Organic synthesis in a modular robotic system driven by a chemical programming language

Sebastian Steiner,[1] Jakob Wolf,[1] Stefan Glatzel,[1] Anna Andreou,[1] Jarosław M. Granda,[1] Graham Keenan,[1] Trevor Hinkley,[1] Gerardo Aragon-Camarasa,[1,2] Philip J. Kitson,[1] Davide Angelone,[1] and Leroy Cronin[1]*

[1]School of Chemistry, [2]School of Computer Science, The University of Glasgow, Glasgow, G12 8QQ, UK. *Correspondence to: Lee.Cronin@Glasgow.ac.uk

One Sentence Summary: A modular platform for synthesis is demonstrated that makes purified organic compounds autonomously without physical reconfiguration and is driven using a chemical programming language.

Abstract: The synthesis of complex organic compounds is largely a manual process that is often incompletely documented. To address these shortcomings, we developed an abstraction that maps commonly reported methodological instructions into discrete steps amenable to automation. These unit operations were implemented in a modular robotic platform using a chemical programming language which formalizes and controls the assembly of the molecules. We validated the concept by directing the automated system to synthesize three pharmaceutical compounds, Nytol, Rufinamide, and Sildenafil, without any human intervention. Yields and purities of products and intermediates were comparable to or better than those achieved manually. The syntheses are captured as digital code that can be published, versioned, and transferred flexibly between platforms with no modification, thereby greatly enhancing reproducibility and reliable access to complex molecules.

The automation of chemical synthesis is currently expanding, and this is driven by the availability of digital labware. The field currently encompasses areas as diverse as the design of new reactions *(1)*, chemistry in reactionware *(2)*, reaction monitoring and optimization *(3, 4)*, flow chemistry *(5)* for reaction optimization and scale up, to full automation of the synthesis

laboratory *(6)*. Established technologies like automated peptide *(7)* or oligonucleotide synthesis *(8)*, flow chemistry *(9)*, or high-throughput experimentation *(10)* are mainstays of modern chemistry, while emerging technologies like automated oligosaccharide synthesis *(11)* or iterative cross coupling *(12)* have the potential to further transform the chemical sciences. Each of these examples, however, relies on a distinct protocol, as there is no current digital automation standard for computer control of chemical reactions *(13)*. Hence, there is no general programming language for chemical operations that can direct the synthesis of organic compounds on an affordable, flexible, modular platform accessible to synthetic chemists, and could in principle encompass all synthetic organic chemistry. This situation is comparable to the *era* before digital programmable computers, when existing computational devices were fixed to a dedicated problem.

A generalized approach to automating chemical synthesis would be beneficial because making compounds is one of the most labor-intensive branches of chemistry, requiring manual execution of a range of unit operations such as reagent mixing, liquid-liquid extractions, or filtrations. Despite the modular nature of the operations, standardization and automation are still severely limited. Furthermore, the ambiguous way in which synthetic protocols are communicated has contributed to a mounting reproducibility crisis (*14*). Syntheses are reported in prose, omitting many details explaining exactly how operations were carried out and making many assumptions about the skill level of the researcher repeating the process. We hypothesized that a more standardized format for reporting a chemical synthesis procedure, coupled with an abstraction and formalism linking the synthesis to the physical operations, could yield a universal approach to a chemical programming language; we call this architecture and abstraction the *Chemputer*.

In developing the Chemputer platform we wanted to build on the two hundred or more years of the chemical literature, and the experience of the many thousands of bench chemists active

in the world today, naturally leading to a standardization embodied in a codified standard recipe or chemical code for molecular synthesis. For this to be possible, it was essential that the approach mirrors how the bench chemist works. We opted to base the implementation of the Chemputer upon the round-bottom flask working in batch as our key reaction module, with well-defined inputs and outputs. We based this choice on the fact that most protocols already published for complex molecule synthesis rely on this type of apparatus. Next, we identified the four key stages of a synthetic protocol from our abstraction: reaction, workup, isolation, and purification. These stages can be subdivided into several unit operations, which in turn are implemented in a specific automated platform.

Developing Code for Chemistry.

By developing control software as well as hardware modules for laboratory-scale synthesis which can be automatically cleaned and re-used in subsequent reaction steps, we were able to define a process for combining individual unit operations into full, multistep syntheses to produce desired products autonomously (Fig. 1A). For the Chemputer to operate as shown in Fig. 1B, the states of the inputs, reactor, and outputs must be defined and controlled programmatically. We therefore created a *Chempiler*, which is a program to produce specific low-level instructions for the modules that comprise the Chemputer architecture. It can run commands used to control the modular platform, from our abstraction layer, so that a typical written scheme can be turned into a specific code to run the modules with ease. Every module was then designed with drivers for the device or equipment and a standardized application programming interface (API) exposing the instruction set to the Chempiler (Fig. 2). The use of a program coupled to the Chemputer architecture allows users to directly run published syntheses without reconfiguration, provided the necessary modules and drivers are present within their system. Thus, the practical implementation of the Chemputer architecture converts digital code to chemical synthesis operations in accord with the standard protocol of a chemical

reaction based on the four processes described in Fig. 1C. To prevent the need for manual reconfiguration, the physical modules and their connections and representation are stored in memory as a directed graph which allows knowledge and control of their states in real-time.

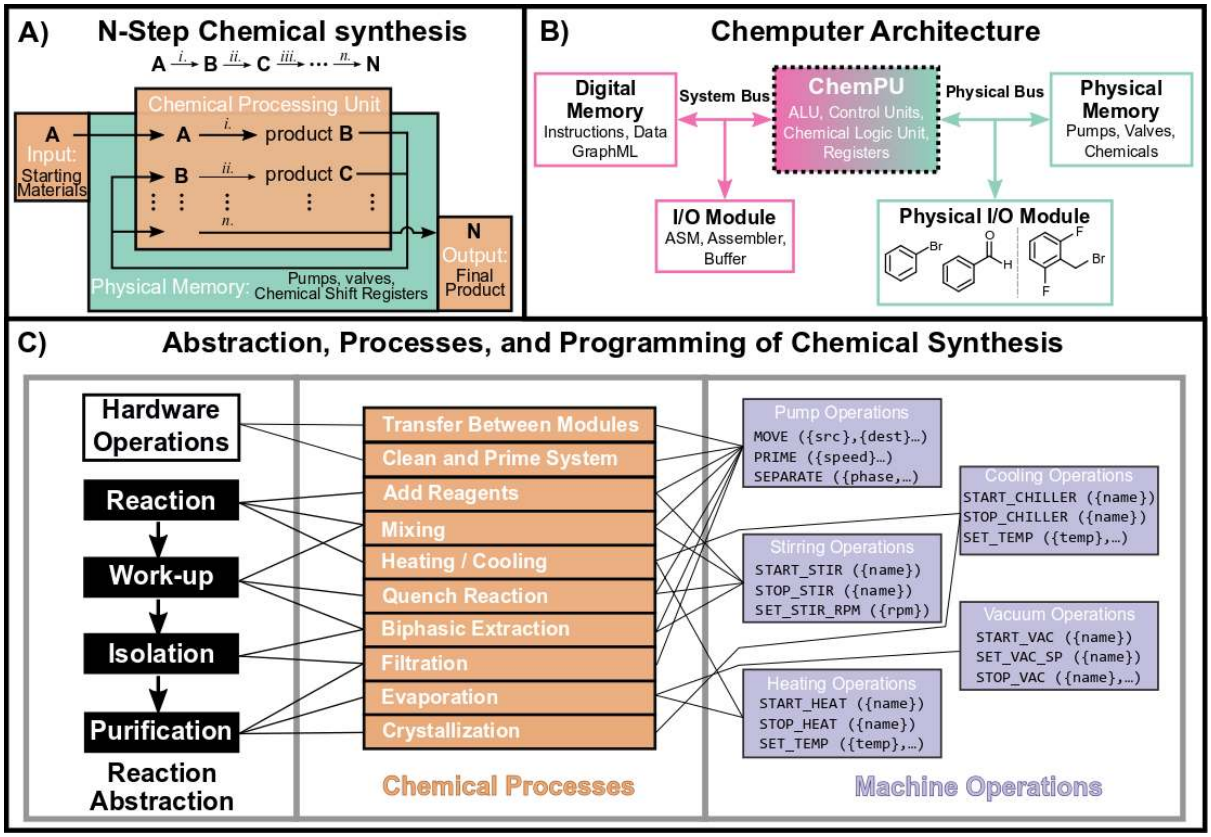

Fig. 1: **Operating principles of the Chemputer.** A) Schematic representation of a stepwise chemical synthesis, formalizing the reagents as inputs and products as output. B) Diagram outlining the computing-architecture of the Chemputer. C) The abstraction of chemical synthesis allowing development of an ontology that can be universally programmed using a machine. Similar to digital computers, the Chemputer has a memory and a bus system, but these are both digital and physical. By considering chemical reagents and products in a memory bus, it is possible to break the process of complex molecule synthesis into steps or cycles that can be run using the physical hardware.

The physical routing that links the connected modules is described as a graph using an open source format, GraphML, which allows the Chempiler to find paths between a source flask and a target flask, as well as address devices like hotplate stirrers based on the vessel they are connected to. GraphML is an open standard, extensible markup language (XML)-based exchange format for graphs *(15)* (example graphs used in the syntheses described herein can be found in supplementary file GraphML Files.zip). Synthetic procedures are codified using a

scripting language named Chemical Assembly (ChASM) which provides instructions for all currently implemented machine operations and supports the definition of functions and variables. To develop the ChASM code, we built a standard procedure, starting with a written synthetic scheme which is formalized using a chemical descriptive language or XDL. The XDL has the advantage of eliminating ambiguous interpretation of the synthesis procedures by explicitly and systematically listing all required information without making any assumptions or inferences (example ChASM files used in the syntheses described herein can be found in supplementary file ChASM Files.zip).

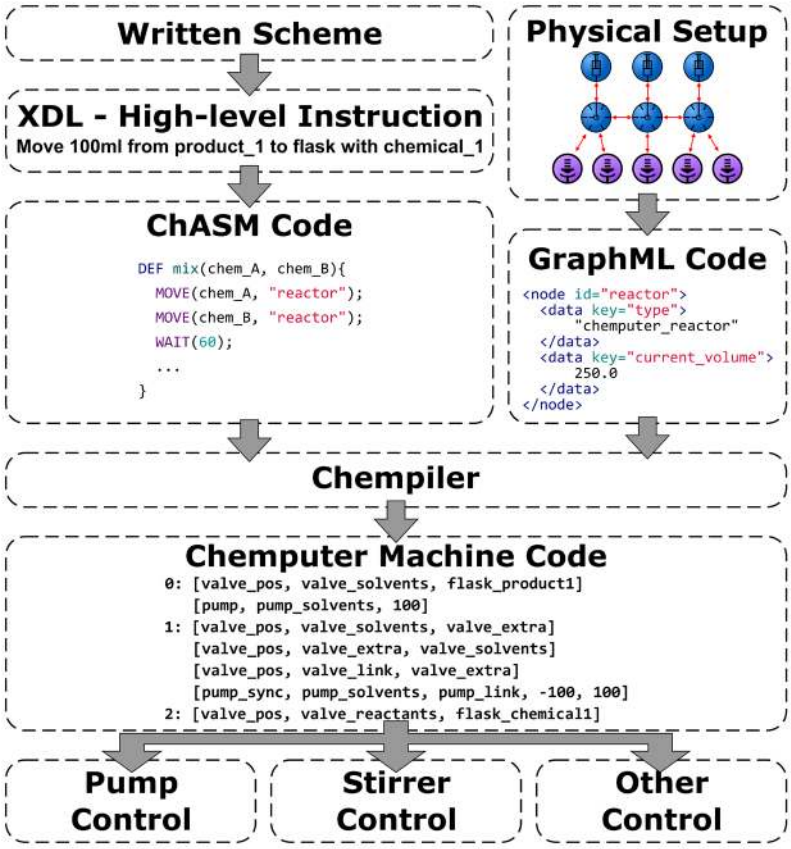

Fig. 2: **Diagram showing the flow of information in the synthesis platform.** To translate experimental steps into a set of pump movements or hotplate stirrer commands, the user provides the synthesis instructions in a text format similar to a written protocol. The graphical representation of the synthesis platform, which contains all necessary information about the fluidic connectivity, is represented using GraphML. The written scheme from a published procedure or lab book entry is translated into a series of ChASM commands. Both the ChASM and the GraphML file are passed to the Chempiler. The Chempiler command dispatcher then uses those two pieces of information to control the physical labware through the respective device drivers, effectively executing the synthesis.

To control the chemistry, the Chempiler was designed to accept ChASM commands, such as "start stirring the reactor", find the module in question in the GraphML definition, and schedule the execution using the appropriate low-level instructions. The modular Chemputer components thereby constitute a versatile and interoperable architecture for chemical synthesis. Also, a given ChASM file could be run on many different platform instances, using hardware of different makes and models, connected in different ways. If the required unit operation modules are present, and all required reagents and solvents are provided, the hardware-agnostic ChASM code can be freely paired with a system-specific GraphML file to synthesize the same product on another, different platform without re-optimization.

During the development of the process and programming language, we found it helpful to visualize the workflow one would follow when manually reproducing the procedure. Writing the XDL and converting it to ChASM then becomes intuitive even for users with no programming experience. Many operations are repetitive enough to be defined as functions that can be re-used many times. For instance, we have defined several functions such as "evaporate to dryness" or "add reagents and heat up to X ˚C", and we expect that this library can be greatly expanded with future use. Once the ChASM file is completed, a graph of the platform is drafted, following a set of rules detailed in the SM. To validate the ChASM and the graph, we implemented two simulation modes in the Chempiler. First, all operations are performed as they would be on the real system, but instead of issuing commands to physical devices, the commands are logged to a text file. Next the system simulates the process without reagents as a dry run. The simulations flag any potential issues such as syntax errors in the XDL representation and the ChASM, inconsistencies in the graph representation, impossible operations, or overfilling of vessels. Once the simulations run without errors, the user can load the reagents and solvents in accordance with the graph and start the synthesis. The Chempiler also has a breakpoint command option for additional safety and compatibility checks as needed.

Modular Synthesis Platform.

To produce a physical platform that could implement our architecture and achieve the syntheses outlined below, we had to step away from thinking about reactions and rather adopt a process centered way of thinking. Although the 20 most commonly used reaction classes in drug discovery *(16)* span a wide range of chemistries, from amide bond formation to Buchwald-Hartwig couplings, experimentally most of them simply involve mixing of several reagents in the correct order, often under heating or cooling. These processes are typically followed by a workup and a purification technique such as distillation, recrystallisation, or chromatography. We therefore concluded that a synthesis platform capable of performing the unit operations of mixing under heating or cooling, liquid/liquid separation, filtration, and solvent evaporation could in principle perform a large fraction of all organic syntheses and embody our abstraction of chemical synthesis (Fig. 1C). As those unit operations do not always occur in the same order, especially where multistep syntheses are concerned, a flexible means of moving material between the modules was required. To that end, we built the physical architecture around a fluidic backbone consisting of a series of syringe pumps and six-way selection valves (Fig. 3). The appeal of this design is that it is expandable: the user can always add more pump/valve units (backbone units) to the ends of the backbone. Material can be moved between modules in an arbitrary order, including multiple uses of the same module at different points in the sequence. The process to transfer liquid from a port on the backbone to another backbone is described in the SM (Fig S8). The pump on the source unit aspirates the appropriate amount, then the valve on the source unit and the adjacent unit switch to the bridge, and the source pump and the adjacent pump move simultaneously to transfer the liquid contents from the source syringe to the next syringe. This process is repeated, and the liquid is moved along the Backbone until it reaches the destination unit, which in turn dispenses it to the destination port. Surplus ports on the six-way valves accommodate solvents and reagents, and additional

backbone units may be entirely dedicated to supplying chemicals. This is in stark contrast with even the most advanced flow chemical setups to date *(17)*, which have to be physically reconfigured for every new synthesis. In flow, the number and sequence of unit operation modules must match the number and sequence of required unit operations, whereas in our system the ability to address modules independently and reuse them as required removes the need for physical reconfiguration. This is achieved using the GraphML, which gives a complete description of the connectivity, allowing the pumps and valves to be dynamically reconfigurable resources, and facilitates the movement of solvents and solutions from reagent flasks to the various components required for a given synthesis step. To ensure modularity we designed our own pumps and valves, controlled and powered by a single ethernet cable plugged into each item, powered from a network switch placed next to the fume hood (see Fig. S11).

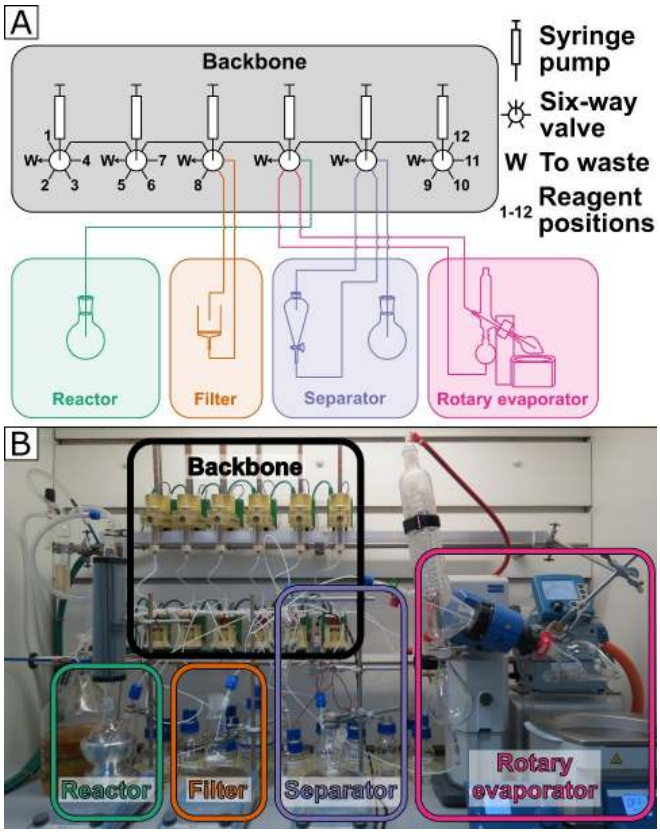

Fig. 3: **Physical implementation of the synthesis platform.** A) Schematic representation of the Chemputer, highlighting modules used for four commonly encountered unit operations. The lines represent fluidic connections. B) Photograph of one Chemputer setup used in this work. The various modules are highlighted in correspondence to the schematic.

In this work we developed modules for the unit operations of mixing, filtration, liquid/liquid separation, evaporation, and chromatographic separation which are all key to practically implement our abstraction of a chemical reaction. Detailed descriptions of the individual modules can be found in Fig. S34-S44 and accompanying supplementary text. However, both the physical architecture and the architecture of the Chempiler (*vide supra*) were specifically designed to allow for future addition of other modules, enabling automation of even more reactions.

The reactor (Fig. S34) consisted of a commercially available two-necked, pear-shaped flask equipped with an air condenser. We decided to use common laboratory glassware rather than a jacketed reactor vessel both for simplicity, and to lower the barrier for reproducing the setup. A pear-shaped flask was chosen over the more common round-bottomed flask to accommodate a wider range of reaction volumes. Heating and stirring were accomplished using a computer controllable stirrer hotplate and a custom manufactured aluminum heating block for the pear-shaped flask. A 1/8" O.D. PTFE tube was held in place by a ground glass joint to GL18 thread adapter with a GL18 screw cap and insert so that its end reached the bottom of the flask. A slight argon overpressure was maintained by the inert gas system (see SM). For the sildenafil synthesis cooling of the reactor was required, and we used a recirculation chiller with a temperature range of -30°C to 160°C, giving precise temperature control of the reactor.

The liquid/liquid separation, one of the most common isolation techniques, was also the most challenging task to automate in a robust fashion. While in continuous flow there are solutions utilizing membrane technology *(18)*, we found that commercially available hydrophobic frits are usually designed to be single use and therefore lack long-term reliability. We then attempted to employ a modified separating funnel and computer vision to directly replicate the way a human chemist would perform liquid/liquid extractions. However, we found that solutions based either on a colored floater *(19)* or on direct recognition of the phase boundary *(20)*

worked well in a test environment but were otherwise unreliable. In real syntheses, imperfections like poor phase separation, strongly colored or cloudy solutions, or unusual extraction solvents often lead to complete failure of the image recognition algorithms. We therefore abandoned the computer vision for a sensor-based approach. Initially we investigated optical and capacitive sensors because they do not require direct contact with the medium. Unfortunately, those sensors also performed poorly in some cases, so ultimately, we built a conductivity sensor from two pieces of stainless-steel tubing inserted into the flow path (see SM). This sensor reliably detected phase boundaries in all test cases and enabled us to perform separations in a robust fashion. The sensor was connected to a custom-made separating funnel with a B45 ground glass joint at the top and two B14 side arms (Fig. S36). Instead of a stopcock it had a glass ¼-28 UNF male threaded connector fitted to the bottom. An Arduino Due was used to read the sensor via a simple voltage divider circuit. The top inlet tube was suspended by a ground glass joint to GL18 thread adapter with a GL18 screw cap and insert. To facilitate efficient extractions through thorough mixing, a computer controlled overhead stirrer was fitted above the separator.

To perform liquid/liquid extractions, the mixture was pumped into the separator through the top inlet, or in case of a wash, the washing solvent was added through either the top or bottom port, depending on whether it constituted the top or bottom layer of the biphasic mixture. The two layers were then stirred vigorously with the overhead stirrer, followed by a period of settling under very slow (50 rpm) stirring. Then, the bulk of the lower phase was usually transferred to the target vessel to speed up the process. The volume moved was determined empirically for every separation. The actual separation commenced with withdrawing the dead volume of the sensor and tubing from the bottom port. The volume removed was dispensed into the lower phase target vessel. Then, a sensor reading was taken and compared against a reference value. If the reading was lower than the reference, it was assumed the lower phase

was organic, else it was assumed the lower phase was aqueous. Then, one milliliter was transferred to the lower phase target. Another sensor reading was taken and compared against a predefined threshold value. This threshold depended on whether the lower phase was aqueous or organic. Either way, if the sensor reading was outside the threshold, it was concluded that a phase change had been detected. If not, another milliliter was withdrawn, and the cycle continued, until a phase change was detected. Then, the dead volume of the sensor and tubing was transferred to the lower phase target vessel. If the upper phase target was specified as the separator, the process was concluded. Else, the upper phase was withdrawn as well and transferred to the target vessel.

For the solvent evaporation, a computer controllable rotary evaporator was modified by routing a piece of PTFE tubing through the vapor duct into the evaporation flask to pump product into and out of the flask (Fig. S37). The receiver flask was fitted with a glass ¼-28 UNF male threaded connector and a PTFE tube, allowing it to be emptied in situ. One complication was that after distillation at reduced pressure, upon venting, oily products were forced back into the tube reaching into the evaporation flask. This problem was solved by affixing a magnetic stirrer bar to the tube using PTFE shrink wrap (Fig. S38). A strong magnet was then positioned in such a way that it would attract the tube and lift it out of the product, allowing the system to be vented. When the flask was lowered into the heating bath, the tube would be released and drop, thereby allowing product to be withdrawn. Solvent evaporation started with pumping the solution to be evaporated into the distillation flask of the rotary evaporator. A cartridge filled with molecular sieve could be switched into the flow path by two six-way selection valves, allowing for the solution to be dried prior to evaporation. The flask was then lowered into the heating bath, and the rotation was started. The vacuum pump was started, lowering the pressure to 900 mbar to degas the solution and avoid excessive foaming later on. The heating bath and the recirculation chiller servicing the condenser were switched on, the target temperatures were

set, and execution of the script was suspended until the target temperatures were reached. The vacuum setpoint was then changed to the target distillation pressure, and the vacuum pump speed was adjusted according to the solvent, to avoid bumping. Then, the execution of script was suspended for a user-defined amount of time to allow the main distillation to finish. After the distillation was complete, the flask was lifted which caused the inlet tube to be attracted by the magnet (Fig. S38), lifting it out of the remaining solution. The vacuum pump was subsequently stopped, and the vacuum was vented. Then, a user-defined amount of distillate was removed from the distillate flask and discarded. The parameters (pressure, timings, volumes) were always chosen empirically, either through experimental trial and error, or from prior experiences with the system.

At this point, we found that proceeding directly to drying the product under maximal vacuum would often lead to a few milliliters of residual solvent distilling over, which decreased the drying efficiency. Thus, the flask was lowered back into the bath and the vacuum pump was set to maximum power for two minutes, drawing out any residual solvent. The sequence of raising the flask, venting the vacuum, and emptying the distillate flask was then repeated. Next, the flask was lowered once again, the vacuum pump was set to maximum power and started, and the cooling of the condenser was switched off. The product was then dried for a user-defined amount of time. After the drying was complete, the flask was lifted once more, the vacuum was vented, and the rotation and heating bath were switched off.

The filtration unit consisted of a custom-made, jacketed, sintered glass Büchner funnel (made in-house, see SM) fitted with a B29 ground glass joint at the top, two B14 side arms, and a glass ¼-28 UNF male threaded connector at the bottom. The top inlet tube was suspended by a ground glass joint to GL18 thread adapter with a GL18 screw cap and insert while the bottom outlet tube was connected to the threaded connector with a straight union piece. Stirring was accomplished by a computer controllable overhead stirrer.

To allow efficient drying of the precipitate, the bottom outlet of the filter was connected to the central inlet of a six-way valve. One outlet of that valve was then connected to the backbone, while another outlet was connected to the laboratory vacuum system via a Woulff bottle. This allowed the user to switch the filter bottom between the backbone (for liquid addition or withdrawal) and vacuum (for drying). The whole platform could be cleaned automatically by pumping suitable washing solvents into the modules. This cleaning cycle would return the system to its initial state, ready for the next reaction stage. Consequently, preparing the system for another synthesis was simply a matter of swapping the reagent bottles.

Proof-of-principle automated syntheses of three drugs.

To highlight the power of this approach we chose three targets, diphenhydramine hydrochloride **6** (Fig. 4), rufinamide **10** (Fig. 5), and sildenafil **17** (Fig. 6). The process of digitizing a synthesis always starts with a traditional, written scheme such as an experimental procedure from a publication or a lab book entry. We chose three published syntheses, all replicated manually to establish benchmark yields and purities for comparison with the automated runs.

Diphenhydramine hydrochloride is an ethanolamine derivative used as antihistamine and mild sleep aid. It is marketed as Nytol™ in the UK and as Benadryl® in the US. The synthesis is a four-step sequence starting with a Grignard reaction. Rufinamide is a triazole derivative used as an anticonvulsant to treat various seizure disorders and its synthesis is a relatively simple process to automate. Sildenafil is prescribed to treat erectile dysfunction and is best known under the brand name VIAGRA®; its industrial synthesis route (*21*) features a chlorosulfonation with highly aggressive chlorosulfonic acid and thionyl chloride. We reasoned that successful automated handling of those aggressive reagents would demonstrate the versatility and robustness of the system, as well as highlighting the safety benefit arising from automating dangerous procedures.

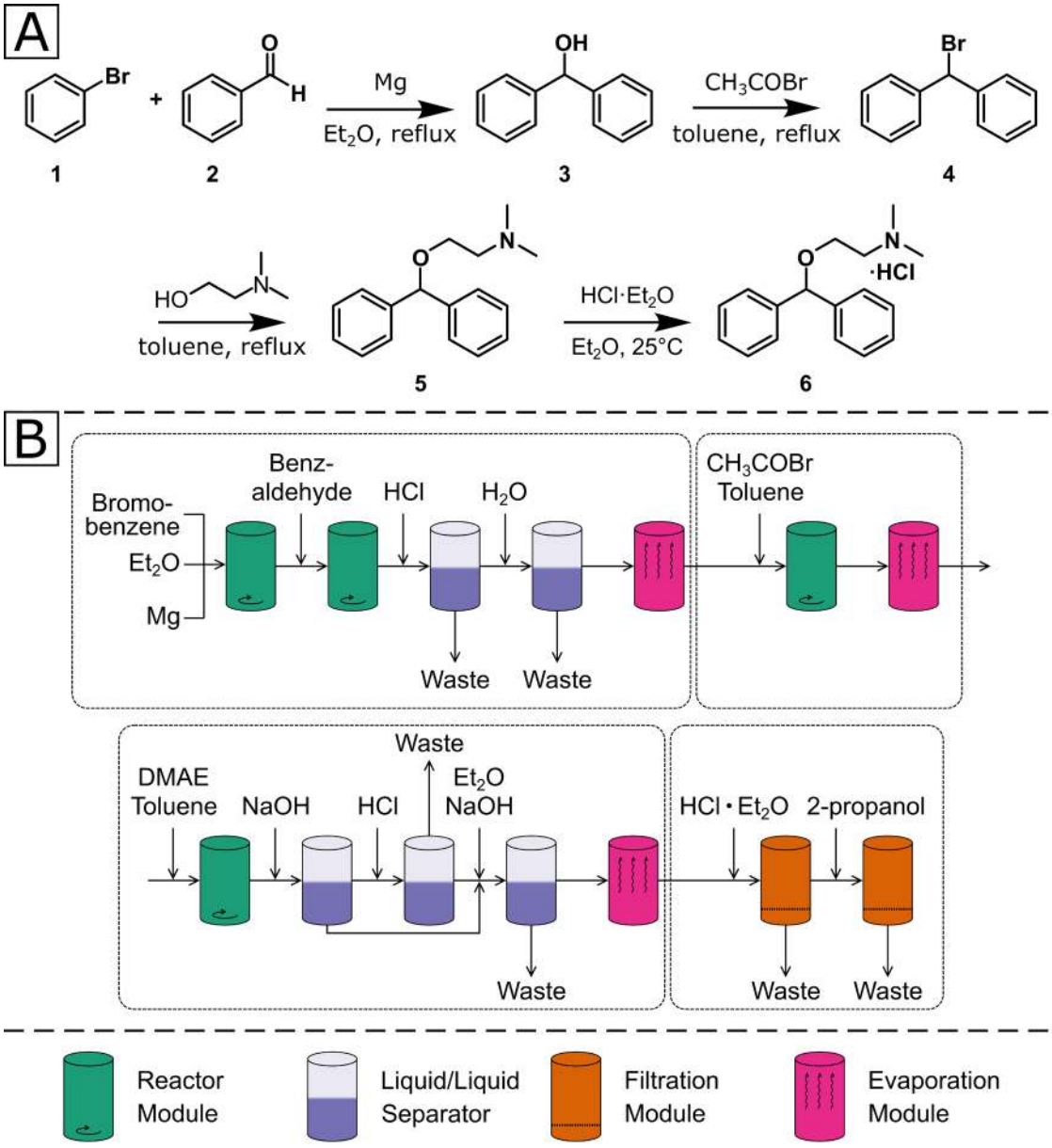

Fig. 4: **Synthesis of Diphenhydramine Hydrochloride 6.** A) modified synthetic route to diphenhydramine hydrochloride. B) Sequence of unit operations required for the synthesis. The dotted boxes denote the four stages of the synthesis. DMAE: dimethylaminoethanol.

After reproducing the synthesis of diphenhydramine hydrochloride **10** (*22*) manually, we made some small modifications and started the process on the platform. The synthesis commenced with Grignard reagent formation, for which the reactor was manually charged with dried magnesium grit. All the other required reagents and solvents were loaded into 100 mL or 250 mL standard GL45 bottles, and all non-aqueous storage bottles were purged with argon

and stored under positive pressure. All the operations described were performed by the Chemputer under full Chempiler control and the program used, as well as a description of the process in prose, can be found in the SM. The automated synthetic procedure was started by automatically priming the tubes to the chemical reservoirs, followed by auto-cleaning the backbone with water, isopropanol and dry diethyl ether. The system then continued autonomously through the whole synthesis of diphenhydramine hydrochloride without human intervention as follows. Diethyl ether and a small portion of bromobenzene were added to the magnesium and the mixture was heated under reflux to initiate the Grignard reagent formation. After cooling, the remaining bromobenzene was added at a rate of 1 mL/min and the mixture was again heated to reflux. Using syringe pumps for moving material allowed us to precisely control addition rates and thus increase reproducibility of synthetic protocols even further. Subsequently, a solution of benzaldehyde in diethyl ether was added at a rate of 1 mL/min and the mixture was held at reflux for another 5h. As the platform presented herein is largely a proof of concept, no process analytical technology (PAT) have been implemented yet, so all reaction times were determined beforehand and hard-coded into the ChASM script. However, the modular architecture of the platform and control software should make adding PAT to future iterations of the platform straightforward. After quenching of the reaction with dilute HCl the layers were separated, the organic layer was washed with water, and concentrated *in vacuo* as described in the previous chapter. The system then cleaned itself and directly proceeded with the bromination. To that end, the reactor was charged with acetyl bromide, the crude diphenylmethanol **3** was transferred from the rotary evaporator flask to the reactor with three portions of toluene, and the mixture was heated to reflux, all without human intervention. After a predetermined reaction time, the mixture was transferred to the rotary evaporator and evaporated to dryness. The subsequent Williamson ether synthesis was initiated in a similar fashion, and after a predetermined time the reaction was quenched with aqueous sodium

hydroxide. The system then automatically conducted an aqueous workup and concentrated the product *in vacuo*. Once again, the system cleaned itself, charged the jacketed filtration module with hydrochloric acid, and transferred the crude free base **5** to the filter with three portions of diethyl ether. To ensure smooth precipitation, the mixture was stirred vigorously and the free base solution was added very slowly. After the addition was completed, the off-white precipitate was collected by automatic filtration and recrystallized from isopropanol, utilizing the heating and cooling capabilities of the jacketed filter. Drying at 60°C in a stream of argon for one hour yielded pure diphenhydramine hydrochloride giving an isolated yield of 58% over four steps or 87% per step on average. While this is slightly less than the 68% overall achieved manually, the average yield per step (87% automated vs. 91% manual) is comparable in our view. Overall the platform performed the synthesis fully automatically in 77 h (see movie S1) while the manual synthesis took four work days.

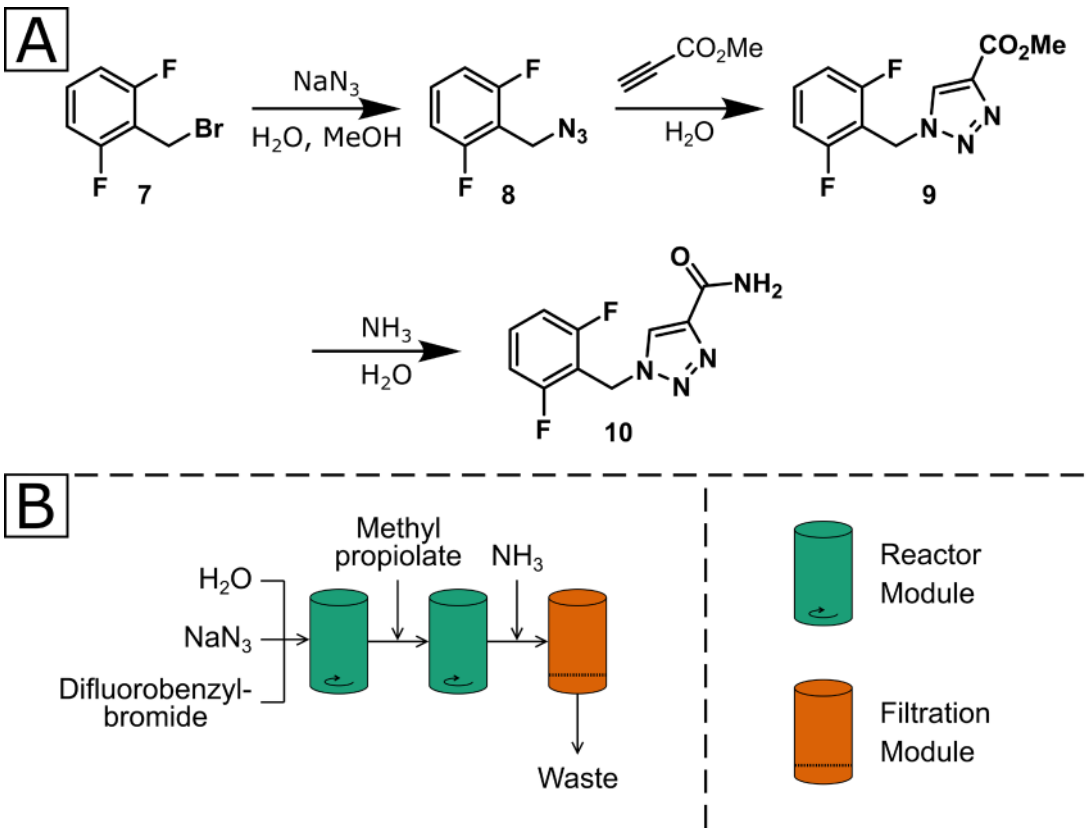

Fig. 5: **Synthesis of Rufinamide 10.** A) synthetic route to rufinamide. B) Sequence of unit operations required for the synthesis.

Next, we conducted an automated synthesis of the antiseizure drug rufinamide **10**, a triazole derivative commonly prepared via click reaction between the corresponding azide **8** and methyl propiolate (Fig. 5) (*23*). The synthesis began with an azide formation for which the reactor was charged manually with 2,6-difluorobenzyl bromide **7**; the remaining reagents were provided in bottles as described above. From here on, unless explicitly stated otherwise, all described operations were performed by the Chemputer under full Chempiler control. An aqueous solution of sodium azide was added to the reactor to prepare the organic azide for the triazole click with methyl propiolate, which was performed inside the jacketed filtration module. Subsequent saponification with aqueous ammonia led to precipitation of the target compound. Filtration followed by three aqueous washes yielded pure rufinamide at 46% isolated yield, which was slightly better than the manual synthesis (38%). The automated synthesis took 38 h. To demonstrate the power of the Chempiler software, and the interoperability of the code, we then went on to run the same ChASM file on a "full scale" platform, equipped with slightly different hardware, which was connected in an entirely different way. The platform produced pure rufinamide in 44% yield without any problems or changes to the code (movie S2).

In the next synthesis we prepared sildenafil, better known under the brand name VIAGRA® as shown in Fig. 6 (*15*). For this synthesis, we fitted the reactor with a heating block connected to the recirculation chiller, allowing us to cool as well as heat. From here on, unless explicitly stated otherwise, all described operations were performed by the Chemputer under full Chempiler control. The reactor was cooled to 15°C and automatically charged with chlorosulfonic acid, thionyl chloride, and molten ethoxybenzoic acid. Chlorosulfonic acid is corrosive, so when writing the ChASM script we took great care to minimize contact times and enacted a strict cleaning regime. Chlorosulfonic acid and thionyl chloride also react violently with trace amounts of water, producing large volumes of gas. Therefore, the backbone was automatically flushed with dry diethyl ether and dried with a small amount of thionyl chloride

before charging the reactor. After a predetermined reaction time the filtration module was charged with water and cooled to 0°C. The reaction mixture was then slowly dripped into the water, quenching the excess thionyl chloride and chlorosulfonic acid and precipitating the product **12**, which was collected by automated filtration. The subsequent sulfonamide formation with N-methylpiperazine **13** in water was performed in the filtration module as well. Unfortunately, the sulfonamide **14** did not crystallise spontaneously, so a slurry of a small amount of product in water was added to seed the crystallisation.

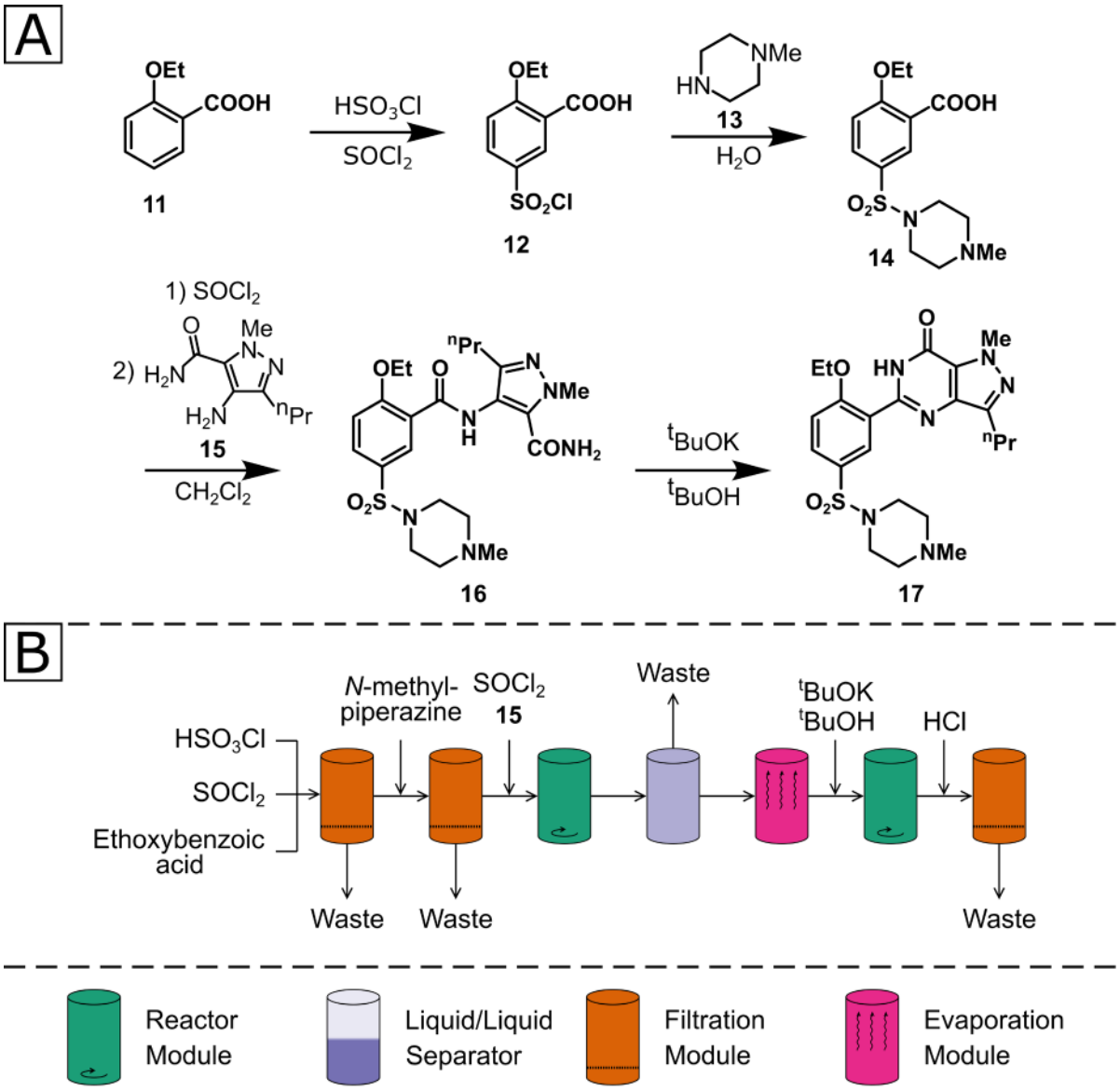

Fig. 6: **Synthesis of Sildenafil 17.** A) synthetic route to sildenafil. B) Sequence of unit operations required for the synthesis.

The industrial process for sildenafil employs *N*,*N*'-carbonyldiimidazole (CDI) for the amide coupling in the next step. However, this reaction did not work in our hands, neither manually nor in automation, thus we decided to go *via* the acid chloride instead. The carboxylic acid **14** was thoroughly dried by flowing argon through the filter cake while at the same time heating the filtration module to 60 °C, followed by acid chloride formation with thionyl chloride in dichloromethane. The reactor module was charged with a solution of 4-amino-1-methyl-3-n-propyl -1H- pyrazole-5-carboxamide **15** in dichloromethane and triethylamine and cooled to 10°C. The crude acid chloride solution was then pumped from the filter module to the reactor, and the reaction was quenched with water, the layers separated, the organic layer was dried over activated molecular sieves and evaporated to dryness, yielding amide **16** as an off-white solid. For the subsequent cyclization, the crude amide was transferred back to the reactor by dissolving it with a solution of potassium *tert*-butoxide in *tert*-butanol, and the mixture was heated at reflux for 8 h. After cooling to 10°C, the reaction was quenched with water, and the solution was transferred to the filter module. To induce precipitation of the sildenafil, the mixture was neutralized with aqueous hydrochloric acid. After filtration, the solid was washed with water and dried under a stream of argon at 50°C, yielding sildenafil at 44% isolated yield over 102 h (movie S3).

Outlook

The complete automation of all of synthesis is an ambitious objective, but this work makes a start towards that goal as the Chemputer architecture presents a general abstraction of the process that works with traditional bench scale techniques. The versatile programming language, use of traditional and inexpensive labware (the total cost of the parts for the robotic modules including non-standard glassware is less than $10 K per system) (*25*) and the payoff in reproducibility after validation of the process, means adoption could be straightforward.

Initially the synthesis of compounds will be validated reaction by reaction, but we imagine that eventually it will be possible to go straight from a reaction database to chemical code that can run the platforms.

Materials and methods summary

The manual and automated syntheses of the three target molecules, as well as detailed descriptions of the platform and the control software are described in detail in the supplementary materials and have been deposited in a repository (*26*) along with the ChASM code and GraphML. Videos of the automated syntheses are available as Supplementary Movies S1 – S3. We will continue to update the ChASM code for the syntheses and updates of this will be available to download (www.chemify.org). A brief summary of the syntheses is provided below.

Synthesis of Diphenhydramine Hydrochloride

The reactor module was charged manually with magnesium grit and dried by heating to 150 °C under a stream of argon for 15 minutes. After cooling to room temperature (approx. 25 °C) diethyl ether and 2 mL of bromobenzene were added to the magnesium and the mixture was heated to reflux for 20 minutes. After cooling below 25 °C, 8.65 mL of bromobenzene were added at a rate of 1 mL/min and the mixture was again heated to reflux for 20 minutes. Subsequently a solution of benzaldehyde in diethyl ether were added at 1 mL/min and the mixture was held at reflux for 5h. After quenching of the reaction with dilute HCl the layers were separated, the organic layer was washed with water and evaporated to dryness, yielding crude diphenylmethanol. After automatic cleaning of the system, the reactor was charged with acetyl bromide and the crude diphenylmethanol was transferred from the rotary evaporator to the reactor with three portions of toluene. The mixture was heated to reflux for 4h and subsequently evaporated to dryness, yielding crude bromodiphenylmethane. The system was

automatically cleaned once more, the reactor was charged with 2-(dimethylamino)ethanol and 10 mL of toluene, and the bromodiphenylmethane was transferred to the reactor with three portions of toluene. The mixture was heated to reflux for 20h. After cooling below 30 °C, the reaction was quenched with aqueous sodium hydroxide. The layers were separated, and the organic layer was extracted with 2M aqueous hydrochloric acid three times. Equimolar aqueous sodium hydroxide was added to the combined aqueous layers and the mixture was extracted with diethyl ether three times. The combined etheric layers were evaporated to dryness, yielding crude diphenhydramine free base. The jacketed filter module was then charged with etheric hydrochloric acid and the crude free base was slowly transferred to the filter with three portions of diethyl ether. The precipitate formed was collected by filtration, dried under a stream of argon, and recrystallized from isopropanol, yielding pure diphenhydramine hydrochloride as white crystalline powder.

Synthesis of Rufinamide

The reactor was charged manually with 2,6-difluorobenzyl bromide. An aqueous solution of sodium azide was added to the reactor, the mixture was heated to 75 °C for 12h and subsequently transferred to the jacketed filter module. Neat methyl propiolate was added and the mixture was heated to 65°C for 4h.An aqueous solution of ammonia was subsequently added and the mixture was held at 75°C for an additional 12h, precipitating the target compound. Filtration followed by three aqueous washes yielded pure rufinamide as white crystalline powder.

Synthesis of Sildenafil

The reactor was automatically charged with chlorosulfonic acid, thionyl chloride, and molten ethoxybenzoic acid. The mixture was stirred at 15 °C for 30 minutes, followed by stirring at room temperature for an additional 12h. Subsequently, the filtration module was charged with

water and cooled to 0°C. The reaction mixture was slowly dripped into the water, quenching the excess thionyl chloride and chlorosulfonic acid and precipitating the product. The supernatant solution was removed, and the precipitate was washed with cold water, yielding 5-chlorosulfonyl-2-ethoxybenzoic acid as white powder. The wet solid remaining in the filter module was subsequently slurried in cold water and N-methylpiperazine was added slowly. After 5 minutes, crystallization was initiated by adding a slurry of a small amount of product. The solid was collected by filtration, washed with cold water, and dried under a stream of argon at 50°C, yielding 2-ethoxy-5-(4-methyl-1-piperazinesulfonyl)-benzoic acid as a white powder. The dry carboxylic acid was slurried in dichloromethane and cooled to 5 °C. Thionyl chloride and a catalytic amount of dimethylformamide were added and the mixture was stirred for 5h at 25 °C. Subsequently, the reactor module was charged with a solution of 4-amino-1-methyl-3-n-propyl -1H- pyrazole-5-carboxamide in dichloromethane and triethylamine and cooled to 10°C. The crude acid chloride solution was pumped from the filter module to the reactor and the mixture was stirred for 16h at 25 °C. After quenching the reaction with water, the layers were separated, the organic layer was dried over activated molecular sieves and evaporated to dryness, yielding 4-[2-ethoxy-5-(4-methyl-1-piperazinylsulfonyl)-benzamido]-1-methyl-3-propyl-1H-pyrazole-5-carboxamide as an off-white solid. For the subsequent cyclization, the crude amide was transferred back to the reactor by dissolving it with a solution of potassium *tert*-butoxide in *tert*-butanol, and the mixture was heated at reflux for 8 h. After cooling to 10°C, the reaction was quenched with water, and the solution was transferred to the filter module. To induce precipitation of the sildenafil, the mixture was neutralized with aqueous hydrochloric acid. After filtration, the solid was washed with water and dried under a stream of argon at 50°C, yielding the title compound as white crystalline powder.

References and Notes:

1. J. Li, S. G. Ballmer, E. P. Gillis, S. Fujii, M. J. Schmidt, A. M. E. Palazzolo, J, W. Lehmann, C. F. Morehouse, M. D. Burke, Synthesis of many different types of

organic small molecules using one automated process. *Science* **347**, 1221-1226 (2015).
2. P. J. Kitson, G. Marie, J.-P. Francoia, S. S. Zalesskiy, R. C. Sigerson, J. S. Mathieson, L. Cronin, *Science,* **359**, 314-319 (2018).
3. C. Rougeot, H. Situ, B. H. Cao, V. Vlachos, J. E. Hein, Automated Reaction Progress Monitoring of Heterogeneous Reactions: Crystallization-Induced Stereoselectivity in Amine-Catalyzed Aldol Reactions. *React. Chem. Eng.* **2**, 226-231 (2017).
4. H. Florian, L. M. Roch, C. Kreisbeck, A. Aspuru-Guzik, *ACS Cent. Sci.* Article ASAP DOI: 10.1021/acscentsci.8b00307 (2018).
5. A. Adamo *et al.*, On-demand continuous-flow production of pharmaceuticals in a compact, reconfigurable system. *Science* **352**, 61-67 (2016).
6. A. G. Godfrey, T. Masquelin, H. Hemmerle, A remote-controlled adaptive medchem lab: an innovative approach to enable drug discovery in the 21st Century. *Drug Discov Today* **18**, 795-802 (2013).
7. R. B. Merrifield, Automated synthesis of peptides. *Science* **150**, 178-185 (1965).
8. G. Alvarado-Urbina *et al.*, Automated synthesis of gene fragments. *Science* **214**, 270-274 (1981).
9. M. B. Plutschack, B. Pieber, K. Gilmore, P. H. Seeberger, The Hitchhiker's Guide to Flow Chemistry. *Chem Rev* **117**, 11796-11893 (2017).
10. G. Godfrey, T. Masquelin, H. Hemmerle, A remote-controlled adaptive medchem lab: an innovative approach to enable drug discovery in the 21st Century. *Drug Discov Today* **18**, 795-802 (2013).
11. O. J. Plante, E. R. Palmacci, P. H. Seeberger, Automated solid-phase synthesis of oligosaccharides. *Science* **291**, 1523-1527 (2001).
12. J. Li, S. G. Ballmer, E. P. Gillis, S. Fujii, M. J. Schmidt, A. M. E. Palazzolo, J. W. Lehmann, G. F. Morehouse, M. D. Burke, Synthesis of many different types of organic small molecules using one automated process. *Science* **347**, 1221-1226 (2015).
13. P. D. Karp, Pathway databases: A case study in computational symbolic theories. *Science* **293**, 2040-2044 (2001).
14. M. Baker, 1,500 scientists lift the lid on reproducibility. *Nature* **533**, 452-454 (2016).
15. http://graphml.graphdrawing.org/ accessed 19/10/2018
16. D. G. Brown, J. Boström, Analysis of Past and Present Synthetic Methodologies on Medicinal Chemistry: Where Have All the New Reactions Gone? *J Med Chem* **59**, 4443-4458 (2016).
17. A. C. Bedard *et al.*, Reconfigurable system for automated optimization of diverse chemical reactions. *Science* **361**, 1220-1225 (2018).
18. A. Adamo, P. L. Heider, N. Weeranoppanant, K. F. Jensen, Membrane-Based, Liquid-Liquid Separator with Integrated Pressure Control. *Industrial & Engineering Chemistry Research* **52**, 10802-10808 (2013).
19. M. O'Brien, P. Koos, D. L. Browne, S. V. Ley, A prototype continuous-flow liquid-liquid extraction system using open-source technology. *Org Biomol Chem* **10**, 7031-7036 (2012).
20. Eppel, S.; Kachman, T. Computer vision-based recognition of liquid surfaces and phase boundaries in transparent vessels, with emphasis on chemistry applications *arXiv.org e-Print archive* [Online], 2014. http://arxiv.org/abs/1404.7174.
21. D. J. Dale *et al.*, The chemical development of the commercial route to sildenafil: A case history. *Organic Process Research & Development* **4**, 17-22 (2000).

22. Ahmadi, M. Khalili, R. Hajikhani, N. Safari, B. Nahri-Niknafs, Anti-inflammatory effects of two new methyl and morpholine derivatives of diphenhydramine on rats. *Medicinal Chemistry Research* **21**, 3532-3540 (2011).
23. R. D. Padmaja, K. Chanda, A Short Review on Synthetic Advances toward the Synthesis of Rufinamide, an Antiepileptic Drug. *Organic Process Research & Development* **22**, 457-466 (2018).
24. M. A. Gouda, W. S. Hamama, Overview of the synthetic routes to sildenafil and its analogues. *Synthetic Communications* **47**, 1269-1300 (2017).
25. The total cost of the parts for robotic modules including non-standard glassware is less than $10K per system but this rises to $30K if the stirrers, evaporator, and chiller system are included although these components might reasonably be expected to be found in most well equipped synthesis laboratories.
26. https://doi.org/10.5281/zenodo.1481731

Acknowledgments: We would like to thank Matthew Craven for help with the XDL development, Hessam Mehr, Sergey Zalesskiy, Mark Symes, and Richard Hartley for comments on the manuscript.

Funding: The authors gratefully acknowledge financial support from the EPSRC (Grant Nos EP/H024107/1, EP/J015156/1, EP/K021966/1, EP/L015668/1, EP/L023652/1), ERC (project 670467 SMART-POM).

Author Contributions: LC conceived the initial concept design approach and coordinated the team. PJK, GAC, GK and LC developed the abstraction. LC, TH and SG developed the fluidic backbone and initial versions of the pumps and valves including the electronics and firmware. TH outlined the initial XDL with LC. SS built the final versions of the rig, developed ChASM, and conducted the syntheses with help from JW, AA, and JMG. DA helped with the sildenafil automation. GK and GAC developed the software system with SS. LC and SS wrote the manuscript.

Competing interests: LC is the founder and director of DeepMatterPLC (deepmatter.io) which aims to digitize chemistry.

Data and materials availability: All the software to run the chemputer to synthesize the compounds described here have been deposited on Zenodo (*26*).

Supplementary Materials

Materials, Methods, and analytical data for the compounds synthesized and full specifications for construction of chemputer 1.0 with XDL and ChASM 1.0. The code (ChASM and GraphML) to synthesize the# compounds.

Figures S1 to S58

Movie S1 – S3

ChASM Files.zip

GraphML Files.zip