Published in final edited form as:

Science. 2015 March 13; 347(6227): 1221–1226. doi:10.1126/science.aaa5414.

Synthesis of many different types of organic small molecules using one automated process

Junqi Li[1,†], **Steven G. Ballmer**[1,†], **Eric P. Gillis**[1], **Seiko Fujii**[1], **Michael J. Schmidt**[1], **Andrea M. E. Palazzolo**[1], **Jonathan W. Lehmann**[1], **Greg F. Morehouse**[1], and **Martin D. Burke**[1,*]

[1]Howard Hughes Medical Institute, Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Abstract

Small molecule synthesis usually relies on procedures highly customized for each target. A broadly applicable automated process could greatly increase the accessibility of this class of compounds to enable investigations of their practical potential. Here we report the synthesis of 14 distinct classes of small molecules using the same fully automated process. This was achieved by strategically expanding the scope of a building block-based synthesis platform to include even Csp^3-rich polycyclic natural product frameworks and discovering a catch-and-release chromatographic purification protocol applicable to all of the corresponding intermediates. With thousands of compatible building blocks already commercially available, many small molecules are now accessible with this platform. More broadly, these findings illuminate an actionable roadmap to a more general and automated approach for small molecule synthesis.

Small molecules perform many important functions in nature, medicine, and technology. However, efforts to discover and optimize new small molecule function are often impeded by limitations in synthetic access to this class of compounds. For peptides (1) and oligonucleotides (2), the development of automated synthesis platforms removed this bottleneck. The resulting expanded access to these molecules permitted widespread exploration and applications of their functional potential. Substantial progress has also been made towards automating the synthesis of oligosaccharides (3). In each of these cases, automation was enabled by the development of a general building block-based synthesis strategy and a common purification process for the corresponding intermediates. Such standardization reduced the number of processes employed and thus decreased the number of challenges involved in automating the synthesis platform. In contrast, despite tremendous progress in the field, small molecule syntheses typically employ strategies and purification methods that are highly customized for each target, thus requiring automation solutions to be developed on an *ad hoc* basis (4–7). To enable the more generalized automation of small

*To whom correspondence should be addressed. mdburke@illinois.edu.
†These authors contributed equally to this work.

Supplementary Materials:
Materials and Methods
Figures S1–S14, Tables S1–S3
References (34–52)

molecule synthesis, we asked whether many different types of small molecules could be prepared using a common building block-based strategy and a common purification process.

Small molecules can be very diverse in structure, as illustrated by representative compounds **1-14** in Fig. 1A. Synthesis of this entire set of targets using a common approach thus represents a major challenge. However, like peptides, oligonucleotides and oligosaccharides, most natural products (such as **1-4**) are biosynthesized via the iterative assembly of a small set of building blocks, such as malonyl coenzyme A, isopentenyl pyrophosphate, and pyruvic acid (8). Many materials and pharmaceuticals (such as **5-9**) comprise collections of aryl and/or heteroaryl components (9). Even topologically complex natural products containing macrocyclic or polycyclic frameworks (such as **10**-**14**) are usually biosynthesized via iterative building block-based assembly of linear precursors, which are then (poly)cyclized to yield more complex molecular architectures (10–12). This analysis suggests that many small molecules might be accessible via a common, biosynthesis-inspired strategy involving the iterative assembly of building blocks. Supporting this notion, we recently demonstrated that more than 75% of all polyene natural product motifs can be prepared using just 12 building blocks and one coupling reaction (13).

That study employed a synthesis platform analogous to iterative peptide coupling that sequentially assembles bifunctional *N*-methyliminodiacetic acid (MIDA) boronates (Fig. 1B) (14). Iterative cycles of coupling and deprotection are enabled by the MIDA ligand, which attenuates the reactivity of boronic acids and thus prevents undesired oligomerization. This approach is both efficient and flexible because all the required functional groups, oxidation states, and stereochemical elements are pre-installed into the building blocks. These features are then faithfully translated into the products using the same mild and stereospecific coupling chemistry (15). Hundreds of MIDA boronates and thousands of additional halide and boronic acid building blocks are now commercially available, and natural products from most major biosynthetic pathways, including **1**-**4**, have been manually synthesized in prior studies using this strategy (16–19).

With this promising starting point, we set out to expand the scope of this platform to include all of the structures shown in Fig. 1A. Csp^3-rich cyclic and polycyclic natural product frameworks such as **10-14** represent especially challenging targets, and thus required a strategic advance. Because many of these molecules are biosynthesized via cyclization of modular linear precursors derived from iterative building block assembly (10–12), we hypothesized that an analogous linear-to-cyclized strategy might enable this platform to access many such structures. In this approach, the same building block assembly process is used to generate a linear precursor, which is then (poly)cyclized into the topologically complex product. The stereochemical information in the building blocks is first translated into linear precursors via stereospecific couplings and then into the targeted products via stereospecific and/or stereoselective (poly)cyclization reactions. To enable such cyclizations, such linear precursors must be suitably flexible and therefore rich in sp^3 hybridized carbons. Building block-based assembly of these precursors thus requires many Csp^3 couplings, which can be challenging.

To test this linear-to-cyclized strategy, we targeted the manual synthesis of the highly complex pentacyclic secodaphnane alkaloid core **14**. This core can in theory be derived from a much simpler linear precursor **15** via a bioinspired cyclization cascade involving amine condensation and intramolecular Diels-Alder and Prins cyclizations (20). In turn, **15** was targeted via the iterative assembly of building blocks **16**-**18** in which all of the required double bond stereochemistry is pre-installed. Csp^3 coupling of **16** and **17** generates MIDA boronate intermediate **19**. Deprotection of **19** yields a free boronic acid which is engaged in another Csp^3 coupling with **18** to complete the iterative assembly of linear precursor **15**. The aforementioned diastereoselective cyclization cascade is then used to transform **15** into the pentacyclic alkaloid **(±)-14**.

Having established a strategy to access even topologically complex small molecules using the same building block-based approach, we next questioned whether a common purification protocol could also be developed and thereby enable this platform to be automated. Small molecule synthesis typically involves purifications customized for each intermediate, such as chromatography with eluents optimized for each compound. Such customization is incompatible with generalized automated purification. Solid-phase synthesis can address this problem for peptides (1), oligonucleotides (2), oligosacharides (3), and some organic polymers (21). In some cases, syntheses of natural products and pharmaceuticals have also been aided by solid-phase methods (22). This approach is well-established, compatible with a wide range of chemistries, and has been employed in industry (23). However, small molecules do not possess a common functional group handle for attachment to solid support which precludes generalized application of this approach. We thus needed a different solution.

We recognized that each iteration of building block assembly in our platform generates a MIDA boronate as the key intermediate (Fig. 1B and C). This led us to question whether the MIDA boronate motif could serve as a surrogate common handle for purification. In this vein, we discovered that MIDA boronates uniformly possess highly unusual binary affinity for silica gel with certain pairs of eluents (Fig. 2A). Specifically, all MIDA boronates **20**-**39**, with appended fragments representing a wide range of sizes, polarities, and functional group content, show minimal mobility on thin layer silica gel chromatography when eluting with MeOH:Et_2O (Fig. 2A, left). However, all of the same MIDA boronates are rapidly eluted with THF (Fig. 2A, right). This phenomenon enabled us to develop a new type of catch-and-release purification protocol applicable to any intermediate that contains a MIDA boronate. A crude reaction mixture is passed over silica gel and the MIDA boronate is temporarily caught while excess reagents and byproducts are removed via washing with MeOH:Et_2O. The MIDA boronate is then cleanly released by switching the eluent to THF. The MIDA boronate motif can thus serve as both a latent reactive functional group for iterative coupling and a traceless common functional group handle for generalized purification.

With both a common building block-based synthesis platform and a common purification method, we designed and built a synthesizer that iteratively assembles MIDA boronate building blocks in a fully automated fashion (Fig. 2B) (24). This device comprises three modules that sequentially execute the deprotection, coupling, and purification steps required for each cycle. All solutions are automatically transferred via computer-controlled syringe

pumps running custom designed software. Thus, each automated synthesis simply requires placing pre-packed cartridges onto the synthesizer and pressing “start”.

The fully automated synthesis commences at the deprotection module, where THF and water are syringed into a cartridge containing the MIDA boronate and NaOH. After deprotection, the reaction is quenched and the resulting THF solution of the freshly prepared boronic acid is separated from the water-soluble MIDA ligand. The coupling module then heats and stirs a solution of the next building block and the coupling reagents. The synthesizer then adds the freshly prepared solution of boronic acid to the coupling reaction. At the end of the reaction, the synthesizer filters and transfers the crude reaction mixture to the purification module, which executes the catch-and-release purification protocol with MeOH:Et_2O followed by THF. The THF solution of the purified product is then transferred directly into the deprotection module to start the next iteration of the synthesis.

To first test the capacity of this synthesizer to execute one cycle of deprotection, coupling, and purification, we subjected a series of commercially available aryl, heteroaryl, vinyl, and alkyl MIDA boronates to automated deprotection and coupling with a model bifunctional building block, 4-bromophenyl MIDA boronate (24) (Table S2). Using a standard set of hydrolysis conditions (NaOH, THF:H_2O, 23 °C, 20 min), and coupling conditions (PdXPhos, K_3PO_4) (25), we obtained the desired cross-coupling products in good yields and purities in all cases (Table S2, entries 1–3). The synthesizer was also capable of executing a Csp^3 coupling using $Pd[P(o\text{-}tol_3)]_2$ and Ag_2O/K_2CO_3 (Table S2, entry 4).

Accessing many pharmaceuticals and materials represented by structures **5**-**9** requires the flexibility to link building blocks via carbon-heteroatom and/or carbon-carbon bonds. The stability of MIDA boronates towards many reaction conditions (26, 27) and the synthetic versatility of boronic acids (28) allowed us to add carbon-heteroatom bond formations to the same platform. The synthesizer successfully executed a series of automated carbon-heteroatom bond formation, including a Buchwald-Hartwig amination, O-alkylation, and amide bond formations (Table S3). Despite the different reagents and byproducts, the same catch-and-release process purified all of the corresponding MIDA boronate products.

Having confirmed the capacity to reliably execute single cycles of deprotection, coupling, and purification, we next targeted the automated synthesis of a wide range of linear small molecules (**1**-**9**) via multiple carbon-carbon and/or carbon-heteroatom bond formations (Fig. 2C). These include natural products from major biosynthetic pathways (**1**-**4**), materials components (**5, 6**), and pharmaceuticals/biological probes (**7**-**9**). Most of the corresponding building blocks are commercially available. Similar to automated peptide, oligonucleotide, and oligosaccharide syntheses, all of the synthesizer-generated final products were purified using standard chromatographic techniques, and any protecting groups other than MIDA were easily removed in a separate step (24). In each case, a single automated run successfully delivered the targeted small molecule in multimilligram quantities, fulfilling the requirements of most functional discovery assays.

The development of small molecules with optimized functions often requires efficient access to many structural derivatives of a parent compound. To test if this platform could enable

such access, we targeted the automated preparation of many derivatives of the complex neolignan natural product ratanhine **4**. In this experiment, we did not optimize of any of the deprotection, coupling, or purification conditions used to construct **41** (Fig. 2C). We input four sets of building blocks representing common substructural elements found throughout the neolignan family and/or other pharmaceutically relevant motifs (Fig. 3). These building blocks included variations in oxidation states, methylation patterns, fluorine content, aromatic ring identity, and size. They also represent pre-programmed oligomer lengths of 3 to 4 units based on whether the third building block was a bifunctional halo-MIDA boronate or a capping halide. In the event, the synthesizer successfully generated 20 out of 20 of the targeted derivatives, collectively representing all possible combinations of this four-component matrix of building blocks (Fig. 3).

Finally, we tested whether a wide range of macro- and polycyclic natural products and natural product-like cores (**10**-**14**, Fig. 1A) could be generated using the same automated building block assembly process and the linear-to-cyclized strategy. The macrocyclic natural product citreofuran posseses both Csp^3 and atropisomerism stereochemical elements (Fig. 4, entry 1). This complex target can be derived from linear precursor **76** (29), which can in theory be assembled from building blocks **72-74**. All of the required stereochemical information for citreofuran is pre-encoded in the chiral non-racemic MIDA boronate building block **72**. On the synthesizer, fully automated deprotection of **72**, Csp^3 coupling with **73**, and purification yielded intermediate **75**. A second round of deprotection and coupling of the resulting 2-furanyl boronic acid with **74** produced linear precursor **76**. This linear precursor was then deprotected and atropdiastereoselectively macrocyclized to generate citreofuran.

Oblongolide is a norsesquiterpene γ-lactone natural product containing a 6,6,5-tricyclic core with five Csp^3 stereogenic centers, one of which is quaternary (Fig. 4, entry 2). The three building blocks **44**, **77**, and **78** were automatically assembled via iterative Csp^2 and Csp^3 couplings to produce linear precursor **80**. After deprotection, the linear precursor was subjected to a cascade intramolecular substrate-controlled diastereoselective Diels-Alder reaction and lactonization process (30) which defined the four contiguous stereogenic centers in oblongolide.

In cases where no Csp^3 stereogenic centers are present in the linear precursors, the enantioselectivity of cyclizations can be controlled using a rapidly expanding toolbox of chiral catalysts (31). This approach allows the stereoselective construction of the natural product-like hexahydroindene and steroid-like core structures **12** and **13** using the same linear-to-cyclized strategy (Fig. 4, entries 3 and 4). Specifically, building blocks **81-83**, all possessing olefins with pre-defined geometries required for cyclization, were assembled on the synthesizer to produce linear precursor **85**. This precursor was then subjected to deprotection and a chiral imidazolidinone-promoted organocatalytic enantio- and diastereoselective Diels-Alder reaction to generate **12** (Fig. 4, entry 3) (32). Similarly, iterative Csp^3 coupling of building blocks **16, 17**, and **86** generated linear precursor **87**, which then underwent catalyst-controlled enantio- and diasereoselective cation-π cyclization (33) followed by reduction to generate **13** (Fig. 4, entry 4). Finally, by simply replacing

building block **86** with **18** and using the same automated platform, even the highly complex pentacyclic secodaphnane core **(±)-14** was readily prepared (Fig. 4, entry 5).

Thus, many different types of small molecules (**1-14** in Fig. 1A) can be synthesized using one automated building block assembly platform. This advance was enabled by standardizing the synthesis and purification processes used to assemble these structures. Importantly, a majority of the building blocks employed herein are already commercially available.

Further expanding the scope of this automated synthesis platform represents an actionable roadmap toward a general and broadly accessible solution to the small molecule synthesis problem. This roadmap includes creating building blocks representing the highly redundant substructural elements found in many small molecules (13), developing better methods for iteratively coupling those building blocks together, and advancing the capacity for biosynthesis-inspired cyclizations of linear precursors to yield complex natural product frameworks. Achieving these objectives stands to better enable the scientific community to bring the substantial power of small molecule synthesis to bear upon many important unsolved problems in society.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We gratefully acknowledge the NIH (GM080436 and GM090153), NSF (0747778), HHMI, and Bristol-Myers Squibb for funding. M.D.B. is an HHMI Early Career Scientist, J.L. was an HHMI International Student Research Fellow, and J.L. and E.P.G. were Bristol-Myers Squibb Graduate Fellows. We thank Dr. Danielle Gray for performing the X-ray analysis on **N-bromoacyl-14**. Metrical parameters for the structure of **N-bromoacyl-14** are available free of charge from the Cambridge Cyrstallographic Data Centre under reference CCDC-1045844. The University of Illinois has filed patents on MIDA boronate chemistry and the automated synthesis platform reported herein. These patents have been licensed to REVOLUTION Medicines, a company founded by M.D.B.

References and Notes

1. Merrifield RB. Science. 1965; 150:178–185. [PubMed: 5319951]
2. Caruthers MH. Science. 1985; 230:281–285. [PubMed: 3863253]
3. Plante OJ, Palmacci ER, Seeberger PH. Science. 2001; 291:1523–1527. [PubMed: 11222853]
4. Ley SV, Fitzpatrick DE, Ingham RJ, Myers RM. Angew Chem Int Ed. 2015; 5410.1002/anie.201410744
5. Fuse, S.; Machida, K.; Takahashi, T. New Strategies in Chemical Synthesis and Catalysis. Pignataro, B., editor. Vol. Chapter 2. Wiley-VCH; Weinheim, Germany: 2012.
6. Godfrey AG, Masquelin T, Hemmerle H. Drug Discov Today. 2013; 18:795–802. [PubMed: 23523957]
7. Newton S, et al. Angew Chem Int Ed. 2014; 53:4915–4920.
8. Dewick, PM. Medicinal Natural Products: A Biosynthetic Approach. Wiley; 2009.
9. Vitaku E, Smith DT, Njardarson JT. J Med Chem. 2014; 57:10257–10274. [PubMed: 25255204]
10. Yoder RA, Johnston JN. Chem Rev. 2005; 105:4730–4756. [PubMed: 16351060]
11. Stocking EM, Williams RM. Angew Chem Int Ed. 2003; 42:3078–3115.
12. Kopp F, Marahiel MA. Nat Prod Rep. 2007; 24:735–749. [PubMed: 17653357]
13. Woerly EM, Roy J, Burke MD. Nature Chem. 2014; 6:484–491. [PubMed: 24848233]

14. Gillis EP, Burke MD. J Am Chem Soc. 2007; 129:6716–6717. [PubMed: 17488084]
15. Miyaura N, Suzuki A. Chem Rev. 1995; 95:2457–2483.
16. Gillis EP, Burke MD. Aldrichimica Acta. 2009; 42:17–27. [PubMed: 22523433]
17. Woerly EM, Cherney AH, Davis EK, Burke MD. J Am Chem Soc. 2010; 132:6941–6943. [PubMed: 20441218]
18. Gray KC, Palacios DS, et al. Proc Natl Acad Sci. 2012; 109:2234–2239. [PubMed: 22308411]
19. Fujita K, Matsui R, Suzuki T, Kobayashi S. Angew Chem Int Ed. 2012; 51:7271–7274.
20. Heathcock CH, Piettre S, Ruggeri RB, Ragan JA, Kath JC. J Org Chem. 1992; 57:2554–2566.
21. Elliot EL, Ray CR, Kraft S, Atkins JR, Moore JS. J Org Chem. 2006; 71:5282–5290. [PubMed: 16808517]
22. Mentel M, Breinbauer R. Top Curr Chem. 2007; 278:209–241.
23. Maechling S, Good J, Lindell SD. J Comb Chem. 2010; 12:818–821. [PubMed: 20879771]
24. Materials and Methods are available as supplementary materials on *Science* online.
25. Kinzel T, Zhang Y, Buchwald SL. J Am Chem Soc. 2010; 132:14073–14075. [PubMed: 20858009]
26. Gillis EP, Burke MD. J Am Chem Soc. 2008; 130:14084–14085. [PubMed: 18837550]
27. Grob JE, et al. Org Lett. 2012; 14:5578–5581. [PubMed: 23092156]
28. Hall, DG., editor. Boronic Acids: Preparation and Applications in Organic Synthesis, Medicine and Materials. Wiley-VCH; Weinheim, Germany: 2011.
29. Bracher F, Schulte B. Nat Prod Res. 2003; 17:293–299. [PubMed: 12822910]
30. Shing TKM, Yang J. J Org Chem. 1995; 60:5785–5789.
31. Corey, EJ.; Kürti, L. Enantioselective Chemical Synthesis: Methods, Logic, and Practice. Direct Book Publishing; Texas: 2010. p. 121-151.
32. Wilson RM, Jen WS, MacMillan DWC. J Am Chem Soc. 2005; 127:11616–11617. [PubMed: 16104734]
33. Snyder SA, Treitler DS, Schall A. Tetrahedron. 2010; 66:4796–4804.
34. Pangborn AB, Giardello MA, Grubbs RH, Rosen RK, Timmers FJ. Organometallics. 1996; 15:1518–1520.
35. Still WC, Kahn M, Mitra A. J Org Chem. 1978; 43:2923–2925.
36. Surendra K, Corey EJ. J Am Chem Soc. 2012; 134:11992–11994. [PubMed: 22780430]
37. Jung ME, Duclos BA. Tetrahedron. 2006; 62:9321–9334.
38. Fujii S, Chang SY, Burke MD. Angew Chem Int Ed. 2011; 50:7862–7864.
39. Vaz B, Domínguez M, Alvarez R, de Lera AR. Chem Eur J. 2007; 13:1273–1290. [PubMed: 17066395]
40. Hayes ME, Shinokubo H, Danheiser RL. Org Lett. 2005; 7:3917–3920. [PubMed: 16119931]
41. Paterson I, Anderson EA, Dalby SM, Lim JH, Genovino J, Maltas P, Moessner C. Angew Chem Int Ed. 2008; 47:3016–3020.
42. Hu M, Li J, Yao SQ. Org Lett. 2008; 10:5529–5531. [PubMed: 19053720]
43. Gerstenberger BS, Konopelski JP. J Org Chem. 2005; 70:1467–1470. [PubMed: 15704987]
44. Burke, MD.; Dick, GR.; Knapp, DM.; Gillis, EP.; Klubnick, JA. PCT/US2011/0201806. Aug 18. 2011
45. Wenderski TA, Marsini MA, Pettus TRR. Org Lett. 2011; 13:118–121. [PubMed: 21138313]
46. Sakaguchi H, Tokuyama H, Fukuyama T. Org Lett. 2007; 9:1635–1638. [PubMed: 17394345]
47. Dipl-Ing TR, Breit B. Chem Eur J. 2009; 15:6345–6348. [PubMed: 19479928]
48. Huang Z, Negishi E-I. Org Lett. 2006; 8:3675–3678. [PubMed: 16898789]
49. Ijuin R, Takashima T, Watanabe Y, Sugiyama Y, Suzuki M. Bioorg Med Chem. 2012; 20:3703–3709. [PubMed: 22607882]
50. Araoz R, et al. J Am Chem Soc. 2011; 133:10499–10511. [PubMed: 21644584]
51. Lifchits O, Reisinger CM, List B. J Am Chem Soc. 2010; 132:10227–10229. [PubMed: 20662497]
52. Otero L, Vaz B, Álvarez R, de Lera AR. Chem Commun. 2013; 49:5043–5045.

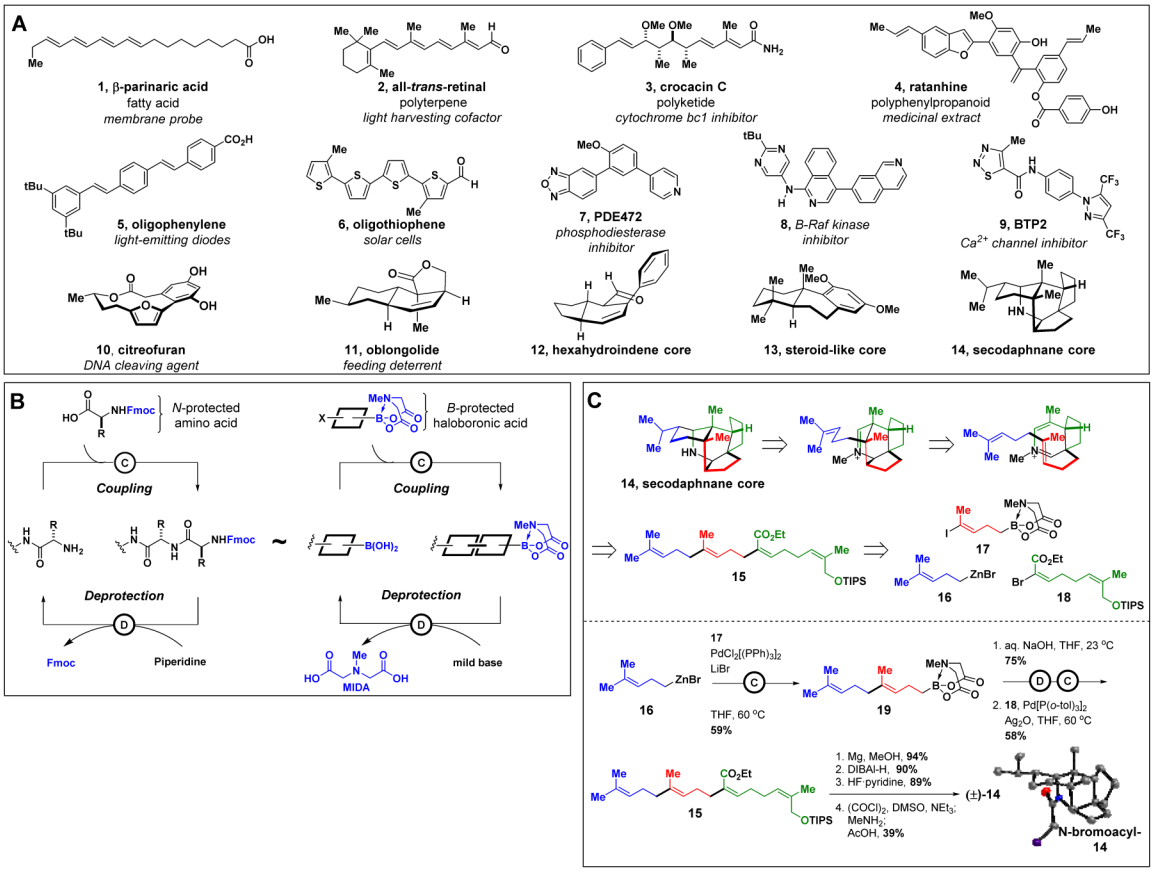

Fig. 1. Common building block-based approach for making many different types of small molecules

(A) A collection of compounds representing the structural and functional diversity of small molecules. **(B)** Analogous building block-based strategies for the synthesis of peptides and small molecules. **(C)** Synthesis of the Csp^3-rich pentacyclic secodaphnane core **(±)-14** via iterative Csp^3 couplings to yield a linear precursor followed by biosynthesis-inspired cascade polycyclization. An X-ray structure of the N-bromoacyl derivative of **14** was obtained to unambiguously confirm the structure of **14**. TIPS = triisopropylsilyl.

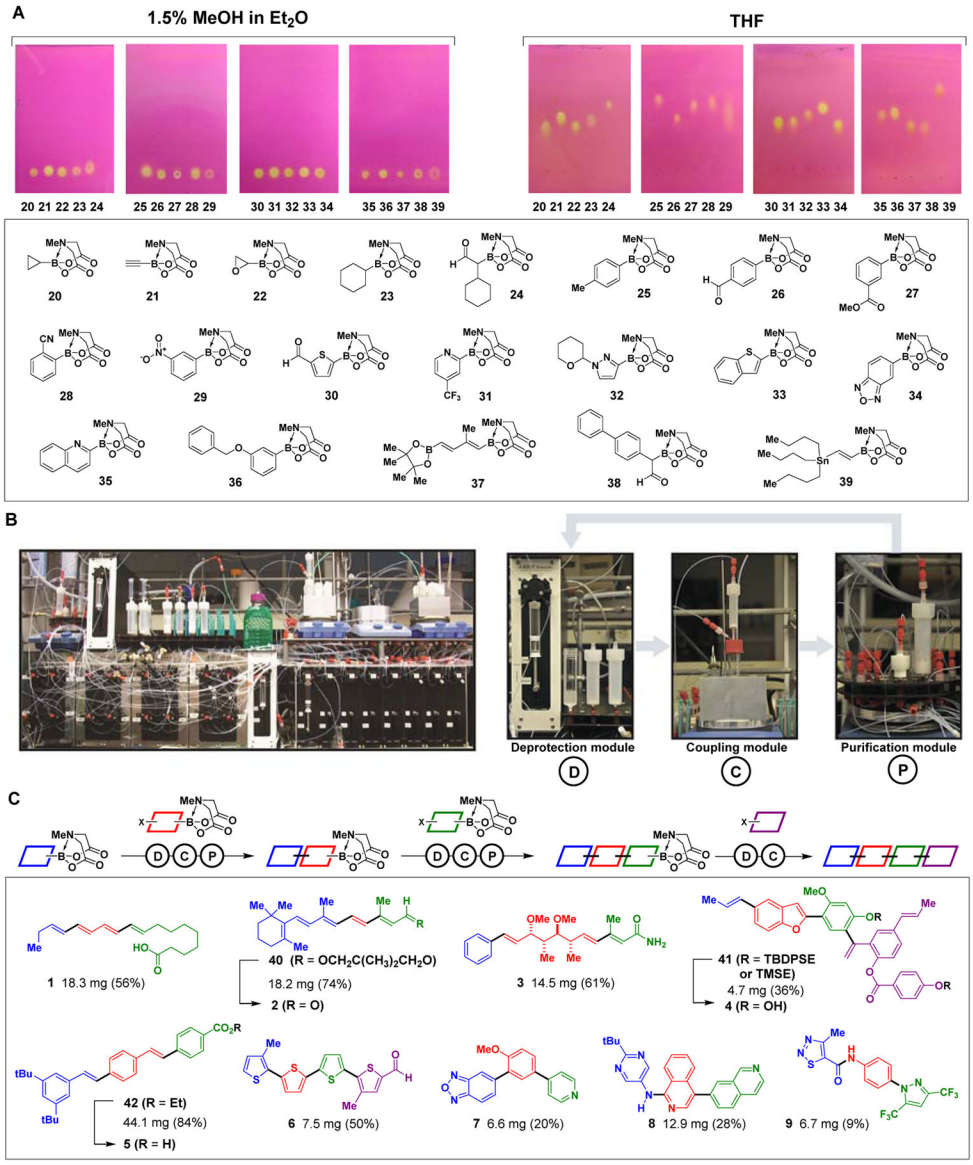

Fig. 2. General purification process to enable automation

(A) MIDA boronates uniformly show binary elution properties on silica gel thin-layer chromatography. **(B)** Photograph of the small molecule synthesizer which comprises three modules that execute the deprotection Ⓓ, coupling Ⓒ, and purification Ⓟ steps. **(C)** Automated synthesis of natural products, materials, pharmaceuticals, and biological probes via iterative coupling of building blocks indicated by different colors (24).

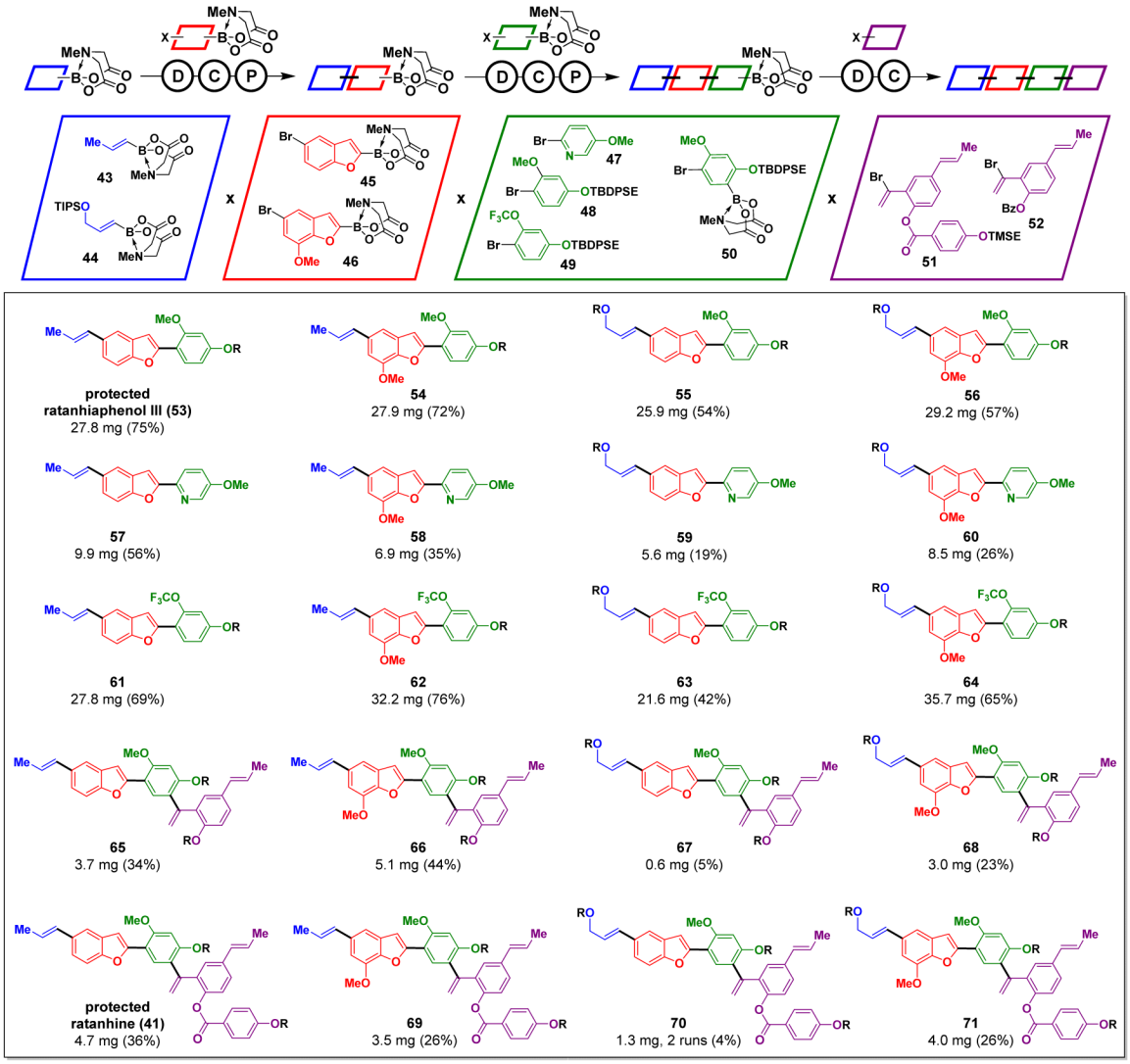

Fig. 3. Automated synthesis of ratanhine derivatives

Conditions: deprotection – NaOH, THF:H_2O; coupling – cycle 1: $Pd(OAc)_2$, SPhos, K_2CO_3, THF, 55 °C, 16h. cycle 2: $Pd(OAc)_2$, XPhos, K_3PO_4, THF, 55 °C, 14 h. cycle 3: $Pd(OAc)_2$, SPhos, K_3PO_4, THF, 55 °C, 24 h; purification – SiO_2, MeOH:Et_2O; THF. All protecting groups other than MIDA (**R** = **TIPS**, **TBDPSE**, **TMSE**, or **Bz**) were successfully removed in a separate step (24).

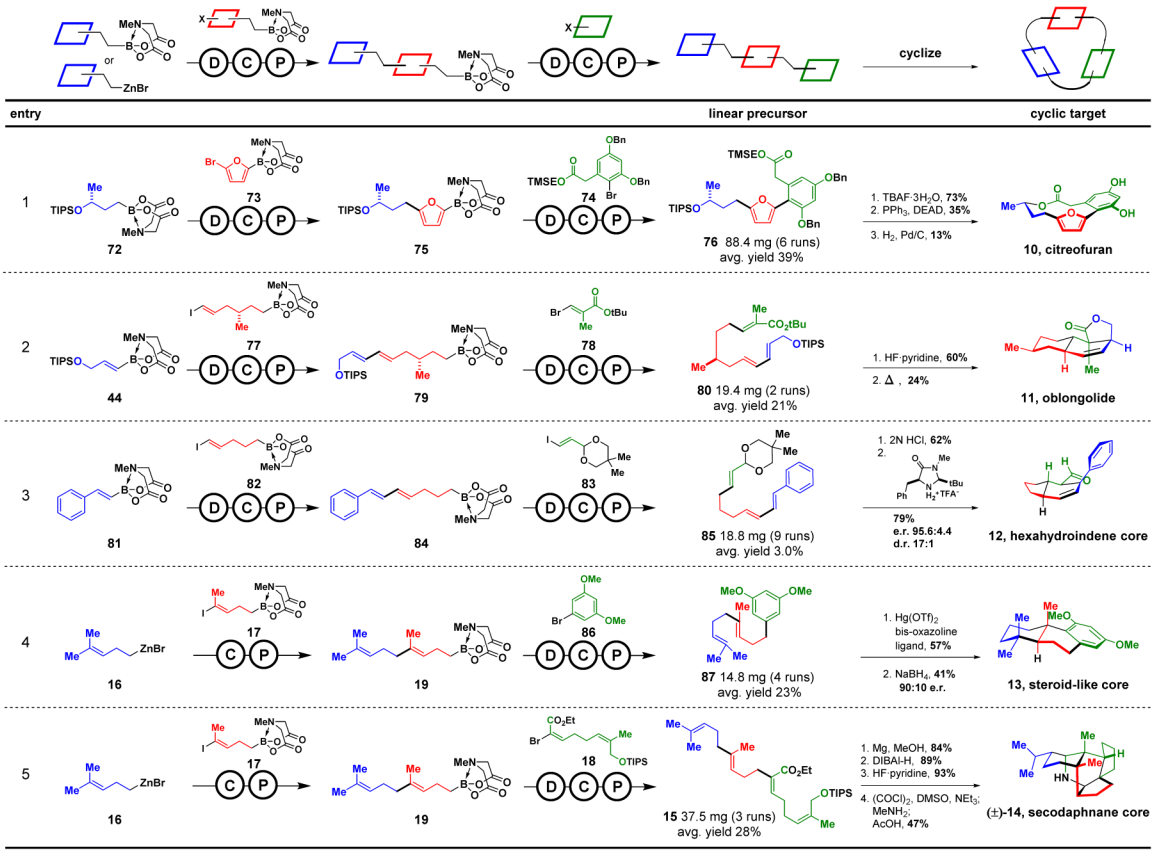

Fig. 4. Synthesis of Csp^3-rich macro- and polycyclic natural products and natural product-like cores

Modular linear precursors assembled via automated Csp^2 and Csp^3 couplings are diastereo- and/or enantioselectively cyclized.