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

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

Small-molecule synthesis usually relies on procedures that are 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 Csp3-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 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 toward 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 synthesis 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-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 to 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 to 4) are biosynthesized via the iterative assembly of a small set of building blocks, such as malonyl coenzyme A, iso- pentenyl pyrophosphate, and pyruvic acid (8). Many materials and pharmaceuticals (such as 5 to 9) comprise collections of aryl and/or heteroaryl components (9). Even topologically complex natural products containing macrocycllic or polycyclic frameworks (such as 10 to 14) are usually biosynthesized via an 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-methylimidic 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 of the required functional groups, oxidation states, and stereochemical elements are preinstalled into the building blocks. These features are then faithfully translated into the products using the same mild and sterースpecific 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 to 4, have been manu-

ally 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. Csp3-rich cyclic and polycyclic natural product frameworks such as 10 to 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 sp3 hybridized carbons. Building block–based assembly of these precursors thus requires many Csp3 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 biospanned 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 to 18 in which all of the required double bond stereochemistry is preinstalled. Csp3 coupling of 16 and 17 generates MIDA boronate intermediate 19. Deprotection of 19 yields a free boronic acid, which is engaged in another Csp3 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 a generalized automated purification. Solid-phase synthesis can address this problem for peptides (1), oligonucleotides (2), oligosaccharides (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, is 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. Thus, we needed a different solution.

Howard Hughes Medical Institute (HHMI), Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.
*These authors contributed equally to this work. †Corresponding author. E-mail: mdburke@illinois.edu
We recognized that each iteration of building block assembly in our platform generates a MIDA boronate as the key intermediate (Fig. 1, B 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 to 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:Et2O (Me, methyl; Et, ethyl) (Fig. 2A, left). However, all of the same MIDA boronates are rapidly eluted with tetrahydrofuran (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. The synthesizer filters and transfers the crude reaction mixture to the purification module, which executes the catch-and-release purification protocol with MeOH:Et2O 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, and vinyl 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:H2O, 23°C, 20 min) and coupling conditions (PdXPhos, K3PO4) (25), we obtained the desired cross-coupling products in good yields and purities in all cases (table S2, entries 1 to 3). 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:Et2O 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.

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 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 prepacked cartridges onto the synthesizer and pressing “start.”

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. 1Bu, tert-butyl. (B) Analogous building block–based strategies for the synthesis of peptides and small molecules. R, amino acid side chains. (C) Synthesis of the Csp3-rich pentacyclic secodaphnane core (T)-14 via iterative Csp3 couplings to yield a linear precursor followed by biosynthesis-inspired cascade polycyclization. An x-ray crystallographic study of the N-bromoacetyl derivative of 14 unambiguously confirmed the structure of 14. TIPS, trisopropylsilyl; Ph, phenyl; DMSO, dimethyl sulfoxide.

We recognized that each iteration of building block assembly in our platform generates a MIDA boronate as the key intermediate (Fig. 1, B 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 to 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:Et2O (Me, methyl; Et, ethyl) (Fig. 2A, left). However, all of the same MIDA boronates are rapidly eluted with tetrahydrofuran (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. The synthesizer filters and transfers the crude reaction mixture to the purification module, which executes the catch-and-release purification protocol with MeOH:Et2O 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, and vinyl 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:H2O, 23°C, 20 min) and coupling conditions (PdXPhos, K3PO4) (25), we obtained the desired cross-coupling products in good yields and purities in all cases (table S2, entries 1 to 3). The synthesizer was also capable of executing a
Csp3 coupling using Pd[P(o-tol)3]2 (tol, tolyl) and Ag2O/K2CO3 (table S2, entry 4).

Accessing many pharmaceuticals and materials represented by structures 5 to 9 requires the flexibility to link building blocks via carbon-heteroatom and/or carbon-carbon bonds. The stability of MIDA boronates toward 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 formations, including a Buchwald-Hartwig amination, O-alkylation, and amide bond formations (table S3). Despite the different reagents...
and by-products, 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 to 9) via multiple carbon-carbon and/or carbon-heteroatom bond formations (Fig. 2C). These include natural products from major biosynthetic pathways (1 to 4), materials components (5, 6), and pharmaceuticals and biological probes (7 to 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 multi-milligram 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 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 preprogrammed 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 to 14 in Fig. 1A) could be generated using the same automated building

Fig. 3. Automated synthesis of ratanhine derivatives. Conditions: Deprotection – NaOH, THF:H2O. Coupling – cycle 1: Pd(OAc)2, SPhos, K2CO3, THF, 55°C, 16 hours; cycle 2: Pd(OAc)2, XPhos, K3PO4, THF, 55°C, 14 hours; cycle 3: Pd(OAc)2, SPhos, K3PO4, THF, 55°C, 24 hours. Purification – SiO2, MeOH:Et2O; THF. All protecting groups other than MIDA [R: TIPS, TBDPSE, TMSE, or Bz (benzoyl)] were successfully removed in a separate step (24). OAc, acetate.
block assembly process and the linear-to-cyclized strategy. The macrocyclic natural product citreofuran possesses both Csp3 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 \(72\) to \(74\). All of the required stereochemical information for citreofuran is pre-encoded in the chiral nonracemic MIDA boronate building block \(72\). On the synthesizer, fully automated deprotection of \(72\), Csp3 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 atropodiastereoselectively macrocyclized to generate citreofuran.

Oblongolide is a norsesquiterpene \(\gamma\)-lactone natural product containing a 6,6,5-tricyclic core with five Csp3 stereogenic centers, one of which is quaternary (Fig. 4, entry 2). The three building blocks \(44\), \(77\), and \(78\) were automatically assembled via iterative Csp2 and Csp3 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 Csp3 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 \(81\) to \(83\), all possessing olefins with predefined 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 Csp3 coupling of building blocks \(16\), \(17\), and \(86\) generated linear precursor \(87\), which then underwent catalyst-promoted enantio- and diastereoselective cation-\(\pi\) 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 \((1\) to \(14\) 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.

REFERENCES AND NOTES
24. Materials and methods are available as supplementary materials on Science Online.

ACKNOWLEDGMENTS
We acknowledge the NIH (grants GM080435 and GM000153), the NSF (grant 0747778), HHMI, and Bristol-Myers Squibb for funding.

LUNAR GEOLOGY
A young multilayered terrane of the northern Mare Imbrium revealed by Chang’E-3 mission

Long Xiao,1,2* Peimin Zhu,1 Guanyou Fang,3a Zhiyong Xiao,1,2* Yongliao Zou,1 Jiannian Zhao,1 Na Zhao,1 Yuefeng Yuan,1 Le Qiao,1 Xiaoping Zhang,2 Hao Zhang,1 Jianguang Wang,1 Jun Huang,1 Qian Huang,1 Qi He,1 Bin Zhou,2 Yicai Ji,3 Qunying Zhang,3 Shaoxiong Shen,3 Yupi Li,3 Yunze Gao3

China’s Chang’E-3 (CE-3) spacecraft touched down on the northern Mare Imbrium of the lunar nearside (340.49°E, 44.12°N), a region not directly sampled before. We report preliminary results with data from the CE-3 lander descent camera and from the Yutu rover’s camera and penetrating radar. After the landing at a young 450-meter crater rim, the Yutu rover drove 114 meters on the ejecta blanket and photographed the rough surface and the excavated boulders. The boulder contains a substantial amount of crystals, which are most likely plagioclase and/or other mafic silicate mineral aggregates similar to terrestrial dolerite. The Lunar Penetrating Radar detection and integrated geological interpretation have identified more than nine subsurface layers, suggesting that this region has experienced complex geological processes since the Imbrian and is compositionally distinct from the Apollo and Luna landing sites.

Hang’E-3 (CE-3) landed at 340.49°E, 44.12°N on the Moon on 14 December 2013, and it released the Yutu (Jade Rabbit) rover the next morning (1). This was the first soft landing on the Moon since the Soviet Union’s Luna 24 mission in 1976 and is a new landing site in the north part of the Mare Imbrium (fig. S1). Yutu is China’s first lunar geologic mission and traveled in total ~114 m on the lunar surface. Following a zigzagging route, Yutu came to a halt about 20 m to the southwest of the landing site (Fig. 1). Yutu explored the lunar surface and subsurface near a young crater (fig. S2) using its four main instruments: Panoramic Camera, Lunar Penetrating Radar (LPR), Visible–Near Infrared Spectrometer (VNIS), and Active Particle-Induced X-ray Spectrometer (APXS). In this study, we report the preliminary results obtained from the cameras and LPR.

*China University of Geosciences, Wuhan 430074, China.
3a Macau University of Science and Technology, Macau, China.
1Institute of Electronics, Chinese Academy of Science, Beijing 100190, China.
2The Centre for Earth Evolution and Dynamics, University of Oslo, Sem Sælandsetvei 24, 0371 Oslo, Norway.
3Hong Kong University of Science and Technology, Shenzhen, Guangdong, China.

**Corresponding author. E-mail: yangl@cup.edu.cn (X.Y.); zhupm@cup.edu.cn (P.Z.); gyfang@mail.ie.ac.cn (G.F.)

SUPPLEMENTARY MATERIALS
www.sciencemag.org/content/347/6227/1221/suppl/DC1
Materials and Methods
Supplementary Text
Fig. S1 to S3
Tables S1 to S3
References (34–52)
22 December 2014; accepted 12 February 2015
10.1126/science.aaa5414
Synthesis of many different types of organic small molecules using one automated process
Junqi Li et al.
Science 347, 1221 (2015);
DOI: 10.1126/science.aaa5414

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of March 12, 2015):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/content/347/6227/1221.full.html

Supporting Online Material can be found at:
http://www.sciencemag.org/content/suppl/2015/03/11/347.6227.1221.DC1.html

A list of selected additional articles on the Science Web sites related to this article can be found at:
http://www.sciencemag.org/content/347/6227/1221.full.html#related

This article cites 46 articles, 4 of which can be accessed free:
http://www.sciencemag.org/content/347/6227/1221.full.html#ref-list-1

This article appears in the following subject collections:
Chemistry
http://www.sciencemag.org/cgi/collection/chemistry