

Chemical Synthesis Optimiser

Related Applications

5 The present case claims the benefit of, and priority to, GB 2315721.7 filed 13 October 2023 (13.10.2023), the contents of which are hereby incorporated by reference in their entirety.

Field of the Invention

The present invention relates to a method of optimising an instruction set for a chemical synthesis and a chemical synthesiser for use in a method of optimising an instruction set for a chemical synthesis.

10 ***Background***

Smart laboratory automation holds promise to accelerate chemical research, eliminate tedious tasks, improve safety and reliability.^{1,3,4} Recently there has been significant progress towards more automated synthesis platforms: systems that can perform a large variety of synthetic processes, giving access to a diverse set of target compounds.^{7,9,13}

15

One of the present inventors has shown how a method of synthesis for a target product can be translated into digital model of that method: see WO 2019/137954. This digital model comprises a digital description of the chemical and/or physical steps within the method of synthesis. The present inventor has also shown how literature methods in natural language
20 may be converted into a machine-readable instruction set: see WO 2021/219999. This instruction set may then be used to generate a hardware index, which is the required reactionware for the performance of a synthetic procedure.

20

Reaction optimisation has more recently become a part of the chemical automation
25 development.¹⁷⁻²¹, but the vast majority of platforms are limited to narrow chemical tasks with few exceptions of systems for flow chemistry.²²⁻²⁴ These systems are bound to specific hardware modules and software, thus making the optimal protocols not transferable across chemical automation robots. The optimisation is limited to normal reaction steps, and focuses on the reaction product to direct optimisation efforts. This restricts the optimisation ability of
30 the systems to a few static parameters.

25

Moreover reaction optimisation traditionally focuses on the use of random or semi-random selections to identify parameters for optimisation. For example, WO 2013/175240 describes a method for generating a product having one or more characteristics which meet or exceed a
35 specification. The combination of inputs into the system are determined by a semi-random selection of parameters, where a parameter may be retained if it is associated with a reaction

35

producing having good characteristics. However, the optimisation by such a system is relatively slow and relies on the system randomly chancing on a promising set of parameters.

Using such machine-readable instruction sets and physical reactionware developed from digital models that are responsive to the machine-readable instructions, it is possible to develop and utilise standardised models for chemical synthesis that can be executed by an automated synthesiser.

While these platforms perform experiments in an automated fashion, they are limited to predetermined, sequential processes, adapted from literature and trivial laboratory operations. There are still challenges where a reactions instructions are dependent on secondary factors like maintaining a certain temperature, starting or stopping a reaction step based on a colour change, scaling up or down reactions, and accounting for the surrounding environmental conditions (so called "dynamic" steps). These dynamic steps often require a varying input, in order to keep the dynamic parameter at the required value. For example, the addition of an exothermic-reaction causing reagent might need to be closely monitored to keep the temperature of the reaction at the required value. However, the exact addition rate needed will need to change depending on the scale of the reaction, outside temperature, solvent volume and other factors.

An operator often knows how to perform these 'dynamic' steps from prior experience, and so such steps are relatively straightforward for an operator to perform through regular monitoring of the reaction. However automation of such dynamic reaction steps has proved challenging. The dynamic steps are particularly difficult to reproduce across different synthesisers, and the optimal performance of the dynamic step may vary. As a result, the repeatability of reactions and quality of the reaction product produced can be variable, depending on different batches of reagents, and different external conditions. The quality of the product produced is particularly variable when such dynamic reaction steps are present in the synthesis. This makes optimisation of such parameters by an automated chemical synthesiser challenging.

Consequently, there is a need for a chemical synthesis platform for use in conducting chemical syntheses under reproducible conditions, which is able to adapt to changes in external conditions, and carry out dynamic reaction steps, with good repeatability. There is also a need for a chemical reaction optimisation platform which can intelligently select and perform a large range of reaction steps, including complex dynamic reaction steps, and optimise a variety of reaction parameters to provide a more comprehensive approach to chemical synthesis optimisation. Such platforms are needed to more fully leverage the potential of robotic systems in the chemical domain.

Summary of the Invention

The present inventors have now devised a method and chemical synthesiser for the optimization of chemical synthesis protocols which looks to address the above-mentioned

problems. Reaction data is used for process state monitoring and, with predefined rules, allows for dynamic procedure execution, self-correction and real-time decision making during a chemical synthesis. Product data is used to quantifying reaction outcomes, and is coupled with the reaction data and dynamic execution to create a closed-loop method and chemical synthesiser for reaction optimization.

Generally, the invention relates to a method of optimising a chemical synthesis. The method may be performed by a chemical synthesiser operating autonomously.

- 10 In a general aspect, the present invention provides a method of optimising a chemical synthesis, the optimisation comprising the steps of:
- (i) performing a chemical synthesis in a reaction vessel of a chemical synthesiser operating according to an instruction set to produce a reaction product,
 - (ii) recording reaction data using an analytical unit of the synthesiser during the
 - 15 chemical synthesis and/or recording product data on the reaction product,
 - (iii) analysing the reaction data and/or product data using a control unit of the chemical synthesiser to determine a characteristic of the synthesis; and
 - (iv) adjusting one or more reaction parameters of the instruction set, to provide an adjusted instruction set, based on a prediction that the adjustment will improve the
 - 20 characteristic of the synthesis.

- In a first aspect of the invention there is provided a method of optimising an instruction set for a chemical synthesis, the method comprising performing a series of optimisation test syntheses, where each optimisation test synthesis includes the steps of:
- 25 (i) performing the chemical synthesis in a reaction vessel of a chemical synthesiser operating according to an instruction set to produce a reaction product,
 - (ii) recording reaction data using an analytical unit of the synthesiser during the chemical synthesis and/or recording product data on the reaction product,
 - (iii) analysing the reaction data and/or product data using a control unit of the chemical
 - 30 synthesiser to determine a characteristic of the synthesis; and
 - (iv) adjusting one or more reaction parameters of the instruction set, to provide an adjusted instruction set,
- wherein the optimisation test synthesis is repeated one or more times, and in each repeat of the test synthesis, the chemical synthesis is performed according to the adjusted
- 35 instruction set, which is adjusted from the previous test syntheses, and
 - the adjustment to the one or more reaction parameters is made following a prediction that the adjustment will improve the characteristic of the synthesis, wherein the prediction is based on the reaction parameters associated with reaction data and/or product data from one or more previous optimisation test syntheses and, where present, one or more previous
 - 40 exploration test syntheses, and
 - wherein the method is performed autonomously

In some embodiments, the method comprises:

performing a series of exploration test syntheses, where each exploration test synthesis includes the steps of:

- (i) performing the chemical synthesis in a reaction vessel of a chemical synthesiser operating according to an instruction set to produce a reaction product,
- 5 (ii) recording reaction data using an analytical unit of the synthesiser during the chemical synthesis and/or recording product data on the reaction product,
- (iii) analysing the reaction data and/or product data using a control unit of the chemical synthesiser to determine a characteristic of the synthesis; and
- (iv) adjusting one or more reaction parameters of the instruction set, to provide an
10 adjusted instruction set,
 wherein the exploration test synthesis is repeated one or more times, and in each repeat of the test synthesis, the chemical synthesis is performed according to the adjusted instruction set, which is adjusted from the previous test syntheses, and
 the adjustment to the one or more reaction parameters is made to explore a reaction
15 space wherein the adjustment is based on the reaction parameters associated with reaction data and/or product data from one or more previous test syntheses.

In the method the adjustment to the one or more reaction parameters is made following a prediction that the adjustment will improve the characteristic of the synthesis. This prediction
20 step allows for the intelligent and rapid optimisation of the instruction set based on data from previous test syntheses.

In some embodiments, for the optimisation test synthesis, the comparison of the reaction parameters to reaction data and product data from previous test syntheses identifies reaction
25 parameters associated with a change in reaction data and/or product data, and preferentially adjusts reaction parameters associated with the change. Preferably, the comparison of the reaction parameters to reaction data and product data from previous test syntheses identifies reaction parameters associated with the highest rate of change in reaction data and/or product data, and preferentially adjusts reaction parameters associated with the highest rate
30 of change.

The present invention provides for a method for optimising a chemical synthesis, including a broader range of chemical procedures and reaction steps, and which uses data on the reaction product and also recorded during the reaction itself to drive the optimisation of the
35 chemical synthesis. Using an array of analytical sensors to monitor the reaction mixture, the reaction environment (e.g., headspace) as well as the surrounding conditions, provides a more detailed picture of the reaction. Leveraging the combination of product data and reaction data allows for more effective reaction optimisation, beyond that of previous manual optimisation efforts and previous automated optimisation attempts. This method is shown to
40 improve reaction product yield, purity of the product (reduction of impurity), increase starting material conversion and is also thought to be helpful in reducing reagent or solvent usage.

The use of reaction data also allows for the identification of important points in the chemical reaction. For example, a change in reaction data and/or a high rate of change in the reaction data indicates an important facet of the chemical synthesis. Identifying these facets, from the reaction data, allows the optimisation to focus on reaction parameters which are important for this point in the synthesis. Accordingly, use of the reaction data allows for more efficient optimisation of chemical synthesis, as regions of interest can be identified and focused on for the optimisation efforts.

Moreover, reaction data taken during the chemical synthesis allows for the use of dynamic reaction parameters, or dynamic reaction steps.

In some embodiments, one or more of the reaction parameters is a dynamic reaction parameter, wherein the dynamic reaction parameter is adjusted during the test synthesis such that reaction data satisfies a different reaction parameter. Preferably, for the optimisation test synthesis, the comparison of the reaction parameters to reaction data and product data from previous test syntheses identifies dynamic reaction parameters associated with a change in reaction data and/or product data, and preferentially adjusts dynamic reaction parameters associated with the change. Preferably, the comparison of the reaction parameters to reaction data and product data from previous test syntheses identifies dynamic reaction parameters associated with the highest rate of change in reaction data and/or product data, and preferentially adjusts dynamic reaction parameters associated with the highest rate of change.

The dynamic reaction parameters provide for enhanced reaction control, and an increased diversity of the chemical synthesis protocols which can be performed using the chemical synthesiser. The dynamic parameters facilitate self-correcting reaction execution using feedback from reaction sensors. For example, reaction time, temperature, or rate of reagent addition can be used as dynamic reaction parameters, as these are adjusted to fulfil a different reaction parameter, such as colour, pressure, or temperature, respectively.

It is also thought that optimisation of dynamic reaction parameters provides a promising opportunity for effective reaction optimisation, as these parameters are typically difficult to optimise in manual synthesis and traditional automated optimisation apparatus. Thus, in some embodiments the method of the invention may identify a dynamic reaction parameter associated with a change in reaction data and/or product data, and preferentially adjusts dynamic reaction parameters associated with the change.

Furthermore, the dynamic reaction parameters allows for optimisation of synthesis at different scales and across different chemical synthesisers. For example, if a reaction is provided at a small scale, different chemical and physical inputs may be needed for larger scale reactions. For example, for an exothermic reaction the rate of addition of a reactant may be needed to be modified to avoid thermal runaway. This rate of addition cannot be scaled linearly, as the rate of heat loss is not linear with the scale. Moreover, the rate of heat loss may change

depending on the chemical synthesiser used (e.g., because of reaction vessel thermal capacity, shape, external temperature etc.). Therefore, the rate of addition will not be the same across different chemical synthesisers. The present inventors have developed a dynamic reaction control that allows for the dynamic adjustment of reaction conditions, and the optimisation of these dynamic reaction parameters at different scales and across different chemical synthesiser apparatus.

In a second aspect of the invention there is provided a chemical synthesiser for use in the method of optimising an instruction set for a chemical synthesis of claims 1 to 19, the synthesiser comprising a reaction vessel, an analytical unit and a control unit, wherein:

- the reaction vessel is for the performance of the chemical synthesis according to an instruction set,
- the analytical unit is for recording reaction data from the reaction vessel during the chemical synthesis and/or for recording product data on the reaction product, and for reporting the reaction data and/or product data to the control unit;
- the control unit is for analysing the reaction data and/or product data received from the analytical unit to determine a characteristic of the synthesis, and for adjusting one or more reaction parameters of the instruction set, to provide an adjusted instruction set, wherein the adjustment to the one or more reaction parameters is made following a prediction that the adjustment will improve the characteristic of the synthesis, wherein the prediction is based on the reaction parameters associated with reaction data and/or product data from one or more previous optimisation test syntheses.

In some embodiments, the chemical synthesiser is for performing the method autonomously.

The chemical synthesiser may be a robotic chemical synthesiser for autonomous performance of the chemical synthesis, and for autonomous control of the analytical device.

The control unit may be suitably programmed with an algorithm to analyse reaction data, product data and reaction parameters from previous test synthesis, and to adjust the reaction parameters for the next test synthesis.

The invention also provides a database, wherein the database is a collection of instruction sets, produced by a method of the invention. The database may be provided with an optimised instruction set. The database may include an instruction set which can be used to obtain a reaction product using an automated chemical synthesiser.

These and other aspects and the embodiments of the invention are described in further detail herein.

Summary of the Figures

The present invention is described with reference to the figures listed below.

Figure 1 shows an overview of the dynamic chemical operations execution. The procedure, illustrated using χ DL syntax, may be executed sequentially one step after another or dynamically, e.g., as highlighted with *AddDynamic* where a *wait* step is inserted if the temperature rises above given threshold. With integrated analysis and a corresponding χ DL step (*Analyze*), the procedure could run iterative with dynamic update of its parameters, thus allowing for closed-loop reaction conditions optimization. The *Setup* step represents the definition of the hardware configuration of the system and the chemical procedure defined digitally in the χ DL code. In the *Execution* step these components are used to operate the physical robotic system to achieve the defines chemical operations, including the collection of analytical data about the execution. In the *Optimization* step the data is analysed and algorithms are used to predict parameters for the next round of experiments.

Figure 2 shows the suite of sensors and analytical instruments which may be integrated in the Chemputer stack, including a colour sensor, pH probe, environmental sensor (pressure, humidity and temperature), conductivity sensor, RTD probe, liquid sensor, camera (for video capture and active failure detection), raman spectrometer, HPLC-DAD and Benchtop NMR. The sensors are connected to a *SensorHub*, which provides a unified interface for the range of sensors.

Figure 3 shows an overview of the framework for chemical discovery and optimization, ChemputationOptimizer, its system architecture and operation. The initial χ DL procedure (*proc_v0.xdl*) is obtained either via text translation from literature or upon algorithmic reaction discovery. Together with the hardware graph (*graph.json*) and a configuration file (*config.json*) are loaded, from which the optimizer framework extracts the parameters to be optimized. If enough resources are available, the optimizer performs the scheduling routine to allocate hardware resources to the corresponding procedure steps for parallel execution, minimizing the total duration. A locking mechanism which ensures error-free execution during runtime and eliminates the risk of unexpected cross-contamination. Upon successful execution, the outcome of the reaction is analyzed quantitatively and the results are fed to the optimization algorithm to suggest the next parameter set and update the initial procedure. The updated procedure, together with the results table is saved, using the database. The architecture and implementation details are given in the *Experimental* section of the present case.

Figure 4 (a) shows passively monitored formazine synthesis, a turbidity reference material, **(b)** shows dynamically executed thioether oxidation and **(c)** shows iodine-mediated nitrile formation with corresponding exemplary data captured within the process. **(a)** Example data was captured via passive monitoring during the formazine synthesis: The liquid detector tracked reagent addition, the colour sensor (RGBC) could detect the increase in turbidity as the colloidal suspension formed, and the environmental (BME) and internal reaction temperature (RTD) sensors could capture data relevant for reproducibility. **(b)** Internal reaction temperature was captured in real-time and used to control the addition of the oxidating agent in the highly exothermic thioether oxidation reaction. The original procedure

demanded to keep the internal reaction temperature below 75°C. After the addition, passive monitoring further revealed an uncontrolled exotherm upon heating the reaction mixture to the temperature (85°C) specified in the literature procedure. (c) A colour sensor captured data during the addition of iodine and its subsequent consumption as the reaction proceeds. The gradient was calculated post hoc, clearly showing the iodine addition and consumption as peaks.

Figure 5 shows results of a closed-loop reaction optimization. (a) Four-component Ugi reaction scheme (top); plot of the parameter space reduced to two dimensions using t-distributed stochastic neighbour embedding algorithm, where colour specifies the target parameter and the shape corresponds to the strategy used for the specific objective (bottom left); example of the ^{19}F NMR spectrum of the reaction mixture (bottom right). (b) Van Leusen Oxazole synthesis scheme (top); optimization results plot, *: *purity* is the relative area of the product peak divided by the sum of all peak areas, *desirability* is calculated as the weighted average between the two objectives ; example of the HPLC chromatogram (bottom right). (c) Manganese-catalysed epoxidation scheme (top); example plot of the reaction monitoring using Raman spectroscopy (bottom left); example Raman spectrum (bottom right). (d) Trifluoromethylation reaction scheme; plot of the parameter space reduced to two dimensions using t-distributed stochastic neighbour embedding algorithm, where colour specifies the target parameter and the shape corresponds to the strategy used for the specific objective (bottom left); examples of the ^{19}F NMR spectra where highlighted peak corresponds to the CF_3 group in the specified product (bottom right).

Figure 6 shows the further optimization examples. General reaction schemes for the a) tosMIC and d) phloroglucinol reactions with denoted parameters being optimized during the optimization cycle; corresponding parallel coordinate plots b) and d) where each vertical axis represent one of above mentioned parameters altered during each reaction with the last axis denoting the amount of product detected. Starting from the left parameters are connected with a line which represents a single reaction run. The reactions have been divided into three arbitrary groups, coloured in red, green and blue, for ease of interpretation and visualization, with respect to the normalized integration obtained for each reaction (amount of product). The coloured bold line within each group denotes a reaction of highest normalized integration; the orange dashed line denotes initial reaction parameters applied for the first reaction. Change of the integration observed during the optimization process for the e) tosMIC and f) phloroglucinol reaction with the second campaign denoted by a dashed line.

Figure 7 shows a plan photograph of an example SensorHub board.

Figure 8 shows a photograph of an example RGBC colour sensor attached to a three necked round bottom flask. The left neck of the flask includes a reaction temperature sensor. The right neck of the flask includes a

Figure 9 shows a photograph of a liquid sensor.

- Figure 10** shows a graph illustrating an example of the endpoint detection method using an RGB sensor. The top graph shows values obtained from the “clear” channel of the RGB sensor (filtered using median filter), with brown points showing values averaged over 300 seconds interval and green points indicating the detected endpoint time. The bottom graph shows the gradient values calculated from the sampled values, with aquamarine horizontal lines indicating the $\pm 1\%$ value of the absolute sensor reading and “no change” orange points falling within this interval.
- Figure 11** shows a schematic of the hardware setup for the Example optimization of the Four component Ugi reaction described in the examples.
- Figure 12** shows a schematic of the hardware setup for the Example optimization of the Van Leusen Oxazole synthesis reaction described in the examples.
- Figure 13** shows a schematic of the hardware setup used for the Raman-monitored catalytic oxidation reaction of the Sodium 4-vinylbenzenesulfonate (Styrene Sulfonate Oxidation Optimization) described in the examples.
- Figure 14** shows a schematic of the hardware setup used for the optimization of the tosMIC and phloroglucinol reaction described in the examples.
- Figure 15** shows a schematic of the hardware setup for the explorative optimization of the trifluoromethylation reaction described in the examples.
- Figure 16** shows an example of an HPLC chromatogram for the tosMIC reaction, obtained after performing optimization for the identified region as described in the examples.
- Figure 17** shows a ^{19}F NMR spectrum obtained after performing optimization of the example trifluoromethylation reaction, for the identified spectrum region. The product shown has not been isolated as pure material and its structure has not been fully confirmed.
- Figure 18** shows a ^{19}F NMR spectrum obtained after performing optimization of the example trifluoromethylation reaction, for the identified spectrum region. The product shown has not been isolated as pure material and its structure has not been fully confirmed.
- Figure 19** shows a ^{19}F NMR spectrum obtained after performing optimization of the example trifluoromethylation reaction, for the identified spectrum region. The product shown has not been isolated as pure material and its structure has not been fully confirmed.
- Figure 20** shows an example of the individual regions scoring calculated using Equation 1. Top: ^{19}F NMR spectrum with each identified region is highlighted with colour with top annotations indicating peak’s size (in number of data points) and score (see Equation 1).

Bottom: bar chart with areas for individual peaks (aligned on the x axis) with their harmonic mean indicated as orange horizontal line.

Figure 21 shows an example of the information score difference for two ^{19}F NMR spectra, calculated using equation 2.

Figure 22 shows an illustration of the novelty coefficient calculation. The spectrum of interested is plotted navy grey, set of individual regions for all previously measured spectrum is plotted as green circles (labelled as "Cumulative regions map") and novel regions are highlighted with orange crosses.

Detailed Description of the Invention

Generally, the invention relates to a method of optimising an instruction set for a chemical synthesis. The method may be performed on a chemical synthesiser. The method is preferably performed autonomously.

Optimisation refers to a change or a series of changes which results in an improvement. That is, a change which results in the achievement of a desired result.

In the context of a chemical synthesis, an optimisation is a change to the synthesis method which may improve a quality of the synthesis itself or a quality of the reaction product. Qualities of the synthesis and qualities of the reaction product may be known as characteristics of the synthesis. A user may determine which characteristics of the synthesis are to be improved.

For the reaction product, optimisation of a chemical synthesis may result in an increase in reaction product yield, reaction product purity, reaction product to starting material ratio, or reaction product to impurity ratio. For the reaction itself, optimisation may result in a reduction in solvent mass consumption, reagent mass consumption. It could also reduce the use of environmentally harmful reagents or solvents, or expensive reagents or solvents.

A chemical synthesis may be defined by an instruction set. The instruction set may be a versionable *executable code* which is capable of execution by a chemical synthesiser, having all the abstract explicit operations in a chemical programming language. The instruction set is an apparatus agnostic description of chemical operations, and the system interprets the instruction set to execute a synthesis on a chemical synthesiser.

Accordingly, in optimising a chemical synthesis the method of the present invention is a method of optimising an instruction set for a chemical synthesis.

In a general aspect, the present invention provides a method of optimising a chemical synthesis, the optimisation comprising the steps of:

- (i) performing a chemical synthesis in a reaction vessel of a chemical synthesiser operating according to an instruction set to produce a reaction product,
- (ii) recording reaction data using an analytical unit of the synthesiser during the chemical synthesis and/or recording product data on the reaction product,
- 5 (iii) analysing the reaction data and/or product data using a control unit of the chemical synthesiser to determine a characteristic of the synthesis; and
- (iv) adjusting one or more reaction parameters of the instruction set, to provide an adjusted instruction set, based on a prediction that the adjustment will improve the characteristic of the synthesis.

10

The work in the present case also exemplifies the optimisation of a chemical synthesis where the instruction set for the synthesis is automatically generated from literature procedures using an intelligent software system based on the open-source universal chemical programming language standard and format, χ DL. This open standard has been designed to

15 allow any chemical transformation to be precisely expressed and run on any compatible robotic platform. Thus, described herein is a method and chemical synthesiser which can generate all the components necessary to execute and optimise an automated synthesis directly from a literature procedure using a universal chemical description language.

20 The chemical synthesiser includes modules connected sequentially containing all the necessary hardware for the chemical operations to obtain the targeted product.

Some synthesis methods and synthesis apparatus have been described previously.

25 EP 4,129,462 describes a method for predicting reaction rate by performing a reaction under varying conditions, and using the results from these reactions to update a predicted reaction rate. Paragraph [0054] of EP 4,129,462 describes a 'reaction condition search step' which involves searching for reaction conditions that improve the predictability of the reaction rate, such that the rate constant converges during the search. The document does not describe an

30 optimisation where there are multiple reaction parameters to be optimised and does not discuss optimisation of dynamic reaction parameters. The document is also silent on the use of reaction data to inform the optimisation.

US 2020/0225251 describes an apparatus for reaction screening and optimisation, with the

35 apparatus having a number of reaction vessels to carry out reactions in parallel. Paragraphs [0061]-[0062] describe the use of feedback on the final product to adjust and optimise reaction conditions. The document does not describe the use of reaction data in the optimisation. There is also no disclosure of the optimisation of multiple reaction parameters or dynamic reaction parameters.

40

US 2022/0331772 describes a material synthesis apparatus, which may carry out a synthesis method calculated using a previously trained synthesis model. The previously trained synthesis model may be updated by receiving results of a synthesis as feedback from the

material synthesis apparatus (see paragraph [0043]). The examples describe a first synthesis, followed by a second synthesis – where results of the first synthesis are used to help determine commands for the second synthesis. Although US 2022/0331772 describes feedback based on the results of a synthesis, it does not describe how reaction data may be used to optimise a synthesis, or how dynamic reaction parameters may be optimised using the reaction data.

WO 2020/007962 describes a method of developing a predictive model for a set of reactions sharing a common element, such as a reagent or mechanism. The predictive model is constructed from a set of completed reactions, and the model is used to enable the exploration of chemical space. The method described in WO 2020/007962 does not optimise an instruction set for a particular chemical synthesis. The method is not iterative, as the predictive model is based on a set of completed reactions. The document does not describe a method where each repeat of the test synthesis is performed according to a newly adjusted instruction set, where the adjusted instruction set is made from a prediction based on reaction data and/or product data from the previous test syntheses.

Method of Optimising a Chemical Synthesis

In a first aspect of the invention there is provided a method of optimising an instruction set for a chemical synthesis, the method comprising performing a series of optimisation test syntheses, where each optimisation test synthesis includes the steps of:

- (i) performing the chemical synthesis in a reaction vessel of a chemical synthesiser operating according to an instruction set to produce a reaction product,
 - (ii) recording reaction data using an analytical unit of the synthesiser during the chemical synthesis and/or recording product data on the reaction product,
 - (iii) analysing the reaction data and/or product data using a control unit of the chemical synthesiser to determine a characteristic of the synthesis; and
 - (iv) adjusting one or more reaction parameters of the instruction set, to provide an adjusted instruction set,
- wherein the optimisation test synthesis is repeated one or more times, and in each repeat of the test synthesis, the chemical synthesis is performed according to the adjusted instruction set, which is adjusted from the previous test syntheses, and
- the adjustment to the one or more reaction parameters is made following a prediction that the adjustment will improve the characteristic of the synthesis, wherein the prediction is based on the reaction parameters associated with reaction data and/or product data from one or more previous optimisation test syntheses and, where present, one or more previous exploration test syntheses, and
- wherein the method is performed autonomously.

The methods of the invention are for optimising an instruction set of a chemical synthesis. That is, the instruction set is optimised to improve one or more characteristics of the synthesis.

The chemical synthesis is performed in a chemical synthesiser, such as an autonomous chemical synthesiser (e.g., a chemical robot). The chemical synthesis is typically performed on a chemical synthesiser, as described herein. The chemical synthesis may be mapped into a reactionware of a chemical synthesiser.

The chemical synthesis may be a single step synthesis. Alternatively, the chemical synthesis may be a multistep synthesis. The multistep synthesis may comprise two or more steps, such as a step of preparing an intermediate compound from one or more reagents, and the subsequent transformation of that intermediate compound into a reaction product. The multistep synthesis may comprise parallel steps, such as a step of preparing two intermediate compounds in parallel, and a step of reacting the intermediate compounds. The chemical synthesis may comprise two or more chemical reactions, such as three or more chemical reactions. The chemical synthesis may comprise a series of chemical reactions, as well as parallel chemical reactions, that are later brought to convergence for the production of a target product.

A downstream chemical reaction may be performed in a reaction vessel different to that of an upstream chemical reaction. Here, the chemical synthesis will require the transfer of material between reaction vessels as the synthesis progresses. The transfer may be achieved by any suitable material transfer method, such as fluidic transfer, and may be undertaken by a chemical synthesiser and may be autonomously controlled.

The chemical synthesis is not particularly limited and may encompass methods for the production of small organic compounds, metal complexes, supramolecular structures and polymers, amongst others. The chemical synthesis may also encompass biological reactions, or reactions where biomolecules participate as reagents or catalysts. Such biomolecules may include polypeptides, such as proteins, including enzymes and antibodies, polynucleotides, polysaccharides, metabolites and cofactors, amongst many others.

The worked examples in the present case exemplify the optimisation of reaction parameters for multicomponent³¹, heterocycle synthesis³², and catalytic reactions³³. The worked examples provided in the present case exemplify the optimisation of a ugi reaction (reaction of benzaldehyde, benzylamine, 2-fluorobenzoic acid and isocyanide), a Van Leusen oxazole reaction (reaction of 4-formylbenzonitrile and TosMIC), a styrene sulfonate oxidation (Mn-catalysed epoxidation of styrene sulfonate with hydrogen peroxide), a tosMIC reaction (reaction of tosMIC and benzylidenemalononitrile), a phloroglucinol reaction (reaction of phloroglucinol, benzylidenemalononitrile and 1,8-bis(dimethylamino)naphthalene) and a trifluoromethylation reaction (reaction of cyclohexanone, 4-fluoroacetophenone, butyrolactone and Ruppert-Prakash reagent).

The chemical synthesis may also include one or more steps of preparing a reaction mixture for a chemical reaction, as well as one or more methods for the work-up of a reaction mixture

following a chemical reaction, and one or more methods for purifying a product of a chemical reaction. The optimisation may focus on the chemical reaction itself, as well as the steps of preparing the reaction mixture, work-up and purification. For each chemical reaction in a chemical synthesis, the chemical reaction may include one or more work-up, purification or preparation steps.

The chemical synthesis may include a preparation step, the preparation step comprising preparing a chemical synthesiser for performance of a reaction, including the steps of generating an instruction set, generating a virtual reaction platform for execution of the instruction set, and generating the physical hardware (such as a physical reactionware) for the reaction according to the instruction set from the virtual platform.

The method comprises performing a series of test syntheses. A test synthesis is one repeat of the chemical synthesis according to an instruction set. Typically, the instruction set is adjusted after a synthesis such that each subsequent test synthesis is performed with a different instruction set. In this way, the present invention provides a method of iteratively optimising an instruction set for a chemical synthesis.

The test synthesis is repeated one or more times, and in each repeat of the test synthesis the chemical synthesis is performed according to an adjusted instruction set which is an adjustment of the instructions set from the previous chemical synthesis. The test synthesis is preferably repeated two or more times, five or more times, ten or more times, twenty or more times, with each synthesis preferably performed according to a different instruction set.

Preferably, the test synthesis is repeated until the instruction set is optimised.

In some embodiments, the method comprises performing a series of exploration test synthesis. This may be known as the exploration phase. Preferably the exploration phases comprises three or more, five or more, seven or more, nine or more, exploration test synthesis.

The exploration test syntheses explore a reaction space. A reaction space is a space defined by a range of possible values for each of the reaction parameters used in an instruction set. The range of possible values for each reaction parameters may be pre-determined (e.g., input by a user). In addition or alternatively the range of possible values may be determined by the chemical synthesiser.

In some embodiments, the exploration phase comprises performing a series of exploration test syntheses, where each exploration test synthesis includes the steps of:

- (i) performing the chemical synthesis in a reaction vessel of a chemical synthesiser operating according to an instruction set to produce a reaction product,
- (ii) recording reaction data using an analytical unit of the synthesiser during the chemical synthesis and/or recording product data on the reaction product,

(iii) analysing the reaction data and/or product data using a control unit of the chemical synthesiser to determine a characteristic of the synthesis; and

(iv) adjusting one or more reaction parameters of the instruction set, to provide an adjusted instruction set,

5 wherein the exploration test synthesis is repeated one or more times, and in each repeat of the test synthesis, the chemical synthesis is performed according to the adjusted instruction set, which is adjusted from the previous test syntheses, and

10 the adjustment to the one or more reaction parameters is made to explore a reaction space wherein the adjustment is based on the reaction parameters associated with reaction data and/or product data from one or more previous test syntheses.

Typically for the exploration test synthesis, the adjustment to the one or more reaction parameters is determined to explore a reaction space defined by the instruction set, wherein the determination is based on the comparison of the reaction parameters to reaction data and
15 product data from previous test syntheses.

The reaction space is the full extent of the reaction parameters appearing in the instruction set. Although each instruction set is a definite set of reaction parameters, the chemical synthesiser will receiver or determine a range over which each reaction parameter may be
20 varied. Thus, the exploration stage looks to test chemical syntheses across the full scope of the reaction parameters.

The adjustment to the one or more reaction parameters which is made to explore a reaction space is made by an exploration algorithm, such as a random search algorithm.

25 The method comprises one or more optimisation test syntheses. This may be known as the optimisation phase. Preferably the exploration phases comprises three or more, five or more, seven or more, nine or more, exploration test synthesis.

30 Typically, the method comprises one or more exploration test syntheses followed by one or more optimisation test syntheses. Preferably, the exploration phases comprises three or more, five or more, seven or more, nine or more, exploration test synthesis followed by one or more optimisation test syntheses.

35 Typically, the method comprises one or more exploration test syntheses followed by one or more optimisation test syntheses, such as three or more, five or more, seven or more, nine or more, optimisation test syntheses.

40 The exploration test synthesis and the optimisation test synthesis may be sequential or separated by other test syntheses.

For the optimisation test synthesis, the adjustment to the one or more reaction parameters is made following a prediction that the adjustment will improve a characteristic of the synthesis,

wherein based on the reaction parameters associated with reaction data and/or product data from one or more previous test syntheses. The previous test synthesis may be previous optimisation test syntheses and, where present, one or more previous exploration test syntheses.

5

In some embodiments, the prediction that the adjustment will improve a characteristic of the synthesis is based on reaction data, and optionally product data, from the one or more previous test syntheses. The previous test synthesis may be one or more previous optimisation test syntheses and, where present, one or more previous exploration test

10

syntheses. The use of reaction data to optimise the instructions set enables optimisation of dynamic reaction parameters.

In preferred embodiments, the prediction that the adjustment will improve the characteristic of the synthesis is based on reaction data and product data from the one or more exploration test syntheses, and optionally any previous optimisation test syntheses.

15

In some embodiments, the prediction that the adjustment will improve the characteristic of the synthesis is based on an optimisation algorithm, such as a genetic or SMBO optimisation algorithm.

20

In preferred embodiments the comparison of the reaction parameters to reaction data and product data from previous test syntheses identifies reaction parameters associated with a change in reaction data and/or product data, and preferentially adjusts reaction parameters associated with the change.

25

In some such embodiments, the comparison of the reaction parameters to reaction data and product data from previous test syntheses identifies reaction parameters associated with the highest rate of change in reaction data and/or product data, and preferentially adjusts reaction parameters associated with the highest rate of change.

30

The method may further comprise a random test synthesis, using random reaction parameters. The random test synthesis is used to provide reaction data and product data which can be used during subsequent tests to adjust the reaction parameters.

35

The method may also comprise test synthesis wherein the adjustment is configured to provide a prediction between an exploration synthesis test and an optimisation synthesis test. The step may therefore be configured with an algorithm operating between an that used for an exploration test and an optimisation test.

Performing the Chemical Synthesis

40

The method comprising performing a series of test syntheses. Each test synthesis comprises the step of performing the chemical synthesis according to an instruction set in a chemical

synthesiser to produce a reaction product. The instruction set comprises one or more, and typically a plurality of, reaction parameters.

5 An instruction set is an executable description of a method for obtaining the reaction product by chemical synthesis using a chemical synthesiser, such as an automated chemical synthesiser. The chemical synthesis may be defined by the instruction set. Thus, the instruction set provides apparatus agnostic description of chemical operations needed to perform a chemical synthesis. The reaction parameters of the instruction set define the reagents, conditions, and processing information that makes up the chemical synthesis. The instruction
10 set is a typically a machine readable instruction set, coded using xDL.

In some embodiments, the method further comprises obtaining an initial instruction set for the chemical synthesis, the instruction set comprising reaction parameters. The chemical synthesis may comprise one or more consecutive chemical reaction steps. Each chemical
15 reaction step may have an instruction set, or the whole chemical synthesis may have an instruction set. The initial instruction set may be a literature method, converted into a machine-readable instruction set (as described in WO 2021/219999).

The instruction set is an instruction set for the chemical synthesis that can be adjusted to
20 provide an adjusted instruction set. Over a series of test syntheses, the instructions set is optimised to provide an optimised instruction set. The optimised instruction set is typically the instruction set where the desired characteristic of the synthesis (e.g., yield, purity, conversion) is improved.

25 The method may result in an optimised instruction set. The optimised instructions set may be determined when further adjustment to the reaction parameters does not provide an improvement to the characteristic of the synthesis. The instruction sets, such as the optimised instructions set, may be saved in a database.

30 Reaction Parameters

The instruction set comprises reaction parameters. The reaction parameters together provide the information needed to perform a chemical synthesis.

35 Some reaction parameters may be fixed reaction parameters, in other words, reaction parameters which are not adjustable and so are not changeable between versions. Some reaction parameters may be variable reaction parameters, in other words, reaction parameters which are adjustable and so are changeable between versions, in order to optimise the chemical synthesis.

40 The instruction set typically comprises fixed reaction parameters and variable reaction parameters, preferably wherein the variable reaction parameters are variable between an upper and lower limit.

The upper limit and lower limit is typically the range within which the reaction parameter can be adjusted to during the adjustment step. The upper limit and lower limit may be predetermined and input by an operator, or from the literature synthesis. Alternatively, the variable range may be automatically determined based on the initial instruction set. For example, the upper limit and lower limit may be taken as +/- 30% of the reaction parameter provided in the initial instruction set. The upper limit and lower limit may also be limited by the boiling and freezing points of the fluid chemical inputs supplied to the reaction space, and the fluid product output.

10

In some embodiments, each of the reaction parameters in the instruction set is a variable reaction parameter. Preferably, all of the reaction parameters in the instructions set are variable reaction parameters.

15 Reaction parameters may also be referred to reaction conditions (such as chemical or physical conditions) or reaction inputs (such as chemical or physical inputs).

The reaction parameter may be a point value or range of values.

20 The reaction parameter may include physical parameters, such as heating, cooling, light, and mixing.

Physical parameters are intended to refer to a parameter that is not a material such as a reagent, catalyst, solvent, or a component. A physical parameter may refer to, for example, a temperature, such as the temperature of a particular chemical input, or the temperature of the reaction mixture, or the temperature of the reaction headspace, or the temperature of the surroundings. A modulation in temperature may refer to a physical parameter that can raise and/or lower temperature. A series of temperature parameters may be provided that is a gradient of temperature increase and/or decreases. The range of temperature parameters may be limited by the boiling and freezing points of the fluid chemical inputs supplied to the reaction space, and the fluid product output. It is noted, however, that the reaction space may be suitably pressurised thereby to effectively alter the boiling and freezing points of the fluid chemical parameters. In this way a greater range of temperature parameters may be used.

35 Temperature parameter may be used to initiate reagents or favour certain reaction pathways. Temperature parameter may also be used to investigate the stabilities of the reagents and products.

The physical parameter may be light. A series of light parameters may be provided that differ in one or more of intensity, wavelength, exposure time and spectrum. Light parameters may be used to initiate reagents or may be used to favour or alter certain reaction pathways. Light inputs may include UV-vis inputs. The physical parameter may be microwave radiation.

40

The physical parameter may be ultrasound. Such may be useful for the generation of reagents or products. Ultrasound may also aid the dissolution of material.

5 The physical parameter may be pressure. Pressure changes may be used to alter, for example, solvent boiling points.

10 The physical parameter to the system may be a process related parameter for the reaction mixture. The physical parameter may be a stirring rate of the reaction mixture. Thus, the parameter may be a time limited feature for reaction or admixture. After a set time, the reaction mixture may be analysed and the product quantified. Thus, reaction time may be a parameter. Similarly, other process features such as concentration and ratio of chemical parameters, such as the reagent and catalyst chemical inputs, may be a physical parameter.

15 The reaction parameter may include chemical parameters, such as the addition or removal of reagents, solvents and catalysts, or the rate of addition or removal of reagents, solvents and catalysts. Typically, the reference to a chemical parameter is a broad reference to the addition or removal of material, which may be a reagent, catalyst, solvent, or a component, that may allow the preparation of a reaction product.

20 The chemical parameter may specify if a chemical material is provided as a solid, or in a fluid for transfer to the reaction vessel.

Where the material is a fluid, it may be supplied in this form to the reaction space. Alternatively, the material may be diluted, dissolved or suspended in a fluid for delivery to the reaction space. Thus the material may be in solution or suspension. The fluid that dissolves or suspends the material is not particularly limited, and may be water or an organic solvent, for example. The fluid may be independently deliverable to the reaction space. The fluid is also used to provide separation between individual combinations of chemical inputs that are supplied to the reaction space thereby preventing contamination between different combinations.

25
30

35 The identity of the material will be dependent upon the reaction and formulation steps that are to be employed, and may be limited – though not necessarily – by the intended chemical synthesis.

A chemical material may be a reagent. A range of reagents may be provided that differ in their structure and functionality.

40 A chemical material may be a catalyst. A range of catalysts may be provided that differ in their activity, selectivity, or morphology.

A chemical material may be an acid or a base. A range of different acids and bases may be provided, where the acidity differs. Organic and inorganic acids and bases may be selected.

Weak and strong acids and bases may be provided. A buffer may be used to maintain the reaction mixture at a pH range.

5 A chemical material may be a solvent. Organic solvents and water may be used. A range of non-polar, protic and aprotic solvents may be provided. In one embodiment, water is provided as a chemical input.

10 A chemical material may be a salt. A range of different salt forms of a particular component may be used. A range of organic and inorganic salts may be provided.

15 A chemical material may also be a gas. In some embodiments, a chemical material may be an inert gas, such as nitrogen or argon, to supply to the reaction space. In other embodiments, the chemical input is a reaction gas, such as hydrogen, oxygen or carbon dioxide.

20 The reaction parameter may specify that the chemical material is only provided for part of the chemical synthesis. For example, the reaction parameter may specify that the chemical material is only added for the work up of a reaction product, or only for quenching a reaction. Such parameters may specify that the material is provided to the reaction space at some time period after the other materials have been combined, or after a reaction data provides a certain output, thereby to quench a reaction or to permit the work up and possible isolation of reaction product.

25 The concentration of a material within a solution or in a suspension may be selected appropriately by the chemical synthesiser. The effective concentration of the material in the reaction space will depend on the concentration of that material within its individual chemical flow and the volume of other chemical materials with which it is combined in the reaction space. These volumes are dictated by the flow rates of each of the materials, which may be varied as appropriate, to alter the effective concentration of a material in the reaction space.
30 Such techniques will be familiar to those with an understanding of flow chemistry techniques.

Where appropriate, a chemical material may be stored under an inert atmosphere, may be stored under anhydrous conditions or may be stored at reduced temperature, as required.

35 In one embodiment, there is provided 2 or more, 4 or more, 6 or more, or 10 or more chemical materials. For example, the flow chemistry system may comprise a number of controllable syringes equal to the number of specified chemical materials.

40 From time to time it may be necessary to replenish a chemical material. The method of the invention need not be halted to allow such replenishment, and the chemical material may be replenished at such a time as it is not required during the chemical synthesis. The control system may be suitably programmed to predict the time at which a chemical material will become depleted. An operator may be warned accordingly. The control system may also be

suitably programmed to factor in to the decision making and control process the unavailability of an input owing to replenishment. The control system can continue to produce products using inputs other than the input that is being replenished.

- 5 The number of chemical reaction parameters may be one, though in this embodiment the number of physical reaction parameters, which may bring about a change in the chemical material, will typically be larger than the number of chemical reaction parameters.

10 In one embodiment, the instruction set may comprise two or more, three or more, four or more, five or more, six or more, ten or more, twenty or more chemical reaction parameters.

In one embodiment, the instruction set may comprise two or more, three or more, four or more, five or more, six or more, ten or more, twenty or more physical reaction parameters.

- 15 In preferred embodiments, the reaction parameter may be a reagent volume, reagent mass, reagent addition rate, solvent volume, solvent addition rate, temperature, rate of change of temperature, stirring rate, reaction time, or a combination thereof.

20 In some embodiments, each of the reaction parameters in the instruction set are able to be adjusted, to provide an adjusted instruction set. Preferably, each of the reaction parameters in the instruction set are adjusted, to provide an adjusted instruction set.

25 In some embodiments, each of the reaction parameters in the instruction set are able to be optimised. Preferably, each of the reaction parameters in the instruction set are optimised.

Dynamic reaction parameters

30 The reaction parameter may be monitored by analysis of reaction data which is recorded during the chemical synthesis. Where the reaction parameter is monitored using reaction data, and action taken for conformity with a different reaction parameter, this is known as a dynamic reaction parameter. A dynamic reaction parameter is a reaction parameter which is adjusted during the chemical synthesis in response to reaction data obtained during the chemical synthesis. The dynamic reaction parameter is adjusted 'live' during the chemical synthesis, using feedback from the reaction data.

35 For example, if the reaction parameter is reaction mixture temperature, the temperature may be recorded as a reaction data to determine if the reaction parameter is fulfilled. Moreover, the adjusting the heating/cooling of the reaction mixture may be used to ensure the reaction parameter is conformed with.

40 In another example, if the reaction parameter is reaction mixture pH, the pH may be recorded as a reaction data to determine if the reaction parameter is fulfilled. Adjusting the addition of acid or base of the reaction mixture may be used to ensure the reaction parameter is obeyed.

The chemical synthesiser may determine if the reaction parameter is met by adjusting how the reaction is carried out. For example, if the reaction parameter is solvent volume, the chemical synthesiser may adjust the amount of solvent added during the reaction to meet the reaction parameter condition.

Such steps may be intended to bring the course of a synthesis into line a different reaction parameter synthesis, for example such that the reaction profile is brought into accordance with the reference profile for the subsequent parts of the chemical synthesis.

Preferably, one or more of the reaction parameters is a dynamic reaction parameter, wherein the dynamic reaction parameter is adjusted during the test synthesis such that reaction data satisfies a different reaction parameter.

In some embodiments, the dynamic reaction parameter is dynamic reagent addition rate, dynamic reaction time or dynamic reaction temperature. In some such embodiments, the reaction data which satisfies a different reaction parameter is temperature, pressure, colour or pH, preferably temperature or colour.

The dynamic reaction parameter can typically be directly changed by the chemical synthesiser. For example, the amount of reagent added, rate of reaction addition, time or heating/cooling. The dynamic reaction parameter is also used to affect a different reaction parameter. The different parameter is a reaction parameter which may be adjusted as a result of a change in a direct reaction parameter. For example, colour, pH, temperature.

Some reaction parameters can be both direct and indirect. If the reaction parameter is direct or indirect may depend on the nature of the reaction. For example, temperature can be directly controlled by increasing or decreasing the temperature of a hot plate. However, temperature can also be controlled by the rate or reagent addition to an exothermic / endothermic reaction.

A dynamic reaction parameter is maintained if the reaction data indicates that the reaction parameter is fulfilled, or the dynamic reaction parameter is adjusted if the reaction data indicates that the reaction parameter is not fulfilled, and the adjustment is such that after the adjustment the reaction data fulfils the reaction parameter.

In some embodiments, the dynamic reaction parameter is dynamic reagent addition rate. The dynamic reagent addition rate refers to a reagent addition rate that may be adjusted during the chemical synthesis in response to reaction data, such as reaction mixture temperature or reaction mixture pressure. The reaction data is subject to a different reaction parameter. For example, the reaction temperature may be subject to a temperature reaction parameter, which requires that the temperature is within a value or range. The dynamic reagent addition rate is adjusted such that the temperature satisfies the temperature reaction parameter.

In some embodiments, the dynamic reaction parameter is dynamic time. The dynamic time may refer to reaction time, that is adjusted during the chemical synthesis in response to reaction data, such as reaction mixture colour or reaction mixture pH. For example, the
5 reaction mixture colour may indicate if the reaction is ongoing or has completed. The colour therefore may be subject to a different reaction parameter (e.g., that the reaction colour is red). The dynamic time may be adjusted when the reaction colour changes to satisfy the reaction parameter (e.g., the reaction mixture turns red). Thus, the reaction time is determined by the colour change.

10

In some embodiments, the dynamic reaction parameter is dynamic temperature.

The dynamic reaction parameter is typically implemented autonomously, using feedback from the reaction data. Typically the reaction data is measured by the analytical unit and reported
15 to the control unit, and the control unit adjusts the dynamic parameter to satisfy the different reaction parameter.

In additional embodiments the method comprises adjusting two or more reaction parameters of the instruction set, such as three or more, four or more, five or more, six or more, ten or
20 more, twenty or more reaction parameters.

In some embodiments the method comprises adjusting each reaction parameter of the instruction set.

Recording Reaction Data and Product Data

25 Each test synthesis comprising a step of recording reaction data during the chemical synthesis and recording product data on the reaction product.

Reaction Data

30 The reaction data is recorded during the chemical synthesis. The reaction data refers to data obtainable from the reaction mixture, reaction environment or surroundings. This refers to each test synthesis performed. The reaction data is recorded through all or some steps of the chemical synthesis.

35 The chemical synthesiser of the invention is provided with an recording reaction data during the chemical synthesis. The analytical unit then reports the reaction data to the control unit.

Reaction mixture may refer to the reaction mixture itself or the reaction environment.

40 Reaction environment may refer to the reaction mixture as held in a reaction vessel, and may include the headspace over the reaction mixture, where such is present. The reaction data may include data on the surroundings during the reaction. Surroundings refers to outside of

the reaction vessel, for example the environment surrounding the chemical synthesiser (e.g., a fume hood).

The reaction data is data generated throughout the chemical synthesis. Thus, reaction data may be generated for each chemical reaction in the chemical synthesis. Additionally, the reaction data may be generated during the reaction itself, but may also be generated during the initial setting of the reaction mixture, and also the work-up of a reaction, including filtering or purification.

5 The reaction data is typically used in the step of adjusting one or more reaction parameters, as described herein. The use of reaction data in predicting how adjustments to the reaction parameter will improve the characteristic of the synthesis provides for more efficient and more effective optimisation of the chemical synthesis.

10 The reaction data may also be used in dynamic reaction steps, as described above.

The reaction data may be recorded at a set frequency, a variable frequency or continuously. The reaction data is preferably recorded at a high frequency, such as at time intervals of 1 second or less, 0.1 seconds or less, 0.01 seconds or less.

20

In some embodiments, one more than one, such as two or three, different reaction data are recorded simultaneously.

Reaction data may include physical or chemical properties. The reaction data may include reaction mixture colour, reaction mixture temperature, reaction mixture pressure, reaction mixture pH, reaction mixture filter rate, surrounding humidity, surrounding pressure, surrounding temperature, or a combination thereof, preferably reaction mixture colour, reaction mixture temperature or a combination thereof.

25

30 In some embodiments, the recording of reaction data changes depending on the reaction parameters. For example, where the reaction parameter includes the addition of a coloured reagent, the reaction data may include reaction mixture colour. However, where the reaction parameter includes the addition of only colourless reagents, the reaction data may not include reaction mixture colour.

35

In some embodiments, the recording of reaction data changes throughout the chemical synthesis. For example, where the chemical synthesis includes a heating step, the reaction data may include reaction mixture temperature. However, if the final work-up is carried out at ambient temperature, the reaction data may not include reaction temperature.

40

In some embodiments, the change to the recording of reaction data comprises a change to the type of reaction data recorded and/or a change to the frequency that reaction data is recorded. For example, where the chemical synthesis includes a reaction parameter for the

addition of a reagent causing an exothermic reaction, the reaction data may include reaction mixture temperature recorded at a high frequency. However, after the addition of the reagent is complete, the frequency that reaction data is recorded may be reduced.

- 5 By changing the type or frequency of the reaction data recording, the amount of reaction data recorded can be reduced. This may reduce noise in the reaction data, enabling the adjustment step to be carried out more accurately. In addition, or alternatively, the reaction data may be filtered to remove reaction data, which is not relevant for the chemical synthesis, in the same way as described above.

10

Product Data

- The product data is typically analytical data on the reaction product. The product data is preferably based on the final reaction product. The product data is typically recorded at the end of the chemical synthesis, such as at the end of each test synthesis.

15

The chemical synthesiser of the invention is provided with an analytical unit for analysing the products and recording the product data. The analytical unit then reports the product data to the control unit.

20

The product data may be spectroscopic data, such as UV-Vis, Infra-red, Raman, NMR, or mass spectrometry data; chromatography data, such as gas or high-performance liquid chromatography data, or a combination thereof. The product data may be measured using any suitable measurement method, such as described in the examples section.

25

Any suitable form of NMR spectroscopy may be used, such as ^{19}F NMR, ^1H NMR, ^{13}C NMR, ^{31}P NMR. Preferably the NMR is ^{19}F NMR.

Preferably, the product data uses an efficient *in-situ* and on-line analytical techniques, such as UV-Vis, Raman, IR or benchtop NMR.

30

Preferably, the product data is obtained using a passive or non-destructive technique. Thus, the sample may be tested without requiring any physical or chemical degradation of the reaction product. In this way a sample may be tested by many different methods, if needed.

- 35 Product data may be recorded continuously, for example, where the property is a physical or chemical property that can be measured in this way. Analytical data may also be recorded at high frequency, under circumstances where it is not practicable to record continuously, or where it is not useful to do so.

- 40 In some embodiments, the product data recorded is changed depending on the reaction parameters.

Analysing Data

Each test synthesis comprising a step of analysing the reaction data and/or product data to determine a characteristic of the synthesis. Preferably the product data is analysed to determine a characteristic of the synthesis. In some embodiments, reaction data, and
5 optionally product data, are analysed to determine a characteristic of the synthesis.

The chemical synthesiser of the invention is provided with control unit for analysing the reaction data and/or the product received from the analytical unit to determine a characteristic of the synthesis.
10

In some embodiments, the characteristic of the synthesis is product yield, residual starting material amount, product purity, impurity amount, side-product amount, preferably product yield or residual starting material amount, more preferably product yield.

15 Preferably, the characteristic of the synthesis is reaction product yield, reaction product purity, reaction product to starting material ratio, reaction product to impurity ratio, or a combination thereof.

The characteristic of the synthesis may be determined from one or more product data. In
20 some embodiments, the characteristic of the synthesis is yield, and the yield is determined by NMR, Raman or HPLC product data. In some embodiments, the characteristic of the synthesis is residual starting material amount, and the residual starting material amount is determined by NMR, Raman or HPLC product data.

25 In some embodiments, the characteristic of the synthesis may be a novelty score for an individual region calculated as the size of the region (as number of data points) multiplied by the inverse logarithm of the difference between area of the current region and harmonic mean of all identified regions (Equation 1).

$$30 \quad Rs_i = Size_i \times \frac{1}{\log_{10}(|area_i - \overline{area}|)} \quad (1)$$

If one of the region's area is equal to overall harmonic mean, that this region's score is equals to its size. Harmonic mean is used to favour smaller peaks and reduce the weight of large spikes and potential false positives (i.e., noise identified as material peaks).
35

In some embodiments, the characteristic of the synthesis may be a novelty coefficient of the reaction product. The novelty coefficient is calculated with respect to all previously measured spectra (e.g., NMR spectra, HPLC spectra), for which identified regions are stored as arrays with their x axis coordinates. All identified regions for the spectrum of interest are checked
40 versus previously identified regions (spectrum-wise) and the spectrum with regions not found previously will have a higher coefficient (Equation 3).

$$Nc_i = \frac{|F_i \setminus P|}{|F_i|} + \frac{1}{|P|}, \text{ where } P = \bigcup_{j=1}^{k \neq i} F_j \quad (3)$$

F_i – set of identified regions for i^{th} spectrum

- Where the characteristic of the synthesis is yield, and the yield is determined by NMR, Raman or HPLC product data, the improvement in the characteristic of the synthesis is maximising the product peak in the NMR, Raman or HPLC, optionally wherein the improvement in the characteristic of the synthesis is maximising the product peak relative to a reference peak.

- Where the characteristic of the synthesis is residual starting material amount, and the residual starting material amount is determined by NMR, Raman or HPLC product data and the improvement in the characteristic of the synthesis is to minimise the starting material peak in the NMR, Raman or HPLC, optionally wherein the improvement in the reaction product is to minimise the starting material peak relative to a reference peak.

Adjusting Instruction Set

- In the method of the invention, the one or more of the test syntheses is an optimisation test synthesis, and wherein the adjustment to the one or more reaction parameters is determined by a prediction that the adjustment will improve the characteristic of the synthesis, wherein the prediction is based on the comparison of the reaction parameters to reaction data and product data from previous test syntheses.

- The chemical synthesiser of the invention is provided with control unit for adjusting the one or more reaction parameters, and determining the adjustment based on a prediction that the adjustment will improve the characteristic of the synthesis, wherein the prediction is based on the comparison of the reaction parameters to reaction data and product data from previous test syntheses.

- The adjustment is typically an increase or decrease to one of the variable reaction parameters. For example, an increase or decrease to the reaction mixture temperature, an increase or decrease to the reaction time, or an increase or decrease to the reagent addition rate.

- The adjustment is made based on a prediction that the adjustment will improve the characteristic of the synthesis. The prediction is based on the comparison of the reaction parameters to reaction data and product data from previous test syntheses.

- In some embodiments, the prediction uses reaction data and product data recorded from previous exploration iterations of the instructions set.

In some embodiments, the prediction uses the reaction data and product data in optimising the instruction set, uses reaction data and/or product data recorded during the chemical synthesis carried out in previous optimisation iterations of the instructions set. Preferably the prediction uses reaction data and product data recorded during the chemical synthesis
5 carried out in previous exploration and optimisation iterations of the instructions set.

In some embodiments, the comparison of the reaction parameters to reaction data and product data from previous test syntheses identifies reaction parameters associated with a change in reaction data and/or product data, and preferentially adjusts reaction parameters
10 associated with the change.

In some embodiments, the comparison of the reaction parameters to reaction data and product data from previous test syntheses identifies reaction parameters associated with the highest rate of change in reaction data and/or product data, and preferentially adjusts reaction
15 parameters associated with the highest rate of change.

In some embodiments, the comparison of the reaction parameters to reaction data and product data from previous test syntheses identifies dynamic reaction parameters associated with a change in reaction data and/or product data, and preferentially adjusts dynamic
20 reaction parameters associated with the change.

In some embodiments, the comparison of the reaction parameters to reaction data and product data from previous test syntheses identifies dynamic reaction parameters associated with the highest rate of change in reaction data and/or product data, and preferentially adjusts
25 dynamic reaction parameters associated with the highest rate of change.

Preferably, analysing the reaction data and product data to suggest an adjustment to one or more reaction parameters is carried out by an optimisation algorithm. In some embodiments, the optimisation algorithm is a random search, genetic or SMBO algorithm.
30

The optimisation stage uses an algorithm with hyperparameters set to favour optimisation over exploration. The exploration stage uses an algorithm with hyperparameters set to favour exploration over optimisation.

35 The optimisation algorithm is typically a machine learning algorithm. Any suitable machine learning algorithm may be used, which is capable of optimizing the desired characteristic of the synthesis by adjusting reaction parameters. The optimisation algorithm may use manual search, grid search, random search, genetic model or bayesian model-based optimization (e.g., SMBO). Preferably, the optimisation algorithm is a random search, genetic model or
40 SMBO algorithm.

Where the algorithm is a random search algorithm, the values of the reaction parameters are determined from a random number array (sampled from a uniform distribution) within the parameters constraints.

- 5 Where the algorithm is a genetic model algorithm, it may uses truncation selection, single point crossover, and random reset mutation while preserving the best solution (elitism). In case of premature convergence, the population is typically reinitialized. The genetic algorithm may include hyperparameters selected from population size (e.g., number of individuals in the population) and probability of change (e.g., probability of mutating a gene).

10

Where the algorithm is SMBO, the algorithm may be implemented using the `scikit-optimize` python library, as described in the examples.

- 15 A different algorithm may be used for different test synthesis. For example, a different algorithm may be used for an exploration phase to an optimisation (or exploitation) phase.

In a specific embodiment, an exploration phase may use a random search algorithm, and an optimisation phase may use a SMBO algorithm.

- 20 In a specific embodiment, five random experiments may be used to initialize the parameter search space, ten experiments may be used to explore this space (i.e. minimize the uncertainty), five experiments with balanced exploration-exploitation approach and six experiments to optimise the reaction space.

25 *Chemical Synthesiser*

In a second aspect of the invention there is provided a chemical synthesiser for use in the method of optimising an instruction set for a chemical synthesis of the first aspect, the synthesiser comprising a reaction vessel, an analytical unit and a control unit, wherein:

- 30 the reaction vessel is for the performance of the chemical synthesis according to an instruction set,

the analytical unit is for recording reaction data from the reaction vessel during the chemical synthesis and/or for recording product data on the reaction product, and for reporting the reaction data and/or product data to the control unit;

- 35 the control unit is for analysing the reaction data and/or product data received from the analytical unit to determine a characteristic of the synthesis, and for adjusting one or more reaction parameters of the instruction set, to provide an adjusted instruction set, wherein the adjustment to the one or more reaction parameters is made following a prediction that the adjustment will improve the characteristic of the synthesis, wherein the prediction is based on the reaction parameters associated with reaction data and/or product data from one or more
40 previous optimisation test syntheses, and

wherein the chemical synthesiser is for performing the method autonomously.

The present invention also provides for the chemical synthesiser, or parts of the chemical synthesiser, to operate autonomously. Thus, the chemical synthesiser may include an autonomous chemical synthesiser, suitably programmed to carry out the chemical synthesis according to the instruction set. The chemical synthesiser may include an autonomous analytical unit, suitably programmed to record reaction data and product data and report the data to a control unit. The chemical synthesiser may include an autonomous control unit, suitably programmed to analyse the reaction data and product data to determine a characteristic of the synthesis, and suggest an adjustment to one or more reaction parameters which is predicted to improve the characteristic of the synthesis, and providing an iterative instruction set to the chemical synthesiser comprising one or more adjusted reaction parameters.

Reaction Vessel

The chemical synthesiser comprises a reaction vessel for the performance of the chemical synthesis according to an instruction set.

The chemical synthesiser is optionally also provided with one or more modules for work-up and purification, and further comprising reagents, optionally together with solvents and catalysts, for use in the chemical synthesis.

The chemical synthesiser may comprise a reactionware cartridge, for the performance of the chemical synthesis according to the instruction set. The reactionware may include multiple reaction vessels which are fluidically connected. The reactionware cartridge may be a reusable 3D printed 'module-monolith' reactionware cartridge. The reactionware cartridge is optionally also provided with one or more modules for work-up and purification, and further comprising reagents, optionally together with solvents and catalysts, for use in the chemical synthesis. The reactionware cartridge may be automatically generated from literature procedures using an intelligent software system based on the open-source universal chemical programming language standard and format, χ DL.

The chemical synthesiser can be equipped with the components suitable for performing a chemical synthesis, preferably an automated chemical synthesis.

For example, the chemical synthesis may comprise a reagent and solvent handling system, such as a fluid handling system, a reaction vessel, a reagent/solvent supply system and waste system.

The chemical synthesiser typically comprises a fluid handling system. The fluid handling system may have any suitable form, and typically includes a network of pipes, valves, and pumps (e.g., syringe pumps) fluidically connected to the reaction vessel. The fluid handling system preferably includes a flow sensor, such as an optical flow sensor to measure the rate of flow of fluid into the reaction vessel.

- The chemical synthesiser typically comprises a reaction vessel. The reaction vessel typically includes a reactor (e.g., a glass flask, flow cell, etc.), an agitator (e.g., sonicator or stirrer), a heater (e.g., a hot plate) and a cooler (e.g., a chiller).
- 5 Chemical synthesiser may also include a purification module, such as a column chromatography module, a filtration module, a washing module. The modules are suitable for performing their respective task.
- 10 The chemical synthesiser is preferably a robotic chemical synthesiser for autonomous performance of the chemical synthesis, and for autonomous control of the analytical device.
- The function of the chemical synthesiser is as described in the performing chemical synthesis section, described above.
- 15 *Analytical Unit*
- The chemical synthesiser comprises an analytical unit is for recording reaction data from the reaction vessel during the chemical synthesis and for recording product data on the reaction product, and for reporting the reaction data and product data to the control unit.
- 20 The analytical unit comprises analytical devices. The analytical device include sensors for measuring reaction data and product data.
- Typically, the analytical unit comprises a reaction analysis unit for recording reaction data during the chemical synthesis and a product analysis unit for recording product data on the
- 25 reaction product.
- The analytical unit typically comprises a hub (referred to as a 'sensor hub') for centrally connecting the sensors (e.g., the sensors of the reaction analysis unit and the product analysis unit).
- 30 In some embodiments, the reaction analysis unit comprises a colour sensor, a temperature sensor, a pressure sensor, a pH sensor, a flow rate sensor, a humidity sensor or a combination thereof.
- 35 The temperature sensor, a pressure sensor, a humidity sensor may be a reaction mixture sensors (sensors configured to measure the reaction mixture) or surrounding sensors (sensor configured to measure the surroundings). For example, the reaction analysis unit comprises a reaction mixture colour sensor, a reaction mixture temperature sensor, a reaction mixture pressure sensor, a reaction mixture pH sensor, a reaction mixture flow rate sensor, a reaction
- 40 mixture humidity sensor or a combination thereof. For example, the reaction analysis unit comprises a surrounding temperature sensor, a surrounding humidity sensor, a surrounding pressure sensor, or a combination thereof.

Preferably, the reaction analysis unit comprises a reaction mixture colour sensor, a reaction mixture temperature sensor or a combination thereof.

- 5 In some embodiments, the product analysis unit comprises an infra-red spectrometer, a Raman spectrometer, an NMR spectrometer, a mass spectrometer, a gas chromatography instrument, a high-performance liquid chromatography instrument, or a combination thereof.

10 The function of the analytical unit is as described in the recording reaction data and product data steps, described above.

Control Unit

- The chemical synthesiser comprises a control unit for analysing the reaction data and/or product data received from the analytical unit to determine a characteristic of the synthesis, and for adjusting one or more reaction parameters of the instruction set, to provide an
- 15 adjusted instruction set, wherein the adjustment to the one or more reaction parameters is made following a prediction that the adjustment will improve the characteristic of the synthesis, wherein the prediction is based on the reaction parameters associated with reaction data and/or product data from one or more previous optimisation test syntheses, and, where present, one or more previous exploration test syntheses.

20 In some embodiments the control unit is for adjusting two or more reaction parameters of the instruction set, such as three or more, four or more, five or more, six or more, ten or more, twenty or more reaction parameters.

- 25 In some embodiments the control unit is for adjusting each reaction parameter of the instruction set.

The instruction sets are typically stored in the control unit.

- 30 In some embodiments the control unit is for analysing the reaction data and product data received from the analytical unit to determine a characteristic of the synthesis, wherein the characteristic of the synthesis is product yield, residual starting material amount, product purity, impurity amount, side-product amount, preferably product yield or residual starting material amount, more preferably product yield.

35 Preferably, the control unit receives reaction data during the chemical synthesis, and the control unit dynamically adjusts one or more dynamic reaction parameters during the chemical synthesis in response to the reaction data. Preferably, the control unit adjusts the dynamic reaction parameter such that the reaction data fulfils a different reaction parameter.

- 40 The function of the control unit is as described in the analysing data and adjusting instruction set steps, described above.

Other Preferences

Each and every compatible combination of the embodiments described above is explicitly disclosed herein, as if each and every combination was individually and explicitly recited.

- 5 Various further aspects and embodiments of the present invention will be apparent to those skilled in the art in view of the present disclosure.

“and/or” where used herein is to be taken as specific disclosure of each of the two specified features or components with or without the other. For example “A and/or B” is to be taken as
10 specific disclosure of each of (i) A, (ii) B and (iii) A and B, just as if each is set out individually herein.

Unless context dictates otherwise, the descriptions and definitions of the features set out above are not limited to any particular aspect or embodiment of the invention and apply
15 equally to all aspects and embodiments which are described.

Certain aspects and embodiments of the invention will now be illustrated by way of example and with reference to the figures described above.

Comments on Examples

- 20 The Optimiser framework is built on top of the existing abstraction of chemputation, which is encoded using the χ DL programming language²⁶, thus enabling iterative optimization on any hardware platform capable of performing the relevant chemical unit operations (e.g., reagent addition, stirring with temperature control, etc.).
- 25 We demonstrate the system’s usability for reaction optimization by improving the product yield and purity for the 4-component Ugi reaction; Van Leusen oxazole synthesis; manganese-catalysed styrene epoxidation and explorative trifluoromethylation using the Ruppert-Prakash reagent. By using a unified format for storing and sharing procedures, process data and results we ensure that every protocol can be reproduced and verified. The
30 key requirement for any autonomous chemical robot is the ability to dynamically execute a given list of instructions with real-time adaptation to changing process parameters. To realize this on the Chemputer platform, the following components were integrated in the overall framework: a) hardware and software support for a range of low-cost sensors, b) dynamic χ DL as a basis for various feedback control chemical operations, c) software package for
35 analytical instrument control and signal processing, d) χ DL-based package for iterative reaction optimization with support for parallel procedure execution. These improvements enabled the first demonstration of an automated tandem of discovery-optimization framework that uses χ DL code as an input and returns optimized XDL as an output, paving the way for fast collaborative exploration of chemical spaces and reaction conditions.

40

We have included a set of low-cost sensors into the existing infrastructure of the Chemputer platform: colour, temperature, conductivity, and pH sensors for monitoring of chemical processes, and a liquid sensor for tracking material transfer and detecting failures of the liquid handling system. An environmental sensor was added to record the ambient conditions –
5 temperature, pressure and humidity – and identify potential reproducibility issues. All sensors are connected to a custom-designed board, the *SensorHub*: an Arduino module, featuring a variety of communication protocols and connected to the Chemputer IP network. Additionally, graphical interface is provided through a web-based dashboard application, which allows the user to control any sensor individually, or change the rate of the background measurements
10 for demanding processes. In addition to the low-cost sensors, a vision-based condition monitoring system was developed to add flexibility and improving the autonomy of Chemputer operations.

To manage the control of the analytical instruments and provide a unified interface for
15 obtaining spectral data, we have developed a stand-alone Python package – *AnalyticalLabware*. Covering a range of several analytical instruments, our library includes control over UV-Vis, near IR (NIR), Raman and NMR spectrometers as well as HPLC-DAD system from various manufacturers, see Figure 2. In addition, the library also has basic methods for spectra pre-processing, such as peak picking or baseline correction, and domain
20 specific techniques, e.g., zero-filling and apodization for NMR spectra. The package is fully integrated into our Chemputer workflow: from hardware graph, in which the instruments are presented as hardware objects with corresponding connection parameters, to dedicated high-level χ DL steps, specifying sampling routine and additional parameters to perform the data acquisition.

25 At the fundamental level χ DL provides a universal ontology for the encoding and execution of chemical synthesis recognizing that all chemical synthesis is based around four abstract properties: the reaction; workup; isolation; purification. This means that well known chemical reactions can be expressed as a process-driven language that focuses on the practical
30 actions needed to allow the reaction to happen. To extend this to dynamic reaction control, a base class for dynamic processes, *AbstractDynamicStep*, is exposing three abstract methods to control the execution flow, where each method returns a list of steps to be executed, based on the current state of the step. Here, we present a set of dynamic χ DL steps to allow self-correcting procedure execution for a range of potential use cases (addition,
35 transfer, execution, monitoring, optimization).

The *ChemputationOptimizer* software is designed to take further advantage of the χ DL dynamic step by leveraging a set of optimization algorithms to dynamically update the procedure parameters based on an end-point measurement obtained from a given analytical
40 instrument (Figure 3). The χ DL procedure, either translated from a literature or obtained followed combinatorial or active learning reaction discovery, is used as a starting point for the optimization cycle. The user only needs to provide a corresponding hardware graph and a configuration file. After robotically executing the procedure, the reaction output (typically, a

spectrum of the quenched reaction mixture) is analyzed and passed to an optimization algorithm to suggest the next set of input conditions. The user can choose from a wide variety of state-of-the-art optimization algorithms, including those implemented in the Summit²⁷ and Olympus²⁸ frameworks. A server-client interaction allows multiple clients to work together
5 towards a joint optimization problem. With the new set of reaction parameters, the procedure is updated and executed, and this cycle is repeated until the maximum number of iterations, or the desired target is reached. All experimental procedures together with the corresponding set of parameters and reaction results (both raw spectral data and the processed output) are saved in a database and can be verified later.

10 With these new hardware and software developments in place, we validated that low-cost sensors can capture relevant process data of the synthesis execution, ensuring safety and stability of operation. Our experience showed that the most encountered critical failure within the Chemputer is related to various types of syringe breakage. For such potential hardware
15 failures, we employed the vision-based condition monitoring system that uses multi-scale template matching, detects anomalies using a holistic approach of structural similarity, and alerts the operator.

We deployed our data-rich reaction development engine to passively monitor the turbidity
20 during the formazine synthesis, an organic colloid used as a turbidity reference material (Figure 4a). The system did not only capture ambient conditions, e.g., the room's temperature, but reagent priming, and addition left distinct signals in the liquid detector trace and the RGBC sensor detected the increase of turbidity as the reaction progressed. The liquid sensor was further used to monitor the consistency of reagent delivery and proved
25 valuable in challenging steps such as filtration, where transferring volumes are not predefined. Here, a simple binary output (i.e., 0 for an empty tubing and 1 for filled) was sufficient to increase the reliability of the process, while quantitative volume data might be obtained via indirect analysis. Overall, combined data could serve as a process fingerprint and may be used for subsequent validation of any reproduced procedure.

30 We have demonstrated the self-correcting execution of two functional group interconversion reactions (thioether oxidation²⁹ and nitrile formation³⁰), using the feedback from the temperature and colour sensors, respectively. The examples were chosen to highlight the benefits of incorporating the selected low-cost sensors in automated synthesis platforms.
35 The first example shows the slow addition of hydrogen peroxide monitored by an internal temperature probe to prevent the thermal runaway. By utilizing the dynamic step, it was possible to carry out the reaction automatically on a 25-gram scale without exceeding the maximum temperature specified in the literature procedure during the oxidant addition step (Figure 4b). Next, the passive temperature monitoring revealed an uncontrolled exotherm
40 during the subsequent heating step. Such insights can be easily discerned through our dashboard, enabling safe process development and scale-up. Our second example demonstrates the use of a simple colour sensor to monitor a nitrile synthesis from an

aldehyde using ammonia and iodine, and dynamically adjust the reaction time as the discolouration indicates complete reagent consumption (Figure 4c).

5 The reaction time of the chosen process varies depending on the aldehyde substrate, which has to be determined using a supervised trial run if in-line feedback is not available. Typically, exact reaction times are unknown and require a more general solution, such as the one outlined in our χ DL step for dynamic execution which performs the respective child step(s) until termination criteria are met. In our experiments, absolute thresholds for sensor readings proved unreliable as termination criteria due to changes in the ambient light, however we
10 found that rate of colour change could be used to detect the end of the process.

Beyond enhanced process control and condition monitoring, we were able to optimize reaction conditions for multicomponent³¹, heterocycle synthesis³², and catalytic reactions³³ using feedback from ¹⁹F NMR, HPLC-DAD, and Raman spectroscopy. The goal was to
15 illustrate a broad change of chemistry that features commonly encountered reaction types in preparative synthesis to showcase the benefits of digitalisation for the traditional organic chemistry community. The algorithms were chosen to highlight the capabilities of our agnostic toolkit (integrated optimization routines using random search, design of experiments, Bayesian optimization and genetic algorithms as well as interface to the Summit and Olympus
20 frameworks) and a user may choose the most appropriate algorithm for their chemical system. In our experience, a wide variety of algorithms were suitable for reaction optimization and even random search provides a strong baseline. As a proof-of-concept for reaction optimization, we have selected the Ugi four-component reaction (Figure 5a). Even though the reaction procedure is well established, it is sometimes not trivial to obtain good results, given
25 that multiple parameters need to be optimized simultaneously. The target parameter for this reaction was selected as the ratio between the area under the curve for the product and reference substance on the ¹⁹F NMR spectrum. In 30 experiments it was possible to achieve a 38% relative improvement for the yield of the product, compared to the starting literature conditions.

30 The optimization was performed using a Sequential Model-Based Optimization algorithm (SMBO) and consisted of four stages: five random experiments to initialize the parameter search space, 14 experiments to explore this space (i.e. minimize the uncertainty), five experiments with balanced exploration-exploitation approach and six experiments to exploit
35 the space (i.e. maximize the outcome, product yield). A Gaussian Process regression model served as the surrogate model. Figure 5a shows a typical ¹⁹F NMR spectrum of the reaction mixture (bottom right) and explored reaction parameters space over the optimization process (bottom left). Showcasing the flexibility of our approach, a Van Leusen oxazole synthesis was optimized using two parallel reactors with independent heating and stirring (Figure 5b). The
40 throughput of the platform can be increased by adding further reactor modules, however there is a trade-off with the number and the quality of updates obtained from the optimization algorithm. We used the SNOBFIT³⁴ algorithm to maximize the area of the product peak in the HPLC chromatogram relative to an internal standard while minimizing impurities and excess

reagents achieving a 10% relative improvement in the weighted objective in 26 iterations. Finally, we optimized starting material conversion relative to an internal standard in a manganese-catalysed epoxidation (5c) using online Raman spectroscopy in conjunction with the Phoenix³⁵ algorithm. Thanks to the fast acquisition time of the Raman instrument, time-series data could be obtained as well, which could be used for further analysis of the reaction kinetics. Not only classic process variables like temperature and time were considered but also often overlooked factors such as addition speed proved crucial in this transformation as they impact the formation of the active catalytic species. 40 iterations were needed to find a robust optimum that leads to full conversion.

We also demonstrated how this closed-loop approach can be extended to facilitate compound discovery and optimization. Our suggested pipeline includes three stages: first, the exploration of the product space achieved by a series of experiments from random search to algorithmic maximization of the heuristic novelty score. Next, the obtained spectra are analyzed to isolate regions of interest and identify potential products. Finally, for each identified product a series of experiments is executed to find reaction conditions, that will maximize the product outcome. As an example of such exploration-optimization strategy, we have selected the trifluoromethylation reaction³⁶ within a small substrate space of four different (Figure 5d). Starting from a complex mixture of substrates our system was able to identify three products of this reaction, which were characterized by corresponding peaks on the ¹⁹F NMR. These peaks were then used to guide the individual optimization experiments, where the target was set to maximize the area under the curve for the main product, while minimizing the areas for the other recognized peaks.

Using combinatorial search in limited chemical space we were able to discover a reaction between toluenesulphonylmethyl isocyanide and benzylidenemalononitrile, referred later as the tosMIC reaction (Figure 6a). The xDL procedure generated upon the discovery was subjected for the optimization using HPLC analysis feedback, set to maximize the product peak. Using the SMBO strategy we were able to increase the yield by 22% in just 32 iterations (Figure 6e). Furthermore, we decided to perform another reaction in a similar manner, this time, between three components: phloroglucinol, benzylidenemalononitrile and 1,8-bis(dimethylamino)naphthalene (Figure 6c). The optimization protocol consisted of two distinct campaigns, the first one comprising 37 reactions and the second campaign, comprising additional 13 reactions in the expanded search space. The reason for such strategy originated from careful analysis of the data obtained for the first campaign: 6 out of the 37 reactions performed at 25°C with maximally allowed 1,8-bis(dimethylamino)naphthalene concentration resulted in yields greater than the yield at initial reaction conditions (27%), suggesting that a more "adventurous" approach may lead to interesting results³⁷. The second campaign was initiated with results from the previous campaign serving as an input and the temperature and constraints allowing for temperatures as low as 0°C while allowing the 1,8-bis(dimethylamino)naphthalene concentration to reach 8 mol. During the first campaign the optimization protocol yielded parameters which resulted

in a 49% yield, an absolute increase of 22% from the 27% obtained for the initial reaction conditions.

However, with the more opened constraints present, the explored space has been expanded into unintuitive areas that would not have been a first choice, or a choice at all, for the experimental chemist: in this example the algorithm moved into temperatures not suited for the solvent used (DMSO freezing point is below 19°C). This shift has led to a further increase in the reaction yield by 50% with respect to the initial reaction conditions, reaching 77%

Low-cost sensors as well as process analytical technology instruments were integrated into the growing Chemputer software stack leading to enhanced process control and insight. This represents a fundamental shift from previous iterations of the Chemputer platform that were open-loop control systems for chemical synthesis to a closed-loop platform with feedback-enabled synthesis, optimization, and discovery capabilities. In practice, this means that new dynamic operations such as temperature-controlled reagent additions or optical endpoint detection that were previously not available can now be used in automated synthesis, increasing safety and reliability of our system. Furthermore, we showed how χ DL procedures derived from the literature can be improved and versioned using optimization and databasing. Executing the optimized, versioned XDL codes on the platform leads to reproducible optimal results. Data thus obtained can be used to understand the influence of process variables on given chemical transformations which in turn may help to populate reaction blueprints with tailored conditions for library synthesis or accelerated molecular discovery projects. In this context, the newly developed χ DL-based parallel execution and intelligent resource allocation have the potential to greatly enhance the throughput of the Chemputer, while maintaining maximum flexibility in varying experimental conditions and avoiding constraints typically encountered in HTE setups. We also demonstrated how this approach can be extended for the exploration of unknown reaction spaces, combining digital discovery and optimization in a single framework. Overall, the reported system provides a universal approach for optimizing digital recipes and can accommodate any further module developments following the χ DL standard. As the number of robotically generated datasets grows, we envisage that real-time telemetry data will provide an important means for data verification. Ultimately, we believe that the toolkit described herein will reduce barriers to automated process development and optimization as well as more complex, autonomous molecular discovery workflows.

Experimental

The following examples are provided to further illustrate the present invention and are not intended to limit the scope of the invention.

Materials and Measurement Methods

Solvents and reagents were used as received from commercial suppliers unless otherwise stated.

- All NMR measurements were recorded on Bruker Avance III HD 600 spectrometer operating at 600 and 151 MHz for ^1H and ^{13}C , respectively, or Bruker Avance III 400 spectrometer operating at 400 and 100 MHz for ^1H and ^{13}C respectively. Spectra were collected at 298 K, chemical shifts are reported in ppm and were calibrated for the (residual) NMR solvent signal (apparent multiplicities are given as s: singlet, d: doublet, t: triplet, q: quartet, m: multiplet, with coupling constants reported in Hz). The spectra were processed with MestReNova 14.0.0. High resolution mass spectra were obtained on a Bruker maXis LC/MS.
- For reaction optimization Magritek Spinsolve 80 Carbon (80 MHz operating frequency for ^1H) benchtop NMR, Agilent 1260 Infinity II LC system (with IDEX MX II sample loop valve) and Ocean Insight QE Pro Raman spectrometer were used.

Software Architecture

- The Chemputer Optimizer was controlled by adapting the previously published Chemical Description Language (χDL) (see <http://xdl-standard.com/> and Mehr).

- Chemputer Optimizer software package has been developed to execute closed-loop synthesis optimization using the Chemputer as automation platform. The core routine utilizes the special methods of the *DynamicStep* of the χDL framework to run the procedure iteratively, updating the user-defined parameters using the feedback from analytical instruments. Additional dynamic steps have been developed to execute chemical operations with process monitoring using various utility sensors: *DynamicAdd* - for dynamic reagent addition, *DynaicTransfer* - for crucial liquid transferring operations and *DoUntil* - for active reaction monitoring. The *AnalyticalLabware* library introduces methods to control several analytical instruments in a unified manner, including HPLC, benchtop NMR and Raman spectrometer. In addition, this library provides operations for basic processing and analysis of the acquired data. All modules are written in Python 3.9, with full source code and documentation available online as supplementary files and on GitHub. The code is compatible with existing ChemPU software stack and can be extended for use in any χDL compatible automation platform.

- The ChemputerOptimizer software package is based on the existing XDL framework, taking the advantage of its routines for procedure preparation (i.e., inserting implied steps and attributes) and execution. The core class, *OptimizeDynamicStep*, is a controlling the dynamic iterative execution of the given procedure and updating its parameters with the new suggestion from the optimization algorithm. Several classes are developed to support the closed-loop experiment execution, mainly: *AlgorithmAPI* – general class to provide interface for existing algorithms for optimization, manages reaction parameters and results, and translates them in an appropriate format; *SpectraAnalyzer* – class to analyze data from analytical instruments (spectra) and calculate the target reaction descriptor (e.g., yield);

OptimizerClient – client part of the server-client based interaction with Summit and Olympus optimization frameworks.

- 5 In addition to common steps to describe synthetic operations, ChemputerOptimizer features several special XDL steps for running closed-loop optimization.

OptimizeStep

- 10 The OptimizeStep step indicates optimization parameters in the XDL procedure. These parameters are then processed during the procedure compilation with the hardware graph. They can be set by the user or are naturally extracted from the procedure, using a predefined set of common optimization steps, e.g., reagent addition or reaction temperature. The OptimizeStep should wrap the step which parameters are subject to the optimization (child step).

- 15 An Example of the OptimizeStep wrapper is shown below, where the optimize_properties attribute indicates the parameter of the child step subjected to the optimization, with its max_value and min_value limits.

```
20 <OptimizeStep
    id="1"
    optimize_properties="
        {'volume':
            {'max_value': 0.3, 'min_value': 0.0}}">
    <Add
25     vessel="reactor"
        reagent="cyclohexenone"
        volume="0.2 mL"
        aspiration_speed="3.0"
        port="0"
        stir="True"
30     priming_volume="0.2 mL" />
```

Analyze

- 35 The Analyze or FinalAnalysis step performs the reaction analysis when executed. The step contains all necessary parameters to run the analysis via Raman, HPLC or NMR instruments, including reaction sampling, sample treatment (e.g., dilution) and instrument control. During the compilation the step's attributes are additionally checked for consistency (e.g., "dilution volume" if dilution is required) and necessary callback functions (i.e., to update the optimization state upon iteration completion) are assigned. The Analyze or FinalAnalysis should be inserted in the procedure after the reaction is complete.

40

An example of the Analyze step to perform the NMR analysis of the reaction mixture upon reaction completion is shown below. The step contains attributes to indicate the vessel where

the analysis should be performed, desired method for the analysis and sample volume (if sample is required). Additional attributes are added for the corresponding analytical method (e.g., shimming for the NMR) and the method_props attribute represents the parameters (e.g., nuclei for the NMR) to run the analysis.

5

When the reaction time is used as a parameter for optimization, quenching steps are typically added to avoid inconsistency between the actual reaction time and the recorded reaction time. Reactions were typically quenched either by rapid dilution or addition of reagents during work-up steps during the analysis. These are part of the high-level Analyze XDL steps. For

10 Raman, no quenching is needed as the measurement is instant. Therefore, time inconsistencies are minimal systematic errors. In this way, executing the optimized, versioned XDL codes on the platform leads to reproducible optimal results.

```
15 <Analyze
    vessel="reactor"
    method="NMR"
    sample_volume="2.5"
    method_props="{
20     'protocol': '1D FLUORINE HDEC',
    'protocol_options': {
        'centerFrequency': -85,
        'PulselengthScale': 1,
        'decouplePower': 0,
25     'AcquisitionTime': 1.64,
        'RepetitionTime': 15,
        'PulseAngle': 90,
        'Number': 64
    }
    }"
30 force_shimming="True"
    cleaning_solvent="THF"/>
```

Background Reaction Monitoring

The StartMonitoring asynchronous step starts the execution of the analysis with the chosen method in a background thread. Every recorded spectrum is saved in the
35 SpectraAnalyzer class and is subjected for the analysis to categorize the found peaks into:

- Product – peaks that appeared during the reaction and remained at the end of the it.
- Starting material – peaks that were present at the start of the reaction and their integration area decreased during its course.
- 40 • Intermediate material – peaks that appeared during the reaction but were absent in the end of it.

- Noise/artifacts – peaks that were persistent during the reaction.

Constrained Step

In some cases, additional constraints beyond upper and lower bounds are imposed on the optimization variables. For example, all factors in mixture experiments must sum up to 100%.

- 5 The ConstrainedStep wraps a child step and adjusts the value of a given parameter as a function of the values of selected OptimizeStep steps and the desired target value.

10 An example of the ConstrainedStep step is given below, that adjusts the volume of a constrained Add step based on two OptimizeStep steps. In this example, the toluene volume would be adjusted to 30 mL. In the given example, the volume of the Add step is calculated as in:

$$V(\text{target}) = V(\text{id1}) + V(\text{id2}) + V(\text{Add})$$

```
15 <OptimizeStep
    id="1"
    optimize_properties="
      {'volume': {'max_value': 20.0, 'min_value': 10.0}}">
    <Add
      vessel="reactor"
      reagent="cyclohexenone"
20     volume="10 mL" />
  </OptimizeStep>
  <ConstrainedStep
    ids="[1, 2]"
    parameter="volume"
25    target="50"
    <Add
      reagent="toluene"
      vessel="reactor"
      volume="99999 mL"
30    <Add/>
  </ConstrainedStep/>
  <OptimizeStep
    id="2"
    optimize_properties="
35    {'volume': {'max_value': 20.0, 'min_value': 5.0}}">
    <Add
      vessel="reactor"
      reagent="pyrrolidine"
      volume="10 mL" />
40  </OptimizeStep>
```

Sensor Control and Execution

- The low-level sensors control is implemented via Arduino CommandTools framework, which provides a unified interface for communication with various devices via Arduino boards¹⁸. The corresponding python library – Commanduino – provides python interface for communication and control of any device interfaced with Arduino CommandTools¹⁹. The sensors package is integrated into Chemputer stack via ChemputerAPI and Chempiler libraries, which provides high-level individual sensor access and overall control of the background recording thread and data i/o respectively.
- 10 To provide access to sensors control and execution for the end user through XDL framework the following steps were developed:
- SetSensor – the starting point to control the background sensor monitoring thread. A callback function could be assigned during compilation, that will be executed with timestamp argument when the sensor will be set.
 - ReadSensor – the step to take a single reading from the sensor and pass it to the callback function (if assigned during the compilation). The reading is acquired from the background monitoring thread, or, if the latter was not started, by direct communication with sensor.
 - GetSensorReadings – the step to take several last readings, filter them (if required) and pass to the callback function (if assigned during the compilation). The readings are requested either by the timestamp or by the number of readings. The following filtering functions are available: median filter and FFT filter.

Software Workflow

- 25 The overall process to run the iterative reaction optimization on a XDL compatible hardware platform (e.g., ChemPU) is as follows:
1. Prepare the chemical procedure as XDL file.
 2. Wrap steps that are subject for the optimization with OptimizeStep step (as described above) to mark the corresponding attribute.
 3. Prepare the graph containing mapping to the hardware.
 4. Prepare the configuration file, containing all necessary parameters to run the optimization (e.g., number of iterations) and load the algorithm (e.g., its name and related parameters), and specifying parameters to perform reaction analysis and calculate the product descriptor (e.g., product peak at the final spectrum).

5. [Optional]. Prepare table with previous data (e.g., experiments run manually) as csv file, containing header with labels, that correspond to the parameters for the optimization.
6. Execute the optimization procedure on the physical platform.

Example code to run the optimization on the ChemPU using *ChemputerOptimizer* and

5 *Chempiler* packages is shown below:

```
from pathlib import Path

from AnalyticalLabware import devices
10 from chemputeroptimizer import ChemputerOptimizer

# ChemPU specific
from chempiler import Chempiler # Control package
15 from AnalyticalLabware.devices import chemputer_devices # Analytical instruments
import ChemputerAPI # Low level API

# Path to the corresponding experiment folder
20 # Should contain the following files:
# procedure.xdl - XDL procedure
# graph.json - hardware graph
# optimization_conf.json - configuration file for the optimization
# (optionally) previous_results.csv - table with previous results
25 path = Path('folder')

# Instantiate main optimizer object
co = ChemputerOptimizer(
30     procedure=path.joinpath('procedure.xdl'),
    graph_file=path.joinpath('graph.json')
)

# Run necessary preparations
35 co.prepare_for_optimization(
    opt_params=path.joinpath(
        'optimization_conf.json').as_posix()
    )

# Upload previous results if given
40 co.load_previous_results(
    results=path.joinpath('previous_results.csv')
    )

# Create chempiler object to control the hardware
45 c = Chempiler(
    experiment_code='test',
```



```
graph_file=path.joinpath('graph.json').as_posix(),
output_dir='test',
simulation=True, # False if running on physical hardware
device_modules=[ChemputerAPI, chemputer_devices]
5 )

# Run the optimization
co.optimize(c)
```

Optimization configuration

10 The configuration file contains all necessary information to run the optimization and calculate the reaction outcome in a json format.¹ The description of each parameter is given below.

- max_iterations – maximum number of experiments to perform within optimization run.
- 15 • target – indicates the calculation for the reaction outcome. Detailed description is given in the section 0. The value for each target attribute dictates when the optimization is terminated.²
- algorithm – indicates the algorithm to be used and its respective attributes. Details of the available algorithms and their parameters are given in the section 0.
- 20 • reference – indicates the peak position of the reference compound (i.e., internal reference for the analysis). Is used to reference the spectrum and calculate the reaction outcome.
- constraints – lists regions on the reaction spectrum to consider when calculating the final reaction outcome. In the current implementation, the intermediate result (e.g., area under the curve for the target peak) is divided by the sum of integration areas for all listed regions.
- 25 • batch_size – number of batches to perform the optimization in parallel.

An example of the configuration file for optimization is shown here:

```
{
30   "max_iterations": 5,
   "target": {
       "spectrum_integration-area_-80.809..-82.049": Infinity
   },
   "algorithm": {
35     "name": "random",
     "random_state": 42
   },
   "reference": -113.15,
   "constraints": [
```

```
5      "-73.214..-73.751",  
        "-73.889..-74.653",  
        "-75.178..-76.334"  
    ],  
    "batch_size": 1  
}
```

Optimiser Compilation

In addition to the compilation process of the XDL procedure into the ChemPU instructions³⁾ ChemputerOptimizer features several operations for internal preparations. Core operations
10 include: identifying or assigning steps for the optimization and analysis, parsing configuration and assigning callback methods to calculate reaction outcome, scheduling the procedure for parallel execution, breakdown of the XDL steps into corresponding low-level substeps and mapping those to the hardware graph.

OptimizeStep Assignment

- 15 If no OptimizeStep is found in the procedure and the optimization is not running in the interactive mode, all steps that may be subject for the optimization will be automatically wrapped with OptimizeStep and the corresponding parameter will be set within the range of $\pm 20\%$ from the initial value.
- 20 XDL steps and their corresponding parameters that may be subjected for the optimization are shown below:

```
25 {  
    "Add": {  
        "volume",  
        "time",  
    },  
    "AddSolid": {  
30        "mass",  
    },  
    "HeatChill": {  
        "time",  
        "temp",  
    },  
35    "HeatChillToTemp": {  
        "temp",  
    },  
    "Stir": {  
40        "time",  
    },  
    "Wait": {  
        "time",  
    }  
}
```

}

Analyze step Assignment

If no Analyze step is found in the procedure and the optimization is not running in the interactive mode, an Analyze step with method="interactive" will be inserted in the procedure after the step that presumably indicate the reaction completion. Such steps are: Stir, Wait, HeatChill, HeatChillToTemp. To avoid any confusion and incorrect reaction analysis, it is recommended to insert the Analyze step manually in the procedure. If no Analyze step is found and the optimization is not running in interactive mode – an exception will be raised.

10 Callback Assignment

Depending on the given Analyze step's method, the callback function will be assigned to correctly calculate and update the result of the reaction. Examples of such functions are listed below, based on the optimization target:

- 15 • final_parameter – a generic target for running optimization in interactive mode, no calculations will be performed on the user input.
- 20 • spectrum_peak-area_XXX – reaction outcome will be calculated as area under the curve for the given peak (where XXX is the x axis coordinate) on the spectrum; if no peak at this position was found, the output will be 0. If reference is given in the optimization configuration, the resulting area will be divided by the area for the reference peak.
- 25 • spectrum_integration-area_LLL..RRR – reaction outcome will be calculated as area under the curve within the given x axis limits (where LLL is the left and RRR is the right borders respectively). This method is insensitive to the peaks found. If reference is given in the optimization configuration, the resulting area will be divided by the area for the reference peak.
- novelty – reaction outcome will be calculated according to the novelty equation (see below)

Scheduling

- 30 If the procedure is executed in a single batch mode, no scheduling is performed. For batched optimization, scheduling the simultaneous execution of several XDL threads in a single run involves establishing an execution plan – i.e., a rough plan of when each step required by the procedure will be executed – ensuring no resource use overlap between XDL threads. We chose to establish an execution plan rather than, for example, just use resources immediately as they become available to preserve as much as possible any necessary structure and
- 35 minimize wait times within a given procedure. To establish a schedule, first all XDL scripts

and the hardware graph are parsed, and all mappings between each of the steps within a given XDL script and the hardware graph are obtained. In other words, for a HeatChill step and a graph containing multiple reactor vessels, a separate XDL object is created where the HeatChill step is attributed to each of these reactor vessels. As a result, we obtain – for

5 each XDL script – a set of possible XDL script-to-hardware maps. Next, each script-to-hardware map is used to establish a set of resource allocations (corresponding to a list of times that each graph node is in use by a given XDL thread). With all these resource allocations in hand, this constitutes an optimization problem, where the objective is to find the XDL script-to-hardware map for each XDL script that minimizes the overall execution time,

10 where no resource is allocated to 2 XDL threads simultaneously at any time. In our case, we implemented several optimization algorithms (genetic algorithm, random search), but found that all combinations could be exhaustively searched given the size of our hardware graph.

Scheduling algorithm pseudocode is shown below. In this example all the schedules are

15 generated and the one with the minimum execution time is selected.

```

G: hardware graph.
{XDLi}: set of XDL scripts to be executed.

20 Mapper: Get all XDL – hardware maps.
Allocate: Get graph resource allocation for a given XDL – hardware map.
Product: Generate all combinations of XDL – hardware maps, taking one possible map
per XDL file.
Schedule: Returns the minimum time execution schedule for a given set of XDL resource
25 allocations.
Evaluate: Calculates the execution time of a schedule.
S: Schedule of XDL execution.
t: Execution time.

30 {{Mi[XDLi]}} ← Mapper(XDLi, G)
{{Ri[Mi]}} ← Allocate({{Mi[XDLi]}} , G)

for combination in Product({{Ri[Mi]}}) do:
    S, t ← Schedule(combination)
35     if t < tmin then:
        tmin ← t
        Smin ← S

return Smin

```

40 Typically, the expected acceleration when using 2 reactors is less than 2x due to multiple processes relying on the same hardware resources. For example, the 2-fold parallel system had an estimated 1.6X acceleration and a 4-fold parallel system exhibited a 2.7x acceleration in simulations. It is important to note that these factors can vary based on the experiments selected by the algorithm (e.g. different combinations of short and long reactions times

45 allowing for different degrees of concurrent execution). In practice, the acceleration depends on the accuracy of the expected durations for the different XDL steps that are used by the

scheduling routine. In our hands, this led to factors of < 1.6X, however this is expected to improve as we collect more data on actual durations.

Hardware mapping and implied steps addition

In addition to the XDL compilation process, further preparations are needed for the Analyze step. Some of the implied steps will be added according to the analytical method used, e.g., shimming for the NMR, however others are general depending on the analytical steps parameters:

- **Sampling.** If sampling has to be performed additional steps to transfer the aliquot to the analytical instrument are added to the list of Analyze step substeps: transfer an aliquot with an excess volume from the reaction vessel to the injection syringe pump (i.e., the closest pump to the analytical instrument); transfer the aliquot from the injection pump to the analytical instrument; transfer the excess volume from the injection syringe back to the reaction vessel.
- **Dilution.** If reaction sample has to be diluted prior to the analysis the following steps are added to the list of Analyze step substeps: the liquid path from the reaction vessel to the dilution vessel is primed with the reaction sample; an aliquot is transferred from the reaction vessel to the dilution vessel; the dilution solvent is then added to the dilution vessel; dilution vessel is stirred for a pre-set time (10 minutes).
- **Cleaning.** If analytical instrument is required to be cleaned after the analysis, the necessary steps to transfer the cleaning solvent to the instrument and flushing the liquid path with gas are appended to the end of the Analyze step substeps.
- **Shimming** (only for the NMR analysis). Shimming is triggered when shimming results are not found on the PC or last shimming was performed more than 24 hours before or `force_shimming` attribute was set. If shimming is required, the solvent for shimming is chosen according to predefined rules, transferred to the instrument and used as a reference for shimming routine.

Spectra Analysis

All spectra acquired during the optimization procedure are managed by the SpectraAnalyzer class, which features several methods to calculate reaction outcome (examples are listed in the *callback assignment* section) and analyze several spectra on a time scale. The analysis of the NMR spectra is based on the SpinsolveNMRspectrum class from AnalyticalLabware module and nmrglue library methods for the analysis. The analysis of the HPLC spectra is based on the AgilentHPLCChromatogram class from AnalyticalLabware module. The analysis of the Raman spectra is based on the

RamanSpectrum class from AnalyticalLabware module. Calculations for the novelty analysis are described below.

Novelty Analysis

The novelty analysis is a reaction descriptor used to drive the optimization towards unexplored product space. It is based on the information gain from the reaction spectrum and the spectrum's novelty with respect to previously measured spectra of the same reaction. The analysis begins with search for regions of interest (i.e., containing meaningful information – peaks)⁴, calculation of their individual scores, total information score for the current spectrum and additional coefficient for the current spectrum novelty.

10 *Individual Regions Score*

Score for individual region is calculated as the size of the region (as number of data points) multiplied by the inverse logarithm of the difference between area of the current region and harmonic mean of all identified regions (Equation 1).

$$Rs_i = Size_i \times \frac{1}{\log_{10}(|area_i - \overline{area}|)} \quad (1)$$

15 If one of the region's area is equal to overall harmonic mean, that this region's score is equals to its size. Harmonic mean is used to favour smaller peaks and reduce the weight of large spikes and potential false positives (i.e., noise identified as material peaks). An example of the individual regions scoring is shown in Figure 20.

20 In Figure 20, the ¹⁹F NMR spectrum (top) is shown with each identified region highlighted with colour with top annotations indicating peak's size (in number of data points) and score (calculated from Equation 1). The bottom shows a bar chart with areas for individual peaks (aligned on the x axis) with their harmonic mean indicated as orange horizontal line.

Information Score

25 Total information on a single spectrum is calculated by multiplying the sum of individual regions by the total number of identified regions (see Equation 2).

$$\sum_{i=1}^r (Rs_i) \times r \quad (2)$$

The score is insensitive to the spectrum phasing (e.g., for NMR spectrum) and will favour spectra with large amount of equally sized regions. An example is shown in Figure 21, which shows the information score for two different ¹⁹F NMR spectra.

Novelty Coefficient

The novelty coefficient is calculated with respect to all previously measured spectra, for which identified regions are stored as arrays with their x axis coordinates. All identified regions for the spectrum of interest are checked versus previously identified regions (spectrum-wise) and the spectrum with regions not found previously will have a higher coefficient (Equation 3).

$$Nc_i = \frac{|F_i \setminus P|}{|F_i|} + \frac{1}{|P|}, \text{ where } P = \bigcup_{j=1}^{k \neq i} F_j \quad (3)$$

F_i – set of identified regions for i^{th} spectrum

The novelty coefficient is insensitive to the regions height and is reduced with the number of obtained data. An example of the novelty coefficient is illustrated in Figure 22, which shows an illustration of the novelty coefficient calculation. The spectrum of interested is plotted navy grey, set of individual regions for all previously measured spectrum is plotted as green circles (labelled as “Cumulative regions map”) and novel regions are highlighted with orange crosses.

Algorithm Interface

The AlgorithmAPI class is designed to provide a unified interface to the algorithmic classes and dynamic XDL execution. The class contains methods to parse the reaction parameters into parameter-agnostic data arrays suitable for the numeric optimization algorithms, initialize algorithm classes, query data for the next setup and save the intermediate results.

Available Algorithms

Random search

Random search algorithm relies on the random submodule of the NumPy library.⁵ The suggestion of this algorithm is a random number array (sampled from a uniform distribution) within the parameters constraints. The following parameters are available to construct the algorithm's class:

- `random_state` – a seed to initialize low-level random number generator, used to preserve reproducibility across multiple optimization runs.

Genetic algorithm

A basic genetic algorithm was written and adopted for the sequential optimization of chemical reactions. It uses truncation selection, single point crossover, and random reset mutation while preserving the best solution (elitism). In case of premature convergence, the population is reinitialized. The key hyperparameters to tune are:

- `pop_size` – Number of individuals in the population.
- `mutation_rate` – Probability (0 to 1) of mutating a gene.

Sequential model-based optimization (SMBO)

The algorithm implementation is based on the `scikit-optimize` python library and provides a simple and efficient library to minimize expensive and noisy black-box function.⁶ The parameters for the underlying algorithm are listed in the official documentation.⁷

5

Design of Experiments

Several experimental designs are available via a wrapper for `pyDOE2`⁸. Available designs include full and fractional factorial designs, Response-surface designs, and randomized designs. The wrapper calls the appropriate `pyDOE2` functions to generate the design matrix and maps the levels onto the appropriate parameter values based on the search space bounds.

10

Dummy methods

The following “dummy” methods are given to the user to run the given reaction iteratively using the respective setup on each iteration:

15

- `fromcsv` – the next setup is read from the given csv file.
- `reproduce` – the next setup is equal to the initial setup from the procedure. Used to check the experimental outcome for reproducibility.

Optimizer Client

The `OptimizerClient` represents the class for communication with the `SummitServer` (see below). The communication is based on the TCP-IP sockets and managed via `selectors` and `socket` modules of the python standard library.⁹ The communication messages are encoded json-like dictionaries containing information about the current procedure, steps and parameters subjected to the optimization and optimization configuration. The class is initialized if the requested algorithm is the one available in the Summit framework and managed via `AlgorithmAPI`.

20

25

Summit Server

The `SummitServer` is the python module that provide access to the Summit framework via TCP-IP communication. The module can be set on a remote server and supports connection for multiple clients with python `selectors` library. The format of the communication messages is JSON with all reaction parameters encoded as “<step name>_<step ID>-<parameter name>”, all parameters for the algorithm and a “hash” field, that allows matching same optimization procedure running on various clients.

30

OlympusWrapper

`OlympusWrapper` is a class that serves as an optimization handler within the `ChemputerOptimizer` server framework, providing access to the state-of-the-art algorithms implemented in the Olympus¹⁰ benchmarking framework. It parses the JSON requests from

35

the client and constructs the appropriate parameter space, observation and planner objects from Olympus. Queried points are returned in JSON format for immediate execution on the ChemPU platform. To use specific algorithms, the user must install the Olympus package and all required dependencies by the algorithm.

5 Data Management

At the end of every iteration, the following data is saved:

- Executed procedure (as .xdl file) – contains the procedure executed at the current iteration with corresponding parameters; could be for the future reproduction.
 - Parameters used (as .json file) – stand-alone file stating all the parameters optimized for the current procedure, their values at the current iteration, and the description and the value for the target for the current iteration.
 - Results list (as .csv file) – stand-alone file, stating all reaction parameters and results used so far for the current procedure. This file is used to set up the algorithm to continue the optimization.
 - Parallel schedule (as .json file) – stand-alone file, stating the schedule of the procedure steps if the optimization is running in parallel using several reaction modules.
 - Spectrum (as .pickle file) – the spectrum obtained during the FinalAnalysis or Analysis step executed at the current batch/iteration (if applicable).
- The data is managed using the NoSQL document database. The entry files are generated using the stand-alone script from the files listed above.

Database Schema

An example XDL Database entry schema is shown in below. This is the basic schema of XDL top-level database entry defining the entry ID and capturing all relevant meta data.

```
Experiment Table
Experiments(
  ID varchar(255) UNIQUE NOT NULL PRIMARY KEY,
  Username varchar(255) NOT NULL,
  ExperimentCode varchar(255),
  Public boolean,
  Description text,
  SMILE text,
  INCHI text,
  Validated boolean NOT NULL,
  Success boolean,
  Keywords text,
  ExperimentSeries varchar(255),
```

```

5      Iteration integer,
      ExactRepeat Boolean,
      ExactRepeatOfExperiment varchar(255),
      EstimatedDuration real,
      ExperimentDuration real,
      ReactionClass varchar(255) NOT NULL,
      Scale real,
      Datetime timestamp,
      Location varchar(255)
10  )

```

The base table 'Experiments' gathers the key information and meta data of the experiment. The meaning of the individual entries in the base table is defined in Table 1.

Table 1: key information and meta data of the experiment

Entry	Comment
ID	Unique identifier
Username	Owner of the entry
Experiment code	Laboratory code (reference to lab notebook) of the experiment
Public	Defines whether the entry is public or private
Description	A short description of what is important about this reaction
SMILE	SMILE of the reaction
INCHI	INCHI of the target product
Success	Defines whether the outcome of the reaction was successful or not
Keywords	Keywords that are characteristic for the reaction (e.g. 'metathesis')
Experiment series	An identifier for all runs that have been performed for this reaction
Iteration	The iteration of this experiment in the experiment series (e.g. the third attempt on a reaction with modified conditions would be iteration 3)
Exact repeat	Number of the exact repeat of an experiment
Exact repeat of experiment	ID of the experiment entry that has been repeated exactly
Estimated duration	The estimated run time of the experiment
Reaction class	The reaction class the experiment belongs to. The available reaction classes are listed in Table 2.
Scale	The amount of the limiting reagent in mol or mmol
Datetime	Starting date of the experiment
Location	Address of the lab in which the experiment was performed

Table 2: List of used reaction classes

Entry	Reaction class
1	C-C bond formation (TM catalysed)
2	C-C bond formation (TM non-catalysed)
3	Heteroatom alkylation and arylation
4	Functional group manipulation
5	Protecting group manipulation
6	Ring and heterocycle formation
7	Reductions and oxidation
8	Multicomponent reactions
9	Other

The 'XDLEntry' table gathers the information associated with the experimental procedure and execution, as illustrated below. The XDLEntry schema captures all entries that are directly or indirectly relevant for the actual procedure and links it to the experiment ID defined in the top-level schema.

```

XDLEntry (
  ExperimentID varchar(255) REFERENCES Experiments (ID),
  XDLVersion real NOT NULL,
  XDL text NOT NULL,
  XDLEXE text NOT NULL,
  Graph text NOT NULL,
  SimulationLog text,
  ExecutionLog text,
  ExecutionScript text,
  LiteratureProcedure text
)

```

The meaning of the entries in this table are defined in Table 3.

Table 3: Explanation of the entries of the XDLEntry table

Entry	Comment
Experiment ID	ID of the database entry as defined in the base table (Scheme S1)
XDL version	Version of the XDL code the experiment has been run with
XDL	The XDL file for the experiment
XDLEXE	The XDLEXE file for the experiment
Graph	The graph file for the experiment
Simulation log	The folder containing the simulation logs of the experiment
Execution log	The folder containing the execution logs of the experiment
Execution script	The Python script used to compile, simulate, and execute the XDL
Literature procedure	The prose procedure text for the experiment (if available)

All entries that are not present in the files generated at the end of iteration execution are filled manually.

Analytical Hardware

- All experiments were executed on the ChemPU platform, equipped with 1/16" PTFE tubing. A
- 5 HPLC instrument (Agilent 1260 Infinity II) was installed with an additional sample loop switching valve (Rheodyne MX Series II™) connected to the liquid handling system. The instrument was triggered after sample loading, with additional control over the experiment achieved using macro commands implemented in the AnalyticalLabware module. A benchtop
- 10 NMR (Magritek Spinsolve 80 Carbon) was equipped with a flow cell and connected to the liquid handling system via threaded fitting and 1/16" PTFE tubing. A Raman spectrometer (OceanInsight QE Pro) is coupled with a 754 nm laser (OceanInsight LASER-785-LAB-ADJ-SMA) and used with a contactless probe (OceanInsight RIP-RPB-785-SMA-SMA) installed at the round bottom flask.
- 15 All analytical experiment were executed via dedicated xDL steps, specifying all necessary protocol options, dilution or quenching steps. Therefore, time inconsistencies between recorded reaction time and actual reaction time are minimal systematic errors. All process sensors were connected to the Chemputer network using the SensorHub - a PCB featuring number of communication protocols and an Ethernet module for control over an IP network.
- 20 The AnalyticalLabware python module contains classes and method to control various analytical instruments and process obtained spectra in a unified way. A generic class for data processing (AbstractSpectrum) provides abstract method for loading the spectral data – redefined in the ancestor classes accordingly – and methods for loading and saving data as a
- 25 binary key-value storage (where keys are relevant spectral data, e.g., x and y axes data, timestamp, peaks, etc.) for subsequent use within AnalyticalLabware framework. Common methods for processing and analysis include baseline correction¹¹ peak picking, integration and spectrum smoothing, with support methods for trimming and displaying the spectrum.
- 30 Further information about the analytical sensors is provided below.

Magritek Spinsolve NMR

- The devices submodule contains classes to control the Spinsolve NMR and process obtained 1D NMR spectra: SpinsolveNMR and SpinsolveNMRSpectrum respectively. The spectrometer control is achieved via client-based communication with the Spinsolve software.
- 35 SpinsolveNMR class provides interface to retrieve available protocols and their options, execute those and store the data within SpinsolveNMRSpectrum class.

The SpinsolveNMRSpectrum inherits from AbstractSpectrum and additionally uses methods from nmrglue library¹² for NMR spectra processing: Fast Fourier transformation

(FFT), baseline correction, zero-filling and apodization. In addition, the class features the methods for region analysis.

Agilent 1260 Infinity II LC

The devices submodule contains classes to control the Agilent 1260 LC and process
5 obtained data: HPLCController and AgilentHPLCChromatogram respectively.

Communication with the instrument is established via a Chemstation macro hplctalk.mac
that monitors and executes commands written to a command file and returns replies by
writing them to a reply file. After downloading the macro from AnalyticalLabware, it is
10 placed in the Chemstation macro directory. This macro must be loaded and activated in the
Chemstation software using the following commands:

```
> Macro hplctalk.mac  
> hplctalk_run
```

15 On the python side, the following methods are available: turning on and off the pump and
lamp, switching the chromatographic method, preparation and start of a run, checking the
instrument status and switching to standby mode. These methods are implemented to write
specific commands to the command file and parse the responses received.

20 The AgilentChromatogram class is inherited from AbstractSpectrum and has additional
methods for loading the raw data from the propriety Chemstation file format.

Ocean Insight QE Pro Raman spectrometer

The devices submodule contains classes to control the Ocean Insight QE Pro Raman
25 spectrometer and process obtained data: OceanOpticsRaman and RamanSpectrum
respectively. The spectrometer class has methods implemented to obtain a (blank) spectrum
and control over integration time and number of scans performed.

The RamanSpectrum class is inherited from AbstractSpectrum and has two additional
30 methods useful for processing of the Raman spectra: subtracting the reference and iterative
peak picking algorithm with sliding threshold, which shows better performance in noisy
spectra.

Sensor Hardware

The enhanced ChemPU features a sensor package to monitor the status of the platform
35 during a run as well as to use real-time data to make decisions for dynamic procedure
execution on the fly.

SensorHub

For integration of process sensors into existing ChemPU stack, we have designed an Arduino shield with built-in Ethernet module for control over an IP network. The board features 12 MOSFET outputs, ethernet-to-serial converter and 8 individually controllable I²C channels on top of the functionality given by the Arduino board¹³. Overall, it allows installation of up to 8 analogue sensors, 8 sensors with I²C communication (achieved using an I²C multiplexer, thus allowing installation of devices with same addresses), several sensors with SPI communication (achieved via “software SPI” Arduino feature) and up to 12 PWM devices.

Individual sensors were selected from commercial manufacturers providing details installation manual and a corresponding software library for seamless integration. Sections below provide detailed descriptions of each individual sensors used in the ChemPU process monitoring framework.

Colour (RGBC) Sensor

The TCS34725 Colour Light-to-Digital Converter with IR Filter¹⁴ was selected as the colour sensor for process monitoring. The main component of the converter is 3 × 4 photodiode array, composed of the red-filtered, green-filtered, blue-filtered, and unfiltered (clear) photodiodes, additionally coated with an IR-blocking filter. The four integrating analogue-to-digital converters (ADCs) simultaneously convert the amplified photodiodes currents to a 16-bit digital value, which is sent over the I²C communication bus. We have chosen a DFRobot board (manufacturers part No. SEN0212) as an evaluation board for the sensor, featuring a unified PH2.0-4P interface for connection to the SensorHub via a crimped cable and a white LED as a light source. The sensor is attached to a round bottom flask as is shown in Figure 7.

Environmental Conditions Sensor

A BME280 digital humidity, pressure, and temperature sensor¹⁵ was selected for the environmental conditions monitoring. We have chosen DFRobot board (manufacturer part No. SEN0236) as an evaluation board for the sensor, featuring a unified PH2.0-4P connection interface for I²C communication. The sensor board is placed inside the fume hood as is shown in Figure 8.

Liquid Sensor

The OPB350 Tube Liquid Sensor¹⁶ was selected to monitor the liquid handling system of the ChemPU. The sensor consists of an IR LED, phototransistor and tube housing, and is designed to recognize the presence of the fluid inside the tube. A custom PCB was designed to simplify the operation and the connection of the sensor to the SensorHub (achieved using analogue input interface). The wired assembly sensor (manufacturer part No. OPB350W125Z for 1/8” tubing and OPB350W062Z for 1/16” tubing) are connected to the board using screwless wire terminal and installed on the tubing connecting syringe pump with the closest switch valve. The liquid sensor is shown in Figure 9.

Reaction Temperature Sensor (RTD)

A MAX31865 RTD-to-Digital Converter¹⁷ was selected to simplify the interfacing the SensorHub with resistance temperature detector (RTD, or temperature probe). The unit has a built-in 15-bit ADC, input protection, a digital controller, and an SPI-compatible interface. We have selected Adafruit PT100 RTD Temperature Sensor Amplifier (manufacturer part No. 3328) as an evaluation board for the MAX31865 converter and 4-wire PTFE-encapsulated PT100 temperature probe (BOLA, part No. P1750-15) installed via swivelling screw fitting (BOLA, part No. D692-24 for GL18 glass thread) on the round bottom flask. An installation example is shown in Figure 8.

10 *IP Camera*

For video recording and computer vision applications, we have selected the UniFi Video Camera G3 Dome (UCV-G3-DOME) which features a wide-angle lens and 1080p high resolution video performance. Following the general ChemPU hardware design principle, the device is compatible with the PoE standard and has an assigned IP address. For a miniaturized ChemPU rig a single camera is sufficient while for a standard ChemPU rig one camera is used to monitor the liquid handling backbone while a second camera monitors the various modules. Passive video monitoring as well as CV-based anomaly detection (details given herein) are available via the `start_recording` method of the `Chempiler` class.

Dynamic Process Steps

- 20 The *ChemputationOptimizer* software is designed to take further advantage of the χ DL dynamic step by leveraging a set of optimization algorithms to dynamically update the procedure parameters based on an end-point measurement obtained from a given analytical instrument or sensor.
- 25 There are three subclasses of dynamic step: `DynamicAdd` – for dynamic reagent addition followed by the feedback from the process sensors; `DynamicTransfer` – for dynamic liquid transfer through pump-valve backbone system with passive feedback from tube liquid sensor; `DoUntil` – for dynamic reaction execution (i.e., stirring or heating) with active feedback from the process sensors. The dynamic step execution is achieved through the use of `on_start`, `on_continue` and `on_finish` method, which each return a list of steps to be executed based on the current state of the step.

Dynamic Addition

- The `DynamicAdd` step is designed to control the reagent addition by the feedback of the process sensor installed on the target vessel. Examples of process sensors include temperature (for exothermic or endothermic reactions), RGB (for reaction involving colour change during the reagent addition) or any other sensor which can be controlled using the XDL framework. The following substeps are executed during `DynamicAdd` execution:

- on_start method returns initial set of substeps: priming pump for addition, starting stirring (if required) and starting the background monitoring thread for the target sensor.
- 5 • on_continue method returns the following steps based on the current state of the execution: 1) filling the syringe (if the volume transferred is less than target volume), 2) adding the aliquot to the target vessel (if the sensor reading is within the safe margin from the threshold) or waiting for settling time (if the sensor reading is outside the safe margin limits), 3) obtaining new reading from the sensor.
- 10 • on_finish method returns the steps in analogy to the Add XDL step: flushing the tubing, stopping the stirring (if required) and stopping the background monitoring thread for the target sensor.

The DynamicStep step shares common arguments with XDL Add step (e.g., reagent and volume) with addition of several arguments respective to the sensor used and execution control.

15

An example of the step is shown below, for dynamic addition of hydrogen peroxide solution to the reactor vessel with temperature (RTD probe) sensor feedback. The following arguments are introduced for sensor and execution control: reading_threshold – sensor reading threshold (degrees Celsius in this example); aliquot_volume – an aliquot for a single addition before the next reading will be acquired; wait_after_addition – delay between addition and sensor reading; settle_delay – delay between several acquisitions if the reading is close to the threshold.

20

25

30

```
<AddDynamic
vessel="reactor"
reagent="30 % hydrogen peroxide solution"
volume="19.8 mL"
sensor="rtd"
reading="temperature"
reading_threshold="40.0"
aliquot_volume="5.0 mL"
wait_after_addition="10"
settle_delay="20"
/>
```

35

Dynamic Transfer

The DynamicTransfer step is designed to monitor the liquid transferring process using tube liquid sensor. In the current implementation this step is for passive monitoring and no decisions are made during its dynamic execution. However preliminary results have shown that it is possible to incorporate the data from the sensor into the workflow and correct the liquid transferring process, when potential errors are encountered, e.g.: blocked tube due to

40

solid precipitation, insufficient reagent, blocked filter frit during filtration process, etc. The following substeps are executed during DynamicTransfer execution:

- on_start method returns substeps in analogy to Transfer XDL step: stop stirring.
- 5 • on_continue method returns the following steps based on the current state of the execution: 1) start background monitoring thread for the liquid sensor and increase the reading frequency to 10 Hz, 2) move calculated volume (based on the target volume, transferred volume and maximum syringe volume) to the aspiration syringe pump, i.e., closest to the starting vessel for this liquid transferring step, 3) calculate the amount of liquid transferred and report the calculation, 4) dispense the calculated volume to the target vessel and set the sensor reading frequency to 0.5 Hz, 5) calculate the new volume to transfer.
- 10
- on_finish method returns the steps in analogy to the Transfer XDL step: flushing the tubing, rinsing (if required) and stopping the background monitoring thread for the liquid sensor.
- 15

The DynamicTransfer step shares attributes with Transfer XDL step, and all parameters for the sensor control are constant values.

- 20 An example of the DynamicTransfer step is shown below, for a dynamic transfer of 13ml from reactor to waste, rinsing with 1 ml of water upon completion.

```
25 <DynamicTransfer
  from_vessel="reactor"
  to_vessel="waste_H"
  volume="13 mL"
  rinsing_solvent="water"
  rinsing_volume="1 mL"
/>
```

Dynamic Execution

- 30 The DoUntil step is designed to allow for dynamic reaction execution (i.e., steps where reaction takes place, stirring or heating) with the feedback from the process sensor. An example for such process sensor is an RGB sensor for the reaction with colour extensive colour change. The step is used as wrapper around the corresponding reaction step and the following substeps are executed during DoUntil execution:

- 35
- on_start method returns all substeps from the child step until the Wait step is encountered and additionally a step to start the sensor in the background thread.
 - on_continue method acquires and analyses readings from the sensor, and returns Wait step if reaction is not finished or no steps if reaction is complete.

- on_finish method returns the remaining substeps from the child step.

An example of the DoUntil step is shown below (represented as Stir step) guided by the feedback from the RGB sensor. The following attributes are used to indicate the sensor parameters: threshold – threshold value indicating that the process is over, reading – the name of the reading to monitor (in this example it indicates the clear channel of the colour sensor).

```
<DoUntil sensor="rgbc" threshold="-3000" reading="clear">  
  <Stir  
    vessel="reactor"  
    time="30 mins" />  
</DoUntil>
```

Endpoint detection methods

Current implementation of the DoUntil step utilizes two possible methods to detect the process endpoint (i.e., the reaction completion):

1. Comparison of the current sensor reading (optionally averaged within indicated time frame) with pre-set threshold value. The threshold value may also be set during the procedure execution, using the specialized XDL step TakeBackgroundReading. In a typical use case, such step should be added before the addition of the last reagent or reactant.
2. Comparison of the current sensor reading with previously obtained to detect no change within the indicated time frame. The algorithm is as follows: 1) apply a median filter to the raw sensor readings (within interval M); 2) calculate the average value for the interval = N ; 3) calculate the gradient for the several last averages (i.e., $N-1$, $N-2$, ...); 4) detect the end point, where the K last points of the gradient lay within $J\%$ of the current absolute reading. Parameters M , N , K and J can be set as the DoUntil step properties.

An example of option 2 using the colour RGB sensor is shown in Figure 11, where the values are averaged over 300 seconds, the gradient for the several last averages are used to detect the end point where the number of points fall within $\pm 1\%$ of the absolute reading.

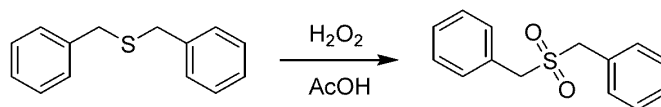
Example Dynamic Execution Reactions

It was found that sensors could be used for dynamic execution of synthetic procedures. This ability was exemplified in two use cases: i) maintaining the internal temperature below a given threshold during an exothermic thioether oxidation ii) waiting for a colour change, thus dynamically adjusting the reaction time, in an iodine-consuming nitrile formation reaction.

Thioether Oxidation

The oxidation of dibenzyl sulphide to dibenzyl sulfone was selected to highlight the benefits of incorporating a temperature (RTD) probe into the ChemPU stack in the context to dynamic addition steps.

5



The literature procedure was translated to XDL, however the key step of this process is the temperature-controlled addition of hydrogen peroxide. The literature procedure states that this step should be performed with the internal temperature not exceeding 75 °C. To emulate this behaviour in the ChemPU the AddDynamic step with the following parameters was used:

10

15

20

25

```
<AddDynamic
  vessel="reactor"
  reagent="30 % hydrogen peroxide solution"
  volume="35.0 mL"
  sensor="rtd"
  reading="temperature"
  reading_threshold="75.0"
  safety_margin="10.0"
  aliquot_volume="2.0 mL"
  wait_after_addition="10.0"
  settle_delay="10.0"
  trend="ascending"
/>
```

Preliminary experiments on a 2.5 g scale showed that the process was strongly exothermic but did not generate enough heat to cross the specified threshold in our setup, however a 10-fold scale up would require dynamic temperature control to prevent the thermal runaway.

30

35

The reactor was manually charged with 25 g of dibenzyl sulphide. The subsequent actions were performed in a fully automated fashion: The starting material was dissolved in glacial acetic acid (100 mL) at 30 °C. After cooling to room temperature, a total of 35 mL of 30% aqueous hydrogen peroxide solution was added in 2 mL portions while maintaining the internal reaction temperature below 75 °C. After the addition, the reaction mixture was stirred at 85°C for 3 hours and subsequently cooled back to room temperature, allowing the product to precipitate.

40

The exothermic addition was controlled using the dynamic XDL step. As soon as the temperature reached the threshold, the addition paused until a safe temperature for continued addition was reached. The passive monitoring further highlighted an uncontrolled exotherm

upon heating the reaction mixture to 85 °C as required by the literature procedure. This example shows the crucial importance of sensor feedback for safe and scalable processes. The crystallized product in the reactor flask was manually filtered, washed with cold water and dried under vacuum (27.26 g, 94.9%).

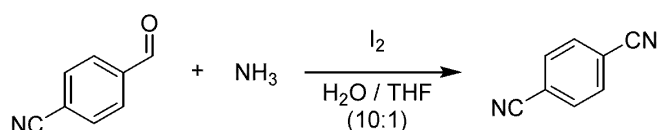
5

¹H NMR (400 MHz, CDCl₃): δ 7.81 – 6.83 (m, 10H), 4.05 (s, 4H).

¹³C NMR (101 MHz, CDCl₃): 131.0, 129.1, 129.1, 127.6, 58.1

Nitrile Formation

10 An iodine-consuming nitrile formation reaction was chosen to illustrate the utility of the colour (RGB) sensor in the context of dynamic procedure execution.



15 The general literature procedure was translated to XDL. Instead of specifying a reaction time, the authors noted that the reaction was stirred until the colour disappeared. This was emulated using the following step sequence featuring the dynamic DoUntil step:

```
20 <TakeBackgroundReading
    sensor="rgb"
    reading="clear" />
<Add
    vessel="reactor"
    reagent="1M iodine thf solution"
    volume="1.1 mL"
25    stir="True"
    priming_volume="1.0 mL"
    aspiration_speed="1"
/>
30 <DoUntil sensor="rgb" threshold="0" margin="100" reading="clear">
    <Stir
        vessel="reactor"
        stir_speed="300"
        time="60 mins" />
</DoUntil>
```

35

The reactor was manually charged with 4-Formylbenzonitrile (131 mg, 1.0 mmol, 1.0 eq.). The subsequent actions were performed in a fully automated fashion: Aqueous ammonia solution (~25%, 10 mL) was added. A background measurement for the RGB sensor was taken. 1 M Iodine in THF solution (1.1 mL, 1.1 mmol, 1.1 eq.) was added. The reaction was
40 stirred until the RGB sensor reading reached the background level within the margin specified

in the procedure, indicating the full consumption of iodine reagent and end point of the reaction. The reaction mixture was quenched using 5% aqueous sodium thiosulfate solution and subsequently extracted using 2 × 15 mL of diethyl ether using the reactor and separator modules. The combined organic phases were collected in an empty holding flask.

- 5 Caution: The reaction of ammonia and iodine forms nitrogen triiodide, an extremely sensitive contact explosive. As such, the process described above should be performed with a minimal excess of iodine and quenched with excess thiosulfate.

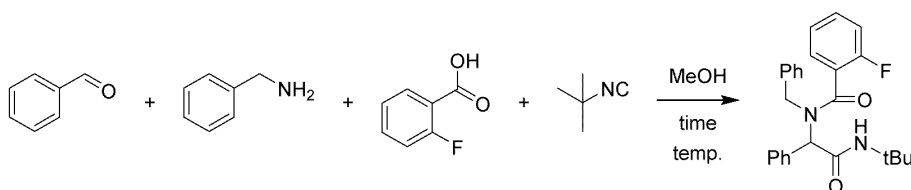
10 The combined organic phases obtained in this process concentrated manually, yielding pure product (90.2 mg, 70.4%).

¹H NMR (400 MHz, d₆-DMSO): δ 8.04 (s, 4H).

¹³C NMR (101 MHz, d₆-DMSO): δ 133.44, 117.80, 115.92

Example Optimisation Reactions

15 Four-Component Ugi Reaction Optimization



20 The four-component Ugi reaction was selected as a complex synthesis with a large number of input parameters. The original procedure²³ was translated into XDL file and amended for running iterative optimization. All reagents were used neat unless otherwise stated. The liquid handling system of the ChemPU was equipped with 1/16" PTFE tubing and corresponding selection valves (IDEX V-240) for precise reagent delivery. 2-Fluorobenzoic acid was selected to allow reaction analysis using ¹⁹F NMR. The hardware setup is shown in Figure 11.

25 The reaction was analyzed using ¹⁹F NMR and the optimization was set to maximize the peak of the Ugi product with respect to the 1,4-difluorobenzene as internal standard.

Synthesis Description

30 Three-neck 25-mL round bottom flask (reactor) equipped with reflux condenser, glass stopper, tubing connector to a liquid handling system, DrySyn© aluminium block and a magnetic stirrer bar. In the beginning of the procedure the liquid handling system was washed with methanol. Benzaldehyde (0.2 mL, 1.96 mmol) was added automatically, following by benzylamine (0.10 - 1.00 mL, 0.92 - 9.20 mmol). The reaction mixture was stirred for 0.0 - 35 30.0 minutes and 2-fluorobenzoic acid (2.0 M solution in methanol, 0.50 - 3.00 mL, 1.00 - 6.00 mmol) was added, following by isocyanide (0.10 - 1.00 mL, 1.63 - 16.3 mmol). The

resulting mixture was stirred for 2.0 - 18.0 hours at 25.0 - 60.0 °C. Thereafter the 1,4-difluorobenzene (0.2 M solution in DCM, 5.00 mL, 1.00 mmol) was added and the sample (2.5 mL) of the resulting mixture was transferred to the NMR for analysis. Upon analysis completion, the sample was transferred back to the flask and all its contents was moved to an empty flask for storage. The reactor was cleaned twice with DCM (15 mL) and used for the next iteration.

Optimisation Description

For the iterative closed-loop optimization the following reaction parameters were set:

- 10 • Benzaldehyde volume – fixed at 0.2 mL (to ensure maximum product amount of 2.0 mmol).
- Methanol volume – variable within 0.0 and 5.0 mL range.
- Benzylamine volume – variable within 0.1 and 1 mL range.
- Imine formation time – variable within 0.0 and 1800.0 seconds range.
- 15 • 2-Fluorobenzoic acid volume (2.0 M in methanol) – variable within 0.5 and 3.0 mL range.
- *tert*-Butylisocyanide volume – variable within 0.1 and 1.0 mL range.
- Reaction temperature – variable within 25.0 and 60.0 °C range.
- Reaction time – variable within 7200.0 and 64800.0 seconds.
- 20 • 1,4-Difluorobenzene (0.2 M solution in DCM) – fixed at 5.0 mL (as 1:2 ratio with benzaldehyde).

The procedure was executed using interactive python environment, each experiment consisted of 5 consecutive reactions with manual reagent refilling in between (or earlier if required by the experiment). The following strategies for optimization were used (full configuration files are provided below): 5 random search, 14 SMBO explorative experiments, 5 SMBO balanced search and 6 SMBO exploitation experiments. The reaction was analyzed using ¹⁹F NMR and the optimization was set to maximize the peak of the Ugi product (at -114.3 ppm), with 1,4-difluorobenzene as internal standard (-119.2 ppm).

30 Configuration file to run four-component Ugi reaction optimization during initial random parameter screening:

```
{
  "max_iterations": 5.0,
  "target": {
    "spectrum_peak-area_-114.3": "inf"
  },
  "algorithm": {
    "name": "random",
```

```
    "random_state": 42
  },
  "reference": -119.2
}
```

5

Configuration file to run four-component Ugi reaction optimization using SMBO algorithm with acquisition function favouring exploration:

```
{
  "max_iterations": 5.0,
  "target": {
    "spectrum_peak-area_-114.3": "inf"
  },
  "algorithm": {
    "name": "smbo",
    "base_estimator": "GP",
    "n_initial_points": 2,
    "acq_func": "PI",
    "acq_func_kwargs": {},
    "random_state": 42
  },
  "reference": -119.2
}
```

10

15

20

Configuration file to run four-component Ugi reaction optimization using SMBO algorithm with acquisition function probabilistically chosen between “LCB”, “EI” or “PI” (referred to as “balanced” strategy, see details in the scikit-optimize API reference⁷):

25

```
{
  "max_iterations": 5.0,
  "target": {
    "spectrum_peak-area_-114.3": "inf"
  },
  "algorithm": {
    "name": "smbo",
    "n_initial_points": 2,
    "acq_func": "gp_hedge",
    "random_state": 42
  },
  "reference": -119.2
}
```

30

35

40

Configuration file to run four-component Ugi reaction optimization using SMBO algorithm with acquisition function favouring exploitation:

```
{
  "max_iterations": 3.0,
  "target": {
```

45

```
    "spectrum_peak-area_-114.3": "inf"  
  },  
  "algorithm": {  
    "name": "smbo",  
    "n_initial_points": 2,  
    "acq_func": "LCB",  
    "acq_func_kwargs": {"kappa": 0.1},  
    "random_state": 42  
  },  
  "reference": -119.2  
}
```

Optimisation Results

The optimization was terminated after 30 iterations, achieving 38% yield increase compared to the literature method.

The product was characterised as follows. The spectroscopic data is in agreement with that previously published²⁴.

¹H NMR (400 MHz, CDCl₃): δ 7.46 – 7.40 (m, 2H), 7.33 – 7.15 (m, 7H), 7.10 – 7.00 (m, 4H), 6.87 – 6.82 (m, 1H), 5.92 (s, 1H), 5.72 (s, 1H), 4.63 (d, J = 16.4 Hz, 1H), 4.41 (d, J = 16.4 Hz, 1H), 1.35 (s, 9H).

¹³C NMR (101 MHz, CDCl₃): δ 168.5, 168.2, 158.3 (d, J_{CF} = 247.3 Hz), 136.9, 135.1, 131.2 (d, J_{CF} = 7.9 Hz), 129.6, 128.8, 128.7, 128.6, 128.5, 128.3, 127.8, 127.1, 124.8 (d, J_{CF} = 17.4 Hz), 124.5 (d, J_{CF} = 3.0 Hz), 115.9 (d, J_{CF} = 21.1 Hz), 64.4, 52.2, 51.7, 28.7.

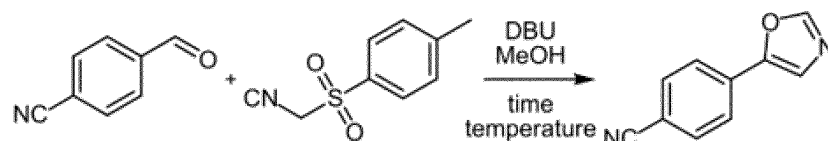
¹⁹F NMR (376 MHz, CDCl₃) δ -115.1.

The optimisation results for each iteration are shown in Table 4.

In Table 4, Objective ("spectrum_peak-area_-114.3") refers to the integration area of the corresponding peak in the spectrum, where 0 indicates no peak found. For lines marked "a", the objective was falsely assigned to the side product peak, leading to incorrect results during "exploitation" approach.

Table 4: Results of the four-component Ugi reaction optimization

Iteration	Add_3- volume	Add_4- volume	Stir_6- time	Add_7- volume	Add_10- volume	HeatChillToTemp_12- temp	Stir_13- time	spectrum_peak-area_- 114.3	Strategy
1	1	0.22	600	1	0.23	20	64800	0.3997	initial
2	4.89	0.41	732.23	1.01	0.77	20	20999.27	0.3513	random
3	0.63	0.49	1226.77	0.93	0.69	20	11612.19	0.3596	random
4	0.08	0.23	1501.71	2.29	0.48	50.34	26902.12	0.2097	random
5	2.17	0.44	1394.91	0.77	0.96	27.24	62775.59	0.2513	random
6	1.82	0.37	1583.79	2.94	0.74	28.65	44660.21	0.3332	random
7	2.17	0.43	1182.48	0.56	0.39	28.83	32768.74	0.1688	exploration
8	1.68	0.98	642.7	2.15	0.81	31.01	42750.15	0.2557	exploration
9	1.34	0.53	179.28	2.28	0.34	37.38	55827.88	0.4	exploration
10	5	1	1800	0.5	1	25	7200	0	exploration
11	4.33	0.67	1453.24	0.51	0.43	25.89	54755.64	0	exploration
12	0.53	0.37	139.83	1.56	0.11	42.38	13629.76	0.2327	exploration
13	0.87	0.25	1342.07	2.22	0.3	56.59	16607.29	0.2652	exploration
14	5	0.1	0	3	0.1	60	64800	0	exploration
15	1.21	0.34	226.62	1.66	0.46	52.66	13751.27	0.3467	exploration
16	3.18	0.39	1171.8	2.18	0.26	52.13	52989.08	0.3723	exploration
17	0.45	0.32	534.94	0.96	0.57	27.51	60105.29	0.2906	exploration
18	1	0.41	1478.19	0.93	0.47	49.16	27644.73	0.2656	exploration
19	5	0.26	438.42	0.53	1	60	49493.84	0.2134	exploration
20	3.51	0.42	28.88	1.87	1	60	16979.32	0.5124	exploration
21	0	0.97	184.89	2.65	0.79	25	54012.88	0.2309	exploitation
22	0	1	0	3	0.1	25	64800	0.2011	exploitation
23	0	1	0	3	0.1	25	64800	0.2048	exploitation
24	5	0.29	1044.02	1.96	0.46	52.69	52023.27	0.4144	balance
25	0	0.15	1137.27	1.1	0.23	25	40311.72	0	balance
26	5	0.44	1484.39	2.12	1	50.48	11831.81	0.5019	balance
27	5	0.46	967.27	3	0.1	26.78	8526.74	0.2113	balance
28	5	0.49	1138.17	1.76	1	57.38	31964.02	0.5035	balance
29	5	0.45	920.79	1.84	1	44.33	22422.12	0.5059	exploitation
30	5	0.43	657.43	1.94	1	55.74	16045.76	0.5506	exploitation
31	4.33	0.42	1245.2	1.89	1	41.11	15034.47	0.5408	exploitation

Van Leusen Oxazole Synthesis Optimization

The Van Leusen oxazole reaction was selected as an example for the synthesis of heterocyclic building blocks, common in the medicinal chemistry literature. The procedure was loosely based on a patent²⁵, translated into XDL and adapted for running iteratively. Notably, it features a small volume addition (as low as 150 μ L) that commonly requires manual calibration or optimization. To make the problem more challenging, the purity of the product mixture was optimized as a secondary objective to implicitly minimize excess reagent used. The liquid handling system of the ChemPU was equipped with 1/16" PTFE tubing and corresponding selection valves (IDEX V-240) for precise reagent delivery. Naphthalene was selected as an internal standard to allow reaction analysis using HPLC-DAD at 254 nm. The hardware setup is shown in Figure 12.

15 Synthesis Description

The process was executed automatically using 2 parallel reactors. 0.25 M TosMIC in MeOH solution (4.10 - 6.15 mL), 0.25 M 4-formylbenzonitrile in MeOH (4.1 mL, containing 0.05 M naphthalene as an internal standard), neat DBU (0.15 - 0.31 mL) and methanol (5 mL) were added to the reactor. The reaction mixture was stirred for 30 - 180 minutes at 25.0 - 75.0 °C. After cooling to room temperature, a sample is withdrawn from the reactor, 40 times diluted in an empty flask, and subsequently loaded onto a 5 mL sample loop and injected into the HPLC. The remaining volume of the reaction mixture was discarded, and the platform reset by cleaning all modules with methanol and/or acetonitrile.

25 The original procedure was translated into χ DL, amended for iterative optimization, and executed on a Chemputer platform equipped with the HPLC and two independent reactor modules, each consisting of three-neck 25-mL round bottom flasks with reflux condenser, glass stopper, tubing connector to a liquid handling system, DrySyn© aluminium block and a magnetic stirrer bar. The optimization target was to maximize the peak area of the product with respect to naphthalene as internal standard while simultaneously minimizing impurities. 30 The SNOBFIT algorithm was used as implemented in the Summit framework through the client-server interface.

Optimisation Description

35 For the iterative closed-loop optimization the following reaction parameters were set:

- TosMIC volume – variable within 4.10 and 6.15 mL range (1.0 - 1.5 eq.)
- DBU volume – variable within 0.15 and 0.31 mL range (1.0 - 2.0 eq.)

- Reaction temperature – variable within 25.0 and 75.0 °C range.
- Reaction time – variable within 1800.0 and 10800.0 seconds.

The optimization target was set as the weighted sum of the peak purity and product peak area (14.5 min retention time) relative to the peak area of the internal standard (naphthalene, 20 min retention time), see **Equation 4** where *AUC* refers to the area under the curve for a given peak, *w* is a weighing factor set to 0.7 and *c* is constant to scale the two objective values to the same order of magnitude set to 0.1. The optimization algorithm was SNOBFIT as implemented in the Summit framework.

$$Target = w * c * \frac{AUC_{product}}{AUC_{reference}} + (1 - w) \frac{AUC_{product}}{\sum_{i=1}^N AUC_{Peak i}} \quad (\text{Equation 4})$$

Configuration file to run Van Leusen oxazole synthesis using SNOBFIT algorithm via the Summit server framework with 2 parallel reactors:

```

15 {
    "max_iterations": 25.0,
    "target": {
        "multie_purity-area_14": "inf"
    },
20 "algorithm": {
    "name": "SNOBFIT",
    },
    "reference": 20.1,
    "batch_size": 2
25 }
```

Optimisation Results

The optimization campaign was concluded after 26 iterations, achieving a 10% increase in the weighted objective when comparing the best to the first iteration or 55% when comparing the best to the worst iteration.

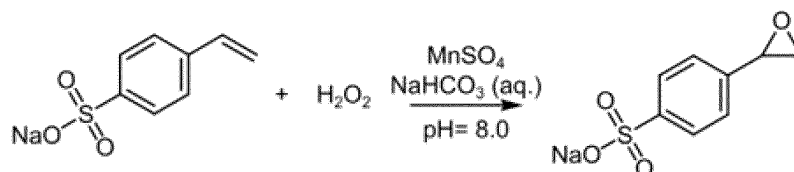
The product was characterised as follows.

¹H NMR (400 MHz, CDCl₃): δ 7.92 (s, 1H), 7.79 – 7.56 (m, 4H), 7.44 (s, 1H).
35 ¹³C NMR (101 MHz, CDCl₃): δ 151.66, 149.87, 132.96, 131.80, 124.81, 124.33, 118.55, 112.10.

The optimisation results for each iteration are shown in Table 5.

Table 5: Results of the Van Leusen Oxazole optimization

Iteration	Add_0-volume	Add_2-volume	HeatChillToTemp_4-temp	Stir_5-time	multie_purity-area_14	Rel. Peak Area	Purity
1	5.930363	0.281536	53.90767	5374.944	0.936862	11.12758	0.69275
2	5.125	0.231	60	1800	1.001846	12.09507	0.742094
3	5.577005	0.170647	56.5265	10761.03	0.960348	10.98935	0.717526
4	4.374598	0.23139	26.359	8101.53	1.014233	11.51263	0.764424
5	5.235639	0.155013	56.21	3598.83	1.018665	12.07292	0.750825
6	4.205514	0.251755	47.3135	9444.6	0.937303	10.6126	0.726961
7	5.552159	0.188565	39.0535	2673.63	1.03046	11.84036	0.754814
8	6.0278	0.252029	62.8975	6748.92	0.931288	10.71362	0.689388
9	4.36937	0.211784	70.9005	4508.82	0.961356	10.89795	0.737969
10	5.751849	0.210262	29.935	6050.52	1.022144	11.72645	0.747777
11	4.878836	0.154	25	1800	0.951958	10.57003	0.755554
12	6.14998	0.281537	25	10761.03	1.001754	11.5991	0.731457
13	6.14998	0.281537	25	1800	1.011689	11.66346	0.754415
14	4.672586	0.259892	53.8685	5742.81	0.946677	10.72659	0.724072
15	4.36937	0.281537	38.442	1800	1.023849	11.6238	0.773037
16	4.672586	0.259892	53.8685	5742.81	0.943857	10.72815	0.725548
17	6.143563	0.154	46.189	1800	1.024607	11.73885	0.753196
18	4.590647	0.202877	61.157	8210.7	0.957861	10.94852	0.729597
19	4.36937	0.189985	47.9565	1800	1.028457	11.62418	0.776575
20	4.675189	0.192694	31.5975	5844.51	1.025144	11.56126	0.769117
21	6.14998	0.202228	50.8665	1800	0.665637	6.807097	0.626507
22	5.987927	0.174625	25	4362.12	0.83316	9.003038	0.698722
23	5.6375	0.1925	67.5	2693.7	0.994121	11.24234	0.739176
24	4.36937	0.281537	25	1800	1.03148	11.59027	0.782197
25	5.253412	0.308	25	3598.83	1.028255	11.57735	0.76337
26	4.6125	0.1925	67.5	2693.7	0.99393	11.11633	0.752316

Styrene Sulfonate Oxidation Optimization

The Mn-catalysed epoxidation of olefins is a reaction of industrial interest²⁶ and has been extensively described by Burgess et al.²⁷ for a variety of substrates. This example was selected because the active catalytic species is formed in situ and under the wrong conditions, manganese gets oxidized and decomposes hydrogen peroxide, leading to poor conversions.²⁸ In our optimization we try to find optimal conversion conditions by changing 4 major parameters of the reaction. The liquid handling system of the ChemPU was equipped with 1/16" PTFE tubing and corresponding selection valves (IDEX V-240) for precise reagent delivery. Acetonitrile was used as an internal standard and the reaction was analyzed Raman spectroscopy. The hardware setup is shown in Figure 13.

Synthesis Description

5 mM MnSO₄ in water solution (0.5 - 5.0 mL), 0.5 M styrene sulfonate in water solution (5.0 mL, 2.5 mmol, 1 eq.), and 0.5 M NaHCO₃ in water solution (0.5 - 5.0 mL) were automatically transferred to a reactor vessel, placed in front of a Raman probe. 30% Hydrogen peroxide in water (1.0 - 5.0 mL) was added at a rate of 0.04 - 10.00 mL/min. The reaction mixture was stirred for 1.0 - 10.0 hours. The process was continuously monitored via Raman and single end-point analysis was taken for calculating the optimization target value. The reaction mixture was discarded, and the platform was reset by cleaning the reactor vessel with water and a cleaning solution (1 × 25 mL cleaning solution and 2 × 25 mL water). The cleaning solution was made by mixing 100 mL 30% H₂O₂, 100 mL 60% nitric acid, 800 mL H₂O).

The original procedure was translated into xDL, amended for iterative optimization, and executed on a Chemputer platform equipped with Raman spectrometer monitoring a single reactor module consisting of a three-neck 25-mL round bottom flask with reflux condenser, glass stopper, tubing connector to a liquid handling system, DrySyn® aluminium block and a magnetic stirrer bar. The optimization target was to minimize the area of the peak of the double bond of the starting material at 1633 cm⁻¹ relative to the area of the peak at 2250 cm⁻¹, with respect to the peak of an internal standard (acetonitrile). The deep Bayesian optimizer Phoenix as available through the Olympus framework was chosen as the algorithm.

Optimisation Description

For the iterative closed-loop optimization the following reaction parameters were set:

- Catalyst volume – variable within 0.5 and 5.0 mL range (0.2 - 1.0 mol%).

- NaHCO₃ volume – variable within 0.5 and 5.0 mL range.
 - H₂O₂ volume – variable within 1.0 and 5.0 mL range (4 - 20 eq.).
 - H₂O₂ addition speed – variable within 0.04 and 10.00 mL/min range.
 - Reaction time – variable within 3600.0 and 36000.0 seconds.
- 5 The optimization target was to minimize the area of the peak corresponding to the double bond of the starting material at 1633 cm⁻¹ relative to the area of the peak at 2261 cm⁻¹, corresponding to the triple bond of acetonitrile which was used as an internal standard. Phoenix²⁹ optimization algorithm was chosen; the optimization was run for 40 iterations.
- 10 Configuration file for the optimization of the Sodium 4-vinylbenzenesulfonate catalytic oxidation using Phoenix optimization algorithm:

```
{  
  "max_iterations": 40,  
  "target": {  
15    "spectrum_peak-area_1633": "inf"  
  },  
  "algorithm": {  
    "name": "Phoenix"  
  },  
20  "reference": 2261,  
  "batch_size": 1  
}
```

Optimisation Results

25

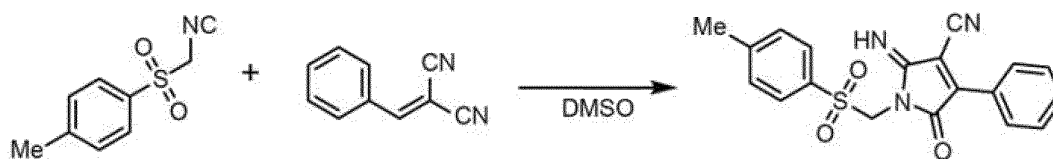
The product was characterised as follows.

- ¹H NMR** (400 MHz, D₂O): δ 7.82 (dd, *J* = 8.3, 1.4 Hz, 2H), 7.46 (d, *J* = 8.4 Hz, 2H), 4.10 (dd, *J* = 4.4, 2.9 Hz, 1H), 3.30 (t, *J* = 4.5 Hz, 1H), 3.05 (dd, *J* = 4.7, 2.9 Hz, 1H).
- 30 **¹³C NMR** (101 MHz, D₂O): δ 142.36, 140.35, 126.36, 125.73, 52.56, 51.49.
- HRMS-ESI**: 199.0076 (Expected: 199.0071)

The optimisation results for each iteration are shown in Table 6.

Table 6: Results for the Styrene Sulfonate Oxidation catalytic oxidation optimization

Iteration	Add_3 volume	Add_4- volume	Add_7- dispense_speed	Add_7- volume	Stir_8- time	spectrum_peak- area_1633
1	3.5	5	0.2	5	1800	-0.4779
2	4.9674	4.71	7.7265	3.3728	2388.626	-0.5818
3	3.4371	4.9868	0.3145	4.9603	1869.3718	-0.6005
4	2.2097	1.958	3.0766	4.7902	5470.981	-0.6115
5	3.4669	4.993	0.2603	4.9791	1836.5115	-0.5505
6	4.2991	4.9903	6.9041	0.8012	3924.3557	-0.4984
7	3.5232	5	0.1512	4.9846	1800	-0.4956
8	2.6742	2.5729	6.9966	3.5736	7151.0801	-0.5769
9	3.4835	4.9965	0.2301	4.9896	1818.2557	-0.4694
10	1.1514	3.5737	4.1572	2.1582	3611.7249	-0.6829
11	3.4905	4.9984	0.2401	4.9955	1827.9069	-0.1065
12	3.4795	4.9887	0.2172	4.9936	1813.7491	-0.487
13	3.4489	0.5034	2.1559	0.723	1832.848	-0.7141
14	3.4959	5	0.2475	4.9875	1830.6141	-0.4692
15	3.5172	1.4056	9.2389	3.4393	6216.7085	-0.3913
16	3.6716	4.9948	0.4108	4.9123	1820.5065	-0.5042
17	1.8718	2.342	3.604	2.7266	6713.3198	-0.5905
18	3.4974	4.9916	0.25	4.9978	1827.2998	-0.4758
19	3.1952	0.7773	1.0382	1.5286	4568.1348	-0.5463
20	3.4868	4.9971	0.2442	5	1833.4457	-0.4422
21	2.2324	3.1281	5.4343	4.0445	6846.0322	-0.5695
22	3.4943	4.9994	0.2325	4.9989	1824.4208	-0.4662
23	0.6164	4.0719	0.3935	1.2007	6975.1602	-0.7717
24	3.4915	4.9935	0.2369	4.9967	1823.3331	-0.4991
25	4.8554	4.3793	4.5024	4.4297	2804.3989	-0.4978
26	3.6246	4.8424	0.0539	4.8127	1829.4752	-0.2498
27	2.9257	1.6921	6.439	1.9837	5734.5981	-0.6364
28	3.5829	4.8928	0.0884	4.8994	1825.9874	-0.3692
29	2.3663	2.667	8.4578	3.866	4308.7773	-0.7029
30	3.4929	4.9978	0.2425	4.9918	1828.5566	-0.5252
31	1.5091	1.1805	3.9048	3.408	5943.9067	-0.8197
32	3.4889	4.9967	0.2389	4.9944	1822.7146	-0.6165
33	3.2902	0.6341	9.7743	1.8064	3343.2942	-0.7428
34	3.4877	4.9986	0.2409	4.9954	1831.5872	-0.6499
35	3.8507	3.0556	2.102	2.8947	5062.3755	-0.5077
36	3.4937	4.9995	0.2456	4.993	1827.9785	-0.6288
37	1.66	3.3253	1.7386	2.2343	2583.7048	-0.7448
38	3.3586	4.8507	0.0881	4.9948	1829.0911	-0.4967
39	2.0985	2.4198	4.9677	1.3539	3030.051	-0.7682
40	3.4878	4.9952	0.2376	4.9974	1832.2788	-0.6978

tosMIC Reaction Optimization

The reaction between tosMIC and benzylidenemalononitrile was selected for optimisation.

- 5 The reaction was subjected to optimization, using HPLC to provide necessary feedback to the decision-making algorithms. All reagents were as 1.0 M solution in DMSO, prepared freshly before a set of 5 consecutive iterations. Naphthalene was used as internal standard for quantitative HPLC analysis. The hardware setup is shown in Figure 14.
- 10 The original procedures was obtained from the algorithmic reaction discovery process, amended for iterative optimization, and executed on a Chemputer platform equipped with HPLC. The tosMIC reaction: for the parameter optimization were used: 10 experiments with random search strategy, and 22 SMBO explorative experiments. The reaction was analyzed HPLC and the optimization was set to maximize the peak of the product with respect to the
- 15 naphthalene as internal standard.

Synthesis Description

- 1.0 M TosMIC in DMSO solution (variable), 1.0 M benzylidenemalononitrile in DMSO (1 mL) and DMSO (variable) were added to the reactor. The reaction mixture was stirred for a
- 20 variable amount of time at a variable temperature. After cooling to room temperature, reference was added to the mixture (1 mL 1.0 M naphthalene) and afterwards a sample (0.5 mL) is withdrawn for the reactor, transferred to an empty flask, diluted to 10.0 mL, and subsequently loaded onto a 5 μ L sample loop and injected into the HPLC. The remaining
- 25 volume of the reaction mixture was discarded, and the platform reset by cleaning all modules with DMSO.

Optimisation Description

- 30 For the iterative closed-loop optimization the following reaction parameters were set:
- Toluenesulphonylmethyl isocyanide – variable within 0.2 and 2.0 ml range.
 - DMSO volume – variable within 0.0 and 5.0 mL range.
 - Pre-heat reaction temperature – variable within 30.0 and 160.0 °C range.
 - Reaction time – variable within 3600.0 and 28800.0 seconds.
- 35 Experiments were run as follows:

1. Initial exploration of the product space using random search within given parameter limits.
2. Exploration of the product space using SMBO algorithm (with parameters setup to favour exploration over exploitation).
- 5 3. Identification of the obtained materials as identified regions of interest on the resulting HPLC chromatograms. For each experiment result was calculated as the integration area of the target region divided by the sum of all other regions identified.
- 10 4. Individual optimization using SMBO algorithm (with parameter setup to favour exploitation over exploration) and product peak integration area as a target. The data from all previous experiments (i.e., set of parameters and corresponding results) were used in each subsequent optimization.

An example of the configuration file used to run the optimization of the tosMIC reaction:

```
15 {  
    "max_iterations": 12,  
    "target": {  
        "spectrum_integration-area_16.7..16.9": Infinity  
    },  
    "algorithm": {  
20     "name": "smbo",  
        "acq_func": "gp_hedge",  
        "n_initial_points": 0  
    },  
    "reference": 18.2,  
25 "constraints": [  
        "14.0..14.4",  
        "15.5..15.7"  
    ],  
    "batch_size": 1,  
30 "control": {  
        "n_runs": 0  
    }  
}
```

35 Optimisation Results

The optimization campaign was concluded after 30 iterations, achieving a 22% increase yield. Yield was calculated from the crude reaction mixture analysis with respect to the internal standard (naphthalene).

- 40 An example of an HPLC chromatogram for the tosMIC reaction is provided in Figure 16, obtained after performing optimization for the identified region.

The product was characterized as follows:

¹H NMR (600 MHz, CDCl₃): δ 8.12 (d, J = 7.2 Hz, 2H), 7.80 (d, J = 8.3 Hz, 2H), 7.62 (t, J = 7.5 Hz, 1H), 7.55 (t, J = 7.7 Hz, 2H), 7.37 (d, J = 7.9 Hz, 2H), 5.15 (s, 2H), 2.46 (s, 3H).

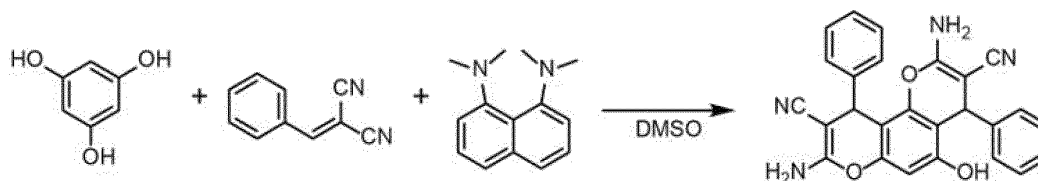
¹³C NMR (151 MHz, CDCl₃): δ 171.31, 165.75, 146.22, 146.02, 130.27, 129.99, 129.53, 128.99, 126.81, 111.59, 60.54, 59.46, 21.91, 21.20, 14.34.

5 **LC-MS** [M+H]⁺ expected: 366.0907, found: 366.0983

The optimisation results for each iteration are shown in Table 7.

Table 7: Results for the tosMIC Reaction Optimization

Iteration	Add_2-volume	Add_3-volume	HeatChillToTemp_4-temp	Wait_5-time	spectrum_integration_area_16.7..16.9	Yield
1	1.35	2.19	102	24100	0.0128	0.296444
2	1.63	2.27	97	5780	0.0253	0.074886
3	1.35	0.49	115	21300	0.00345	0.116744
4	1.4	2.09	108	22600	0.00432	0.217785
5	1.59	0.7	108	20800	0.00684	0.246672
6	1.65	1.97	109	22300	0.00816	0.284063
7	1.34	2.04	36	21000	0.0056	0.29567
8	0.62	2.44	105	21300	0.00367	0.209791
9	1.5	2.52	107	20500	0.00679	0.260611
10	1.56	1.77	53	21400	0.0135	0.286446
11	1.83	2.14	115	21000	0.00956	0.295236
12	1.21	2.06	107	20900	0.0114	0.308481
13	1.61	2.07	111	20400	0.0164	0.333734
14	1.67	2.32	107	21300	0.0171	0.342697
15	1	1	140	14400	0.0093	0.26773
16	1.59	2.19	154	21200	0.00148	0.024947
17	1.2	5	144	26400	0.0112	0.025476
18	0.51	3.47	150	3820	0.0124	0.043266
19	1.88	4.49	42	12200	0.0135	0.036389
20	0.77	4.28	90	9730	0.0259	0.047131
21	1.48	4.14	102	28800	0.0517	0.045649
22	0.35	4.11	85	10500	0.00909	0.03232
23	1.6	2.13	108	28700	0.0422	0.065295
24	1.4	4.22	69	11300	0.0629	0.069571
25	1.61	2.14	106	13100	0.0482	0.07256
26	1.33	4.08	65	17900	0.0619	0.072275
27	1.35	4.06	71	16300	0.0656	0.074072
28	1.36	4.09	73	19300	0.0646	0.085893
29	1.36	4.11	70	9330	0.031	0.369044
30	1.35	4.06	71	16300	0.0329	0.436853
31	1.35	4.06	71	16300	0.0321	0.432427
32	1.35	4.06	71	16300	0.0287	0.431138

Phloroglucinol Reaction Optimization

The reaction between phloroglucinol, benzylidenemalononitrile and 1,8-bis(dimethylamino)naphthalene was selected following the methodology described herein. The reaction was subjected to optimization, using HPLC to provide necessary feedback to the decision-making algorithms. All reagents were as 1.0 M solution in DMSO, prepared freshly before a set of 5 consecutive iterations. Naphthalene was used as internal standard for quantitative HPLC analysis. The optimization protocol consisted of two distinct campaigns (see parameters below). The second campaign was initiated after careful analysis of the results of the first one – with increased limits for 1,8-bis(dimethylamino)naphthalene concentration and temperature. The hardware setup is shown in Figure 14.

Synthesis Description

1.0 M phloroglucinol in DMSO solution (variable), 1.0 M benzylidenemalononitrile in DMSO (1 mL), 0.5 M 1,8-bis(dimethylamino)naphthalene in DMSO solution (variable) and DMSO (variable) were added to the reactor. The reaction mixture was stirred for a variable amount of time at a variable temperature. After cooling to room temperature, reference was added to the mixture (1 mL 1M naphthalene) and afterwards a sample (0.5 mL) is withdrawn for the reactor, transferred to an empty flask, diluted to 10.0 mL, and subsequently loaded onto a 5 µL sample loop and injected into the HPLC. The remaining volume of the reaction mixture was discarded, and the platform reset by cleaning all modules with DMSO.

Optimisation Description

The phloroglucinol reaction: for the parameter optimization were used: 10 experiments with random search strategy, and 40 SMBO explorative experiments. The reaction was analyzed by HPLC and the optimization was set to maximize the peak of the product with respect to the naphthalene as internal standard.

Two campaigns were carried out, having the following reaction parameters.

First campaign, 37 reactions:

- benzylidenemalononitrile – variable within 0.2 and 2.0 ml range.
- 1,8-bis(dimethylamino)naphthalene – variable within 0.4 and 4.0 ml range.
- DMSO volume – variable within 0.0 and 6.0 mL range.
- Pre-heat reaction temperature – variable within 25.0 and 160.0 °C range.

- Reaction time – variable within 3600.0 and 28800 seconds.

Second campaign, 13 reactions:

- benzylidenemalononitrile – variable within 0.2 and 4.0 ml range.
- 5
- 1,8-bis(dimethylamino)naphthalene – variable within 0.4 and 8.0 ml range.
 - DMSO volume – variable within 0.0 and 6.0 mL range.
 - Pre-heat reaction temperature – variable within 0.0 and 160.0 °C range.
 - Reaction time – variable within 3600.0 and 86400.0 seconds.
- 10
- The reason for performing two optimizations originated from careful analysis of the data obtained for the first campaign: 6 out of the 37 reactions performed at 25 °C with maximally allowed 1,8-bis(dimethylamino)naphthalene concentration resulted in yields greater than the yield at initial reaction conditions (27%), suggesting that a more “adventurous” approach may lead to interesting results. The second campaign had been initiated with results from previous
- 15
- campaign serving as input and the temperature and constraints allowing for temperatures as low as 0 °C while allowing the 1,8-bis(dimethylamino)naphthalene higher concentrations. During the first campaign the optimization protocol yielded parameters which resulted in a 49% yield, an increase of 22% from the 27% obtained for the initial reaction conditions. However, with the more opened constraints, the explored space has been successfully
- 20
- expanded into unintuitive areas that would not have been a first choice, or a choice at all, for the experimental chemist: in this example the algorithm moved into temperatures not suited for the solvent used (DMSO freezing point is below 19 °C).

An example of the configuration file used to run the optimization of phloroglucinol reaction:

25

```
{  
  "max_iterations": 12,  
  "target": {  
    "spectrum_integration_area_15.0..15.2": Infinity  
  },  
30  "algorithm": {  
    "name": "smbo",  
    "acq_func": "gp_hedge",  
    "n_initial_points": 0  
  },  
35  "reference": 18.2,  
  "batch_size": 1,  
  "control": {  
    "n_runs": 0  
  }  
40 }
```

Optimisation Results

During the first campaign the optimization protocol yielded parameters which resulted in a 49% yield, an increase of 22% from the 27% obtained for the initial reaction conditions.

5

During the second campaign the optimization protocol yielded parameters which led to a further increase in the reaction yield by 50% with respect to the initial reaction conditions, reaching a yield of 77%.

10 The product was characterized as follows:

¹H NMR (600 MHz, DMSO-*d*₆): δ 10.06 (s, J = 7.2 Hz, 1H), 7.26 (m, 7H), 7.16 (m, 1H), 7.08 (d, J = 7.54 Hz, 2H), 6.98 (s, 2H), 6.71 (s, 2H), 6.28 (s, 1H), 4.64 (s, 1H), 4.45 (s, 1H).

¹³C NMR (151 MHz, DMSO-*d*₆): δ 161.07, 159.91, 154.50, 148.78, 146.45, 146.06, 145.14, 129.02, 128.91, 128.74, 127.68, 127.54, 127.16, 127.09, 126.84, 120.95, 120.72, 109.42, 108.23, 103.71, 98.76, 57.37, 44.61, 40.55, 36.99, 36.67.

15

LC-MS [M+H]⁺ expected: 435.137, found: 434.900

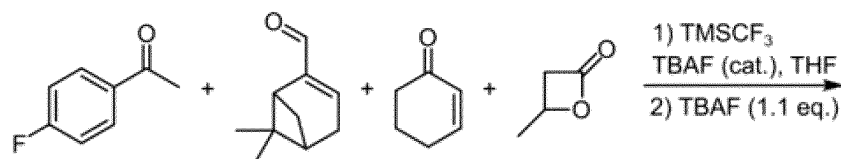
The optimisation results for each iteration are shown in Table 8.

20

Table 8: Results for the Phloroglucinol Reaction Optimization

Iteration	Add_2-volume	Add_3-volume	Add_4-volume	HeatChillToTemp_5-temp	Wait_6-time	spectrum_integration_area_15.0.15.2	Yield
1	1.00	2.00	2.00	25.00	86400.00	3.23	0.27
2	1.00	2.00	2.00	25.00	14400.00	2.28	0.19018
3	1.50	1.11	4.51	98.00	14800.00	2.42	0.20262
4	1.85	3.24	5.68	100.00	8250.00	2.28	0.19031
5	1.54	2.45	0.22	42.00	20100.00	3.72	0.31088
6	1.54	1.37	4.49	132.00	3870.00	1.63	0.13662
7	1.42	0.43	2.71	59.00	17800.00	1.72	0.14338
8	1.13	0.79	4.69	149.00	13500.00	0.651	0.05444
9	1.41	2.16	1.98	63.00	18500.00	2.11	0.17669
10	0.46	3.54	5.99	39.00	28700.00	0.636	0.05313
11	1.85	4.00	0.00	25.00	20500.00	4.85	0.40523
12	2.00	4.00	0.00	108.00	28800.00	5.26	0.43968
13	2.00	4.00	0.00	160.00	3600.00	1.97	0.16475
14	1.00	2.00	2.00	25.00	14400.00	2.28	0.19018
15	1.50	1.11	4.51	98.00	14800.00	2.42	0.20262
16	1.85	3.24	5.68	100.00	8250.00	2.28	0.19031
17	1.54	2.45	0.22	42.00	20100.00	3.72	0.31088
18	1.54	1.37	4.49	132.00	3870.00	1.63	0.13662
19	1.42	0.43	2.71	59.00	17800.00	1.72	0.14338
20	1.13	0.79	4.69	149.00	13500.00	0.651	0.05444
21	1.41	2.16	1.98	63.00	18500.00	2.11	0.17669
22	0.46	3.54	5.99	39.00	28700.00	0.636	0.05313
23	1.85	4.00	0.00	25.00	20500.00	4.85	0.40523
24	2.00	4.00	0.00	108.00	28800.00	5.26	0.43968
25	2.00	4.00	0.00	160.00	3600.00	1.97	0.16475

Iteration	Add_2-volume	Add_3-volume	Add_4-volume	HeatChillToTemp_5-temp	Wait_6-time	spectrum_integration_area_15.0.15.2	Yield
26	2.00	2.55	1.79	44.00	28800.00	3.77	0.3149
27	1.98	4.00	0.00	89.00	27700.00	3.13	0.26183
28	0.20	2.80	0.00	36.00	28800.00	0.352	0.02946
29	2.00	4.00	6.00	25.00	28800.00	3.57	0.29854
30	2.00	4.00	0.00	25.00	28800.00	5.86	0.48977
31	2.00	0.40	0.00	25.00	28800.00	3.8	0.31763
32	2.00	4.00	0.00	25.00	28800.00	5.57	0.46544
33	2.00	4.00	0.00	25.00	28800.00	0.0507	0.00423
34	2.00	0.40	6.00	160.00	28800.00	1.41	0.11797
35	2.00	0.40	0.00	25.00	3600.00	0.304	0.02538
36	2.00	4.00	6.00	160.00	28800.00	1.01	0.08448
37	2.00	0.40	0.00	160.00	28800.00	2.33	0.19453
38	2.00	3.96	0.00	25.00	23900.00	5.72	0.47788
39	2.00	4.00	0.00	25.00	22100.00	5.68	0.47441
40	2.00	4.00	0.00	25.00	20800.00	5.62	0.46982
41	2.00	4.00	0.00	25.00	20400.00	4.77	0.39833
42	2.00	4.00	0.00	25.00	20200.00	4.85	0.40522
43	2.00	4.00	0.00	25.00	20200.00	0.0247	0.00207
44	2.00	4.00	0.00	25.00	28800.00	5.97	0.49939
45	3.00	6.00	0.00	20.00	36000.00	7.09	0.59301
46	3.00	6.00	0.00	10.00	21600.00	8.03	0.67103
47	3.00	6.00	0.00	5.00	28800.00	9.25	0.77287
48	3.00	6.00	0.00	15.00	50400.00	7.15	0.59755
49	4.00	8.00	0.00	0.00	86400.00	4.72	0.39484
50	3.88	8.00	0.00	0.00	3600.00	4.88	0.40791

Trifluoromethylation Reaction Exploration

A trifluoromethylation reaction using the Ruppert-Prakash reagent was setup to demonstrate the capability of the Optimizer framework to discover new product in the product space and optimize each individual product in a closed-loop fashion using feedback from benchtop NMR. The original procedure³⁰ was amended to include an alternative work-up process³¹ to reduce overall experiment time. Reagents were selected to provide several sides of attack as well as a scope for novel materials.

The hardware setup is shown in Figure 15.

The initial exploration phase of the experiment was set to maximize the *novelty* of the product space, i.e. maximize number of novel peaks on the ^{19}F NMR spectrum. In the second phase a set of individual optimization experiments were run with the objective set to maximize the integration area of the regions of interest on the spectrum, as identified during initial phase.

Synthesis Description

Three-neck 25-mL round bottom flask (reactor) equipped with reflux condenser, temperature probe, tubing connector to a liquid handling system, DrySyn© aluminium block (connected to a chiller) and a magnetic stirrer bar. In the beginning of the procedure the liquid handling system was washed with THF. Cyclohexenone (0.20 mL, 1.0 mmol) was added automatically, following by 4-fluoroacetophenone (0.23 mL, 1.0 mmol), butyrolactone (0.16 mL, 1.0 mmol), myrtenal (0.30 mL, 1.0 mmol) and THF (5 mL). The reaction mixture was adjusted to 22 °C and trifluoromethyltrimethylsilane (0.9 mL, 6.1 mmol) was added, following by 0.1 M solution of TBAF in THF (0.4 mL, 0.04 mmol). The resulting mixture was stirred for 5 minutes at maintained temperature. Thereafter the reaction mixture was adjusted to 22 °C, 1.0 M TBAF solution in THF (3.0 mL, 3.0 mmol) was added to cleave TMS group and the reaction mixture was stirred for another 5 minutes. The solution of fluorobenzene (1.0 M in DCM, 2.0 mL, 2.0 mmol) was added and the sample (2.5 mL) of the resulting mixture was transferred to the NMR for analysis. Upon analysis completion, the sample was transferred back to the flask and all its contents was moved to an empty flask for storage. The reactor was cleaned twice with THF (20 mL) and used for the next iteration.

Optimisation Description

Experiments were run as follows:

- 5 1. Initial exploration of the product space using random search within given parameter limits.
2. Exploration of the product space using SMBO algorithm (with parameters setup to favour exploration over exploitation) and a *novelty* parameter as the target.
3. Exploitation of the product space using SMBO algorithm and a *novelty* parameter as the target.
- 10 4. Identification of the obtained materials as identified regions of interest on the resulting ^{19}F NMR spectra. For each experiment result was calculated as the integration area of the target region divided by the sum of all other regions identified.
- 15 5. Individual optimization using SMBO algorithm (with parameter setup to favour exploitation over exploration) and product peak integration area as a target. The data from all previous experiments (i.e., set of parameters and corresponding results) were used in each subsequent optimization.

20 The parameters for the iterative closed-loop optimization were selected to allow for product formation from individual starting materials as well as theoretical mixtures. The following parameters were set:

- Cyclohexenone volume – variable within 0.0 and 0.3 mL range.
- 4-Fluoroacetophenone volume – variable within 0.0 and 0.37 mL range.
- Beta-butyrolactone volume – variable within 0.0 and 0.23 mL range.
- Myrtenal volume – variable within 0.0 and 0.45 mL range.
- 25 • THF volume – variable within 0.0 and 10.0 mL range.
- Pre-heat reaction temperature – variable within 0.0 and 50.0 °C range.
- Trifluoromethyltrimethylsilane volume – variable within 0.36 and 1.42 mL range.
- TBAF (0.1 M in THF solution) volume – variable within 0.2 and 0.8 mL range.
- Reaction time – variable within 300.0 and 18000.0 seconds.
- 30 • TBAF (1 M in THF solution) volume for work-up – variable within 2.0 and 10.0 mL range.
- Work-up time – variable within 300.0 and 3600.0 seconds.
- Fluorobenzene (1 M solution in DCM) – fixed at 2.0 mL.

35 Configuration file to run explorative optimization of trifluoromethylation reaction during initial random screening experiments:

```
40      {  
        "max_iterations": 10,  
        "target": {  
            "novelty": Infinity  
        },  
        "algorithm": {
```

```
    "name": "random",  
    "random_state": 43  
  },  
  "reference": -113.15,  
  "batch_size": 1  
}
```

Configuration file to run explorative optimization of the trifluoromethylation reaction during second phase – exploration of the parameter space:

```
10 {  
    "max_iterations": 11,  
    "target": {  
        "novelty": Infinity  
    },  
15    "algorithm": {  
        "name": "smbo",  
        "base_estimator": "GP",  
        "n_initial_points": 15,  
        "initial_point_generator": "grid",  
20        "acq_func": "LCB",  
        "acq_func_kwargs": {"kappa": 10000}  
    },  
    "reference": -113.15,  
    "batch_size": 1  
25 }
```

Configuration file to run explorative optimization of the trifluoromethylation reaction during second phase – exploitation of the parameter space to maximize the Novelty parameter.

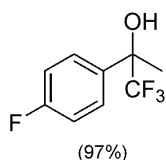
```
30 {  
    "max_iterations": 6,  
    "target": {  
        "novelty": Infinity  
    },  
35    "algorithm": {  
        "name": "smbo",  
        "base_estimator": "GP",  
        "n_initial_points": 1,  
        "initial_point_generator": "grid",  
40        "acq_func": "LCB",  
        "acq_func_kwargs": {"kappa": 0.001}  
    },  
    "reference": -113.15,  
    "batch_size": 1  
45 }
```

An example of the configuration file to run the optimization of the individual product of the trifluoromethylation reaction. Parameter constraints lists all other identified products that should be considered, when calculating the result:

```
{
  "max_iterations": 5,
  "target": {
    "spectrum_integration-area_-75.623..-76.08": Infinity
  },
  "algorithm": {
    "name": "smbo",
    "base_estimator": "GP",
    "acq_func": "LCB",
    "acq_func_kwargs": {
      "kappa": 0.001
    },
    "random_state": 42
  },
  "reference": -113.15,
  "constraints": [
    "-73.319..-73.533",
    "-74.317..-74.41",
    "-75.304..-75.55",
    "-76.629..-76.734",
    "-78.4..-78.868",
    "-79.992..-80.412",
    "-81.12..-81.698",
    "-106.519..-106.721",
    "-113.071..-113.225",
    "-116.401..-116.911"
  ],
  "batch_size": 1
}
```

Optimisation Results

All products were isolated from the collection vessel at the end of the corresponding product optimization campaign. The material was concentrated *in vacuo* and purified by preparative HPLC to give the following products. Yields indicated are calculated from the crude reaction mixture analysis with respect to the internal standard (fluorobenzene).

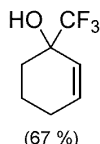


The product was characterized as follows:

¹H NMR (400 MHz, CDCl₃): δ 7.61 – 7.51 (m, 2H), 7.13 – 7.02 (m, 2H), 1.78 (q, *J* = 1.1 Hz, 3H).

5 **¹³C NMR** (101 MHz, CDCl₃): δ 163.0 (d, *J*_{CF} = 247.7 Hz), 134.3 (d, *J*_{CF} = 3.1 Hz), 128.5 – 128.1 (m), 125.6 (q, *J*_{CF} = 285.2 Hz), 115.3 (d, *J*_{CF} = 21.5 Hz), 74.7 (q, *J*_{CF} = 29.5 Hz), 24.1.

¹⁹F NMR (376 MHz, CDCl₃): δ -81.3, -113.7 – -113.8 (m).



10

The product was characterized as follows:

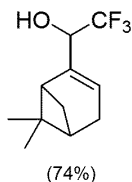
¹H NMR (400 MHz, CDCl₃): δ 6.2 (ddd, *J* = 10.1, 4.9, 2.8 Hz, 1H), 5.7 (ddt, *J* = 10.1, 2.8, 1.6 Hz, 1H), 2.2 – 2.0 (m, 2H), 1.9 – 1.7 (m, 4H).

15

¹³C NMR (101 MHz, CDCl₃): δ 136.5, 126.2 (q, *J*_{CF} = 284.2 Hz), 123.0 (q, *J*_{CF} = 1.7 Hz), 70.6 (q, *J*_{CF} = 28.8 Hz), 29.1 – 29.0 (m), 24.9, 17.3.

¹⁹F NMR (376 MHz, CDCl₃): δ -82.8.

20



The product was characterized as follows:

25 **¹H NMR** (400 MHz, CDCl₃), mixture of isomers: δ 5.81 – 5.73 (m, 1H), 4.41 – 4.30 (m, 1H), 2.50 – 2.30 (m, 4H), 2.16 – 2.09 (m, 1H), 2.05 – 1.98 (m, 1H), 1.30 (s, 3H), 1.18 (dd, *J* = 8.8, 6.0 Hz, 1H), 0.87 – 0.81 (m, 3H).

¹³C NMR (101 MHz, CDCl₃), mixture of isomers: δ 141.6, 141.4, 125.9, 125.1, 124.5 (q, *J*_{CF} = 283.1 Hz), 124.4 (q, *J*_{CF} = 283.1 Hz), 73.7 (q, *J*_{CF} = 31.4 Hz), 73.1 (q, *J*_{CF} = 31.4 Hz), 42.4, 41.8, 40.8, 40.7, 38.2, 37.9, 31.9, 31.8, 31.6, 31.6, 26.3, 26.1, 21.1, 20.9.

30

¹⁹F NMR (376 MHz, CDCl₃), mixture of isomers: δ -77.0, -77.0.

In addition to the isolated products, the optimization was carried out to maximize the peaks of the side materials. Figures 17, 18 and 19 show examples of NMR spectra with highlighted peaks selected for the optimization.

35

Figure 17 shows an ^{19}F NMR spectrum obtained after performing optimization for the identified spectrum region. The product shown has not been isolated as pure material and its structure has not been fully confirmed.

- 5 Figure 18 shows an ^{19}F NMR spectrum obtained after performing optimization for the identified spectrum region. The product shown has not been isolated as pure material and its structure has not been fully confirmed.

- 10 Figure 19 shows an ^{19}F NMR spectrum obtained after performing optimization for the identified spectrum region.

The reaction parameters for each iteration of the second phase are shown in Table 9.

- 15 The optimisation results for each iteration of the second phase are shown in Table 10. The Table headers refer to the chemical shift in the ^{19}F NMR spectrum, and the data is the peak intensity.

Table 9: Reaction parameters used in the Trifluoromethylation Reaction Optimization

Iteration	Add_1- volume	Add_2- volume	Add_3- volume	Add_4- volume	Add_5- volume	HeatChill ToTemp_ 7-temp	Add_8- volume	Add_10- volume	Stir_12- time	Add_14- volume	Stir_16- time
0	0.2	0.02	0	0.38	5.87	11.24	1.16	0.36	7733.61	3.8	3452.54
1	0.18	0.35	0.23	0.02	1.19	38.06	0.49	0.41	5610.34	4.2	467.05
2	0.14	0.21	0.03	0.15	5.05	15.2	0.8	0.45	1951.66	4.35	544.38
3	0.17	0.03	0.12	0.38	2.89	4.14	0.7	0.45	11135.33	2.61	313.07
4	0.075	0.0925	0.115	0.225	2.5	12.5	0.89	0.35	9150	3	1950
5	0.075	0.2775	0.1725	0.225	2.5	12.5	1.155	0.65	13575	3	2775
6	0.075	0.185	0.1725	0.1125	2.5	37.5	1.155	0.65	4725	3	2775
7	0.075	0.185	0.0575	0.3375	5	37.5	0.89	0.65	4725	3	1125
8	0.15	0.2775	0.115	0.1125	5	25	0.89	0.35	13575	5	1125
9	0.075	0.0925	0.115	0.225	2.5	37.5	0.89	0.35	13575	5	1950
10	0.15	0.0925	0.115	0.1125	2.5	37.5	0.625	0.65	13575	3	1125
11	0.225	0.185	0.0575	0.225	2.5	37.5	1.155	0.5	13575	5	1125
12	0.15	0.0925	0.115	0.3375	5	12.5	0.89	0.65	4725	5	1125
13	0.225	0.0925	0.1725	0.1125	2.5	37.5	1.155	0.5	13575	4	1950
14	0.075	0.185	0.0575	0.3375	5	25	0.625	0.5	4725	3	1950
15	0	0	0.2033	0.45	10	0	1.42	0.8	18000	6	3600
16	0.3	0.37	0	0.45	0	50	0.36	0.2	18000	6	3600
17	0	0.37	0.23	0	10	0	1.42	0.2	300	6	300
18	0.3	0.37	0	0	10	50	0.36	0.8	300	2	3600
19	0.3	0	0	0	10	0	1.42	0.8	300	6	300
20	0	0.37	0	0.45	0	50	0.36	0.8	18000	2	3600
21	0.3	0.2361	0	0.45	10	0	0.36	0.8	300	6	3600
22	0.278	0.1257	0.23	0.4485	10	46.4547	1.42	0.2	17033.14	2.3497	3600
23	0.3	0	0	0	0	0	1.42	0.8	18000	6	3600
24	0	0.0073	0.23	0.4318	7.3884	50	0.6393	0.7332	300	2	3600

Iteration	Add_1- volume	Add_2- volume	Add_3- volume	Add_4- volume	Add_5- volume	HeatChill ToTemp_ 7-temp	Add_8- volume	Add_10- volume	Stir_12- time	Add_14- volume	Stir_16- time
25	0.3	0.37	0.23	0.45	0	50	0.36	0.2	18000	6	300
26	0.2372	0.277	0.1004	0.3637	9.6538	8.611	0.8656	0.5804	6231.506	5.307	2368.209
27	0.0038	0.3542	0.23	0.0075	7.5622	0	1.4027	0.7963	18000	2.0344	300
28	0.2901	0.0339	0.0266	0	3.8379	50	1.1133	0.7158	17219.21	4.8051	3600
29	0.0565	0.3554	0.0487	0.0218	8.0995	27.68	0.8693	0.4565	4982.456	5.0804	2571.033
30	0.2637	0.1174	0.1701	0.3444	8.8769	1.9052	1.3426	0.3086	4411.009	5.6418	336.2277
31	0.3	0	0.0126	0	0	0	1.42	0.2	18000	6	3247.71
32	0.0051	0.3364	0.2221	0.0035	7.5653	0	1.3032	0.7983	18000	2.6423	300
33	0	0.2214	0.23	0	10	2.689	1.42	0.8	9738.93	2	300
34	0.1829	0.1246	0.0737	0.0075	7.5814	13.6651	1.4027	0.7963	18000	2.4345	300.0008
35	0.035	0.2975	0.1758	0.1265	7.6666	35.9542	1.0589	0.7999	18000	2.0339	669.7364
36	0.0924	0.3542	0.2146	0.3076	7.5021	0	0.7836	0.793	13514.36	4.2811	414.8101
37	0.2113	0.114	0.2146	0.3826	7.5027	24.3047	0.7836	0.7932	7280.613	4.2779	1503.354
38	0.1327	0.3473	0.2152	0.3546	7.5452	44.3654	0.8886	0.7851	5177.951	4.3175	420.7506
39	0.3	0.0002	0	0	0.0006	0	1.42	0.8	18000	6	3600
40	0.1583	0.3466	0.1638	0.0399	7.5592	34.7032	1.4027	0.6237	11739.65	3.3397	300
41	0.2311	0.1916	0.0185	0.107	0	0	1.0839	0.8	17916.68	4.6267	2714.998
42	0.2027	0.2305	0.0414	0.0971	0.6613	0	1.0743	0.8	17829.98	4.1059	2227.308
43	0.3	0	0	0	4.490189	39.77515	1.02661	0.58786	7327.404	3.866757	2270.768
44	0.3	0	0	0	9.948695	49.10743	1.016392	0.330502	300	2	537.9938
45	0.3	0	0	0	3.644377	27.88199	0.955637	0.567561	300	2.603835	300
46	0.3	0	0	0	3.010489	9.103317	0.922228	0.588549	300	2	300
47	0.3	0	0	0	2.854585	32.1703	0.917143	0.629172	300	2.96925	300
48	0	0.37	0.23	0	0	50	0.54765	0.2	18000	4.252653	3600
49	0	0.37	0.23	0	0	50	0.36	0.2	18000	3.443402	3600
50	0	0.37	0.23	0	0	50	0.541411	0.2	18000	4.102209	3600

Iteration	Add_1- volume	Add_2- volume	Add_3- volume	Add_4- volume	Add_5- volume	HeatChill ToTemp_7-temp	Add_8- volume	Add_10- volume	Stir_12- time	Add_14- volume	Stir_16- time
51	0	0.37	0.23	0	0	50	0.568638	0.2	18000	3.191623	3600
52	0	0.37	0.23	0	1.112906	50	0.589825	0.2	18000	2	3600
53	0.059886	0	0.05998	0.393794	0	8.496476	0.423968	0.2	8675.748	2	300
54	0.2367	0.01234	0.080616	0.388762	0	1.35394	0.660831	0.281183	18000	2	300
55	0	0	0.197377	0.413064	8.380848	29.25047	0.36	0.776153	11971.66	2	945.5722
56	0	0	0.199944	0.418024	7.74733	15.00305	0.36	0.702736	11943.13	2.585061	2603.807
57	0.024245	0	0.23	0.432863	1.839167	0	0.532426	0.730457	11791.81	3.29267	907.0798
58	0.3	0.086836	0	0	2.372057	27.3889	1.327589	0.574897	4396.9	3.804945	1045.072
59	0.3	0.108147	0	0	3.361527	30.58552	0.457217	0.676836	10767.93	4.252617	2070.023
60	0.3	0.232246	0	0	4.812268	39.0591	1.014403	0.628306	5125.857	3.820014	1518.177
61	0.3	0	0	0	1.627027	18.95629	1.086774	0.521459	6580.949	4.063436	1890.513
62	0.3	0	0	0	5.666691	44.75601	1.114701	0.51893	14128.56	4.150016	1269.458
63	0.3	0	0	0	0.558368	0	0.702877	0.2	300	2	300
64	0.3	0	0	0	1.529958	0	0.749839	0.2	300	2	300
65	0.3	0	0	0	0	29.16563	0.611402	0.8	300	2	300
66	0.3	0	0	0	0.375438	20.24044	0.717494	0.8	300	2	300
67	0.3	0	0	0	0	14.77393	0.668792	0.8	300	2	300
68	0.3	0	0	0	6.978	36.2612	1.1331	0.5781	13024.99	4.2031	1288.411
69	0.1788	0.1533	0	0.0906	6.9438	41.2809	1.0869	0.2351	13491.75	2.3058	1147.581
70	0.099411	0.308388	0	0.237783	6.896688	36.37782	1.043806	0.2	12127.08	5.037321	1449.423
71	0.178127	0.154098	0	0.142892	6.91627	44.09105	0.761888	0.233305	13629.27	2	1487.288
72	0.240633	0.055215	0.110086	0.014404	8.732324	46.07651	1.144411	0.225953	13467.95	4.443958	1490.177
73	0.261994	0.3357	0.181466	0.009672	8.788054	46.67213	1.184446	0.295729	13443.88	4.214546	2061.496
74	0.092298	0.205309	0.118327	0.031974	4.814851	45.09937	1.226145	0.225269	13454.57	5.564882	376.7512
75	0.3	0.37	0.23	0.45	0	50	0.36	0.2	300	6	300
76	0.3	0.37	0.23	0.45	10	18.45753	0.36	0.2	18000	6	300

Iteration	Add_1- volume	Add_2- volume	Add_3- volume	Add_4- volume	Add_5- volume	HeatChill ToTemp_ 7-temp	Add_8- volume	Add_10- volume	Stir_12- time	Add_14- volume	Stir_16- time
77	0.3	0.37	0.23	0.45	10	0	0.36	0.8	18000	6	300
78	0.3	0.37	0.23	0.45	10	0	0.36	0.8	18000	6	300
79	0.3	0.37	0.23	0.45	10	0	0.36	0.8	18000	6	300
80	0.101993	0.104922	0.23	0.084776	7.831853	30.42868	0.36	0.221733	1882.467	2	2503.105
81	0	0	0.198333	0.45	8.327648	0	0.36	0.607982	300	3.460318	3600
82	0.026552	0	0.202983	0.45	10	12.79337	0.36	0.50523	9821.36	2.017791	3600
83	0	0	0.23	0.45	10	1.62732	0.36	0.2	9805.984	2.955296	3600

Table 10: Peak intensity at target chemical shifts for Phase 2 of the Trifluoromethylation Reaction Optimization

Iteration	-73.319	-74.317	-75.304	-75.623	-76.629	-78.4	-79.992	-81.12	-106.519	-113.071	-116.401
	-73.533a	-74.41	-75.55	-76.08	-76.734	-78.868	-80.412	-81.698	-106.721	-113.225	-116.911
0	0.026931	0.00807	0.002971	0.445805	0.007102	0.302055	0.04809	0.363223	0.001185	0.102573	0.016347
1	0.00094	0.009541	0.004244	0.046386	0.000931	0.106624	0.827693	0.038715	0.076143	0.184034	0.146073
2	0.027752	0.002456	0.012855	0.148416	0.00558	0.192891	0.355283	0.249336	0.005203	0.134312	0.082326
3	0.02013	0.005185	0.016818	0.939715	0.000999	0.14083	0.073294	0.152789	0.000727	0.154533	0.014469
4	0.01252	0.000731	0.026094	0.336337	0.002723	0.495453	0.12161	0.115449	0.001995	0.163276	0.021963
5	0.004393	0.000453	0.038487	0.2472	0.002989	0.287555	0.492673	0.033372	0.003742	0.111706	0.071846
6	0.0074	0.00021	0.034078	0.104854	8.50E-06	0.590698	0.329655	0.077302	0.001375	0.127612	0.06271
7	0.012679	0.000368	0.005568	0.475924	0.000104	0.246552	0.268785	0.083348	0.005704	0.126322	0.057775
8	0.013562	0.001987	0.01331	0.092497	0.007747	0.288049	0.520699	0.116637	0.005939	0.124066	0.101779
9	0.00567	0.005399	0.026825	0.321798	0.006014	0.54744	0.128593	0.072311	0.002991	0.169844	0.031057
10	0.025828	0.002395	0.022695	0.165243	0.004296	0.242517	0.191296	0.289377	0.002649	0.232915	0.033493
11	0.031351	0.000201	0.007414	0.212891	0.023131	0.262222	0.247272	0.260417	0.003547	0.105051	0.055635
12	0.015538	0.001017	0.016914	0.499698	0.003459	0.328254	0.098611	0.187059	0.000466	0.132692	0.018659
13	0.03343	0.002213	0.034397	0.096894	0.017261	0.573336	0.122443	0.259653	0.002467	0.137017	0.024901

Iteration	-73.319 -73.533a	-74.317 -74.41	-75.304 -75.55	-75.623 -76.08	-76.629 -76.734	-78.4 -78.868	-79.992 -80.412	-81.12 -81.698	-106.519 -106.721	-113.071 -113.225	-116.401 -116.911
14	0.007189	0.005232	0.001982	0.602139	0.000686	0.122655	0.31748	0.057719	0.007518	0.153555	0.068131
15	0.002656	0.005772	0.041705	0.518892	0.001651	0.97603	0.010474	0.00232	0.002556	0.103041	0.005756
16	0.003343	0.001196	0.004481	0.322733	0.002605	0.021159	0.392804	0.004489	0.163685	0.265423	0.095051
17	0.000374	0.000725	0.009511	0.003626	0.000316	1.065229	0.449046	0.004027	0.003589	0.075005	0.089932
18	0.004735	0.001957	0.000957	0.0074	0.001835	0.225859	0.715106	0.081718	0.095298	0.253707	0.01739
19	0.036852	0.000714	0.00096	0.000477	0.031314	1.05318	0.000225	0.502466	0.00103	0.086546	0.003675
20	0.000762	0.000501	0.001097	0.34051	0.000765	0.013483	0.489516	0.006217	0.145487	0.215159	0.099707
21	0.002724	0.003487	0.006295	0.536713	0.001728	0.197231	0.173985	0.013999	0.109195	0.26262	0.003392
22	0.008404	0.004715	0.009333	0.506539	0.001758	0.521677	0.125638	0.079668	0.004264	0.090206	0.02514
23	0.016156	0.000203	0.002818	0.010638	0.019702	0.686884	0.001575	0.677108	0.000885	0.153337	0.004929
24	0.001889	0.036985	0.004354	0.904579	0.00129	0.285038	0.102332	0.003641	0.006038	0.1728	0.010368
25	0.004209	0.0003	0.003218	0.244419	0.001542	0.039368	0.350685	0.01863	0.238494	0.264902	0.082902
26	0.007477	0.014467	0.004675	0.437983	0.00155	0.104838	0.399779	0.110493	0.014178	0.10433	0.085885
27	0.003568	0.001095	0.044824	0.038362	0.000808	0.646166	0.707454	0.004511	0.006673	0.104355	0.002227
28	0.058844	0.001702	0.008197	0.000445	0.043134	0.456031	0.076565	0.59197	0.000574	0.136182	0.016468
29	0.007545	0.000762	0.005213	0.028837	0.002882	0.266123	0.791957	0.066572	0.010283	0.115372	0.146119
30	0.016108	0.00952	0.014236	0.303131	0.007598	0.667077	0.107555	0.136849	0.001235	0.084551	0.024272
31	0.024434	0.000646	0.002751	0.001393	0.025522	0.561601	0.003004	0.74639	0.000843	0.180919	0.002426
32	0.006302	0.001021	0.034352	0.042783	0.002301	0.506954	0.557055	0.048849	0.003437	0.090872	0.097346
33	0.000227	0.0023	0.04659	0.00023	0.00057	1.312141	0.35125	0.002824	0.001182	0.100279	0.030483
34	0.018894	0.00079	0.010694	0.04511	0.003708	0.739151	0.163782	0.294392	0.001053	0.109913	0.0309
35	0.004242	0.000864	0.014121	0.090767	0.001243	0.391997	0.598554	0.035119	0.007937	0.11858	0.102145
36	0.002109	0.023906	0.006266	0.321493	0.000953	0.104398	0.602624	0.020991	0.024035	0.119229	0.114798
37	0.008268	0.045214	0.015427	0.600199	0.00146	0.212933	0.154772	0.094202	0.006551	0.142381	0.030925
38	0.002644	0.022617	0.004289	0.358655	0.00165	0.139952	0.501183	0.029184	0.031227	0.110372	0.099726
39	0.018839	0.000447	0.004291	0.011991	0.02022	0.60678	0.0215	0.722576	0.000984	0.139282	0.00374

Iteration	-73.319 -73.533a	-74.317 -74.41	-75.304 -75.55	-75.623 -76.08	-76.629 -76.734	-78.4 -78.868	-79.992 -80.412	-81.12 -81.698	-106.519 -106.721	-113.071 -113.225	-116.401 -116.911
40	0.015408	0.001338	0.020358	0.035429	0.000152	0.471112	0.502055	0.112417	0.006878	0.083007	0.099082
41	0.035675	0.000191	0.00731	0.034192	0.034565	0.218196	0.217735	0.589328	0.001138	0.131298	0.047909
42	0.030842	0.000538	0.01364	0.079706	0.012073	0.14267	0.369772	0.41443	0.002504	0.119162	0.079393
43	0.0713	0.001826	0.001032	0.002241	0.057747	0.355026	0.009599	0.868988	0.000103	0.154082	0.00368
44	0.0459	0.002067	0.000817	0.003452	0.001438	0.962873	0.000784	0.516171	0.002633	0.126997	0.001295
45	0.0793	0.001333	0.001946	0.003414	0.008486	0.355658	0.003712	1.071054	0.000499	0.1432	0.002514
46	0.058	0.00041	0.003115	0.001431	0.000883	0.288345	0.005747	1.367022	0.000602	0.149348	0.001719
47	0.088	5.36E-05	0.001612	0.001226	0.032118	0.29471	0.011188	1.088354	0.000135	0.139541	0.002657
48	0.001362	0.002587	0.006281	0.000543	0.003614	0.070259	1.195571	0.00788	0.044072	0.202649	0.186179
49	0.002936	0.001261	0.003716	0.000845	0.000709	0.015793	1.068248	0.005618	0.119551	0.25501	0.166787
50	0.000322	0.001923	0.005837	0.000899	0.004253	0.073874	1.244337	0.002973	0.046098	0.193939	0.182214
51	0.000454	0.000409	0.002372	0.003258	3.87E-05	0.083064	1.263569	0.006625	0.043069	0.188766	0.179117
52	0.00084	0.001212	0.003889	0.000442	0.000315	0.103459	1.0955	0.014152	0.070606	0.189544	0.159456
53	0.016634	0.001719	0.016732	0.793561	0.004027	0.138917	0.001262	0.185789	0.002368	0.30772	0.001421
54	0.040607	0.001293	0.014665	0.481211	0.000562	0.068786	0.065395	0.507409	4.94E-05	0.165274	0.015808
55	0.003657	0.104541	0.003604	1.040364	0.000624	0.128106	0.022725	0.001901	0.003192	0.325109	0.001402
56	0.001445	0.14072	0.011406	0.921136	0.002895	0.125702	0.010063	0.001058	0.00097	0.342307	0.002897
57	0.00177	0.088912	0.009961	1.298531	0.001599	0.120717	0.02298	0.013117	0.001715	0.237427	0.003485
58	0.06489	0.000217	0.003485	0.001432	0.03883	0.392216	0.062053	0.732507	0.000484	0.128923	0.019298
59	0.019921	0.000813	0.000647	0.000868	0.004201	0.073595	0.243954	0.88191	0.003142	0.265512	0.028525
60	0.064077	1.74E-05	0.00992	0.00159	0.016466	0.138561	0.326443	0.594806	0.002162	0.107624	0.077661
61	0.053984	0.000671	0.000455	0.002696	0.043093	0.3605	0.00729	1.031689	0.000479	0.139007	0.001122
62	0.065598	0.000611	0.001086	0.003003	0.05737	0.540818	0.008498	0.680715	0.001121	0.127341	0.000983
63	0.077766	0.001036	0.000163	0.009004	0.019434	0.193272	0.003081	1.379115	0.000897	0.178449	0.001625
64	0.001675	0.01308	0.077039	0.107675	0.005981	0.527583	0.033185	0.059698	0.001356	0.594837	0.002586
65	0.069298	0.001278	0.002348	0.001837	0.005182	0.145421	0.140705	0.78133	0.00134	0.249344	0.035989

Iteration	-73.319 -73.533a	-74.317 -74.41	-75.304 -75.55	-75.623 -76.08	-76.629 -76.734	-78.4 -78.868	-79.992 -80.412	-81.12 -81.698	-106.519 -106.721	-113.071 -113.225	-116.401 -116.911
66	0.10295	0.000504	0.001593	0.003553	0.002691	0.144169	0.005991	1.524694	2.09E-05	0.192146	0.001274
67	0.087147	0.001186	0.000234	0.004505	0.008762	0.17284	0.003338	1.29019	0.000792	0.232795	0.001574
68	0.062246	0.000925	0.002244	0.000804	0.043295	0.506	0.003896	0.723594	0.000984	0.151329	0.003858
69	0.029191	0.001016	0.007682	0.055327	0.004453	0.5886	0.197234	0.273115	0.0023	0.123589	0.046113
70	0.008162	0.000451	0.000677	0.228447	0.00758	0.303062	0.457948	0.073087	0.002344	0.104147	0.093896
71	0.038211	0.00067	0.004996	0.171539	0.00473	0.247632	0.24029	0.288309	0.001676	0.155737	0.057077
72	0.028547	0.003006	0.013142	0.055954	0.020742	1.094311	0.084893	0.192772	0.002893	0.125187	0.006481
73	0.016082	0.0046	0.006557	0.041822	0.006145	0.4129	0.536109	0.096164	0.012574	0.099289	0.105119
74	0.010323	0.000786	0.015074	0.043346	0.007609	0.813846	0.329731	0.078278	0.001299	0.107618	0.060831
75	0.003643	0.047311	0.00208	0.125964	0.001074	0.15261	0.02432	0.35608	0.143774	0.273899	0.083117
76	0.002404	0.070188	0.001496	0.186155	0.00201	0.09199	0.03099	0.239419	0.282605	0.25405	0.042957
77	0.000762	0.06593	0.00309	0.142652	0.003277	0.161471	0.01113	0.21099	0.310873	0.254268	0.043913
78	0.000183	0.064404	0.002041	0.144332	0.004349	0.137474	0.02428	0.245369	0.271876	0.255579	0.050324
79	0.001868	0.063778	0.005163	0.167917	0.002887	0.12175	0.013102	0.237106	0.277701	0.262225	0.050606
80	0.003982	0.058216	0.006345	0.05995	0.004268	0.235186	0.174697	0.019966	0.054443	0.752838	0.035284
81	0.003837	0.12127	0.003064	0.848478	0.004678	0.207269	0.002794	0.00016	0.000796	0.324183	0.001057
82	0.002979	0.174337	0.00115	0.795634	0.00201	0.141378	0.012129	0.000196	0.001974	0.355473	0.002064
83	0.002055	0.18681	0.005423	1.003405	0.000786	0.085797	0.004686	0.001123	0.006111	0.317595	0.0016

References

A number of publications are cited above in order to more fully describe and disclose the invention and the state of the art to which the invention pertains. Full citations for these references are provided below. The entirety of each of these references is incorporated
5 herein.

EP 4129462

US 2020/0225251

US 2022/0331772

10 WO 2013/175240

WO 2019/137954

WO 2020/007962

WO 2021/219999

15 The following citations are taken from the *Background and Comments* sections of the application:

1. Trobe, M. & Burke, M. D. The Molecular Industrial Revolution: Automated Synthesis of Small Molecules. *Angew. Chem. Int. Ed.* **57**, 4192-4214 (2018).
- 20 2. Christensen, M. *et al.* Automation isn't automatic. *Chem. Sci.* **12**, 15473-15490 (2021).
3. Wilbraham, L., Mehr, S. H. M. & Cronin, L. Digitizing Chemistry Using the Chemical Processing Unit: From Synthesis to Discovery. *Acc. Chem. Res.* **54**, 253-262 (2021).
4. Shi, Y., Prieto, P. L., Zepel, T., Grunert, S. & Hein, J. E. Automated Experimentation Powers Data Science in Chemistry. *Acc. Chem. Res.* **54**, 546-555 (2021).
- 25 5. Stach, E. *et al.* Autonomous experimentation systems for materials development: A community perspective. *Matter* **4**, 2702-2726 (2021).
6. Davies, I. W. The digitization of organic synthesis. *Nature* **570**, 175-181 (2019).
7. Steiner, S. *et al.* Organic synthesis in a modular robotic system driven by a chemical programming language. *Science* **363**, eaav2211 (2019).
- 30 8. Jiang, T. *et al.* An integrated console for capsule-based, automated organic synthesis. *Chem. Sci.* **12**, 6977-6982 (2021).
9. Chatterjee, S., Guidi, M., Seeberger, P. H. & Gilmore, K. Automated radial synthesis of organic molecules. *Nature* **579**, 379-384 (2020).
10. Shields, B. J. *et al.* Bayesian reaction optimization as a tool for chemical synthesis.
- 35 *Nature* **590**, 89-96 (2021).
11. Li, J. *et al.* Synthesis of many different types of organic small molecules using one automated process. *Science* **347**, 1221-1226 (2015).
12. Coley, C. W. *et al.* A robotic platform for flow synthesis of organic compounds informed by AI planning. *Science* **365**, eaax1566 (2019).
- 40 13. Adamo, A. *et al.* On-demand continuous-flow production of pharmaceuticals in a compact, reconfigurable system. *Science* **352**, 61-67 (2016).

14. Chanda, A. *et al.* Industry Perspectives on Process Analytical Technology: Tools and Applications in API Development. *Org. Process Res. Dev.* **19**, 63-83 (2015).
15. Simon, L. L. *et al.* Assessment of Recent Process Analytical Technology (PAT) Trends: A Multiauthor Review. *Org. Process Res. Dev.* **19**, 3-62 (2015).
- 5 16. Hammer, A. J. S., Leonov, A. I., Bell, N. L. & Cronin, L. Chempuration and the Standardization of Chemical Informatics. *JACS Au* (2021).
17. Schweidtmann, A. M. *et al.* Machine learning meets continuous flow chemistry: Automated optimization towards the Pareto front of multiple objectives. *Chem. Eng. J.* **352**, 277-282 (2018).
- 10 18. Gromski, P. S., Henson, A. B., Granda, J. M. & Cronin, L. How to explore chemical space using algorithms and automation. *Nat. Rev. Chem.* **3**, 119-128 (2019).
19. Mateos, C., Nieves-Remacha, M. J. & Rincón, J. A. Automated platforms for reaction self-optimization in flow. *React. Chem. Eng.* **4**, 1536-1544 (2019).
20. Clayton, A. D. *et al.* Automated self-optimisation of multi-step reaction and separation processes using machine learning. *Chem. Eng. J.* **384**, 123340 (2020).
- 15 21. Christensen, M. *et al.* Data-science driven autonomous process optimization. *Commun. Chem.* **4**, 112 (2021).
22. Bédard, A.-C. *et al.* Reconfigurable system for automated optimization of diverse chemical reactions. *Science* **361**, 1220-1225 (2018).
- 20 23. Collins, N. *et al.* Fully Automated Chemical Synthesis: Toward the Universal Synthesizer. *Org. Process Res. Dev.* **24**, 2064-2077 (2020).
24. Breen, C. P., Nambiar, A. M. K., Jamison, T. F. & Jensen, K. F. Ready, Set, Flow! Automated Continuous Synthesis and Optimization. *Trends Chem.* **3**, 373-386 (2021).
25. Angelone, D. *et al.* Convergence of multiple synthetic paradigms in a universally programmable chemical synthesis machine. *Nat. Chem.* **13**, 63-69 (2021).
- 25 26. Mehr, S. H. M., Craven, M., Leonov, A. I., Keenan, G. & Cronin, L. A universal system for digitization and automatic execution of the chemical synthesis literature. *Science* **370**, 101 (2020).
27. Felton, K. C., Rittig, J. G. & Lapkin, A. A. Summit: Benchmarking Machine Learning Methods for Reaction Optimisation. *Chem. Methods* **1**, 116-122 (2021).
- 30 28. Häse, F. *et al.* Olympus: a benchmarking framework for noisy optimization and experiment planning. *Mach. learn.: sci. technol* **2**, 035021 (2021).
29. Tietze, L. F. & Eicher, T. in *Reaktionen und Synthesen im organisch-chemischen Praktikum und Forschungslaboratorium* Ch. 2.5 - 2.7, 70-119 (1991).
- 35 30. Talukdar, S., Hsu, J.-L., Chou, T.-C. & Fang, J.-M. Direct transformation of aldehydes to nitriles using iodine in ammonia water. *Tetrahedron Lett.* **42**, 1103-1105 (2001).
31. Xiang, Z. *et al.* Concise Synthesis of Isoquinoline via the Ugi and Heck Reactions. *Org. Lett.* **6**, 3155-3158 (2004).
32. van Leusen, A. M., Hoogenboom, B. E. & Siderius, H. A novel and efficient synthesis of oxazoles from tosylmethylisocyanide and carbonyl compounds. *Tetrahedron Lett.* **13**, 2369-2372 (1972).
- 40 33. Lane, B. S., Vogt, M., DeRose, V. J. & Burgess, K. Manganese-catalyzed epoxidations of alkenes in bicarbonate solutions. *J. Am. Chem. Soc.* **124**, 11946-11954 (2002).

34. Huyer, W. & Neumaier, A. SNOBFIT -- Stable Noisy Optimization by Branch and Fit. *ACM Trans. Math. Softw.* **35**, Article 9 (2008).
35. Häse, F., Roch, L. M., Kreisbeck, C. & Aspuru-Guzik, A. Phoenix: A Bayesian Optimizer for Chemistry. *ACS Cent. Sci.* **4**, 1134-1145 (2018).
- 5 36. Johnston, C. P. *et al.* Anion-Initiated Trifluoromethylation by TMSCF₃: Deconvolution of the Siliconate-Carbanion Dichotomy by Stopped-Flow NMR/IR. *J. Am. Chem. Soc.* **140**, 11112-11124 (2018).
37. Torres J. A. G. *et al.* A Multi-Objective Active Learning Platform and Web App for Reaction Optimization. *J. Am. Chem. Soc.* **144**(43): 19999-20007 (2022).

10

The following citations are taken from the *Experimental* sections of the application:

1. ECMA-404. The JSON data interchange 7), [https://www.ecma-](https://www.ecma-international.org/publications-and-standards/standards/ecma-404/)
15 [international.org/publications-and-standards/standards/ecma-404/](https://www.ecma-international.org/publications-and-standards/standards/ecma-404/)
2. Despite ECMA-404 standard for JSON format not allowing "Infinity", this value is correctly parsed with python and therefore is accepted as the target value.
3. Mehr, S. H. M., Craven, M., Leonov, A. I., Keenan, G. & Cronin, L. A universal system for digitization and automatic execution of the chemical synthesis literature. *Science* **370**, 101
20 (2020).
4. Zorin, V., Bernstein, M. A. & Cobas, C. A robust, general automatic phase correction algorithm for high-resolution NMR data. *Magnetic Resonance in Chemistry* **55**, 738-746 (2017).
5. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357-362 (2020).
- 25 6. Head, T., Kumar, M., Nahrstaedt, H., Louppe, G. & Shcherbatyi, I. *scikit-optimize/scikit-optimize* v. v0.8.1 (Zenodo, 2020).
7. *API reference to run bayesian optim* [https://scikit-](https://scikit-optimize.github.io/stable/modules/generated/skopt.Optimizer.html)
[optimize.github.io/stable/modules/generated/skopt.Optimizer.html](https://scikit-optimize.github.io/stable/modules/generated/skopt.Optimizer.html) (accessed August 2021).
8. *pyDOE2: An experimental design packa* <https://github.com/clicumu/pyDOE> (accessed
30 December 2021).
9. *The Python Stay*, <https://docs.python.org/3/library/> (accessed August 2021).
10. Häse, F. *et al.* Olympus: a benchmarking framework for noisy optimization and experiment planning. *Machine Learning: Science and Technology* **2**, 035021 (2021).
11. Eilers, P. & Boelens, H. *Baseline Correction with Asymmetric Least Squares Smoothing*. (2005).
- 35 12. Helmus, J. J. & Jaroniec, C. P. Nmr glue: an open source Python package for the analysis of multidimensional NMR data. *Journal of Biomolecular NMR* **55**, 355-367 (2013).
13. *Arduino Mega 2560 Rt*,
<https://docs.arduino.cc/static/2a5af72674da26b40f0312c1b3e499a2/A000067-datasheet.pdf>
40 (accessed January 2022).
14. *TCS34725 or*, <https://ams.com/en/tcs34725> (accessed September 2021).
15. *Humidity 80*, <https://www.bosch-sensortec.com/products/environmental-sensors/humidity-sensors-bme280/> (accessed December 2021).

16. OPB350 - Tube or,
<https://www.ttelectronics.com/products/categories/optoelectronics/optoelectronics/opb350/>
 (accessed December 2021).
17. MAX31865 RTD-to-Digier,
- 5 <https://www.maximintegrated.com/en/products/interface/sensor-interface/MAX31865.html>
 (accessed December 2021).
18. Comman <https://github.com/croningp/commanduino> (accessed December 2021).
19. Arduino <https://github.com/croningp/Arduino-CommandTools> (accessed December 2021).
- 10 20. Wang, Z., Simoncelli, E. P. & Bovik, A. C. in *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*. 1398-1402 Vol.1392.
21. Dash Pythe, <https://dash.plotly.com/> (accessed January 2021).
22. Angelone, D. *et al.* Convergence of multiple synthetic paradigms in a universally programmable chemical synthesis machine. *Nature Chemistry* **13**, 63–69 (2021).
- 15 23. Xiang, Z. *et al.* Concise Synthesis of Isoquinoline via the Ugi and Heck Reactions. *Organic Letters* **6**, 3155-3158 (2004).
24. Singh, K., Kaur, A., Mithu, V. S. & Sharma, S. Metal-Free Organocatalytic Oxidative Ugi Reaction Promoted by Hypervalent Iodine. *The Journal of Organic Chemistry* **82**, 5285-5293 (2017).
- 20 25. Campos, S. A., Harling, J. D. M., Afjal Hussain & Smith, I. E. D. Benzothiophene derivatives as estrogen receptor inhibitors WO 2015/000867 A1 (2013).
26. Hage, R. & Lienke, A. Applications of Transition-Metal Catalysts to Textile and Wood-Pulp Bleaching. *Angewandte Chemie International Edition* **45**, 206-222 (2006).
27. Lane, B. S., Vogt, M., DeRose, V. J. & Burgess, K. Manganese-catalyzed epoxidations of alkenes in bicarbonate solutions. *J. Am. Chem. Soc.* **124**, 11946-11954 (2002).
- 25 28. Abdolazadeh, S. *Catalysis in complex media: Analytical approaches and mechanisms in manganese catalysed oxidations*, University of Groningen, (2015).
29. Häse, F., Roch, L. M., Kreisbeck, C. & Aspuru-Guzik, A. Phoenix: A Bayesian Optimizer for Chemistry. *ACS Cent. Sci.* **4**, 1134-1145 (2018).
- 30 30. Prakash, G. K. S., Krishnamurti, R. & Olah, G. A. Synthetic methods and reactions. 141. Fluoride-induced trifluoromethylation of carbonyl compounds with trifluoromethyltrimethylsilane (TMS-CF₃). A trifluoromethide equivalent. *Journal of the American Chemical Society* **111**, 393-395 (1989).
31. Johnston, C. P. *et al.* Anion-Initiated Trifluoromethylation by TMSCF₃: Deconvolution of the Siliconate-Carbanion Dichotomy by Stopped-Flow NMR/IR. *J. Am. Chem. Soc.* **140**, 11112-11124 (2018).
- 35

Claims

1. A method of optimising an instruction set for a chemical synthesis, the method comprising performing a series of optimisation test syntheses, where each optimisation test synthesis includes the steps of:

- (i) performing the chemical synthesis in a reaction vessel of a chemical synthesiser operating according to an instruction set to produce a reaction product,
- (ii) recording reaction data using an analytical unit of the synthesiser during the chemical synthesis and/or recording product data on the reaction product,
- (iii) analysing the reaction data and/or product data using a control unit of the chemical synthesiser to determine a characteristic of the synthesis; and
- (iv) adjusting one or more reaction parameters of the instruction set, to provide an adjusted instruction set,

wherein the optimisation test synthesis is repeated one or more times, and in each repeat of the test synthesis, the chemical synthesis is performed according to the adjusted instruction set, which is adjusted from the previous test syntheses, and

the adjustment to the one or more reaction parameters is made following a prediction that the adjustment will improve the characteristic of the synthesis, wherein the prediction is based on the reaction parameters associated with reaction data and/or product data from one or more previous optimisation test syntheses and, where present, one or more previous exploration test syntheses, and

wherein the method is performed autonomously.

2. The method of claim 1, further comprising performing a series of exploration test syntheses, where each exploration test synthesis includes the steps of:

- (i) performing the chemical synthesis in a reaction vessel of a chemical synthesiser operating according to an instruction set to produce a reaction product,
- (ii) recording reaction data using an analytical unit of the synthesiser during the chemical synthesis and/or recording product data on the reaction product,
- (iii) analysing the reaction data and/or product data using a control unit of the chemical synthesiser to determine a characteristic of the synthesis; and
- (iv) adjusting one or more reaction parameters of the instruction set, to provide an adjusted instruction set,

wherein the exploration test synthesis is repeated one or more times, and in each repeat of the test synthesis, the chemical synthesis is performed according to the adjusted instruction set, which is adjusted from the previous test syntheses, and

the adjustment to the one or more reaction parameters is made to explore a reaction space wherein the adjustment is based on the reaction parameters associated with reaction data and/or product data from one or more previous test syntheses.

3. The method of claim 2, wherein the adjustment to the one or more reaction parameters to explore a reaction space is made by an exploration algorithm of the chemical synthesiser, such as a random search algorithm.
4. The method of claim 2 or 3, wherein the method comprises a series of exploration test syntheses followed by a series of optimisation test syntheses.
5. The method of any one of claims 2 to 4, where in an optimisation test synthesis, the prediction is based on the reaction parameters associated with reaction data and/or product data from one or more previous optimisation test syntheses and one or more previous exploration test syntheses.
6. The method of any one of claims 1 to 5, where in an optimisation test synthesis, the prediction that the adjustment will improve the characteristic of the synthesis is made by an optimisation algorithm of the chemical synthesiser, such as a genetic or SMBO optimisation algorithm.
7. The method of any one of claims 1 to 6, where in an optimisation test synthesis, the prediction that the adjustment will improve the characteristic of the synthesis is based on the comparison of one or more reaction parameters with reaction data and/or product data from one or more previous test syntheses, to identify one or more reaction parameters associated with a change in reaction data and/or product data, and preferentially adjusting the one or more reaction parameters associated with the change.
8. The method of any one of claims 1 to 7, where in an optimisation test synthesis, the prediction that the adjustment will improve the characteristic of the synthesis is based on the comparison of one or more reaction parameters with reaction data and/or product data from one or more previous test syntheses, to identify one reaction parameters associated with the highest normalised rate of change in reaction data and/or product data, and preferentially adjusting the reaction parameters associated with the highest normalised rate of change.
9. The method of any one of claims 1 to 8, wherein the reaction data is selected from reaction colour, reaction temperature, reaction vessel pressure, reaction pH, reaction filter rate, surrounding humidity, surrounding pressure, surrounding temperature, or a combination thereof, preferably reaction colour, reaction temperature or a combination thereof.
10. The method of any one of claims 1 to 9, wherein the type of reaction data and/or the sample frequency of reaction data changes depending on the reaction parameters.
11. The method of any one of claim 10, wherein the type of reaction data and/or the sample frequency of reaction data changes throughout the chemical synthesis.

12. The method of any one of claims 1 to 11, wherein the product data is spectroscopic data, such as UV-Vis, Infra-red, Raman, NMR, or mass spectrometry data; chromatography data, such as gas or high-performance liquid chromatography data; or a combination thereof.
13. The method of any one of claims 1 to 12, wherein the characteristic of the synthesis is reaction product yield, reaction product purity, reaction product to starting material ratio, reaction product to impurity ratio, solvent mass consumption, reagent mass consumption or a combination thereof.
14. The method of any one of claims 1 to 13, wherein a characteristic of the synthesis is reaction product yield, and the yield is determined by NMR spectroscopy, Raman spectroscopy or high-performance liquid chromatography product data.
15. The method of any one of claims 1 to 14, wherein a reaction parameter for the instruction set is selected from reagent volume, reagent mass, reagent addition rate, solvent volume, solvent mass, solvent addition rate, reaction temperature, rate of change of reaction temperature, reaction pressure, rate of change of reaction pressure, pH, stirring rate, reaction mixture colour and reaction time.
16. The method of any one of claims 1 to 15, wherein one or more of the reaction parameters is a dynamic reaction parameter, wherein the dynamic reaction parameter is a reaction parameter which is determined by the need to satisfy another reaction parameter.
17. The method of claim 16, wherein the dynamic reaction parameter is selected from one or more of:
- (i) reagent addition rate, such as reagent addition rate determined by reaction temperature,
 - (ii) reaction time, such as reaction time determined by reaction colour, and
 - (iii) reaction temperature, such as reaction temperature determined by reaction pressure.
18. The method of either claim 16 or 17, where in an optimisation test synthesis, the prediction that the adjustment will improve the characteristic of the synthesis is based on the comparison of one or more dynamic reaction parameters with reaction data and/or product data from one or more previous test syntheses, to identify one or more dynamic reaction parameters associated with a change in reaction data and/or product data, and preferentially adjusting the one or more dynamic reaction parameters associated with the change.
19. The method of any one of claims 16 to 18, where in an optimisation test synthesis, the prediction that the adjustment will improve the characteristic of the synthesis is based on the comparison of one or more dynamic reaction parameters with reaction data and/or product data from one or more previous test syntheses, to identify a dynamic reaction parameter associated with the highest normalised rate of change in reaction data and/or product data,

and preferentially adjusting the dynamic reaction parameters associated with the highest normalised rate of change.

20. A chemical synthesiser for use in the method of optimising an instruction set for a chemical synthesis of claims 1 to 19, the synthesiser comprising a reaction vessel, an analytical unit and a control unit, wherein:

the reaction vessel is for the performance of the chemical synthesis according to an instruction set,

the analytical unit is for recording reaction data from the reaction vessel during the chemical synthesis and/or for recording product data on the reaction product, and for reporting the reaction data and/or product data to the control unit;

the control unit is for analysing the reaction data and/or product data received from the analytical unit to determine a characteristic of the synthesis, and for adjusting one or more reaction parameters of the instruction set, to provide an adjusted instruction set, wherein the adjustment to the one or more reaction parameters is made following a prediction that the adjustment will improve the characteristic of the synthesis, wherein the prediction is based on the reaction parameters associated with reaction data and/or product data from one or more previous optimisation test syntheses, and

wherein the chemical synthesiser is for performing the method autonomously.

21. The chemical synthesiser of claim 20, wherein the analytical unit is for recording reaction data during the chemical synthesis and for recording product data on the reaction product.

22. The chemical synthesiser of either claim 20 or 21, wherein the analytical unit comprises a colour sensor, a temperature sensor, a pressure sensor, a pH sensor, a flow rate sensor, a humidity sensor or a combination thereof.

23. The chemical synthesiser of any one of claims 20 to 22, wherein the analytical unit comprises an infra-red spectrometer, a Raman spectrometer, a NMR spectrometer, a mass spectrometer, a gas chromatography instrument, a high-performance liquid chromatography instrument, or a combination thereof.

24. The chemical synthesiser of any one of claims 20 to 23, wherein one or more of the reaction parameters is a dynamic reaction parameter, and the control unit is for adjusting the dynamic reaction parameter during a test synthesis such that reaction data satisfies another reaction parameter.

25. The chemical synthesiser of claim 24, wherein the control unit adjusts a dynamic reaction parameter selected from one or more of:

- (i) reagent addition rate, such as reagent addition rate determined by reaction temperature,
- (ii) reaction time, such as reaction time determined by reaction colour, and
- (iii) reaction temperature, such as reaction temperature determined by reaction.