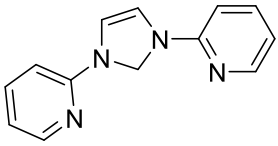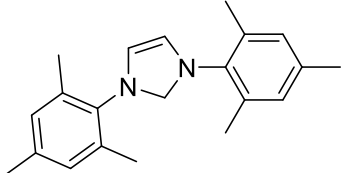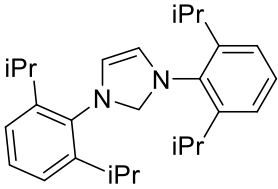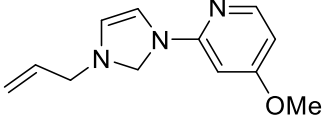
*Figure 101 Mass spectrum of the complex [Fe^{II}(MesPy)_2(MeCN)_2]2PF_6^-. Ligand precursor impurity can be seen at m/z 264 ([**5.26**-PF_6^-]) whilst the peak with an m/z 332 arises from [**5.27**-2PF_6^-]^{2+} species and m/z 695 arises from and adduct with MeO^- ([**5.27**-2PF_6^-+MeO^-]).*

Several other air and moisture stable complexes were tested and analysed and are summarised in Table 36. Each of these were used to provide evidence that MS could be used to detect multiple ligand precursors and complexes. However, only three of the final complexes were observed *via* MS analysis as two formed solid particles so could not be analysed. Of the other three, the I*mes* ligand precursor was very troublesome to dissolve so is not as ideal and using the remaining two ligand precursors, *mes*Py and b*is*Py.

*Table 36 The 5 different NHC ligand precursors used in the ECF reactor for the synthesis of Fe-NHC complexes.*

| Entry | Ligand precursor | Ligand Structure | *m/z* Observed | Observations |
|---|---|---|---|---|
| **5.26** | *mes*Py | | Ligand: 264, [**5.26**-PF$_6^-$]$^+$; Complex: 695, [**5.27**-2PF$_6^-$+MeO$^-$]$^+$ | Red-orange solution formed. Ligand and complex *m/z* seen. |
| **5.28** | *bis*Py | | Ligand: 223, [**5.28**-PF$_6^-$]$^+$; Complex: 613, [**5.29**-2PF$_6^-$+MeO$^-$]$^+$ | Red-orange solution formed. Ligand and complex *m/z* seen. |
| **5.30** | I*mes* | | Ligand: 305, [**5.30**-PF$_6^-$]$^+$ | Would not completely dissolve. Only ligand seen *via* MS. |
| **5.31** | *i*Pr | | Ligand: 390, [**5.31**-Cl$^-$]$^+$; | Solid precipitate formed after reactor. Only ligand seen *via* MS. |
| **5.32** | Vinyl-Py(*p*-OMe) | | - | Orange precipitate formed after reactor. No ligand or complex detected. |

Preliminary runs made use of a batch electrochemical cell with iron electrodes as close together as possible to maximise the current flow across – *ca*. 5 V & 50 mA. Reactions were carried out for approximately 45-60 minutes with final yields varying from 20% to 39%. After 60 minutes, a dark brown insoluble material was produced, presumed to be a mixture of iron oxides and iron (III) hydroxide that deposit after over-potential had occurred. The formation of these could mainly be due to oxygen in the system oxidising any remaining free iron (II) produced from the plates. It is necessary to try and keep the

reaction under nitrogen as well as using dry and degassed solvent. As the cell plates are so far away from one another high voltages are necessary in order to keep a constant current flowing through the solution. A new cell was developed in attempts to improve the reaction process.

### 5.5.2. Batch Cell Improvements
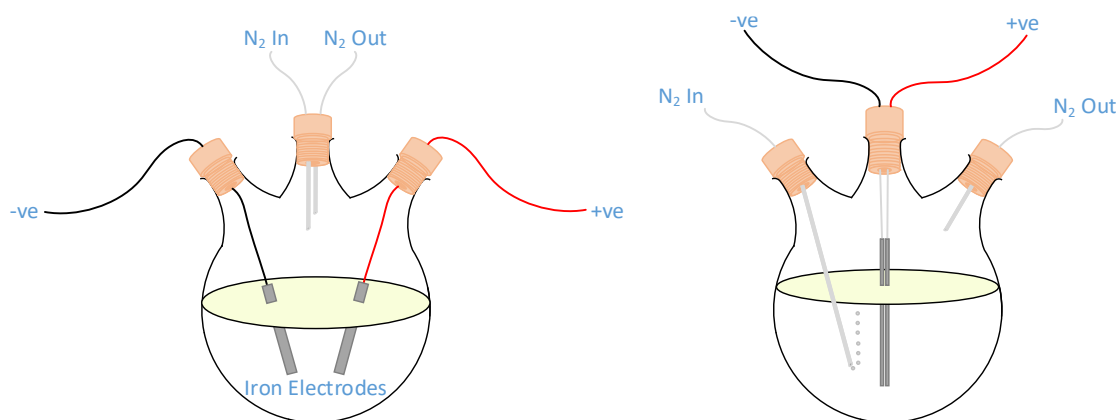


*Figure 102 A comparison between the initial batch cell (left) and the adapted batch cell (right) to bring the metal plates closer together for more efficient synthesis of the Fe-NHC complexes.*

The main improvement to the batch cell necessitated the reduction in the gap between electrodes. Rather than mounting each electrode separately through opposite necks of the round bottomed flask, both were mounted in the central neck attached through the rubber Suba-Seal septum (Figure 102). This ultimately allowed reduction in the voltage improving the efficiency of the cell. It was also necessary to keep the liquid moving between the two closely mounted metal plates. The gap of 1-2 mm reduced the mixing between plates, meaning mixing was more reliant on diffusion of molecules, causing a build-up of reacted species between the plates. Initially a magnetic stirrer bar was used in attempts to agitate the reaction. However, this caused the metal plates to move along with the magnets, creating an unstable set of electrodes and breakage. Rather than using a magnetic stirrer, a long needle was placed into the reactor and nitrogen bubbled through close to the electrodes to disturb the mixture and increase mixing. If the reaction was conducted without this mixing, a large build-up of brown, in soluble material was created. With the $N_2$ mixing, the reaction occurred immediately (5 V) with a visible red/orange species forming around the electrodes. It must be noted, however, that with a high voltage solvent may begin to degrade, MeCN has a wide stability window but will begin to degrade about ca. 2.5 V. Off-line MS of the reaction showed a large proportion

of ligand precursor and the desired complex. Following evidence of complex formation, the Fe-NHC synthesis was transferred into the ECF cell. Previous work within the group has demonstrated the synthesis of a range of Fe-NHC s and their application in a C-H hydroxylation reaction. The most catalytically active Fe-NHC complexes which were easily detected via on-line MS and were also air stable (Figure 103) were chosen to be taken into flow.
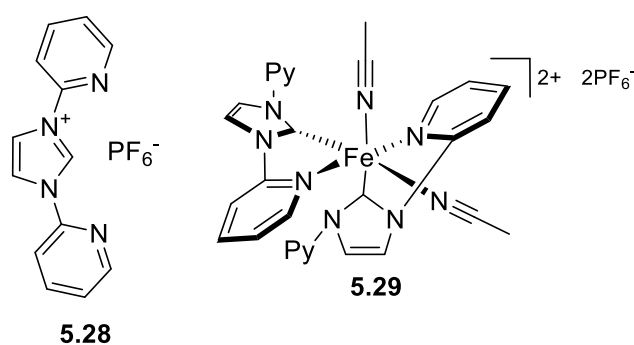


*Figure 103 The formation of complex **5.29** (from **5.28**) was chosen as it is one of the most effective complexes tested in batch and an air and moisture stable complex.*

### 5.5.3. Transition to Continuous Flow
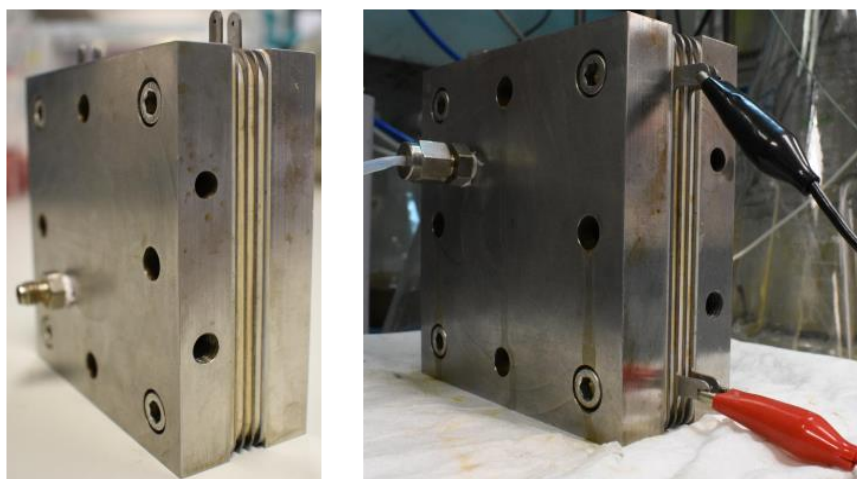


*Figure 104 Images of the electrochemical cell designed within the group to synthesise Fe-NHC complexes in continuous flow. The left image shows the assembled cell with alternating PTFE gaskets and Fe plates with no connections. The right image shows the cell connected to the flow system and anode and cathode of the power pack.*

The reaction from the adapted batch cell was moved into flow using ligand precursor **5.28** and complex **5.29** above and the flow cell developed within the group (Figure 104). Pure iron plates (99.999%) were used for both anode and cathode electrodes and the flow pathway or channels were laser cut out of PTFE spacers to allow the flow stream to pass

over the surface of each electrode. It was essential to keep the iron plates clean and rust free before continuing with any reaction. Any particulate matter on the plates or gaskets caused the flow cell to leak which either stopped flow from reaching the end of the cell or short circuited the electrodes, stopping the reaction (Figure 105).

With initial gaskets, the path width was 4 mm which caused dead volume within the reactor when channels were placed vertically.[vii]
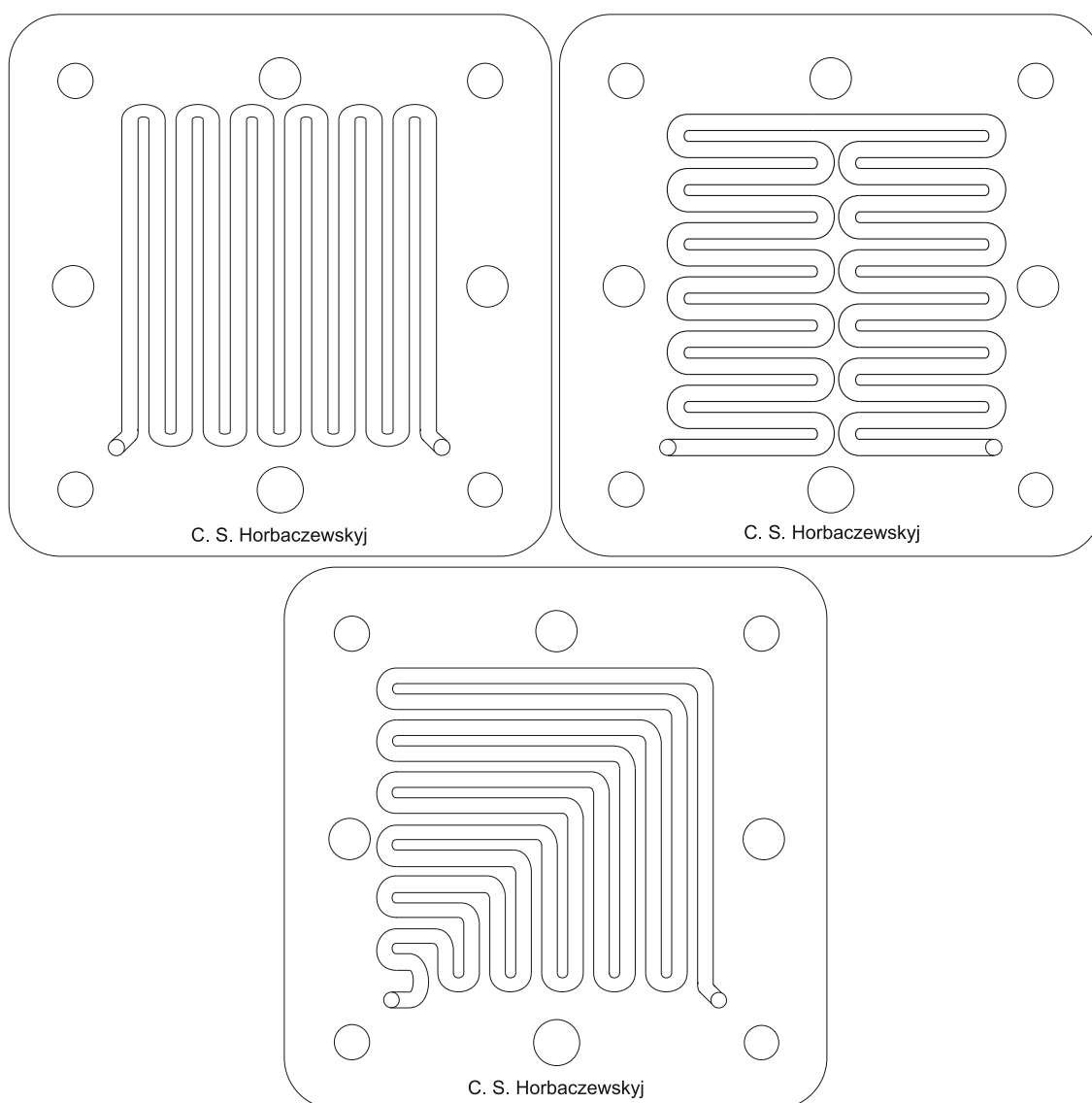


*Figure 105 The three types of PTFE gaskets designed for the ECF cell used to generate Fe-NHC complexes.*

---

[vii] Dead volume occurred due to the gap in the bottom of the flow channel. Once filled the easiest route for the flow stream was above the volume of liquid below it.

After using the cell for a prolonged time, the dead volume caused large levels of etching to occur, degrading the iron plates. To overcome this, new flow gaskets were developed that reduced the path width, increased the path length and increased the path depth. This changed the section volume from *c.a.* 2 mL to 3.9 mL allowing a higher throughput of reaction to increase the cells productivity. Several gasket versions were designed to change the extent of mixing within the cell. This was done by increasing the number of turns the reaction stream would undergo (Figure 105). It must also be noted that the reactor can collect bubbles of gas, if stood with the pathway vertical, so these must be flushed through otherwise residence time is affected. The three new gaskets were used in series to help improve the mixing of the reactor as a whole.

A continuous flow run was completed using both *mes*Py, **5.26**, and b*is*Py, **5.28**, ligand precursors at selected flow rates and voltages to see how the gaskets would perform in the reactor. The power pack was initially set to 0.1 V with a residence time of 2 hours $(0.1 \text{ mL min}^{-1})$ and allowed to reach steady state. On-line MS detected only ligand precursor in each case. The flow rate was kept constant whilst the voltage was increased to 1 V, again leaving the reactor to reach steady state. This time complex formation in both cases was detected in small quantities *via* on-line MS measurements. However, the usual peak seen for complex formation, $[\textbf{5.27}\text{-}2PF_6^-+MeO^-]^{2+}$ at 695 *m/z* and $[\textbf{5.29}\text{-}2PF_6^-+MeO^-]^{2+}$ at *m/z* 613 for both complexes respectively, were not detected. Observed peaks corresponding to ligand precursor $[\textbf{5.28}\text{-}PF_6^-]^+$ and complex $[\textbf{5.29}\text{-}2PF_6^-+MeO^-]^{2+}$, $[\textbf{5.29}\text{-}2PF_6\text{-}MeCN]^{2+}$, $[\textbf{5.29}\text{-}2PF_6\text{-}(MeCN)_2]^{2+}$ are shown in Figure 106. The major peak was $[\textbf{5.29}\text{-}2PF_6\text{-}MeCN]^{2+}$ making it the most advantageous in monitoring the system as it could potentially provide a linear relationship against concentration for calibration of the system.
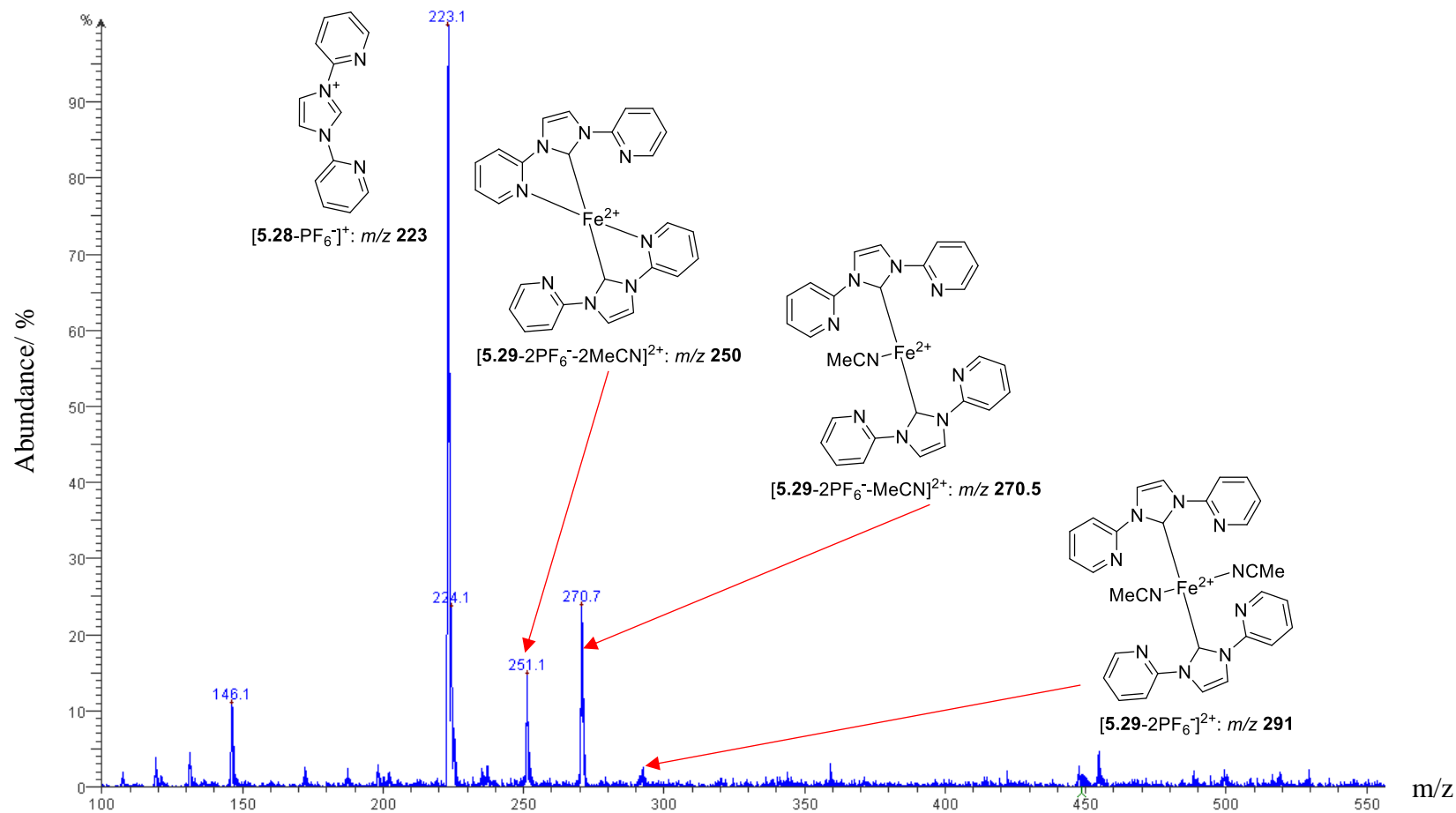
*Figure 106 Mass spectrum outlining the different species observed during a continuous flow reaction. The ligand precursor [**5.28**-PF$_6$] is seen at m/z 223 whilst the complex gives the peaks for: [**5.29**-2PF$_6$-(MeCN)$_2$]$^{2+}$ , [**5.29**-2PF$_6$-MeCN]$^{2+}$ and [**5.29**-2PF$_6$]$^{2+}$ at m/z 251, m/z 270.5 and m/z 291 respectively.*

### 5.5.4. Experimental Understanding Using Design of Experiments

To determine optimum conditions and gain an understanding of the electrochemical reaction, a DoE was performed. The DoE used two factors; residence time and applied voltage. The set boundary conditions for each are summarised in Table 37 with concentration being kept constant at 5 mM.

*Table 37 Table of the boundary conditions used during the DoE for the formation of Fe-NHCs. High voltages were used to gain an understanding of the reaction system if components begin to break down.*

| Boundary | Flow Rate (mL min$^{-1}$) | Voltage (V) |
|----------|---------------------------|-------------|
| **Lower** | 0.15 | 1.50 |
| **Middle** | 1.58 | 4.25 |
| **Upper** | 3.00 | 7.00 |

The first set of conditions 3 mL min$^{-1}$ and 1.5 V ligand precursor dominated the spectra with very small amounts of complex formation. At 7 V with a slower flow rate of 0.15 mL min$^{-1}$, detected ligand precursor level dropped drastically along with a small decrease in complex mass intensity (Figure 107). A peak, *m/z* 187, corresponding to degradation of the ligand precursor *via* loss of the pyridyl substituent increased in intensity. At the slower flow rate (0.15 mL min$^{-1}$) and middle level of voltage (4.25 V) a much higher amount of complex was detected *via* MS, with lower levels of ligand precursor and only small amounts of ligand degradation. Part way through the DoE, 7 V and 0.15 mL min$^{-1}$, the reactor became blocked with a dark brown residue, thought to be insoluble Fe(OH)$_3$ and other products of Fe and NHC degradation. Although this cut the DoE short it still gave enough detail to estimate areas to operate. From this DoE the best conditions were found to be 3 mL min$^{-1}$ using 7 V, though determination cannot be made from an incomplete DoE. A high reaction voltage along with a high flow rate, consequently low residence time, the reaction medium does not spend enough time in the ECF reactor to break down too much, although some precursor breakdown is observed through the peak at *m/z* 187 (Figure 108). The peak at *m/z* 187 was only observed in the MS when using the ECF reactor at high voltages, further verification will need to be completed with a greater amount of species and analysed using NMR spectroscopy.
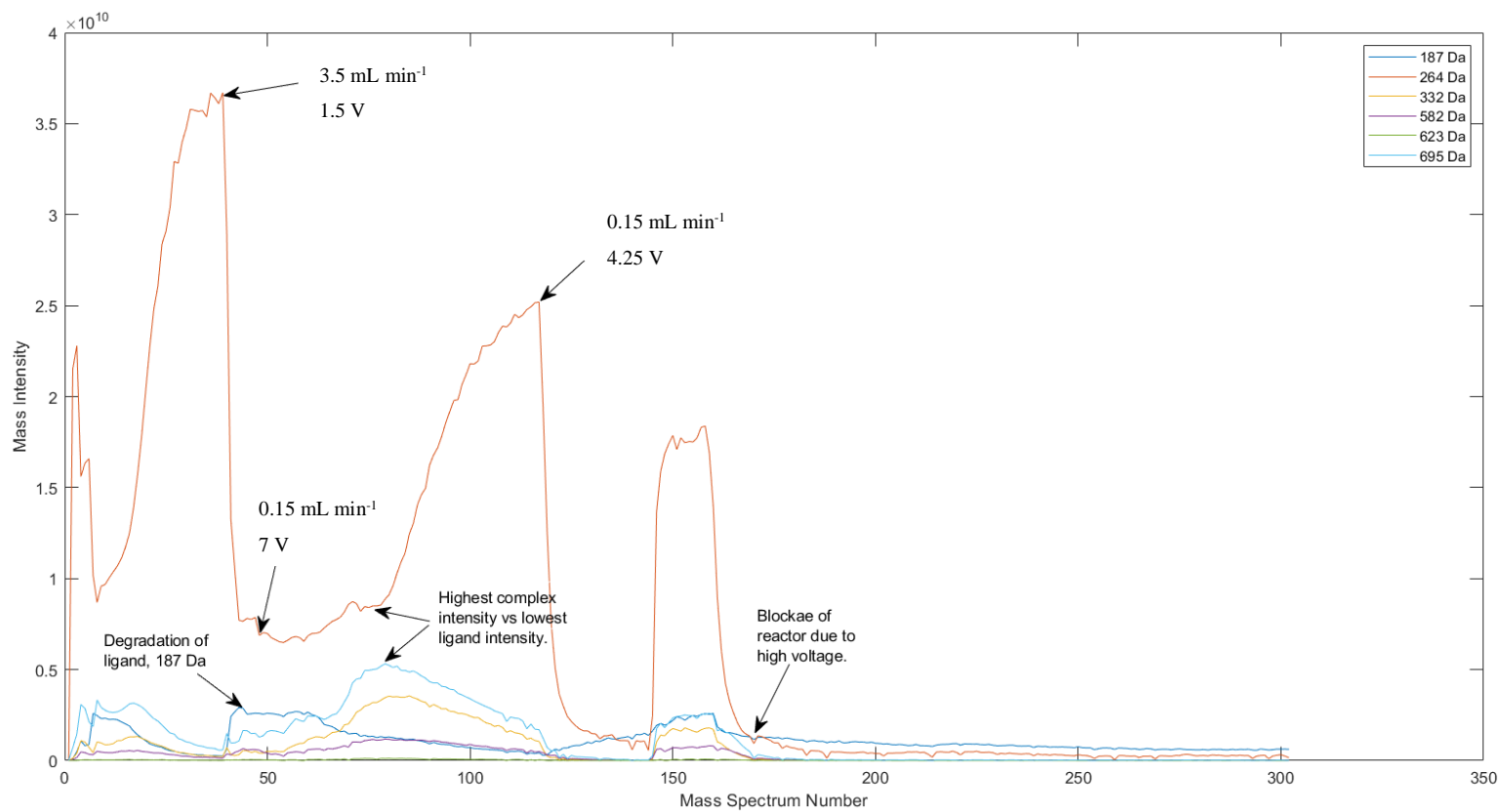
*Figure 107 Initial DoE performed on the MesPy precursor to generate Fe-NHC complexes in continuous flow. At high voltages precursor degradation occurred. The best results were seen at 3 mL min$^{-1}$ and 7 V.*
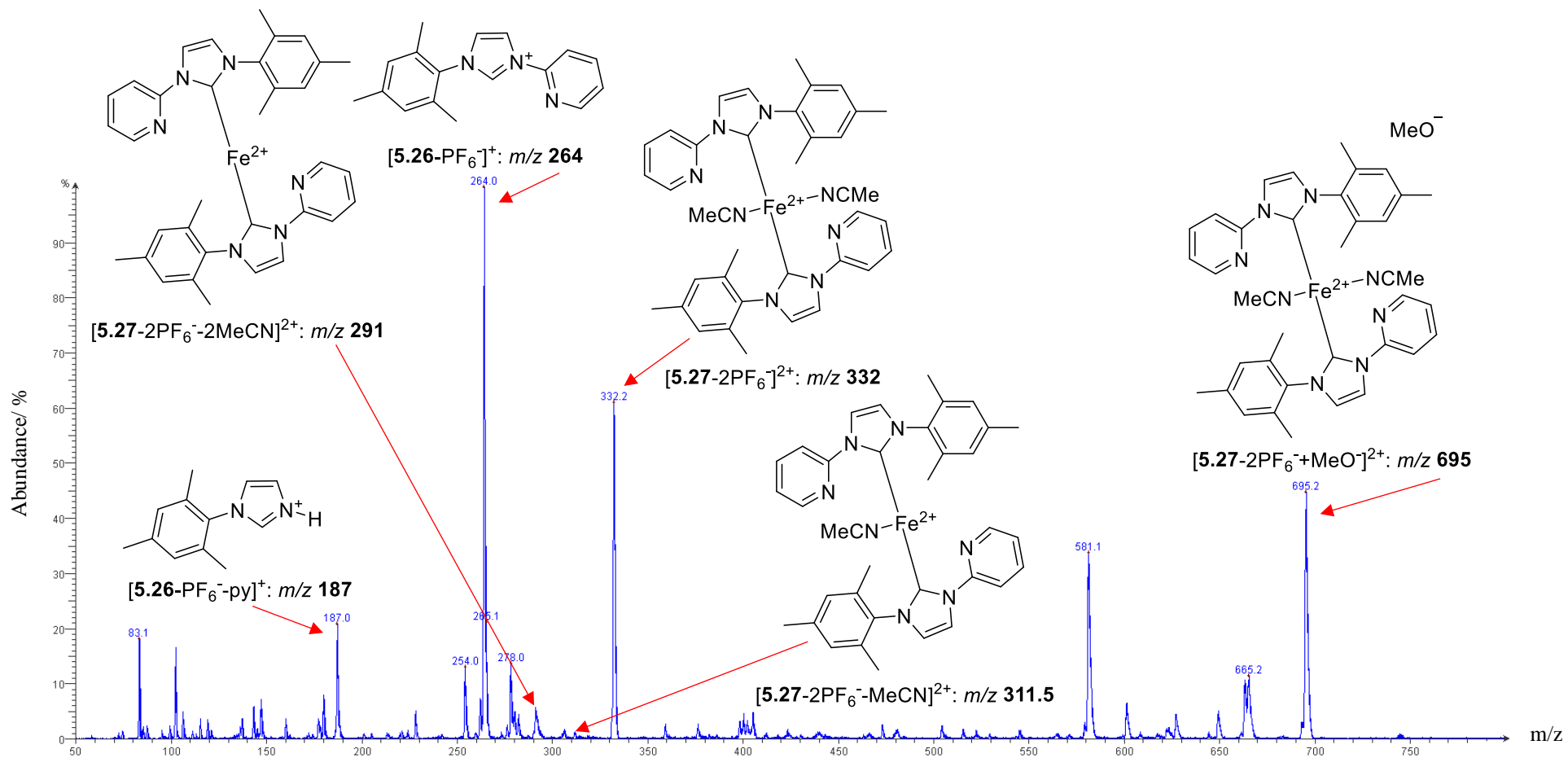
*Figure 108 Mass spectrum showing the precursor and complex synthesised in continuous flow using the ECF reactor.*

Use of the ECF cell at high voltage is not ideal as it causes over potential as well as causing reagent and solvent degradation to occur. The next DoE reduced the total size of experimental area to reduce the chances of blockage and breakdown of the precursor. Again, the voltage and residence time were changed accordingly whilst keeping the number of plates and spacers the same. Voltage levels spanned $0.05 - 0.15$ V with residence times from $10 - 60$ minutes ($0.2 - 1.2$ mL min$^{-1}$). The low voltages were insufficient at providing enough potential energy for complex formation at all levels of residence time, with only ligand precursor being seen by on-line MS and off-line HPLC-MS. The voltage levels were increased to span $0.1 - 0.25$ V with the same residence times of $10 - 60$ minutes. As well as this a subsequent expanded DoE was carried out with voltages from $0.1 - 0.5$ V. However, the used voltages were still too low to induce complex synthesis, far below the reduction potential of the ligand precursor which is *ca.* -2.32 V (versus Fc/Fc$^+$).[128] A voltage level of 2.5 V should be sufficient to begin efficient complex formation.

### 5.5.5. Calibration of the Mass Spectrometer

Once the reactor and gaskets were fully established the MS was calibrated for the ligand precursor and desired complex for quantitation of future reactions for the formation of complex [Fe(*bis*Py)$_2$(MeCN)$_2$]$^{2+}$ (**5.29**) (Figure 103). Samples of each concentration were taken three time, so an average result could be taken and so the reproducibility in the MS could be observed. Each calibration plot is shown in Figure 116 and Figure 117. Samples were manually injected into the MS and ionised using ESI+. The ligand precursor calibration plot shows an initial fit to the linearity of the MS, but at higher concentrations it deviates away from linearity. ESI generally requires samples which are ionisable in solution where acid-base charge transfer processes can occur. With regards to NHC ligand precursors and metal-NHC complexes, each molecule exhibits a charge, either from imidazolium salt or due to the cationic metal species, making ESI ideal. However, this ionisation mode relies on the concentration of the analyte in the reaction medium; the lower the concentration the more efficient ionisation will occur. The calibration concentration varied between 1 - 20 mM for both the ligand precursor and the complex. Loss of linearity occurs after 5 mM meaning any concentrations exceeding this would be out of the linear dynamic range of the MS. A sample containing both the ligand precursor and complex, at 5 mM, was injected into the MS as this mimicked the reaction media closest. The output graph shows a large level of ligand precursor and a very low

level of the desired complex. This could be due to two reasons; either suppression of the complex by the ligand precursor or complex fragmentation *i.e.* ligand coming off the complex in the MS.

Because of this, samples would be taken and analysed using off-line LCMS to look at the amount of complex formed compared to the ligand precursor as well as performing on-line MS. Suppression in the MS is not ideal when trying to observe and develop on-line MS as a quantitative technique. As a comparison in-line UV-Vis was adopted and added to the reaction stream just after the reactor.

### 5.5.6. UV-Vis Measurements of Fe-NHC Complexes

In-line UV-Vis was used to detect the ligand precursor remaining in the reactor after generation of the complex. Each species in the reaction contains chromophores making them ideal to be monitored *via* this technique. The Fe-NHC complex synthesised and monitored was $[Fe(bisPy)_2(MeCN)_2]^{2+}$ which can be observed as a red/orange solution when dissolved in MeCN. The ligand precursor gives a characteristic peak at ~295 nm (Figure 109) whilst the complex gives a characteristic peak at ~405 nm, both monitored by off-line UV-Vis (Figure 110).

When moving to flow and the in-line spectrometer, both the ligand precursor and pure complex were manually injected through to monitor the peaks formed of species during the reaction and to determine desired reaction concentration. The in-line spectrometer and its control user interface have been developed within the group to run using LabView with a plugin to connect through MATLAB. This way it can be triggered to begin sampling when the sample switching device triggers. It was necessary to run the spectrometer in continuous acquisition mode, measuring the absorbance of material passed through. The in-line spectrometer was much more prone to the detector being overloaded by the ligand precursor at lower concentrations. Samples varying from 0.1 to 5 mM (the area best for the MS) tended to overwhelm the UV-Vis detector over a wide range of the chosen concentrations particularly at above 1 mM. The spectrometer itself was also prone to base line change from varying conditions within the laboratory (this was combated by completing references at the beginning of every use) but had a direct effect on the total sample concentration detected. The most reliable maximum concentration which could be used was 0.47 mM before absorbance from the SM caused a large over saturation of the detector and proceeding down to a limit of 0.28 mM. This

made it difficult to monitor concentration during the reaction, particularly of the formation of the desired Fe-NHC complex.
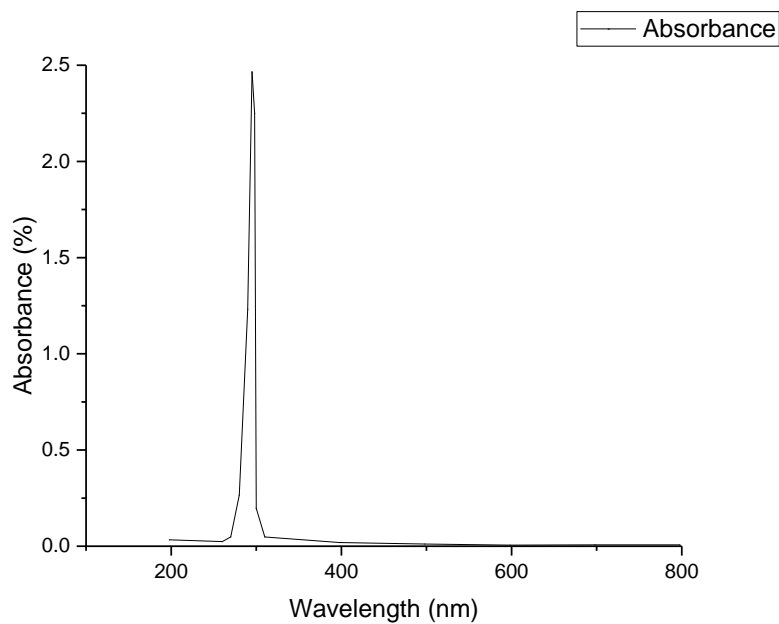


*Figure 109 Recreated off-line UV-Vis spectrum (from receipt printout) of ligand precursor, m/z 223 at 1 mM.*
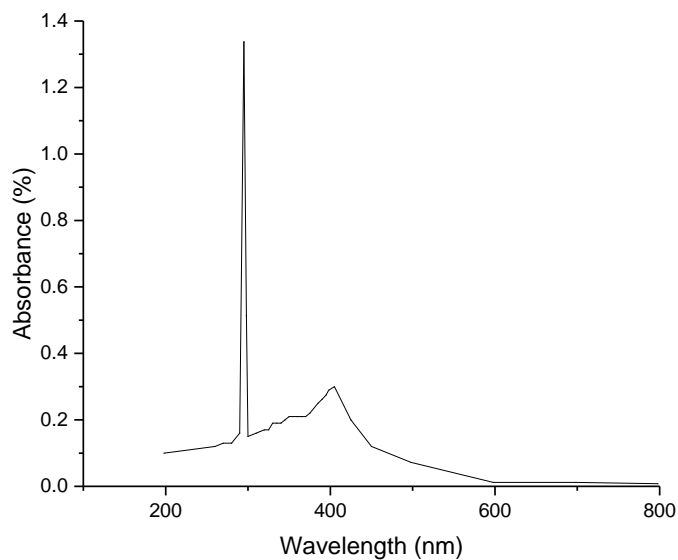


*Figure 110 Recreated off-line UV-Vis spectrum (from receipt printout) of complex, m/z 664 at 0.5 mM.*

### 5.5.7. **Monitoring Fe-NHC Complexes** *via* **MS and UV-Vis**

After integration of the in-line UV-Vis into the system the ligand precursor, 1,3-dipyridylimidazolium hexafluorophosphate, was passed through the electrochemical flow reactor 1 mL min$^{-1}$ with an applied potential of 2.5 V and concentration of 0.25 mM to monitor the formation of Fe-NHC complex using both UV-Vis and MS. Concentration estimation can be calculated from the shoulder of the UV-Vis peak when the precursor is giving a maximum absorption *i.e.* any shoulder movement suggests a change in concentration and therefore conversion of the ligand precursor to the complex. This can then be correlated to the MS. Using a solution of concentration, 0.25 mM, a peak was detected, using UV-Vis, at ~285 nm, for the ligand precursor. When increasing the concentration to that which was required for the MS (5 mM), UV-Vis absorption sharply increased with the peaks shoulder residing at 400 nm, so the overall peak spanned a large majority of the UV wavelength (from 230 nm to 400 nm). The desired complex gave a peak at ~410 nm similar to the off-line UV-Vis measurement (Figure 111) although at the low concentration needed for complex absorption to occur, the ligand precursor peak saturated the detector from 225 nm to 300 nm. MeCN solvent contributes an absorption at ~195 nm just below the lowest wavelength of the detector. However, as it is a bulk constituent to the reaction, some overlap does occur past 200 nm but this does not interfere with the ligand precursor absorption.
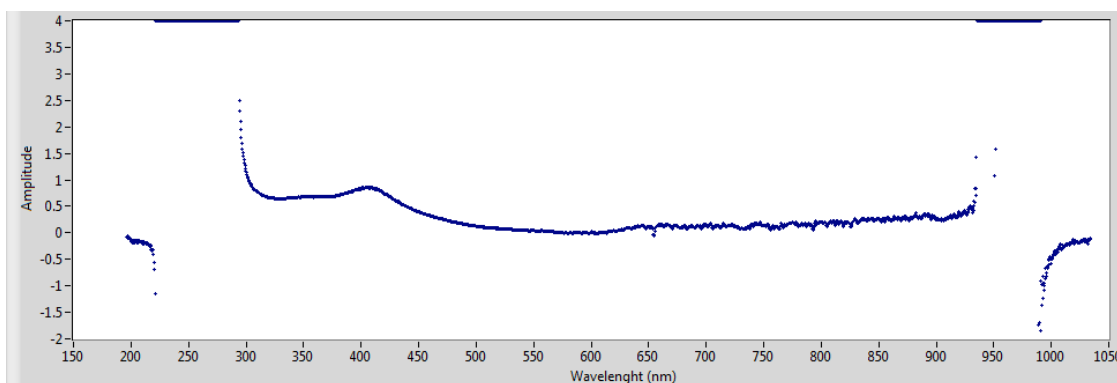


*Figure 111 On-line UV-*Vis *measurement for pure complex, [Fe(bisPy)$_2$(MeCN)$_2$]$^{2+}$, where a ligand precursor absorption is seen at ~285 nm and complex at ~410 nm.*

Rectifying this problem led to attempts to split the reactor flow stream in two so that one stream could be diluted for the UV-Vis and the second could be added directly to the sample-loop for MS detection (Figure 112). However, splitting the flow stream from this

continuous flow reactor proved problematic as the it was usually diverted to sample collection, the path of least resistance.
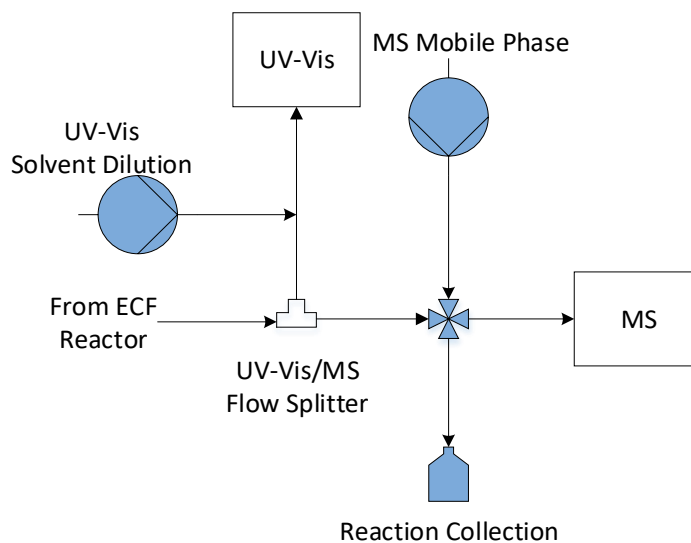


*Figure 112 Splitting of the reaction pathway from the ECF reactor to the UV-Vis where this stream can be diluted leaving the MS stream the correct concentration for detection.*

The pathway to reactor collection only passes through PTFE tubing and the sample switching device whereas the alternate route connects with a dilution stream and passes through the UV-Vis detector which comprises a 5 mm path length cell which increases the pressure drop of this pathway dramatically. It was decided that monitoring the complex could be done by measuring the shoulder of the peak in the on-line UV-Vis, but again loses its quantitative ability similarly to the MS.

## 5.6. Application of Fe-NHC Complexes to a C-H Activation Reaction

Fe-NHC complexes have been generated in flow using an electrochemical reactor and process analytical techniques, MS and UV-Vis, to monitor their formation. Previous work within the group by Chapman has examined several Fe-NHC complexes in a catalytic C-H hydroxylation reaction (Scheme 39).
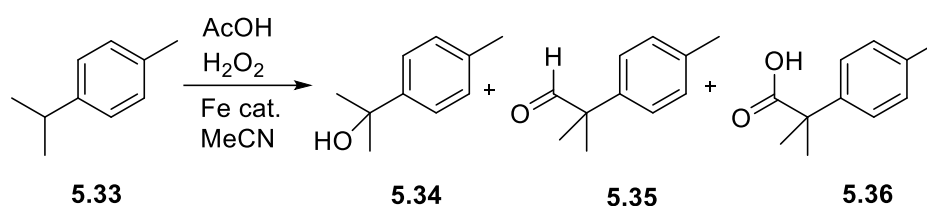
*Scheme 39 C-H Hydroxylation reaction of p-cymene to the desired alcohol on the tertiary carbon. Side products form from over oxidation on the same position to produce aldehyde and carboxylic acid.*

Electrochemical generation of Fe-NHCs in flow was connected to the hydroxylation reaction, also in flow. Fe-NHC generation would be carried out until sufficient volume of the complex had been collected, then this solution would be pumped into the catalytic reaction. The first stage would still include detection *via* on-line MS and in-line UV-Vis spectroscopy whilst the second stage would include on-line separation and detection using HPLC-UV. Reagents (**5.33**) and products (**5.34**, **5.35** and **5.36**) of the hydroxylation reaction were mixed together and manually sampled *via* HPLC in order for a method to be developed.

The ligand precursor solution was pumped through the ECF reactor at 1 mL min$^{-1}$, 2.5 V and 5 mM. The starting precursor was initially a pale-yellow colour changing to an orange colour after exiting the reactor and finally changed to a dark brown suggesting formation of other species *e.g.* oxidation of the NHC or iron salts. UV-Vis spectroscopy gave a peak at ~285 nm evidence of the ligand precursor and a very small broad peak at ~400 nm suggesting evidence of complex formation (Figure 113). The complex feed was allowed to build up before being mixed with *p*-cymene and AcOH and $H_2O_2$ all in MeCN each at 0.1 mL min$^{-1}$ (residence time of 10 mins) set to 50 °C and analysed using on-line HPLC. HPLC chromatograms show peaks at 2.57 minutes for the internal standard and 1.66 minutes for the starting material, **5.33**. Given this, some development still needs to be carried out to join the two reactions together and more reactions need to be carried out with respect to the catalyst to determine what volume is generated electrochemically and how much is needed for the hydroxylation reaction.
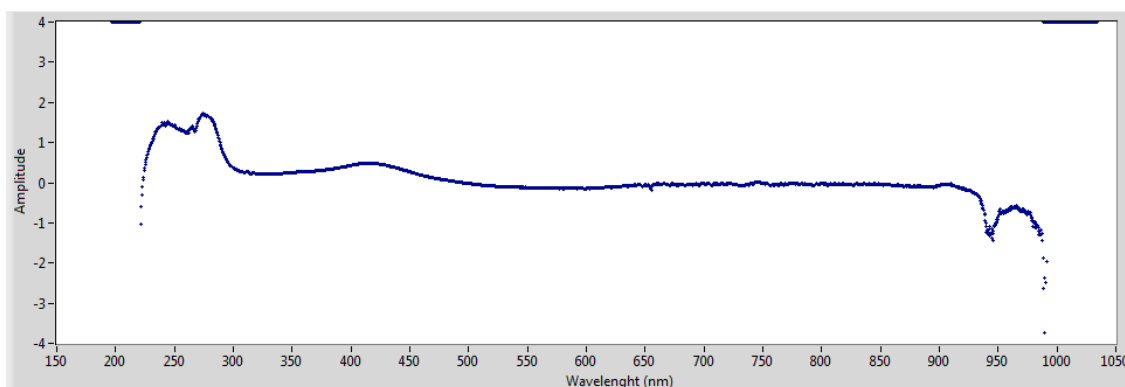
*Figure 113 On-line UV-*Vis *taken during reaction at 5 mM producing a pale-yellow solution. Peaks can be observed at ~285 nm and ~400 nm for ligand precursor and complex respectively.*

## 5.7. Summary

Several imidazolium salts have been used to generate Fe-NHC complexes in continuous flow using an electrochemical reactor, developed within the group. Several complexes were synthesised initially with each one detected using on-line mass spectrometry. The most catalytically active complexes synthesised using the ECF reactor were; [Fe(*mes*Py)$_2$(MeCN)$_2$]$^{2+}$ and [Fe(*bis*Py)$_2$(MeCN)$_2$]$^{2+}$ with MS spectra showing ligand precursor and complex in different forms. However, the MS may sometimes provide details on the formation of the complex but often succumbs to suppression effects when in ESI+ mode. The ligand precursor ionises sufficiently at a range of concentrations, but when the complex is in the same solution as the ligand precursor only ionises either when the ligand precursor level has depleted or when there is sufficient complex formed. This makes quantitation of the complexes very difficult showing that, in this case, use of the MS in ESI+ mode is not adequate. Another technique which may be used to help with quantitation is inline UV-Vis. However, similarly to the MS, this method of analysis succumbs to problems with concentration of the solution. At low concentrations a peak is observed at ~295 nm for the ligand precursor and a peak at ~405 nm can be seen for the complex. Upon increasing the concentration to one which is more sufficient for the MS, the UV-Vis detector is saturated. These issues are not entirely problematic for complex detection and identity determination as species can still be detected but not accurately quantified. Adaption of the system may need to be developed to endure the correct concentration for both MS and UV-Vis to ensure accurate and reliable analysis. There are two sets of conditions which can be used to generate Fe-NHC complexes in

electrochemical flow; 2.5 V with a 15 minute residence time or 7 V with a 4 minute residence time. If using 7 V, more ligand precursor fragmentation will be observed, but in comparison to the first set, a shorter residence time. A more in-depth DoE will be performed using other variables such as reagent concentration and number of electrodes to provide more information on the reaction and also to provide robust reaction operation conditions.

Linking together the flow reactors for both the catalyst generation and a catalytic hydroxylation reaction was attempted. Use of the catalysts in flow may improve the selectivity of the reaction with greater process control and separation of process parameters. HPLC data did not show any conversion of SM, **5.33**, to desired product, **5.34**. This seems likely due to an insufficient levels of active catalyst formed during the reaction even though conditions used have been shown to generate the desired species. This reaction needs to be looked at in much more detail to understand the process in continuous flow with use of a DoE. Subsequently the reaction can be carried out using a self-optimisation algorithm for optimum verification.

### 5.8. Experimental

### 5.8.1. General Reaction Setup

The general setup for flow work performed is described in section 2.7.1 where details on pumps, reactors, analysis and programming can be found. Specific details for each reactions setup are provided in this section for electrochemical batch runs, flow. Other details outlining reagents and solvents with purities and grades can be also found in this section along with details of how analyses of samples were performed.

### 5.8.2. Electrochemical Batch Reactions

The electrochemical cell consists of a power pack (BST PSD30/3B) with adjustable voltage and current settings and built in displays. Iron electrodes (approximately; LxWxD, 40 mm x 10 mm x 1 mm) were prepared and a small hole drilled into the top where wire was placed and pushed through a rubber Suba-Seal to allow contact with the clips from the power pack. The electrodes were spaced with a small section of PTFE and wrapped in parafilm, keeping them ridged, with an approximate gap of 1-2 mm. An external multimeter (TENMA 72-7745, set as an ammeter) was used to accurately monitor the current as the power pack display is limited to 0.01 A. A schematic of the reactor is shown in Figure 114.
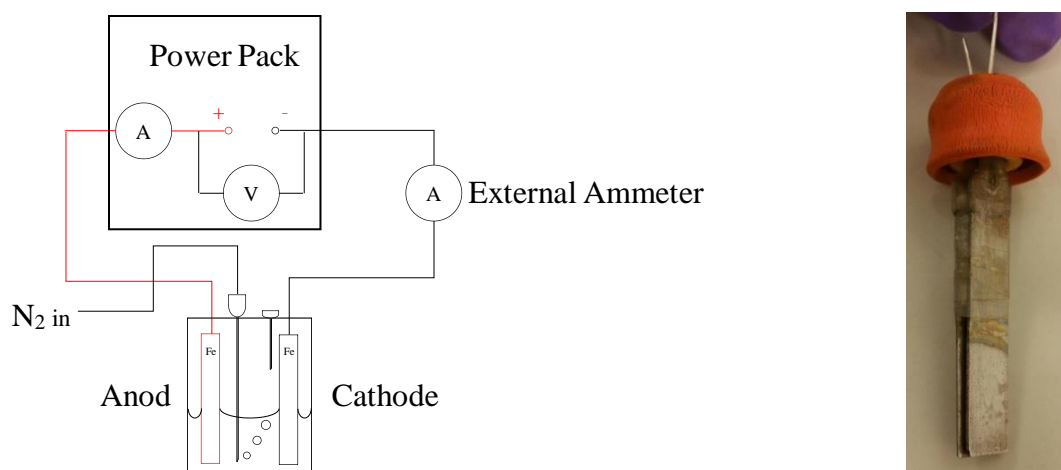


*Figure 114 Left; A schematic of the electrochemical batch cell used in the synthesis of iron NHCs, right; a picture of the electrodes used for the reactor.*

A solution of mesityl pyridyl imidazolium hexafluoro phosphate, [*mes*Py]PF$_6$ (20.0 mg, 48.9 µmol) was dissolved in anhydrous MeCN (20 mL) to give a solution with a concentration of 2.44 mM in an inert atmosphere (final reaction twice the mass). The

flask (a three-necked round bottomed flask) also contained two iron electrodes placed together, supported by a PTFE spacer, as close as possible in the central neck. The left neck contained a purge needle dispelling nitrogen to degas and stir the solution and the left neck a gas exit needle. The power pack voltage and current were set and applied to the solution whilst nitrogen was bubbled through to keep the solution agitated and to keep the vessel flushed. The initial pale yellow-orange solution turned brown after the reaction had started. After the allotted time, the solution was reduced *in vacuo* before the remaining solvent was left to slowly evaporate and allowed to crystallise. A summary of experiments performed can be seen in Table 38. MS: $m/z$ 264 [**5.26**-PF$_6$]$^+$, $m/z$ 695 [**5.27-**2PF$_6$+MeO$^-$]$^+$.

| Reaction | Mass/ mg | Volume/ mL | Voltage/ V | Current/ mA | Distance/ mm | Time/ h |
|---|---|---|---|---|---|---|
| **1** | 21 | 20 | 21.4 | 0.01 | 15 | 4 |
| **2** | 21 | 20 | 21 | 0.01 | 15 | 3 |
| **3** | 21 | 20 | 10 | 0.03 | 2 | 2 |
| **4** | 40.8 | 20 | 5.13 | 0.02 | 1 | 2 |

*Table 38 A summary of the reaction conditions used for the electrochemical synthesis of [Fe(mesPy)$_2$(MeCN)$_2$]2PF$_6$*

### 5.8.3. Electrochemical Flow Reactions
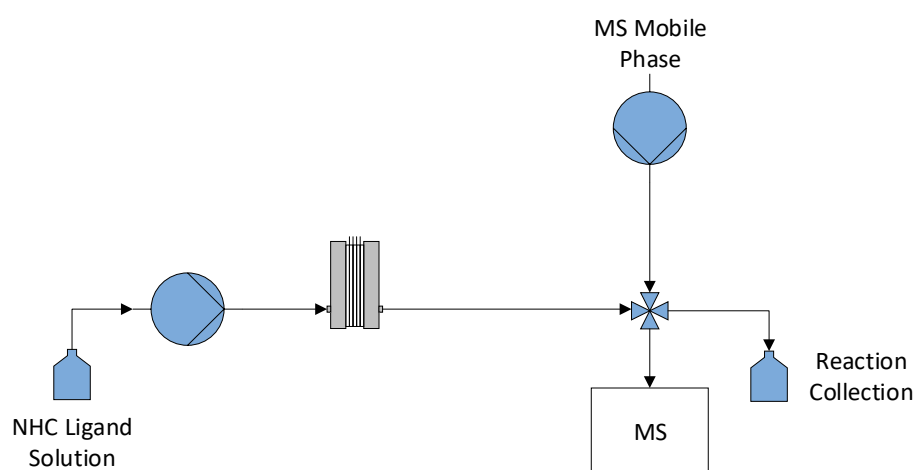


*Figure 115 Schematic diagram of the electrochemical flow reaction to generate Fe-NHC complexes monitored by on-line MS and in-line UV-Vis.*

The ECF reactor was thoroughly cleaned ensuring that all rust and particulate debris was removed. This was done by washing all electrodes and PTFE spacers in dilute HCl (0.1 M) by submerging in an acid bath before being rinsed with acetone and air dried before being assembled; 2x outer casing, 2x casing PTFE spacer, 4x iron electrode and 3x PTFE flow spacer. The ECF cell was connected to pump and sample actuator using 1/16" PTFE tubing and flushed with dry, degassed MeCN. Each NHC ligand precursor, (Table 36, 0.2 g, 0.5 mmol), was weighed and dissolved in MeCN (50 mL) to create a 10 mM concentration solution, this was then primed through a HPLC pump. The battery pack was connected to the ECF reactor so that each electrode alternated in terms of anode and cathode and the required voltage set. The ligand precursor was then pumped at the set flow rate before being analysed using on-line MS.

Mass spectra: Compound numbers given in bold. (**5.26**) $m/z$ 264, $m/z$ 695 (required; $m/z$ 264, $m/z$ 695, found; $m/z$ 264, $m/z$ 695); (**5.28**) $m/z$ 223, $m/z$ 613 (required; $m/z$ 223, $m/z$ 613, found; $m/z$ 223, $m/z$ 613); (**5.30**) $m/z$ 305 (required; $m/z$ 305, $m/z$ 779, found; $m/z$ 305); (**5.31**) $m/z$ 390, $m/z$ 947 (required; $m/z$ 390, $m/z$ 947, found; $m/z$ 390); (**5.32**) none.

### 5.8.4. Mass Spectrometry Method

Mass spectra were gathered in ESI+ mode from $m/z$ 50 to 800 using a scan time of 938 ms and an acquisition time of 60 seconds. Other settings of the MS are summarized below. Mobile phase consisted of 100% MeCN pumped at 0.05 mL min$^{-1+.}$

Polarity: Positive

Capillary Temperature: 100 °C

Capillary Voltage: 180 V

Source Voltage Span: 0 V

Source Voltage offset: 20 V

Source Gas Temperature: 200 °C

ESI Voltage: 3500 V

### 5.8.5. Calibration Using Mass Spectrometry

The NHC ligand precursor, 1,3-dipyridylimidazolium chloride, (**5.28**, 0.205 g, 0.5 mmol), was weighed out and dissolved in MeCN (25 mL) so make a 20 mM stock solution. Aliquots of this solution were taken and diluted to make standards for calibration of the MS instrument; 10 mM, 5 mM, 2.5 mM and 1 mM. The same procedure for complex [Fe(*bis*Py)$_2$(MeCN)$_2$]2PF$_6$ (0.205 g, 0.5 mM) in MeCN (25 mL) to create a

20 mM solution. Aliquots of this standard were taken and diluted to make 10 mM, 5 mM, 2.5 mM and 1 mM. Each solution in turn was manually injected through the sample loop and then analysed using MS to form a calibration curve (Figure 116 and Figure 117). Table 39 provides a summary of the concentrations required and the volumes for each sample needed for calibration of each species *via* MS.

Mass spectra: *m/z* 223, *m/z* 613 (required; *m/z* 223, *m/z* 613, found; *m/z* 223, *m/z* 613).

*Table 39 Volumes and concentrations required for the calibration of 1,3-dipyridylimidazolium chloride and [Fe(bisPy)$_2$(MeCN)$_2$]2PF$_6$.*

| Nº | Concentration / mM | Stock needed/ mL | Final Vol/ mL | Nº | Concentration/ mM | Stock needed/ mL | Final Vol/ mL |
|---|---|---|---|---|---|---|---|
| 1 | 20 | 25 | 25 | 1 | 20 | 25 | 25 |
| 2 | 10 | 5 | 10 | 2 | 1 | 5 | 2 |
| 3 | 5 | 2.5 | 10 | 3 | 0.5 | 2.5 | 2 |
| 4 | 2.5 | 1.25 | 10 | 4 | 0.25 | 1.25 | 2 |
| 5 | 1 | 0.5 | 10 | 5 | 0.1 | 0.5 | 2 |
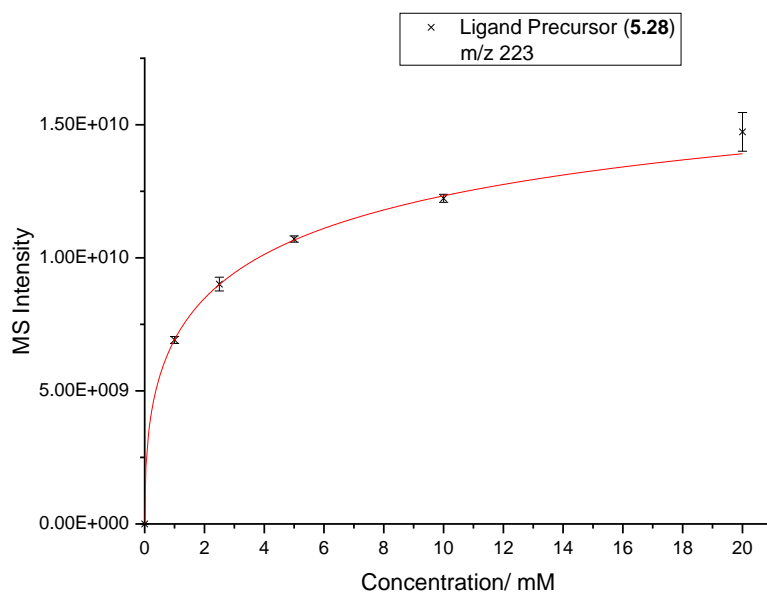


*Figure 116 Calibration curve of ligand precursor, 1,3-dipyridylimidazolium chloride, analysed using mass spectrometry.*
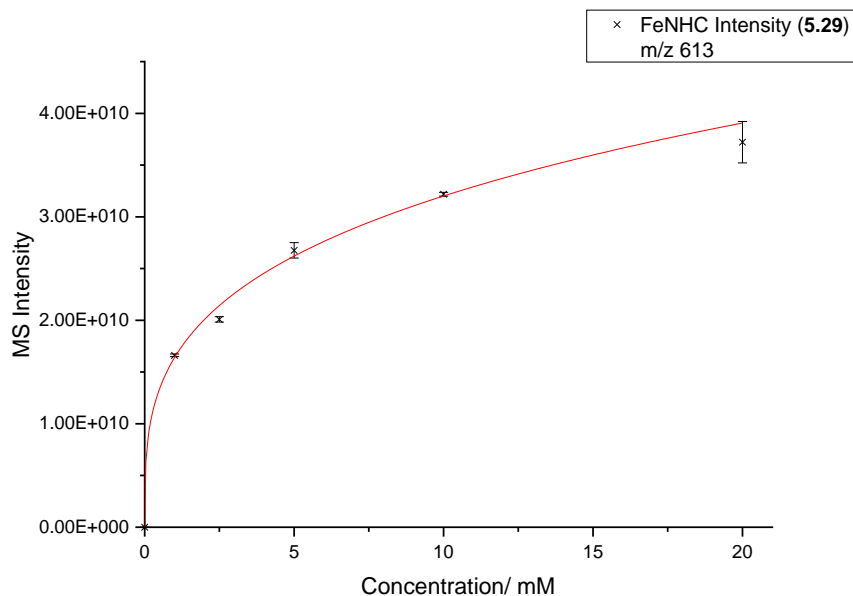
*Figure 117 Calibration curve of Fe-NHC complex, [Fe(bisPy)$_2$(MeCN)$_2$]2PF$_6$, analysed using mass spectrometry.*

### 5.8.6. Calibration Using UV-Vis

The same stock solution as prepared in section **5.8.5** was used to calibrate the UV-Vis. Aliquots of this solution were taken and used to prepare calibrant solutions of 0.25 mM, 0.5 mM, 1 mM, 2.5 mM, 5 mM before being injected through the sample actuator and into the in-line UV-Vis spectrometer. These calibrant solutions were also used for off-line calibration of both ligand precursor and Fe-NHC complex (Figure 118 and Figure 119).

UV-Vis: $\lambda_{max}$ = 295 nm.
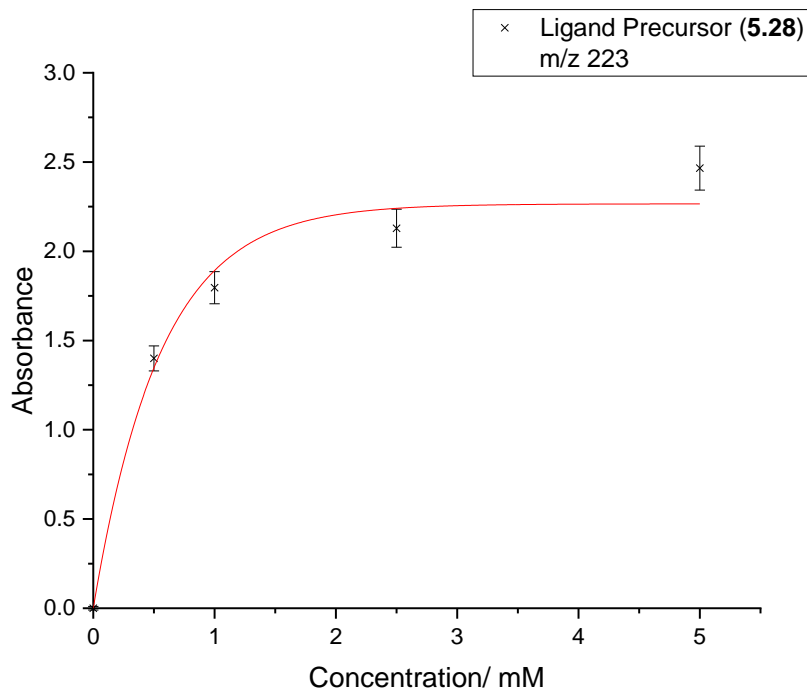
*Figure 118 Calibration curve of ligand precursor (**5.28**), 1,3-dipyridylimidazolium chloride, analysed using off-line UV-*Vis.
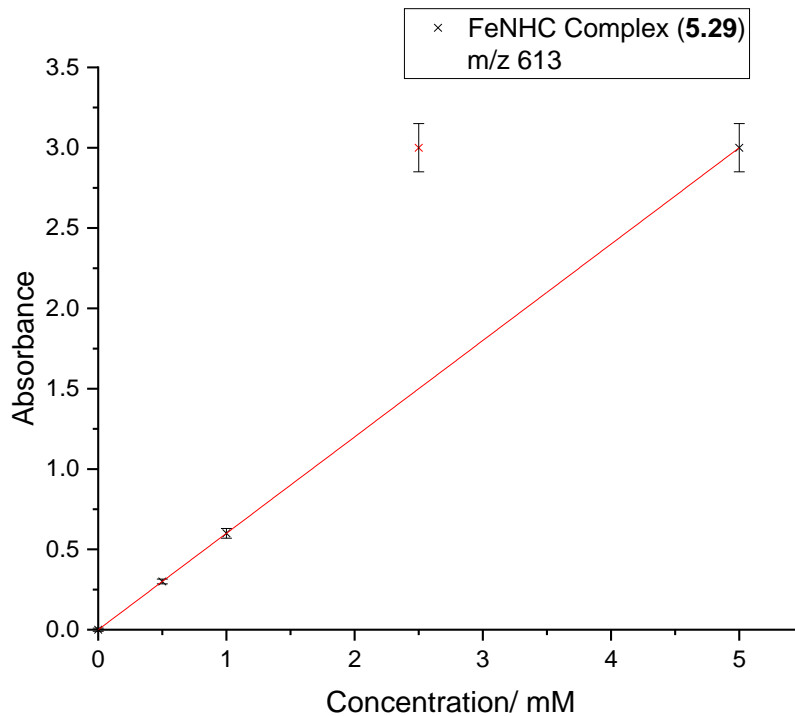


*Figure 119 Calibration of FeNHC complex (**5.29**) and analysed using off-line UV-Vis Spectroscopy.*

### 5.8.7. Continuous Flow DoE
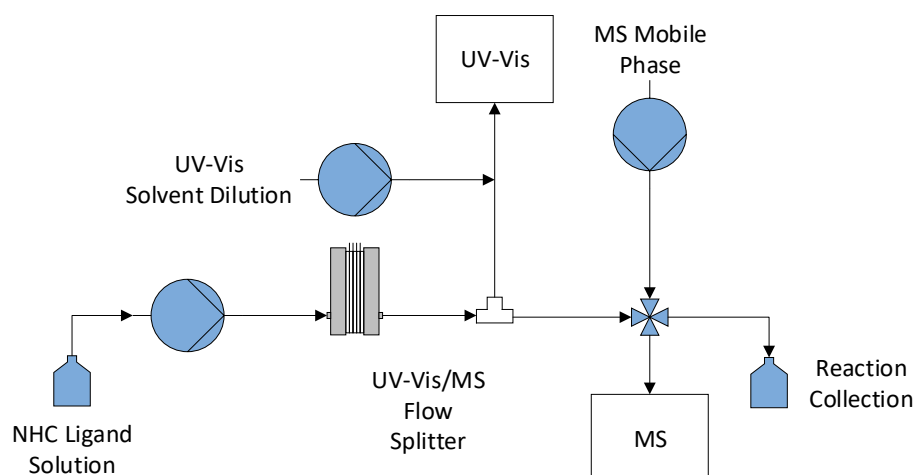


*Figure 120 Schematic diagram of the electrochemical flow reaction to generate Fe-NHC complexes monitored by on-line MS and in-line UV-*Vis.

NHC ligand precursor, (0.31 g, 0.76 mmol), was dissolved in dry MeCN (150 mL), which was then degassed for 30 minutes, to create a solution of concentration 5 mM which was then primed through a HPLC pump. The ECF cell was fully cleaned and flushed with dry, degassed MeCN. The DoE experimental list is summarised in Table 40.

Mass spectra: 187.0 *m/z*, 228.0 *m/z*, 264.0 *m/z*, 332.2 *m/z*, 581.1 *m/z*, 665.1 *m/z*, 695.2 *m/z* (required; 264.0 *m/z*, 332.2 *m/z*, 695.2 *m/z*, found; 264.0 *m/z*, 332.2 *m/z*, 695.2 *m/z*).

*Table 40 DoE conditions used to optimise the generation of [Fe(mesPy)$_2$(MeCN)$_2$]2PF$_6$ using an electrochemical flow cell.*

| Entry | Voltage/ V | Residence Time/ min |
|-------|-----------|---------------------|
| **1a** | 0.25 | 5 |
| **1b** | 0.25 | 5 |
| **2** | 10 | 60 |
| **3** | 5.125 | 60 |
| **4** | 5.125 | 32.5 |
| **5** | 5.125 | 5 |
| **6** | 10 | 5 |
| **7** | 5.125 | 32.5 |
| **8** | 5.125 | 32.5 |

| 9 | 0.25 | 32.5 |
|---|---|---|
| **10** | 10 | 32.5 |
| **11** | 0.25 | 60 |

### 5.8.8. Electrochemical Flow Synthesis of [Fe(*bisP*y)$_2$(MeCN)$_2$]2PF$_6$

The ECF reactor (Figure 104 & Figure 115) was thoroughly cleaned ensuring that all rust and particulate debris was removed. This was done by washing all electrodes and PTFE spacers in dilute HCl, rinsed with acetone and air dried before being assembled; 2x outer casing, 2x casing PTFE spacer, 4x iron electrode and 3x PTFE flow spacer. The ECF cell was connected to pump and sample actuator using 1/16" PTFE tubing and flushed with dry, degassed MeCN. The NHC ligand precursor, (0.205 g, 0.5 mmol), was weighed and dissolved in MeCN (50 mL) to create a 5 mM concentration solution, this was then primed through a HPLC pump. The battery pack was connected to the ECF reactor so that each electrode alternated in terms of anode and cathode and the required voltage set. The ligand precursor was then pumped at the set flow rate before being analysed using on-line MS.

Mass spectra: *m/z* 223, *m/z* 613 (required; *m/z* 223, *m/z* 613, found; *m/z* 223, *m/z* 613)

### 5.8.9. Telescoped Catalyst Generation and C-H Hydroxylation
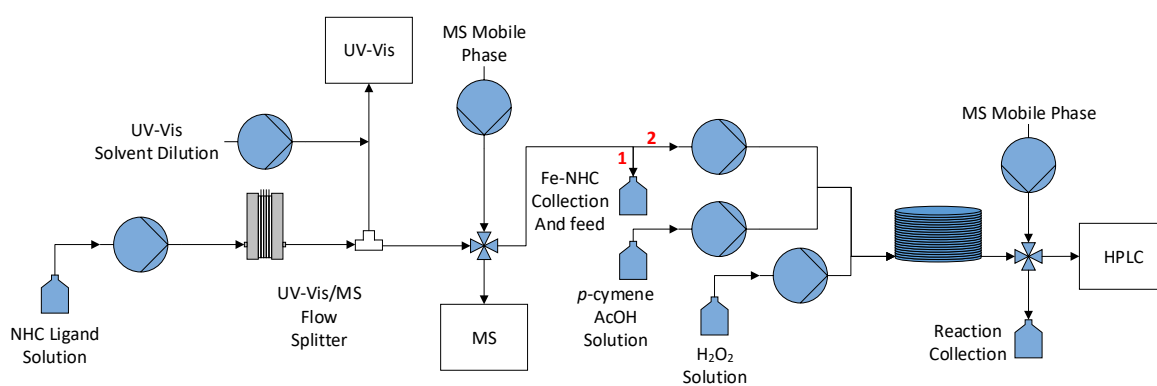


*Figure 121 Schematic of the telescoped process to generate Fe-NHC complexes in electrochemical flow monitored by on-line MS and in-line UV-Vis followed by catalytic C-H hydroxylation of p-cymene in flow monitored by on-line HPLC.*

The ECF reactor (Figure 104 & Figure 115) was thoroughly cleaned ensuring that all rust and particulate debris was removed. This was done by washing all electrodes and PTFE spacers in dilute HCl (0.1 M) by submerging in an acid bath before being rinsed with acetone and air dried before being assembled; 2x outer casing, 2x casing PTFE spacer, 4x iron electrode and 3x PTFE flow spacer. The ECF cell was connected to pump and sample actuator using 1/16" PTFE tubing and flushed with dry, degassed MeCN. The NHC ligand precursor, (0.205 g, 05 mmol), was weighed and dissolved in MeCN (100 mL) to create a 5 mM concentration solution, this was then primed through a HPLC pump. The battery pack was connected to the ECF reactor so that each electrode could be employed as an anode and cathode (*i.e.* electrode 1 = anode, electrode 2 = cathode, electrode 3 = anode *etc.*) and the required voltage set (2.5 V). The ligand precursor was then pumped at the set flow rate (1 mL min$^{-1}$) before being analysed using on-line MS.

A solution of *p*-cymene (1.71 mL, 11 mmol), was prepared in MeCN (50 mL) with AcOH added (0.3 mL, 0.5 eq.) and primed through a HPLC pump. A second solution of $H_2O_2$ (0.2 mL, 1.2 eq) in MeCN (50 mL) was prepared and primed through a HPLC pump. Once sufficient catalytic solution had been prepared (50 mL) both other reagent pump was set flowing at 0.1 mL min$^{-1}$ to give a total flow rate of 0.3 mL min$^{-1}$ and a residence time of 10 mins (reactor volume of 3 mL).

Mass spectra: *m/z* 223, *m/z* 613 (required; *m/z* 223, *m/z* 613, found; *m/z* 223, *m/z* 613).

HPLC: (IS) 2.57 min and (SM) 1.66 minutes at 254 nm.

**Thesis Summary and Project Outlook**

Each reaction carried out throughout this thesis has added valuable knowledge of MS to monitor, model and optimise continuous flow reactions. MS has always been an invaluable tool to gather qualitative data on compounds, whether known or unknown, and determine their identity. The development of techniques that are both qualitative and quantitative add a further dimension to their use.

A wide variety of reactions from simple organic deprotections to metal-based catalyst synthesis were used to explore quantitative analysis. Each case study has employed MS as a standalone technique where the entire reaction system is ionised as a whole. Normally, MS is used as a detector from preceding separation, so each peak can be mass analysed giving an understanding of component concentration and identity *i.e.* both quantitative and qualitative data. Analysing everything together becomes more challenging as ion suppression becomes dominant. The first study, N-Boc deprotection of an API, showed that MS can be used both quantitatively and qualitatively as the reaction could be analysed for its structural composition and reaction conversion. Ion suppression for this reaction was minimal as the compound used was thermally stable generating few species to compete with both SM and DP for ionisation. This success led the reaction system to be explored in more detail to find that temperature, residence time and water concentration were all important for high SM conversion.

It must be noted that any compounds which are less thermally stable may not be viable to be analysed with MS quantitatively. The high temperature of APCI can cause degradation of reaction components which, along with ion suppression, minimise the chances of accurate concentration determination and therefore inaccurate yield and conversion data. This was shown in the second study with N-Boc deprotection of Boc-Arg(Pbf)-OH. Suppression was observed in mass spectra failing IS to be observed but allowing both SM and DP to be seen. It is still possible to see each component in the reaction and use the ratio between peaks to understand and estimate conversion, which is still very useful when generating large data sets. Hybrid DoE and kinetic model design still need some improvement to ensure there are enough data points for the model to be fit. Five residence time points would be an absolute minimum to get a good set of kinetic data and keep the number of experiments to a minimum whilst maximising data. It may be beneficial to create an algorithm which runs all necessary reactions for a full surface

data set *i.e.* DoE, kinetic run and self-optimisation. Each of these methods extracts different information from a reaction, all essential to process understanding optimisation ready for scale-up or academic novel research. This idea was developed by running the DoE, self-optimisation and kinetic run in tandem, along with calibration for the $S_NAr$ reaction for AZD4547. Although not encompassed in the same algorithm, the concept of combining each method together reduces time by a large amount, all three could be completed in three days. Although MS could not quantify the reaction, it still played an important role in understanding the reaction and reinforcing the data gathered by HPLC. Once optimised, reactions can then be further characterised and using multiple analytical techniques. MS and UV-Vis were used to monitor and characterise the generation of Fe-NHC complexes whilst HPLC for the catalytic application of these catalysts. MS offered a valuable insight to monitor the generation of each catalyst allowing ligand precursor and complex to be observed to find conditions for electrochemical synthesis. The use of ESI limited quantitation as this method of ionisation is much more prone to ion suppression than APCI. The DoE model of the reaction, also using MS. This made it easier to telescope a subsequent reaction using MS to monitor and validate the generation of catalyst which could then be fed into a C-H hydroxylation reaction. Although there was no observed reaction in flow for the catalytic hydroxylation (still needing some fine tuning), this concept has been shown to work as mentions for the previous AZD5634 N-Boc deprotection.

Overall MS still holds up to its inherent qualitative nature by being able to monitor reactions extremely well as well as being valuable when modelling reactions particularly DoE methods.

MS can optimise certain reactions but only if they contain thermally stable compounds. APCI is far more quantitative than ESI. Benchtop MS still needs development but is a powerful and versatile technique to be used for on-line analysis of a wide range of reactions.

Although MS is a powerful analytical tool to use for monitoring a variety of on-line reactions more work still needs to be completed to create a robust system. This is not necessarily related to this PhD work but from the benchtop system itself. The CMS is limited to only a single sample from a single reaction every time but is capable of running multiple samples from different reactions or multiple samples from one reaction. A simple multichannel switch in built into the system would be very useful and would help

gain much more essential data *e.g.* a sample before and a sample after the reaction. It would also be very useful of the instrument could acquire data in both ESI mode and APCI mode simultaneously rather than having to switch sources whenever the other is needed. This would be advantageous to gain data on components in the reaction which don't ionise *via* one source or contain multiple groups for ionisation as with the work in Chapter 5. Even with these limitation the CMS is a great instrument to use to analyse qualitatively and quantitatively (reaction dependent). Prior knowledge of reactions is essential to quantitate by this technique or more in-depth method development to determine advantageous conditions without (or minimising) the decomposition of molecules. Further work using thermally stable systems may help with the and then proceeding to system with are no thermally stable (particularly with regards to APCI).

The caveats of MS are well established and have been further outlined in this work. Finding molecules and reactions which avoid these would be relatively simple but will need to align with the nature of each ionisation method. Reactions involving more polar groups *i.e.* amines, amides, carboxyl *etc.* are well …. To APCI whilst functional groups which comprise salts, easily dissociate in solution or contain many charged section *e.g.* amino acids, fit well with ESI. The system could be functionalised to monitor and optimise catalytic reactions where MS could be employed to not only quantitate the reaction be look deeper into catalytic species or intermediates formed during catalysis.

This technique has huge potential, but it should not be forgotten that the best reactions and research make use of a variety of analytical techniques and this should be translated to flow and optimisation reactions.

## References

1       F. E. Valera, M. Quaranta, A. Moran, J. Blacker, A. Armstrong, J. T. Cabral and D. G. Blackmond, *Angew. Chemie Int. Ed.*, 2010, **49**, 2478–2485.

2       R. L. Hartman, J. P. McMullen and K. F. Jensen, *Angew. Chemie Int. Ed.*, 2011, **50**, 7502–7519.

3       I. Kovacs, R. V Jones, K. Niesz, C. Csajagi, B. Borcsek, F. Darvas and L. Urge, *JALA J. Assoc. Lab. Autom.*, 2007, **12**, 284–290.

4       B. P. Mason, K. E. Price, J. L. Steinbacher, A. R. Bogdan and D. T. McQuade, *Chem. Rev.*, 2007, **107**, 2300–2318.

5       D. F. H. V. D. G, *Flow Chemistry*, 1st edn., 2014, vol. 1.

6       M. B. Plutschack, B. Pieber, K. Gilmore and P. H. Seeberger, *Chem. Rev.*, 2017, **117**, 11796–11893.

7       J. H. Bannock, S. H. Krishnadasan, M. Heeney and J. C. de Mello, *Mater. Horizons*, 2014, **1**, 373–378.

8       D. L. Browne, M. Baumann, B. H. Harji, I. R. Baxendale and S. V Ley, *Org. Lett.*, 2011, **13**, 3312–3315.

9       P. R. D. Murray, D. L. Browne, J. C. Pastre, C. Butters, D. Guthrie and S. V Ley, *Org. Process Res. Dev.*, 2013, **17**, 1192–1208.

10      K. J. Hartlieb, M. Saunders, R. J. J. Jachuck and C. L. Raston, *Green Chem.*, 2010, **12**, 1012–1017.

11      C. J. Mallia and I. R. Baxendale, *Org. Process Res. Dev.*, 2016, **20**, 327–360.

12      B. R. Holtz, B. R. Berquist, L. D. Bennett, V. J. M. Kommineni, R. K. Munigunti, E. L. White, D. C. Wilkerson, K.-Y. I. Wong, L. H. Ly and S. Marcel, *Plant Biotechnol. J.*, 2015, **13**, 1180–1190.

13      U. of York, The Essential Chemical Industry Online.

14      R. Porta, M. Benaglia and A. Puglisi, *Org. Process Res. Dev.*, 2016, **20**, 2–25.

15      V. Capriati, S. Florio, R. Luisi, F. M. Perna and A. Spina, *J. Org. Chem.*, 2008, **73**, 9552–9564.

16      D. Brown-Brulant, D. Depeyre and A. Isambert, *Comput. Chem. Eng.*, 1995, **19**, 537–542.

17      L. Macků and D. Novosad, *Procedia Eng.*, 2015, **100**, 360–369.

18      J. Worstell, ed. J. B. T.-B. and S.-B. R. Worstell, Butterworth-Heinemann, 2015, pp. 51–84.

19      P. F. Carneiro, B. Gutmann, R. O. M. A. de Souza and C. O. Kappe, *ACS Sustain. Chem. Eng.*, 2015, **3**, 3445–3453.

20      J. L. Quon, H. Zhang, A. Alvarez, J. Evans, A. S. Myerson and B. L. Trout, *Cryst. Growth Des.*, 2012, **12**, 3036–3044.

21      D. R. Snead and T. F. Jamison, *Chem. Sci.*, 2013, **4**, 2822–2827.

22      J. S. Moore and K. F. Jensen, *Angew. Chemie Int. Ed.*, 2014, **53**, 470–473.

23    M. Rasheed and T. Wirth, *Angew. Chemie Int. Ed.*, 2011, **50**, 357–358.

24    S. K. Teoh, C. Rathi and P. Sharratt, *Org. Process Res. Dev.*, , DOI:10.1021/acs.oprd.5b00001.

25    V. Sans, L. Porwol, V. Dragone and L. Cronin, *Chem. Sci.*, 2015, **6**, 1258–1264.

26    H. S. Fogler, *Elements of Chemical Reaction Engineering*, 2006, vol. 1.

27    L. Malet-Sanz and F. Susanne, *J. Med. Chem.*, 2012, **55**, 4062–4098.

28    J. Britton and T. F. Jamison, *Nat. Protoc.*, 2017, **12**, 2423.

29    O. Reynolds, *Philos. Trans. R. Soc. London*, 1883, **174**, 48.

30    H. S. Fogler, in *Elements of Chemical Reaction Engineering*, University of Michigan, 5th edn., 2005, pp. 767–794.

31    E. Mielke, P. Plouffe, N. Koushik, M. Eyholzer, M. Gottsponer, N. Kockmann, A. Macchi and D. M. Roberge, *React. Chem. Eng.*, 2017, **2**, 763–775.

32    D. D. Joye, *Ind. Eng. Chem. Res.*, 1996, **35**, 2399–2403.

33    C. S. Horbaczewskyj, A. A. Lapkin, R. A. Bourne and C. E. Willans, in *Green Chemical Engineering*, eds. P. T. Anastas and A. A. Lapkin, WILEY-VCH Verlag, 2018, pp. 329–369.

34    C. Houben and A. A. Lapkin, *Curr. Opin. Chem. Eng.*, 2015, **9**, 1–7.

35    Mettler-Toledo, FlowIR For Continuous Flow Chemistry.

36    R. A. Skilton, A. J. Parrott, M. W. George, M. Poliakoff and R. A. Bourne, *Appl. Spectrosc.*, 2013, **67**, 1127–1131.

37    C. F. Carter, H. Lange, S. V Ley, I. R. Baxendale, B. Wittkamp, J. G. Goode and N. L. Gaunt, *Org. Process Res. Dev.*, 2010, **14**, 393–404.

38    L. Malet-Sanz, J. Madrzak, S. V Ley and I. R. Baxendale, *Org. Biomol. Chem.*, 2010, **8**, 5324–5332.

39    Z. Qian, I. R. Baxendale and S. V Ley, *Chem. – A Eur. J.*, 2010, **16**, 12342–12348.

40    P. Koos, U. Gross, A. Polyzos, M. O'Brien, I. Baxendale and S. V Ley, *Org. Biomol. Chem.*, 2011, **9**, 6903–6908.

41    T. Brodmann, P. Koos, A. Metzger, P. Knochel and S. V Ley, *Org. Process Res. Dev.*, 2012, **16**, 1102–1113.

42    S. Mozharov, A. Nordon, D. Littlejohn, C. Wiles, P. Watts, P. Dallin and J. M. Girkin, *J. Am. Chem. Soc.*, 2011, **133**, 3601–3608.

43    K. Buckley and A. G. Ryder, *Appl. Spectrosc.*, 2017, **71**, 1085–1116.

44    D. Šahnić, E. Meštrović, T. Jednačak, I. Habinovec, J. Parlov Vuković and P. Novak, *Org. Process Res. Dev.*, 2016, **20**, 2092–2099.

45    G. Wypych and G. Wypych, *Handb. Mater. Weather.*, 2013, 1–25.

46    P. M. Günther, F. Möller, T. Henkel, J. M. Köhler and G. A. Groß, *Chem. Eng. Technol.*, 2005, **28**, 520–527.

47    D. A. Foley, E. Bez, A. Codina, K. L. Colson, M. Fey, R. Krull, D. Piroli, M. T. Zell and B. L.

Marquez, *Anal. Chem.*, 2014, **86**, 12008–12013.

48  E. Danieli, J. Perlo, A. L. L. Duchateau, G. K. M. Verzijl, V. M. Litvinov, B. Blümich and F. Casanova, *ChemPhysChem*, 2014, **15**, 3060–3066.

49  Nmrblog, 60 MHz NMR Comparison to 300 MHz NMR, https://nmrblog.wordpress.com/2010/11/27/60-mhz-nmr-of-essential-oils-from-benchtop-system-comparison-to-300-mhz-nmr-data/, (accessed 3 February 2018).

50  A. J. Parrott, R. A. Bourne, G. R. Akien, D. J. Irvine and M. Poliakoff, *Angew. Chemie Int. Ed.*, 2011, **50**, 3788–3792.

51  J. P. McMullen and K. F. Jensen, *Org. Process Res. Dev.*, 2011, **15**, 398–407.

52  J. P. McMullen and K. F. Jensen, *Org. Process Res. Dev.*, 2010, **14**, 1169–1176.

53  E. Naegele, *Making your LC method Compatible with Mass Spectrometry*, Agilent Technologies, 2011.

54  P. Gates, mechanism of ion formation, http://www.bris.ac.uk/nerclsmsf/techniques/hplcms.html, (accessed 28 January 2016).

55  Chromacademy, Mass Spectrometry - Fundamental LC-MS Electrospray Ionisation Theory, www.chromacademy.com.

56  Chromacademy, Mass Spectrometry - Fundamental LC-MS Atmospheric Pressure Chemical Ionisation (APCI), www.chromacademy.com, (accessed 28 January 2016).

57  P. Gates, Mechanism of APCI, http://www.bris.ac.uk/nerclsmsf/techniques/hplcms.html, (accessed 28 January 2016).

58  T. W. T. Bristow, A. D. Ray, A. O'Kearney-McMullan, L. Lim, B. McCullough and A. Zammataro, *J. Am. Soc. Mass Spectrom.*, 2014, **25**, 1794–1802.

59  D. L. Browne, S. Wright, B. J. Deadman, S. Dunnage, I. R. Baxendale, R. M. Turner and S. V Ley, *Rapid Commun. Mass Spectrom.*, 2012, **26**, 1999–2010.

60  S. Krishnadasan, R. J. C. Brown, A. J. deMello and J. C. deMello, *Lab Chip*, 2007, **7**, 1434–1441.

61  M. Friedman and L. J. Savage, *"Planning Experiments Seeking Maxima," in Techniques of Statistical Analysis*, McGraw-Hill, United States of America, NY, 1947.

62  Media_Eclipse_Business_Ltd, Separation Science, http://www.sepscience.com/Sectors/Miscellaneous/Articles/160-/Optimization-of-Analytical-Methods-Using-Factorial-Designs----Part-1-An-Introduction#content.

63  L. A. Reddy, G. C. Malakondaiah, K. S. Babu, A. Bhattacharya, R. Bandichhor, V. Himabindu, P. P. Reddy and R. V. Anand, *Org. Process Res. Dev.*, 2008, **12**, 66–68.

64  P. Goos and B. Jones, Exploring best practice in Design of Experiments, http://www.rsc.org/chemistryworld/2014/12/best-practice-experiments-webinar.

65  R. Leardi, *Anal. Chim. Acta*, 2009, **652**, 161–172.

66    L. Sanchez, *Statistical Design of Experiments Applied to Organic Synthesis*, Michigan State University, 2006.

67    C.-F. Mandenius and A. Brundin, *Biotechnol. Prog.*, 2008, **24**, 1191–1203.

68    M. Meeuwse, J. van der Schaaf and J. C. Schouten, *AIChE J.*, 2012, **58**, 247–255.

69    I. A. Gargurevich, *J. Chem. Eng. Process Technol.*, 2016, **07**, 1–6.

70    J. Rostrup-Nielsen, *J. Mol. Catal. A Chem.*, 2000, **163**, 157–162.

71    G. Donati and R. Paludetto, *Catal. Today*, 1997, **34**, 483–533.

72    W. Spendley, G. R. Hext and F. R. Himsworth, *Technometrics*, 1962, **4**, 441–461.

73    M. H. Wright, *Doc. Math.*, 2012, **Extra Volume ISMP**, 6.

74    M. W. Routh, P. A. Swartz and M. B. Denton, *Anal. Chem.*, 1977, **49**, 1422–1428.

75    W. Huyer and A. Neumaier, *ACM Trans. Math. Softw.*, 2008, **35**, 1–25.

76    N. Holmes, R. A. Bourne, T. M. Letcher., J. L. Scott. and D. A. Patterson, *Chemical Processes For a Sustainable Future*, Royal Society of Chemistry, Cambridge, 1st edn., 2015.

77    J. J. Tomick., The Pennsylvania State University, 1995.

78    L. R. Parker, M. R. Cave and R. M. Barnes, *Anal. Chim. Acta*, 1985, **175**, 231–237.

79    H. Schichl, in *A Personal View on Global Optimisation*, Cambridge, 2014.

80    PHRMA, *Biopharmaceutical Research & Development: The Process Behind New Medicines*, PHRMA, http://www.phrma.org/sites/default/files/pdf/rd_brochure_022307.pdf, 2015.

81    M. E. Kopach, D. J. Roberts, M. D. Johnson, J. McClary Groh, J. J. Adler, J. P. Schafer, M. E. Kobierski and W. G. Trankle, *Green Chem.*, 2012, **14**, 1524–1536.

82    N. Kockmann, M. Gottsponer, B. Zimmermann and D. M. Roberge, *Chem. – A Eur. J.*, 2008, **14**, 7470–7477.

83    M. Damm and C. O. Kappe, *Anal. Chim. Acta*, 2011, **707**, 76–83.

84    I. R. Baxendale, R. D. Braatz, B. K. Hodnett, K. F. Jensen, M. D. Johnson, P. Sharratt, J.-P. Sherlock and A. J. Florence, *Int. Symp. Contin. Manuf. Pharm.*, 2014, 19.

85    S. Borukhova, T. Noël and V. Hessel, *ChemSusChem*, 2016, **9**, 67–74.

86    D. M. Roberge, B. Zimmermann, F. Rainone, M. Gottsponer, M. Eyholzer and N. Kockmann, *Org. Process Res. Dev.*, 2008, **12**, 905–910.

87    D. E. Fitzpatrick, C. Battilocchio and S. V Ley, *Org. Process Res. Dev.*, 2016, **20**, 386–394.

88    P. Yaseneva, P. Hodgson, J. Zakrzewski, Fal, R. E. Meadows and A. A. Lapkin, *React. Chem. Eng.*, 2016, **1**, 229–238.

89    M. Saber, J. M. Commenge and L. Falk, *Chem. Eng. Sci.*, 2010, **65**, 372–379.

90    C. Amador, A. Gavriilidis and P. Angeli, *Chem. Eng. J.*, 2004, **101**, 379–390.

91    M. S. Mettler, G. D. Stefanidis and D. G. Vlachos, *Ind. Eng. Chem. Res.*, 2010, **49**, 10942–10955.

92    P. Strazzolini, N. Misuri and P. Polese, *Tetrahedron Lett.*, 2005, **46**, 2075–2078.

93      B. Li, M. Berliner, R. Buzon, C. K. F. Chiu, S. T. Colgan, T. Kaneko, N. Keene, W. Kissel, T. Le, K. R. Leeman, B. Marquez, R. Morris, L. Newell, S. Wunderwald, M. Witt, J. Weaver, Z. Zhang and Z. Zhang, *J. Org. Chem.*, 2006, **71**, 9045–9050.

94      D. S. Coffey, M. K. N. Hawk, S. J. Ghera, P. G. Marler, P. N. Dodson and M. L. Lytle, *Org. Process Res. Dev.*, 2004, **8**, 945–947.

95      N. Srinivasan, A. Yurek-George and A. Ganesan, *Mol. Divers.*, 2005, **9**, 291–293.

96      J. Wang, Y.-L. Liang and J. Qu, *Chem. Commun.*, 2009, 5144–5146.

97      V. H. Rawal and M. P. Cava, *Tetrahedron Lett.*, 1985, **26**, 6141–6142.

98      A. R. Bogdan, M. Charaschanya, A. W. Dombrowski, Y. Wang and S. W. Djuric, *Org. Lett.*, 2016, **18**, 1732–1735.

99      J. Choy, S. Jaime-Figueroa, L. Jiang and P. Wagner, *Synth. Commun.*, 2008, **38**, 3840–3853.

100     U. Jacquemard, V. Bénéteau, M. Lefoix, S. Routier, J.-Y. Mérour and G. Coudert, *Tetrahedron*, 2004, **60**, 10039–10047.

101     G. Pavan Kumar, D. Rambabu, M. V Basaveswara Rao and M. Pal, *J. Chem.*, 2013, **2013**, 5.

102     N. Ravindranath, C. Ramesh, M. Ravinder Reddy and B. Das, *Adv. Synth. Catal.*, 2003, **345**, 1207–1208.

103     J. M. Lopez-Soria, S. J. Perez, J. N. Hernandez, M. A. Ramirez, V. S. Martin and J. I. Padron, *RSC Adv.*, 2015, **5**, 6647–6651.

104     A. Laval, Alfa Laval ART Plate Reactor 37.

105     A. Laval, Alfa Laval ART Plate Reactor 49.

106     B. Li, R. Li, P. Dorff, J. C. McWilliams, R. M. Guinn, S. M. Guinness, L. Han, K. Wang and S. Yu, *J. Org. Chem.*, , DOI:10.1021/acs.joc.8b02909.

107     C. for P. A. and C. Technology, No Title, www.cpact.com.

108     S. Luo, Y. Zhang, J. Cao, B. He and S. Li, *Colloids Surfaces B Biointerfaces*, 2016, **148**, 181–192.

109     R. H. Böger and S. M. Bode-Böger, *Annu. Rev. Pharmacol. Toxicol.*, 2001, **41**, 79–99.

110     R. H. Böger, *J. Nutr.*, 2007, **137**, 6.

111     R. G. Knowles and S. Moncada, *Biochem. J.*, 1994, **298**, 249–258.

112     R. Brinkman, R. Margaria and F. J. W. Roughton, *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, 1934, **232**, 65–97.

113     N. Holmes, G. R. Akien, R. J. D. Savage, C. Stanetty, I. R. Baxendale, A. J. Blacker, B. A. Taylor, R. L. Woodward, R. E. Meadows and R. A. Bourne, *React. Chem. Eng.*, 2016, **1**, 96–100.

114     S. K. Sythana, S. R. Naramreddy, S. Kavitake, V. K. Ch and P. R. Bhagat, *Org. Process Res. Dev.*, 2014, **18**, 912–918.

115     F. Terrier, *Modern Nucleophilic Aromatic Substitution*, Wiley-VCH Verlag GmbH & Co.

KGaA, Weinheim, Germany, 2013.

116    B. J. Reizman and K. F. Jensen, *Org. Process Res. Dev.*, 2012, **16**, 1770–1782.

117    C. A. Hone, R. A. Bourne and F. L. Muller, *Chim. Oggi*, 2015, **33**, 7.

118    H. P. Gavin, *The Levenburg-Marqurdt Algorithm For Nonlinear Least Squares Curve-Fitting Problems*, 2019.

119    H. A. Taylor and E. E. Juterbock, *J. Phys. Chem.*, 1934, **39**, 1103–1110.

120    M. N. Hopkinson, C. Richter, M. Schedler and F. Glorius, *Nature*, 2014, **510**, 485.

121    A. J. Arduengo, R. L. Harlow and M. Kline, *J. Am. Chem. Soc.*, 1991, **113**, 361–363.

122    V. Nesterov, D. Reiter, P. Bag, P. Frisch, R. Holzner, A. Porzelt and S. Inoue, *Chem. Rev.*, 2018, **118**, 9678–9842.

123    M. Fèvre, J. Pinaud, Y. Gnanou, J. Vignolle and D. Taton, *Chem. Soc. Rev.*, 2013, **42**, 2142–2172.

124    E. Peris, *Chem. Rev.*, 2018, **118**, 9988–10031.

125    M. N. Hopkinson and F. Glorius, in *N-Heterocyclic Carbenes in Organocatalysis*, ed. A. T. Biju, Wiley-VCH Verlag GmbH & Co. KGaA, 2019, pp. 1–35.

126    S. Fantasia, J. L. Petersen, H. Jacobsen, L. Cavallo and S. P. Nolan, *Organometallics*, 2007, **26**, 5880–5889.

127    C. E. Willans, *N-Heterocyclic Ligands (NHCs)*, Leeds, 2017.

128    M. R. Chapman, Y. M. Shafi, N. Kapur, B. N. Nguyen and C. E. Willans, *Chem. Commun.*, 2015, **51**, 1282–1284.

129    X. Cui, J. W. Ogle and K. Burgess, *Chem. Commun.*, 2005, 672–674.

130    R. A. Batey, M. Shen and A. J. Lough, *Org. Lett.*, 2002, **4**, 1411–1414.

131    J. Huang and S. P. Nolan, *J. Am. Chem. Soc.*, 1999, **121**, 9889–9890.

132    L.-A. Schaper, E. Tosh and W. A. Herrmann, in *N-Heterocyclic Carbenes: From Laboratory Curiosities to Efficient Synthetic Tools*, The Royal Society of Chemistry, 2011, pp. 166–195.

133    Y. Suzuki, T. Toyota, F. Imada, M. Sato and A. Miyashita, *Chem. Commun.*, 2003, 1314–1315.

134    M. C. Jahnke and F. Ekkehardt Hahn, in *N-Heterocyclic Carbenes: From Laboratory Curiosities to Efficient Synthetic Tools*, The Royal Society of Chemistry, 2011, pp. 1–442.

135    L. Delaude and A. Demonceau, in *N-Heterocyclic Carbenes: From Laboratory Curiosities to Efficient Synthetic Tools*, The Royal Society of Chemistry, 2011, pp. 196–227.

136    A. T. Normand and K. J. Cavell, in *N-Heterocyclic Carbenes: From Laboratory Curiosities to Efficient Synthetic Tools*, The Royal Society of Chemistry, 2011, pp. 252–283.

137    V. Cesar, L. H. Gade and S. Bellemin-Laponnaz, in *N-Heterocyclic Carbenes: From Laboratory Curiosities to Efficient Synthetic Tools*, The Royal Society of Chemistry, 2011, pp. 228–251.

138    N. Marion*, in *N-Heterocyclic Carbenes: From Laboratory Curiosities to Efficient Synthetic Tools*, The Royal Society of Chemistry, 2011, pp. 317–344.

139    Y. Fort and C. Comoy, in *N-Heterocyclic Carbenes: From Laboratory Curiosities to Efficient Synthetic Tools*, The Royal Society of Chemistry, 2011, pp. 284–326.

140    J. Louie and R. H. Grubbs, *Chem. Commun.*, 2000, 1479–1480.

141    G. Hilt, P. Bolze and I. Kieltsch, *Chem. Commun.*, 2005, 1996–1998.

142    F. Roudesly, J. Oble and G. Poli, *J. Mol. Catal. A Chem.*, 2017, **426**, 275–296.

143    M. S. Chen and M. C. White, *Science (80-. ).*, 2007, **318**, 783–787.

**Appendices**

**Appendix I**

This appendix highlights the main MATLAB codes which were used throughout the work presented in this thesis. These include timer restart codes, HPLC extraction code and MS extraction code. Finally, links to the other codes are proved which are all open source, these include the full CO2GUI optimisation code, SNOBFIT code and the MATLAB repository where other useful codes can be found.

Timer code restart:

```
>>a=timerfindall

  Timer Object Array

  Index:  ExecutionMode:  Period:  TimerFcn:              Name:

  1       fixedSpacing      10      1x2 cell array         timer-1

  2       fixedSpacing      10      1x2 cell array         timer-2

  3       fixedRate          5      rheodyneTimerCallback  timer-3

  4       fixedRate       1200      1x2 cell array         timer-4

  5       singleShot         1      1x2 cell array         timer-5

  6       fixedSpacing      12      1x2 cell array         timer-6

  7       fixedSpacing      13      1x2 cell array         timer-7

  8       fixedSpacing      19      1x2 cell array         timer-8

  9       fixedSpacing      21      1x2 cell array         timer-9

>>a(6)

  Timer Object: timer-6

  Timer Settings

          ExecutionMode: fixedSpacing

                 Period: 12

               BusyMode: drop

                Running: on

  Callbacks

                TimerFcn: {@stabletimcallback [1x1 struct]}

                ErrorFcn: ''

                StartFcn: {@stablestartcallback [1x1 struct]}

                 StopFcn: ''

>>Start (a(6))
```

Timer code restart alternative:

> *>> stop(timerfindall('Tag','snobfit_tim'))*

> *>> set(timerfindall('Tag','snobfit_tim'),'StartDelay',10)*

> *>> start(timerfindall('Tag','snobfit_tim'))*

Auto-restart Code:

> *ttt = allowedDeviation>0;*

> *if ttt==0*

>   *allowedDeviation = 1;*

> *else*

> *end*

> *ttt2 = allowedNoise>0;*

> *if ttt2==0*

>   *allowedNoise = 1;*

> *else*

> *end*

HPLC data extraction code:

% user input command which specifies the span or the chromatographic method for extraction (HPLC method length in mins) and the time spacing interval between each point. This creates a column of numbers from start to end of the method.

> *>>rtable=autortable(0,4,0.1);*

> *>>[dataoutput,nameoutput] = HPLCmultiv3(rtable,1);*

> *>>nameoutput=nameoutput'*

> *>>rtable=rtable'*

% function for the rtable using the autortable built in program from MATLAB. MinRT and maxRT both correspond to the set limits in the chromatographic method whilst timeint is the time spacing interval between each point. MinRT, MaxRT and timeint are all specified from user input.

> *Function rtable=autortable(minRT,maxRT,timeint);*

>   *rtable=(minRT:timeint:maxRT);*

>   *rtable(2,:)=rtable+timeint;*

>   *rtable=rtable';*

> *end*

% Program accesses the 'rtable' file and gets a directory from the user to access HPLC data and concatenates them in the correct order. The program then extracts relevant data *i.e.* peak areas at each time interval specified in 'rtable' along with the filename and time. 'Del' corresponds to the wavelength number of the file if multiple wavelengths have been selected using a diode array detector (DAD).

```
function [dataoutput,nameoutput,timeoutput] = HPLCmultiv3(rtable,del)
        path=uigetdir;
        path=strcat(path,'\');
```

% gets all the files with ".D" into the dircom subject, one folder for each report file from Agilent so can pull path data.

```
        dircom = strcat(path,'*.D*');
        gcdata = dir(dircom);
```

% gets the last GCs taken, should match up with the last number of conditions.
% NOTE: sorts by name not date, fine with autoHPLC naming convention changed as HPLC pre-makes folder.

```
        nameoutput={'a'};
        dataoutput=zeros(1,size(rtable,1));
        timeoutput=zeros(1,6);
        gcdata=gcdata(1:end-del,:);
        for i = 1:size(gcdata)
                file = strcat(path, gcdata(i).name,'\REPORT01.xls');
        end
        dataoutput=dataoutput(2:end,:)
```

MS data extraction code:

```
    function [fileno,filename,MS_Date,MSint,T]=MSextraction(mass);
```

% Function to extract the file number, file name, date MS file was created,
% Total integration of each MS mass given and displayed in a table. Only the mass (or m/z) needs to be specified for extraction of data... Masses
% can be given singularly [m1], as multiples [m1, m2, m3... etc] or as a range [m1:m100... etc].
% e.g.[filename,MSint134]=MSgraballCHmulti([76,116,134])
% Extraction function to be linked with GUI for automated mass extraction

```
    directory=uigetdir;
```

% ask user for a directory

% If the same directory is to be accessed the this can be placed in brackets with a full domain description e.g.

% (C:\Users\userprofile\Documents\Advion Mass Spectrometer)

```
    fileloc=strcat(directory,'\*.cdf');
```

%concatenate file location and type of files

```
    listofMS=dir(fileloc);
```

%List all files in that directory ending with cdf - these are not in date order but alphanumeric i.e. 10 is before 2

```
    for i=1:size(listofMS,1);
      sourcetime(i,1)=(listofMS(i).datenum);
      sourcetime(i,2)=(i);
```

%list of file no and date time

```
    end
      B=sortrows(sourcetime,1);
```

%sort by datetime to order files when taken

```
    for j=1:size(mass,2);
    for i=1:size(listofMS,1);
      source=strcat(directory,'\',listofMS(B(i,2)).name);
```

%move down list based on date time from B

```
        [MSout,MSout2]=MSintCH(source, mass(j));
```

%extract total integrated using MSintCH file

```
      fileno(i)=i;
      MSint(i,j)=MSout;
      filename{i}=listofMS(B(i,2)).name;
```

% filedate_time=listofMS(B(i,3)).name

% list of filenames

```
    end
    end
    figure
```

% Creates an empty figure for the extracted data to be visualised in.

% This can be useful if monitoring steady state or multiple injections are

% made in a single file.

```
    fileno=fileno';
```

% Transposes file number to column from row

```
    filename=filename';
```

% Transposes filename from row to column

```
    valid_file_names= find(~[listofMS.isdir]);
```

% Finds all valid file names in List of MS

```
    file_date=[listofMS.datenum];
```

% Extracts and lists the datnum information from the structure

```
    dv=datevec(file_date);
```

% Changes the datnum to a date vector format (YYYY-M-DD HH:MM:SS)

```
    t=datetime(dv,'InputFormat','YYYY-MM-dd HH:mm:ss');
```

% Changes each row input format of the date vector to calendar format

```
    MS_Date=sortrows(t,1);
```

% Sorts the date time information from earliest to latest

```
     plot(fileno,MSint);
```

% Plots the extracted data in the opened figure

```
    T=table(fileno,filename,MS_Date,MSint,...
       'VariableNames',{'File_Number','File_Name','Date','mz'});
```

% Puts all the extracted data into a table in the setup stated

```
    savdir='C:\Users\Christopher\Documents\PhD Documents\Reports etc\';
```

% Saves the table of data to the specified file (doesn't yet do this)

```
    file_name = 'MS_Extraction_Data.xlsx';
```

% Save name and format of the table of data

```
    writetable(T,file_name,'Sheet',1,'Range','A1');
```

% Writes the table of data to an Excel spreadsheet in the; sheet number and cell range stated. Writes to excel sheet works but not in the specified save directory. Instead, it currently saves in the active MATLAB folder!!

Self-optimisation code:

For Snobfit code and drivers and the code used to control all reactor peripherals, access the Snobfit algorithm and optimise a reactions using HPLC; the original version of each code can be found at :–

- https://www.mat.univie.ac.at/~neum/software/snobfit/
- https://uk.mathworks.com/MATLABcentral/profile/authors/1795689-geoffrey-akien
- https://uk.mathworks.com/MATLABcentral/fileexchange/26288-co2gui-lab-control-and-automation?focused=efaf85fd-331d-c5f4-b8ee-f389c1faee95&tab=function