nature synthesis

Article

https://doi.org/10.1038/s44160-023-00473-6

Universal chemical programming language for robotic synthesis repeatability

Received: 31 March 2023

Accepted: 11 December 2023

Published online: 11 January 2024

Check for updates

Robert Rauschen [1,3], Mason Guy [2,3], Jason E. Hein [2] & Leroy Cronin [1]

The amount of chemical synthesis literature is growing quickly; however, it takes a long time to share and evaluate new processes among laboratories. Here we present an approach that uses a universal chemical programming language (χDL) to encode and execute synthesis procedures for a variety of chemical reactions, including reductive amination, ring formation, esterification, carbon–carbon bond formation and amide coupling on four different hardware systems in two laboratories. With around 50 lines of code per reaction, our approach uses abstraction to efficiently compress chemical protocols. Our different robotic platforms consistently produce the expected synthesis with yields up to 90% per step, allowing faster and more secure research workflows that can increase the throughput of a process by number-up instead of scale-up. Chemputer-type platforms at the University of Glasgow and the University of British Columbia Vancouver were used, as well as Opentrons robots and multi-axis cobotic robots to distribute and repeat experimental results. Protocols for three case studies involving seven reaction steps and three final compounds were validated and disseminated to be repeated in two international laboratories and on three independent robots.

Repeatability and falsifiability are crucial for scientific research[1], but the ever-increasing amount of published data makes it increasingly difficult to ensure the validation of published results[2,3]. Furthermore, the development and integration of automation and machine learning is currently transforming the field of chemistry[4–12]. Currently, the absence of an open standard for encoding and recording both successful and failed experiments adversely affects research progress by preventing the use of elaborated digital tools. Concerning studies even talk about a 'reproducibility crisis' in science[13], highlighting the importance of developing new strategies for the efficient validation of data. This includes chemical reactions, where a characteristic—almost technical—language style has emerged for communicating experimental protocols. Despite this common language, reproducing reactions remains challenging and often requires the expertise of human chemists to interpret prosaic protocols and infer implied information. Without effective methods to capture the tacit knowledge in chemical protocols, reliable communication and validation of experiments will become intractable, and advancements in chemistry will fail to meet their full potential.

The challenge of standardizing automated chemistry can be addressed by capturing chemical knowledge in χDL, a machine-readable, universal, chemical description language[14–17] that allows the user to store the experimental protocols in a standardized way that can be understood by both humans and robots. χDL is platform-agnostic by design and capable of interfacing with any automated chemical hardware. It was prototyped on the Chemputer platform[17–20] but is flexible enough to allow for rapid integration onto systems designed for completely different chemical purposes. We showcase this by executing the same χDL protocols on an Opentrons platform, a Kinova multi-axis cobotic robot and two Chemputers, emphasizing that the same code is equally executable on multiple hardware architectures. Unfettered by the hardware limitations of the individual platforms, χDL is used to combine different machines into one workflow and enhance the overall capabilities of the chemist designing the experiment.

[1]Advanced Research Centre, University of Glasgow, Glasgow, UK. [2]Department of Chemistry, The University of British Columbia, Vancouver, British Columbia, Canada. [3]These authors contributed equally: Robert Rauschen, Mason Guy. e-mail: jhein@chem.ubc.ca; lee.cronin@glasgow.ac.uk

https://doi.org/10.1038/s44160-023-00473-6

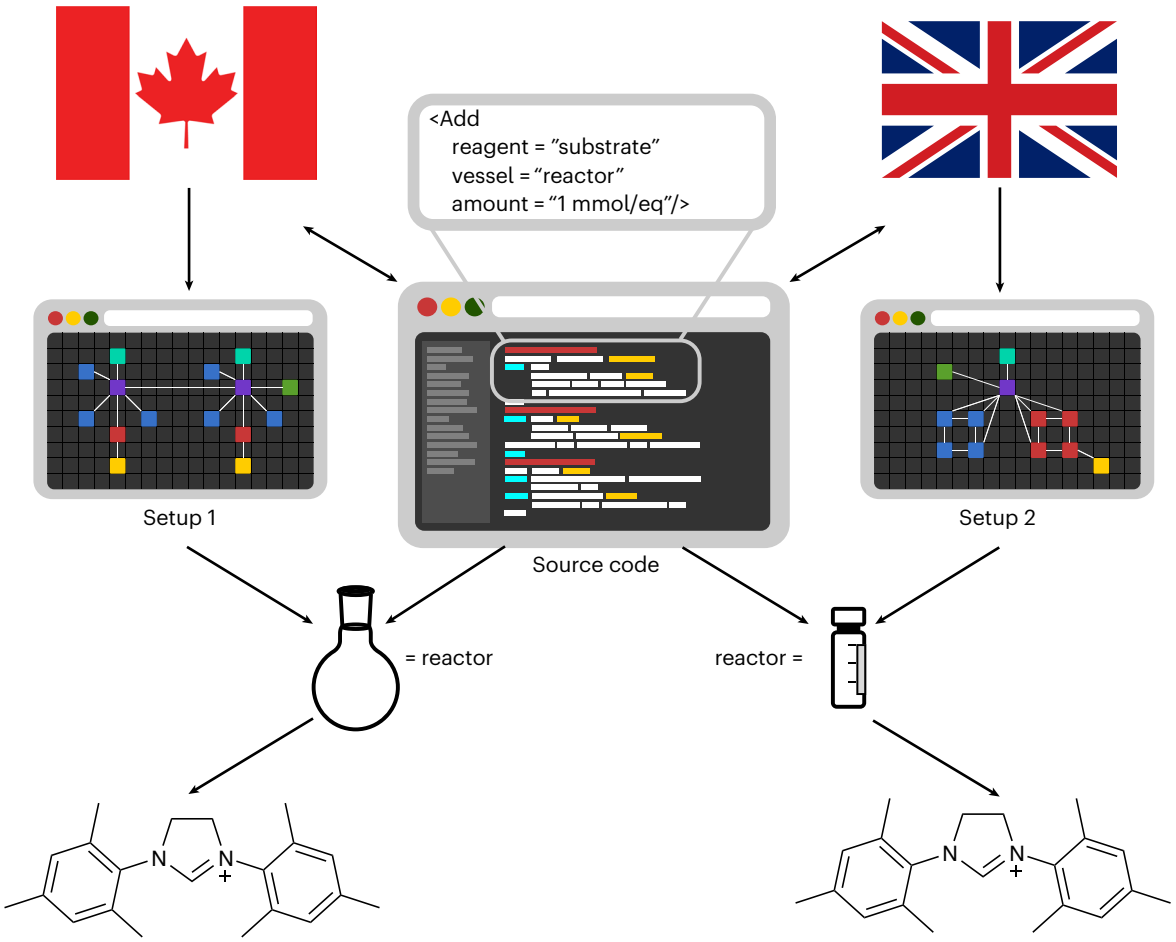

Fig. 1 | Universality of χDL. The design principle of χDL allows multiple researchers to collaborate on the same synthetic problem by sharing the χDL source code. If the same χDL code is executed on different hardware configurations (represented as graphics in the figure), the molecular output will be the same.

When captured as χDL protocols, automated procedures provide the added advantage of being almost instantaneously verifiable. For example, if two laboratories have congruent robotic hardware in place that has been equipped with an interface to execute χDL protocols, sending χDL code from one location to the other allows the immediate execution of the protocol on a physically different robotic system. Herein, 'congruent robotic hardware' explicitly does not entail having the same modules in place but modules that can execute the same chemical unit operations as defined in the step overview of the χDL standard[21].

Sharing digital procedures in the form of χDL facilitates collaborative research projects wherein the merits and labour forces of different platforms and laboratories can be joined to optimize a synthesis protocol, collectively achieving higher impact (Fig. 1). Traditionally, protocol and method descriptions are often compressed, reformatted or visualized in the supplementary information of a publication to fulfil layout and readability requirements, leading to a loss of information. This loss is minimized with the platform-agnostic χDL approach, because the abstract version of the code is designed to capture in adequate detail all the chemical information that is needed to repeat the synthetic protocol. The information about the low-level hardware operations to complete the protocol is not contained in the abstract χDL protocol itself but in the software package that is responsible for 'chempiling' the protocol on a robotic platform. Thus, a χDL file might be translated into a different sequence of hardware-specific commands on a different platform (for example an 'Add' step could be completed as a sequence of aspirate-dispense cycles or with a peristaltic pump) but the chemical output of the protocol remains the same. When chempiled on the same type of platform, the inferred hardware operations will be identical. By this means, χDL minimizes ambiguities in interpretation.

As part of this new way of collaborating, larger projects are divided into subtasks, such as breaking a multistep synthesis into individual reactions. These subtasks can be easily distributed to different robots within or between laboratories because they have a common software language to communicate. Applying this methodology to larger scale collaborations or in industry can result in delocalized supply chains for a central research facility and efficient knowledge distribution in divergent research projects. Chemical results can be recorded[17] and spread for validation and further usage by a 'host' to different 'peers', as well as from peer to peer with the χDL protocol allowing a decentralized cloud-like architecture for joint work (Fig. 2). We showcase both host-to-peer and peer-to-peer χDL transfers in the context of the tetramethyl *N*-methyliminodiacetic (TIDA) boronate protocol described herein. If desired, one instance can synthesize all the components needed for a specific target by requesting protocols for the different building blocks from arbitrary locations. Alluding to the BitTorrent software, we dub this concept for efficient collaboration 'ChemTorrent'.

By that naming, we do not imply a direct map to the BitTorrent software that is used to efficiently share large data files in a collaborative network of host computers. Instead, we expand the concept of torrenting beyond the digital realm, because each χDL protocol represents a physical molecule. Unlike data packages, it is not helpful to divide χDL code into pieces arbitrarily but only into segments that constitute a reaction step after which an isolatable substance is formed. Those pieces can then be distributed just as they would in a BitTorrent network. The host will be the platform that initially

https://doi.org/10.1038/s44160-023-00473-6

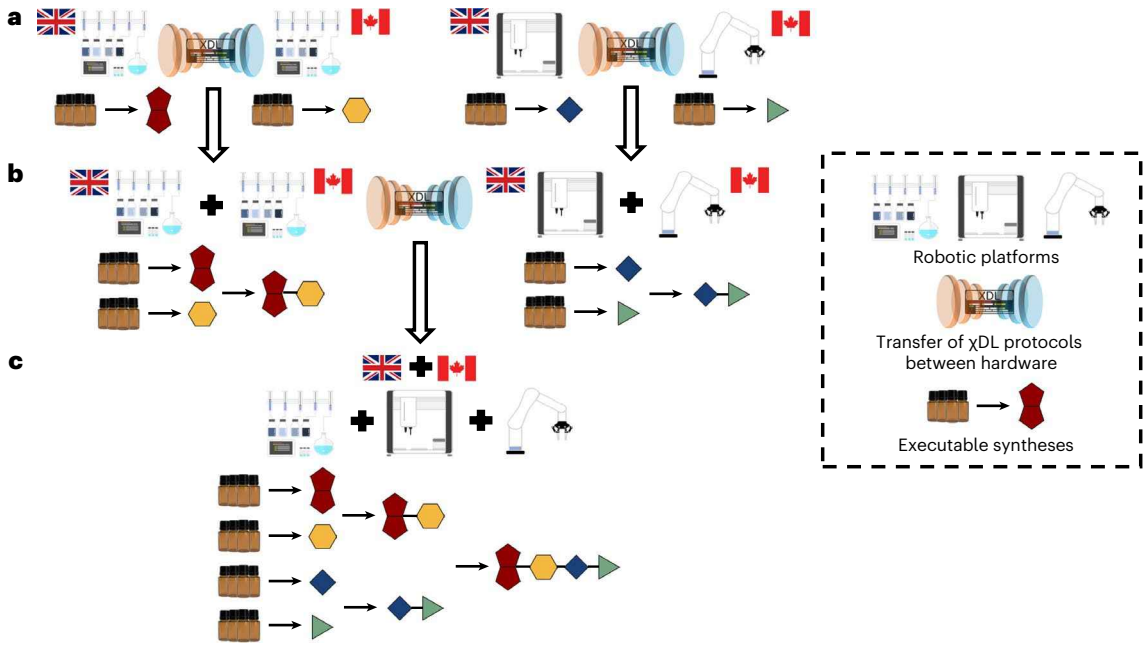

Fig. 2 | Convergent synthesis of complex molecules with the ChemTorrent approach. a, Building block synthesis can be split across different machines and laboratories before being validated and shared with others. **b**, Once building blocks can be made in both laboratories, the coupling reactions can be split between platforms and optimized. **c**, The protocols for all building blocks and couplings are shared between all systems and can be combined into the final product anywhere.

developed and optimized a χDL protocol. Peers will need to validate a χDL piece against the reported characterization data before they are allowed to share it with other peers, just as the credibility of a data piece in a torrent is validated with a cryptographic hash. After validation, the peer can choose to optimize and update the χDL protocol and become a host for the updated protocol. In this manner, the dynamic allocation of chemical resources in different laboratories facilitates the assembly of elaborated χDL sequences to achieve the synthesis of complex targets.

Results and discussion

In the following sections we describe how our laboratories in Canada and Scotland worked on χDL protocols together. The process usually involved repeating a sequence of the following kind:

(1) Develop and optimize a protocol and become a host for sharing that protocol.
(2) Share it with the other laboratories that take the role of peers.
(3) The peer runs the protocol as it is for validation purposes.
(4) If desired, the peer can go back to step 1 and become a host itself.

By validation, we mean matching the characterization data that is relevant to the synthetic task. For syntheses performed with the intent of obtaining usable material, we compare yields and purities just as Rohrbach et al.[17] did in an earlier publication (see their section 'reproducibility of the ChemPU synthesis'), additionally including typical characterization data for organic compounds, such as nuclear magnetic resonance (NMR) spectroscopy, high performance liquid chromatography (HPLC) and mass spectrometry. For protocols performed where yield and purity are not the primary objectives, we instead compare conversion by HPLC or by NMR spectroscopy.

Transferring protocols between platforms of the same type

Synthesis of H_2IMes•HBF_4 (compound 4). To demonstrate that communication of χDL protocols enables repeatable chemical procedures regardless of location or hardware configuration, we chose a three-step synthesis of the carbene precursor 1,3-Bis(2,4,6-trimethylphenyl) imidazolinium tetrafluoroborate, compound **4** (refs. 22–24) (Fig. 3).

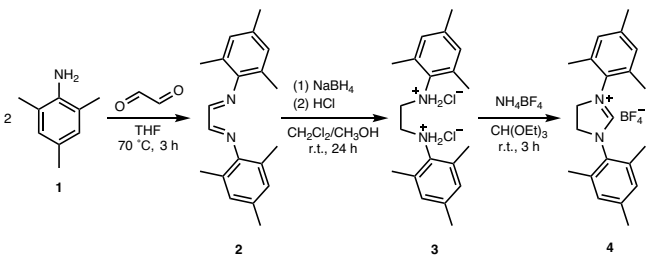

Fig. 3 | Synthesis of H_2IMes•HBF_4 (4). Following standard synthetic operations, a carbene precursor was made via reductive amination of glyoxal with 2,4,6-trimethylaniline, followed by a ring-closing reaction in triethyl orthoformate. r.t., room temperature; THF, tetrahydrofuran.

As the chemistry is robust and well studied, it provided a useful starting point to investigate the efficacy of our χDL-assisted communication procedure.

The original protocol was designed on the Chemputer in Canada where three executable χDL protocols were responsible for the three major synthetic transformations: bis-imination, reduction and ring formation with salt exchange. Initially, these scripts contained manual interventions due to hardware limitations. These included rewiring poly tetrafluoroethylene-tubing connections to rearrange modules or to change the liquid-handling backbone architecture, transferring solid materials and crystallization of the final product in ethanol. The scripts with interventions provided compound **3** in 45% yield, which was then transformed to **4** with 82% yield from **3**. These χDL protocols were then transferred to Scotland for validation on their Chemputer platform and subsequent optimization.

The three scripts were consolidated into a single χDL protocol, and the final crystallization was automated, thus achieving an entirely automated synthetic sequence with minimal human intervention. The only necessary interaction left for the operator was the manual addition

https://doi.org/10.1038/s44160-023-00473-6

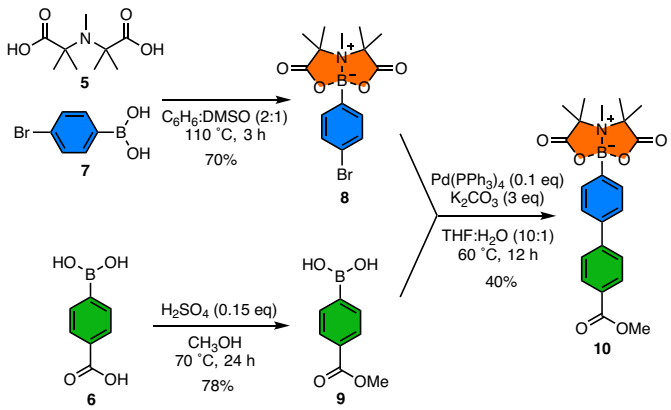

Fig. 4 | Convergent synthesis workflow using the ChemTorrent approach. Two boronic acids (compounds **6** and **7**) were chosen as building blocks to demonstrate the synthesis approach that was illustrated in Fig. 2. Protocols were developed for protecting them orthogonally with the TIDA protection group (for **6**) and by esterification of the carboxylic acid group (for **7**) before encoding the combining Suzuki-type coupling reaction. No chemical knowledge about the building block protection is required for the instance that implements the combining synthesis, which allowed us to develop the two boronic acids independently in different laboratories. DMSO, dimethyl sulfoxide; eq, equivalents.

of NH_4BF_4 to the jacketed filter reactor before starting the final step. This optimized χDL protocol returned **3** in 88% yield and **4** in 47% yield from **3**. The yield differences stem from the different crystallization methods that were used in the two laboratories. The success of the fully automated crystallization in the jacketed filter reactor is often sensitive to factors like the exact gas flow used to purge the reaction mixture to avoid unwanted leaking through the filter frit. The filtration efficiency can easily vary with the grain size of the precipitate and the stability of the vacuum pressure that is applied to initiate the suction of the filtrate. Crystallizations are also among the most difficult operations to automate without substantial time investment. Manual crystallizations often use language like 'dissolved in a minimal amount of solvent' where the chemist can adjust the volumes to ensure the ideal amount is added. In the current version of χDL, fixed volumes must be used, which can result in over-dilution and lower yield. Work is currently underway to develop dynamic applications of χDL protocols to fine tune such reaction parameters.

Repeating a process between laboratories using χDL protocols is much easier than via 'traditional' approaches of communicating synthetic operations in prose or oral communication. Misinterpreting a procedure from another laboratory may result in the same synthetic steps being performed differently. These discrepancies may initially be insignificant, but over a large project can compound into major differences in results between laboratories. Recreating the synthesis using χDL protocols entails confidence that, along each step of the synthesis, all collaborating groups will perform the same synthetic unit operations. This reproduction of synthetic unit operations also makes troubleshooting easier when diagnosing potential reasons a workflow might fail on a new system. Before rerunning a χDL protocol, the peer will still need to ensure the purity of starting materials and solvents is equal to that of the host. The quality of intermediate products must be equally confirmed before proceeding with a multistep synthesis.

TIDA-boronate formation, esterification, then Suzuki reaction

Having established the successful transfer of χDL protocols between laboratories, we then developed a fully digital cross-laboratory optimized synthesis. Following the convergent approach depicted in Fig. 4, we took advantage of the capabilities in the different laboratories that were working on the synthesis project in a decentralized fashion. This particularly included using a feedback-controlled system in Canada to optimize the reaction time for the first building block, while the synthesis for the second building block was independently developed in Scotland.

Then, we utilized the Suzuki–Miyaura coupling, a reliable method for carbon–carbon bond formation in organic synthesis, to combine the building blocks. We exploited TIDA-protected boronate esters, recently reported by the Burke laboratory[25], due to their greater stability toward hydrolysis over the *N*-methyliminodiacetic acid protected analogues. This permitted the presence of water during the coupling reaction, which increased the solubility of the reagents and facilitated the development of an automated process.

The protection of boronic acid **7** with the TIDA protection group **5** (Fig. 4, top left) was initially developed in Canada (59% yield) then sent to Scotland. Optimization of the reaction duration was achieved by utilizing a feedback-controlled system[26]. The digital procedure for this boronic acid esterification was divided into three subsections according to the best-practice guidelines for χDL protocols[14]: preparation, reaction and workup with isolation. The feedback program begins by running the preparation χDL protocol, which loads the appropriate starting materials and solvents into the reactor. As the reaction begins and continues, there is automated sampling every 10 min via an online HPLC that tracks the boronic acid conversion. Meanwhile, the Dean–Stark trap, connected to the refluxing reaction mixture, is automatically emptied at 60 min intervals. These operations are repeated until the feedback control system decides to stop the reaction and isolate the product. The algorithm makes this decision by comparing the change in the amount of remaining boronic acid in the sample against a user-dependent threshold value. For a more detailed explanation of the decision-making program, see the 'TIDA-boronate esterification with feedback control' section of the Supplementary Information. This decision-making program was able to determine sensible time points to stop the reaction for three boronic acids (one of them exemplarily shown in Supplementary Fig. 3).

From the data obtained for the described protection method and the synthesis of boronic acid, we extracted the distillation procedure for 4-bromophenyl boronic acid and created one hardcoded χDL protocol containing setup, distillation of an appropriate duration and workup. This script was sent to Scotland for validation, where the target boronic ester was synthesized on the first attempt (59% yield) with no modification of the reaction parameters. The combination of χDL and graph file provide explicit detail not only of the operation but also of the exact method of its execution. This clarity allowed the Scotland team to spot a suboptimal detail in the drying of the organic phase. The original procedure used an additional round bottom flask containing the drying agent in combination with an inline filter; this drying flask was replaced with an inline drying cartridge, reducing the amount of lost material during the workup. Such a detail would have normally been hidden by the interpretation of a common phrase like 'the organic layer was dried using sodium sulphate' in the experimental protocol. Upon this optimization of the workup, the yield of compound **8** increased to 71%, which was then counter-validated in Canada (72% yield). Comparing the spectral data in Fig. 5 shows that **8** was produced in comparable purity. After validating the functionality of the χDL protocol, it was shared with other robotic operators in a peer-to-peer fashion in both Canada and Scotland to produce a variety of TIDA-protected boronic acid building blocks for another collaborative synthesis project. The exact compounds are detailed in the 'TIDA-boronate esterification' sections of the Supplementary Information.

While the synthesis of the first building block was being developed in Canada, the synthesis of the second building block, (4-(methoxycarbonyl)phenyl)boronic acid, compound **9**, was developed in Scotland (78% yield) and counter-validated in Canada (72% yield). Being a generic esterification of a carboxylic acid, the synthesis of **9** was

https://doi.org/10.1038/s44160-023-00473-6

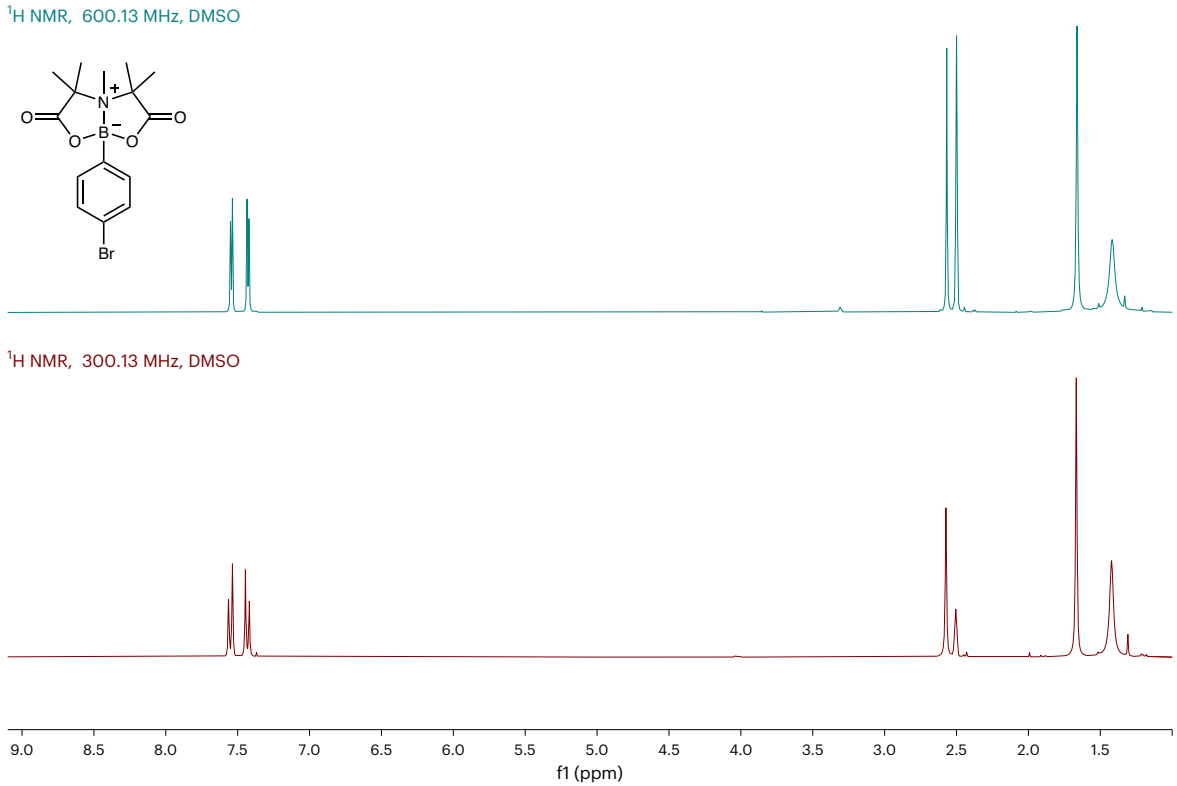

Fig. 5 | Comparison of spectral data for compound 8. Spectral analysis of product from the Cronin laboratory (top) and the Hein laboratory (bottom). This NMR spectroscopy data, collected after running the optimized χDL procedure for the synthesis of compound **8** in the two different laboratories, show excellent agreement.

straightforward and no noteworthy effort was invested in optimizing the reaction conditions.

The Scotland team developed the χDL script for the Suzuki–Miyaura reaction in three sections. The first section consisted of the automated coupling procedure under an inert gas atmosphere; the second handled the workup of the crude mixture using the separator and rotary evaporator modules; the third section performed a catch and release purification protocol as described by the Burke laboratory[25]. All sections were performed using one Chemputer capable of performing each subtask (Fig. 6).

Transferring protocols between different robotic platforms

The execution of chemical synthesis protocols has been focused on Chemputer-type platforms, which are primarily used for prototyping hardware, χDL steps and chemistry[20]. Next, we demonstrate the platform independence of the χDL standard by executing the same code on platforms of different types. Four robots were used in this demonstration—a Kinova multi-axis cobotic robot arm (Fig. 7a), an Opentrons pipetting robot (Fig. 7b) and two Chemputer platforms (Fig. 7c,d) (one for prototyping and experience and the other one for final validation). Each platform has strengths suited for different synthetic tasks.

The Opentrons robot operates on small volumes (down to 1 µl) with a capacity of 120 reaction vials on a two-dimensional grid. The vials can be heated to 99 °C and agitated by shaking the vial racks. Depending on the transfer volume, the χDL software on the Opentrons robot decides which of the attached pipette arms is best suited for completing the desired transfer in the quickest way. With volatile solvents, the software automatically adds prewetting cycles of the pipette tip to ensure that the transfer is accurate. If a library of reagent combinations is to be realized on this robot, the end user need not explicitly hardcode the addition of every reagent individually—if a list of reagents is given, the χDL software can automatically infer all reagent combinations. Moreover, χDL allows the definition of blueprints, which encode a generic reaction procedure and can be called various times within a χDL protocol, each time with different reaction conditions (for example, with a new temperature or substrate concentration). These features allow the concise encoding of chemical procedures for high-throughput screenings. All the tedious, low-level unit operations that are necessary to complete the procedure will be automatically inferred by the χDL interpreter for a given platform. This interpreter is ultimately responsible for translating them into machine-readable code and communicating them to the hardware's Application Programming Interface (API).

The Kinova robot arm also operates on small volumes, but it has the added benefit of unrestricted movement in three-dimensional space. The high versatility of the system, which is primarily run on a Python interface, allows it to adapt to a wide variety of hardware modules, making it easy to integrate the robot arm into the χDL software standard.

The Chemputer, conversely, is fixed in position but operates as a universal synthesizer on batch scale, using larger volumes to transfer reagents and products along a liquid-handling backbone. As outlined in previous sections, the Chemputer is characterized by its high adaptability, which stems from its modular architecture that allows connecting to the liquid-handling backbone any module that can execute χDL commands. It was developed as a robotic equivalent to a human chemist operating on gram scale and performs most of the common bench-chemistry unit operations like liquid reagent addition, filtration, separation, evaporation, heating, cooling and drying. These unit operations can be combined to accomplish typical laboratory tasks, such as recrystallization or extraction of aqueous mixtures with subsequent drying over a drying agent and removal of the extraction solvent under reduced pressure. As the 'gold-standard' for running χDL protocols, the Chemputer platform is ideal to cross-validate any χDL protocol that was developed on a different χDL platform or even independently with

https://doi.org/10.1038/s44160-023-00473-6

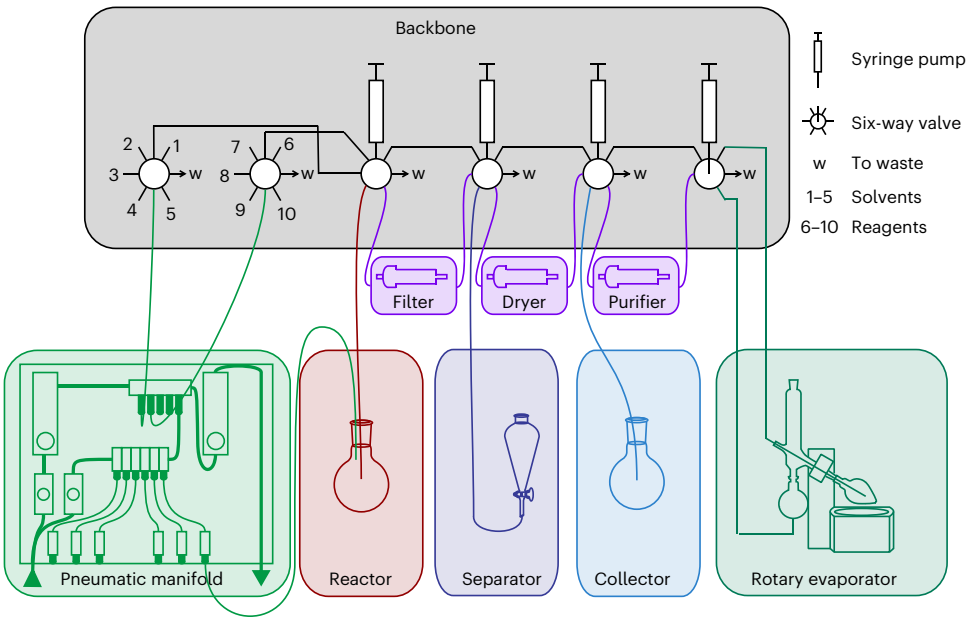

Fig. 6 | Schematic representation of an advanced Chemputer setup. The reactor is equipped with a magnetic stirrer and a reflux condenser. Argon atmosphere is applied to the reactor flask via the top port of the attached reflux condenser using the pneumatic manifold. The manifold also provides passive inert gas supply for the reagent and solvent reservoirs.

stand-alone hardware and protocols. When combined with a robot that can run high-throughput screenings, such as those mentioned above, the Chemputer can facilitate a workflow where successful reaction conditions arising from the small-scale, high-throughput screenings are scaled up seamlessly by just running the same χDL blueprint on the Chemputer platform with adjusted scaling factors. For further specifications and abilities of the Chemputer platform, we refer the reader to related publications[18].

Creating bindings for a robot to execute χDL protocols

The χDL syntax was designed as an abstract and platform-agnostic representation of chemical knowledge, so it can, in theory, be used to interface with any kind of platform that has the means to complete the necessary chemical unit operations. A smooth execution of χDL protocols usually requires software work in advance to implement bindings between the abstract χDL code and the concrete hardware operations available on the platform. As an overarching guideline, the χDL documentation[21] provides a canonical collection of abstract χDL steps listed with the set of mandatory parameters for each step. This collection can then be compared to the library of predefined hardware operations provided in the documentation of the platform's API to determine which χDL steps can be completed by the platform. Next, platform-specific χDL base steps should be defined that include parameters that are inferred by the software upon chempilation of the protocol. For example, a 'transfer' step on the Chemputer platform would involve a backbone movement base step that takes the move speed of the pump plungers as an inferred parameter. Conversely, a Transfer step on the Opentrons platform would need a movement of the pipette arm and the move speed of the gantry would need to be inferred as a parameter. After breaking down the top-level χDL steps into platform-specific base steps, a mapper to the hardware API commands can be implemented that is only responsible for translating the syntax but is not dealing with additional operational logic. If the researcher is sufficiently skilled in coding and is aiming for an implementation that is smoothly interwoven with the surrounding χDL software stack, a good starting point implementating χDL bindings is a clone of the chemputerxdl repository[27] that was implemented by the Cronin Group to interpret χDL protocols on the Chemputer platform[17].

Synthesis using carbonyl diimidazole

To demonstrate the χDL execution on these platforms, we chose to perform a carbonyl diimidazole (CDI)-assisted amide coupling of 2-methoxybenzoic acid with different amines (Fig. 8). First, a coupling protocol for 2-(2-chlorophenyl)ethylamine was developed in Canada on the Kinova. The protocol was repeated on the Chemputer platforms in both Canada and Scotland. It was then applied to a variety of substrates on the Opentrons.

Execution of the amide coupling was originally achieved using the Kinova robotic arm in Canada; however, the procedure was hardcoded in Python[26]. Unfortunately, this approach requires the operator to have a strong understanding of the Python language before making any reaction modifications. We were able to translate the programmed workflow of the platform and convert the defined actions into χDL executable steps. We first determined which steps of the χDL standard[21] were required to execute this synthesis using χDL. This χDL standard is a list of steps detailed by the Cronin Group as 'χDL supported'. When developing a robotic platform that can execute χDL procedures, there are two options. The first is to build a system from the beginning intending that it run χDL protocols. The designer would determine which steps of the standard their platform will be capable of executing and equip it with the hardware to do so. The second option is to take a previously built robotic platform and translate its operations into the framework of the χDL standard. 'Add', 'stir' and 'wait' were the only steps required for this amide coupling procedure. To enable χDL execution on the Kinova platform, we took our previously developed arm and syringe pump movements and imported them into the Kinova-specific chempiler. Once the χDL procedure calls for an addition to be made, the chempiler executes the functions that already existed on the system from when it was originally developed. The system operator no longer requires an understanding of Python and can instead use χDL to design any synthesis. Using the χDL-driven Kinova platform and following the reaction progress via HPLC analysis allowed the synthesis of the target amide **13** with 83% conversion but without final purification. Having demonstrated that a χDL protocol was able to replicate a previously optimized synthetic procedure, we then scaled up the reaction and ran the same χDL on the Chemputer in Canada. Endpoint HPLC analysis shows complete conversion of starting acid to product amide on the

https://doi.org/10.1038/s44160-023-00473-6

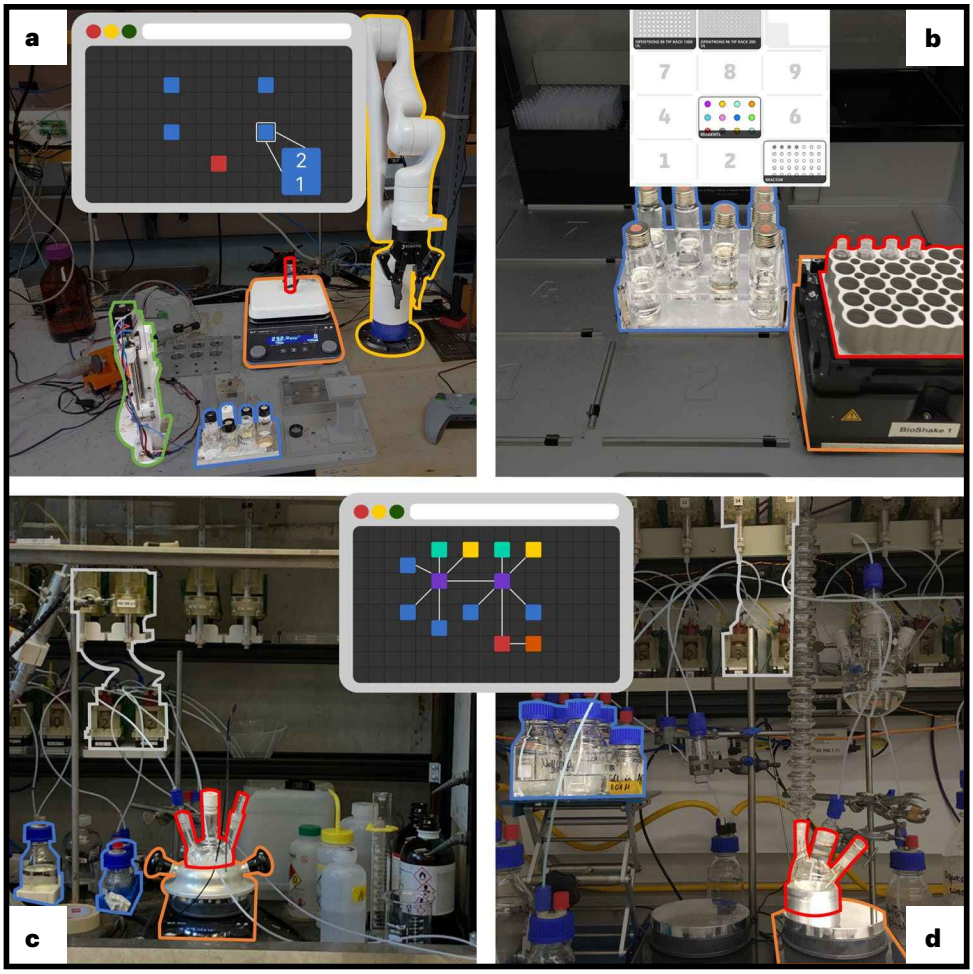

Fig. 7 | Carbonyl diimidazole assisted amide coupling. a, Using the Kinova robotic platform. **b**, Using the Opentrons platform, **c**, Using the Chemputer in Canada. **d**, Using the Chemputer in Scotland. The respective hardware graphs are represented as insets in the photographs. Congruent modules are contoured with equivalent colours: red, reactors; blue, reagent bottles; green, Kinova needle; orange, heating and/or mixing equipment; yellow, Kinova arm; grey, Chemputer backbone.

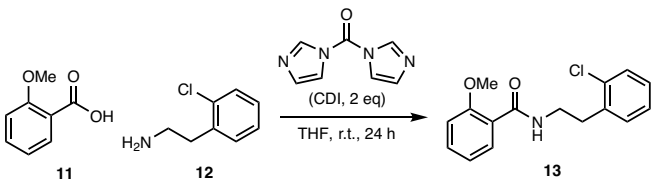

Fig. 8 | CDI-mediated amide coupling of 2-methoxybenzoic acid to 2-(2-chlorophenyl)ethylamine. The shown model reaction was used to develop a χDL-protocol on the Kinova and Chemputer platforms. The protocol was then applied to other amines on the Opentrons platform.

Chemputer, outcompeting the Kinova robot's 83% conversion. For a discussion on the differences between the two platforms in conversion, see the 'CDI coupling' section of the Supplementary Information.

The χDL protocol was then validated on the Chemputer in Scotland, producing amide **13** in 93% yield. Next, the team in Scotland framed the χDL in a 'blueprint' template. This χDL feature leaves the chemical information in the code unchanged but wraps it into a coding construct that can be called numerous times as a unit following a functional programming paradigm. The blueprint feature is usually used to indicate that a χDL procedure has been validated adequately and is considered to work for different substrates or reaction conditions. This framing of the code was crucial to allow for meta-χDL features, like iterating over different values of an input variable.

By this means, the Opentrons robot was programmed to iterate over five different amines, including the already validated compound **12**, applying the χDL blueprint for the CDI-assisted amide coupling to each of the reagent combinations. All the amines were successfully converted into the corresponding amides using a common χDL blueprint (see the 'CDI coupling' section in the Supplementary Information). The χDL blueprint was then used again with upscaled parameters for benzylamine on a Chemputer platform in Scotland to counter-validate that the χDL protocol works equally well on different platforms. The desired amide was obtained in 74% yield on the Chemputer.

Conclusions

By utilizing χDL scripts, we have demonstrated the ability to precisely capture the exact procedure of chemical syntheses and rapidly exchange information between research groups without ambiguity or misinterpretation. While the development of the χDL protocols themselves requires knowledge of both automation and synthesis, their easy implementation in another laboratory highlights the value of the system as a means for non-chemists to synthesize materials typically only available to those with advanced synthetic capabilities. This could

allow teams of researchers access to materials of interest they would otherwise need to outsource the synthesis, expanding their pool of potential research targets.

The concept of Chemputation, manifested in the χDL-driven execution of diverse chemical reactions on various robots, opens the door for faster and more fruitful collaboration projects. Synthetic tasks can be easily distributed between different laboratories with this approach because chemical knowledge can be communicated as ready-to-run source code that does not require a deeper understanding of the underlying chemical principles. A research group can, for example, seamlessly work on the functionalization of a core molecule that a collaborator has synthesized without investing time into manual synthesis or delaying a project with shipping times. We have shown the ease of this workflow for seven chemical reactions, including multistep synthesis, convergent synthetic approaches and divergent explorations of reactivity, demonstrating how yields and purities of χDL-coded syntheses are excellently repeated in different iterations, geographical locations and hardware setups.

Methods

Materials

Solvents and reagents were used as received from commercial suppliers unless otherwise stated. TIDA was provided by the Burke group at the University of Illinois at Urbana-Champaign.

Robotic hardware

χDL procedures were executed in Scotland on a Chemputer-type platform and an Opentrons pipetting robot. In Canada, χDL procedures were also executed on a Chemputer-type platform in addition to a multi-axis cobotic platform using a Kinova 6-axis robotic arm.

NMR spectroscopy

Measurements were performed with a Bruker Avance III HD 600 and Bruker AV-300 MHz spectrometers in Scotland and Canada, respectively. Spectra were collected at 298 K; chemical shifts are reported in ppm and were calibrated for the (residual) NMR solvent signal (2.50 ppm for DMSO-d_6 and 7.26 ppm for $CDCl_3$).

HPLC-UV/Vis (mass spectrometry)

Analysis in Scotland was performed using a Thermo-Dionex-Ultimate 3000 HPLC connected to a Bruker MaXis Impact quadrupole time-of-flight mass spectrometer with an electrospray source, operating exclusively in positive mode. Analysis in Canada was performed using an Agilent 1260/1290 infinity HPLC connected to an Agilent single quadrupole mass spectrometer. For more detailed information regarding the methods of analysis on these systems, see Supplementary Tables 1 and 2.

χDL execution

All χDL files utilized in this study are presented within the Supplementary Information. The chemical information in the χDL files is repeated as human-readable output within the 'Synthesis protocols' section for clarity.

Data availability

The experiment data that support the findings of this study are available in the manuscript files and from the corresponding author upon reasonable request. The source data underlying Supplementary Figs. 2 and 3 are provided in Supplementary Data 1.

Code availability

χDL files (.xdl) and Chemputer graph files (.json) can be opened and edited with the ChemIDE app on https://croningroup.gitlab.io/chemputer/xdlapp/. The χDL software standard is linked here: https://croningroup.gitlab.io/chemputer/xdl/standard/index.html. A complete docker image of the synthetic platforms and hardware can be made available by request.

References

1. Popper, K. R. *Logik der Forschung: Zur Erkenntnistheorie der modernen Naturwissenschaft* (Springer, 1934).
2. Gushee, D. E. Factors affecting dissemination of chemical information. *J. Chem. Doc.* **11**, 201–204 (1971).
3. Arveson, M. H. Economic aspects in the dissemination of chemical knowledge. *J. Chem. Doc.* **1**, 1–3 (1961).
4. Boga, S. B. et al. Selective functionalization of complex heterocycles via an automated strong base screening platform. *React. Chem. Eng.* **2**, 446–450 (2017).
5. Burger, B. et al. A mobile robotic chemist. *Nature* **583**, 237–241 (2020).
6. Chatterjee, S., Guidi, M., Seeberger, P. H. & Gilmore, K. Automated radial synthesis of organic molecules. *Nature* **579**, 379–384 (2020).
7. Godfrey, A. G., Masquelin, T. & Hemmerle, H. A remote-controlled adaptive medchem lab: an innovative approach to enable drug discovery in the 21st Century. *Drug Discov. Today* **18**, 795–802 (2013).
8. Legrand, M. & Bolla, P. A fully automatic apparatus for chemical reactions on the laboratory scale. *J. Autom. Chem.* **7**, 513591 (1985).
9. MacLeod, B. P. et al. Self-driving laboratory for accelerated discovery of thin-film materials. *Sci. Adv.* **6**, eaaz8867 (2020).
10. Okamoto, H. & Deuchi, K. Design of a robotic workstation for automated organic synthesis. *Lab. Robotics Automat.* **12**, 2–11 (2000).
11. Orita, A., Yasui, Y. & Otera, J. Automated synthesis: development of a new apparatus friendly to synthetic chemists (MEDLEY). *Org. Process Res. Dev.* **4**, 333–336 (2000).
12. Tanaka, Y., Fuse, S., Tanaka, H., Doi, T. & Takahashi, T. An efficient synthesis of a cyclic ether key intermediate for 9-membered masked enediyne using an automated synthesizer. *Org. Process Res. Dev.* **13**, 1111–1121 (2009).
13. Baker, M. 1,500 scientists lift the lid on reproducibility. *Nature* **533**, 452–454 (2016).
14. Steiner, S. et al. Organic synthesis in a modular robotic system driven by a chemical programming language. *Science* **363**, eaav2211 (2019).
15. Hammer, A. J. S., Leonov, A. I., Bell, N. L. & Cronin, L. Chemputation and the standardization of chemical informatics. *JACS Au* **1**, 1572–1587 (2021).
16. 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**, 101–108 (2020).
17. Rohrbach, S. et al. Digitization and validation of a chemical synthesis literature database in the ChemPU. *Science* **377**, 172–180 (2022).
18. Gromski, P. S., Granda, J. M. & Cronin, L. Universal chemical synthesis and discovery with 'The Chemputer. *Trends Chem.* **2**, 4–12 (2020).
19. Wilbraham, L., Mehr, S. H. M. & Cronin, L. Digitizing chemistry using the chemical processing unit: from synthesis to discovery. *Accounts Chem. Res.* **54**, 253–262 (2021).
20. Angelone, D. et al. Convergence of multiple synthetic paradigms in a universally programmable chemical synthesis machine. *Nat. Chem.* **13**, 63–69 (2021).
21. XDL documentation. *Cronin Group, UoG* https://croningroup.gitlab.io/chemputer/xdl/standard/index.html (2022).
22. Arduengo, A. J. et al. Imidazolylidenes, imidazolinylidenes and imidazolidines. *Tetrahedron* **55**, 14523–14534 (1999).
23. Buckley, B. R. & Neary, S. P. Thiadiazolidine 1-oxide systems for phosphine-free palladium-mediated catalysis. *Tetrahedron* **66**, 7988–7994 (2010).
24. Higgins, E. M. et al. pK_as of the conjugate acids of *N*-heterocyclic carbenes in water. *Chem. Commun.* **47**, 1559–1561 (2011).

25. Blair, D. J. et al. Automated iterative Csp3–C bond formation. *Nature* **604**, 92–97 (2022).
26. Liu, J., Sato, Y., Yang, F., Kukor, A. J. & Hein, J. E. An adaptive auto-synthesizer using online PAT feedback to flexibly perform a multistep reaction. *Chem. Methods* **2**, e202200009 (2022).
27. Cronin Group, UofG. χDL code. *Zenodo* https://zenodo.org/record/6534009 (2022).

Acknowledgements

Financial support for this work was provided by the EPSRC (grants EP/L023652/1, EP/R01308X/1, EP/S019472/1 and EP/P00153X/1), Defence Advanced Research Projects Agency (DARPA) under the Accelerated Molecular Discovery Program (Cooperative Agreement HR00111920027, dated 1 August 2019). Additional support was provided by the Canada Foundation for Innovation (CFI) (CFI-35883) and the Natural Sciences and Engineering Research Council of Canada (NSERC) (RGPIN-2021-03168, Discovery Accelerator Supplement RGPAS-2021-00016). Student support was provided by the German Academic Scholarship Foundation (to R.R.) and an NSERC CGS-D scholarship (to M.G.). We thank D. Thomas, M. Siauciulis and E. Trushina from the Cronin Laboratory at the University of Glasgow and S. Rohrbach from Chemify Ltd for proofreading the manuscript and for support in the laboratory.

Author contributions

L.C. and J.H. conceived the idea of utilizing χDL protocols between different laboratories and named the concept of this paper 'ChemTorrent'. L.C. and J.H. together coordinated the research project and mentored R.R. and M.G. All experimental work was completed by R.R. and M.G. in equal contribution with help from the respective research groups in Glasgow and Vancouver. The body of this manuscript and the Supplementary Information were written by R.R. and M.G. with input from all the authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44160-023-00473-6.

Correspondence and requests for materials should be addressed to Jason E. Hein or Leroy Cronin.

Peer review information *Nature Synthesis* thanks Linjiang Chen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Peter Seavill, in collaboration with the *Nature Synthesis* team.