Trends in Chemistry

CellPress REVIEWS

Review

Modularity: Adding New Dimensions to Total Synthesis

Alexander W. Sun,[1,2] Sebastian Lackner,[1,2] and Brian M. Stoltz[1,*]

As the field of synthetic chemistry seeks to tackle new frontiers, total synthesis is primed to address significant medical challenges such as the rise of antibiotic resistance and cancer. One emerging frontier focuses on increasingly concerted efforts to utilize modular total synthesis as a strategy to generate analogs of natural product targets for biological studies, with the ultimate goal of new therapeutic development. This new frontier is enabled by a confluence of human ingenuity in synthetic design, newly developed reactions that facilitate advances in total synthesis strategies, and emerging technologies. In this review, we highlight the evolving trend of modular total synthesis, including new reactions and automated technologies. This trend should lead to an increasingly important source of new medicines to improve human health.

Highlights

Total synthesis has made explosive strides since the 1800s. Whereas in the past, complex natural products were achieved in dozens of steps (sometimes up to 40 steps for relatively low-molecular-weight structures), today's total syntheses are significantly more concise, attaining complex targets often in as few as ten steps.

Such concise syntheses, when coupled with a modular strategy, enable the preparation of sufficient quantities of natural-product analogs for biological studies. Thus, natural product synthesis is increasingly used as a means of addressing current challenges in medical research and drug development.

However, the evolution of this emerging trend depends on the continued development of new technologies and new reactions that facilitate increasingly efficient and concise total syntheses.

Total Synthesis: Modern Challenges

In the 21st century, **total synthesis** (see Glossary) is primed to address significant medical challenges as the field seeks to tackle new frontiers. In addition to achieving ever-shorter, more concise syntheses, the primary focus is increasingly shifting toward efforts to generate analogs of natural product targets for biological studies, most commonly with the ultimate goal of new drug development. Natural products (with their complex 3D architectures) are often evolutionarily refined to exhibit pinpoint biological effects; however, they often have poor drug-like properties, such as off-target toxicities and low bioavailability. Thus, to produce a natural product-based therapeutic, structural optimization is often necessary to overcome such liabilities while maintaining on-target efficacy. Lovering and colleagues [1] suggested that moving from the current drug 'flat land' of predominantly aromatic compounds to more saturated, chiral targets featuring greater three-dimensionality may significantly increase clinical success rates in the search for new medicines, as increasing sp^3 character is associated with increased target selectivity [1,2]. However, the architectural intricacy of natural product-based medicines is double-edged: their complexity often precludes direct synthetic manufacturing routes, instead relegating their production to biosynthetic and semisynthetic means [3–5]. In recent years, this limitation has been increasingly overcome by changing tides in the field, as a combination of chemical ingenuity in synthetic design and new reaction development have resulted in increasingly efficient total syntheses that not only provide targets in sufficiently high yield to enable further biological investigations, but also enable the synthesis of analogs.

An important strategy to access such natural product analogs lies in modular total synthesis. Based on the design principle of **modularity** that involves dividing a system into smaller 'modules' that can be interchanged, modular total synthesis can be defined as a convergent approach to a target using interchangeable building blocks of intermediate complexity (Figure 1A). Therefore, by varying the component building blocks, diverse analogs of the target can be achieved. Modular approaches to natural products are desirable because they enable **structure–activity relationship (SAR)** interrogations of every structural component of the target (theoretically) while also providing access to a diverse chemical space. Suitable

[1]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA
[2]These authors contributed equally.

*Correspondence:
stoltz@caltech.edu (B.M. Stoltz).

modular analogs may additionally be used as chemical probes to interrogate biological pathways.

This modular approach to natural products is complementary to the rich history of identifying natural-product-like drugs through a set of related strategies comprising **semisynthesis**, **divergent synthesis**, and **diversity-oriented synthesis (DOS)** [6–10]. In semisynthesis, an advanced natural product isolated from a natural source such as a plant or bacterium is then divergently and chemically converted to a variety of analogs [3–5]. However, the structural diversity achievable through semisynthesis is inherently limited by the structure of the isolated natural product. In divergent approaches [11–18], structural intermediates along the way to a target natural product serve as branch points for the divergent synthesis of other molecules that are not necessarily natural products. Compared with modular synthesis, divergent approaches similarly access a wide swath of chemical space; however, the structural variations achievable through divergent synthesis are also limited by the complexity of the point of divergence. In this review, we focus exclusively on the trend of modular synthesis, which may lead to biologically efficacious natural product analogs and, by virtue of identifying the structural components necessary for bioactivity, could also result in structurally simplified analogs that are more amenable to commercial synthesis.

The use of the word 'modular' to describe convergent total syntheses is relatively new, appearing with significant frequency only in the past two decades (Figure 1B). This terminology has further experienced a rapid expansion in the total synthesis lexicon during the past 3 years of 2016–2018, appearing at least twenty times in 2016 in common organic chemistry journals. The increasing use of the term modularity coincides with an increase in analog-oriented total syntheses. In this review, we highlight a trend in the field toward the concise, modular total synthesis of pharmaceutically relevant natural products in a manner that enables scalability and SAR studies. We further demonstrate how the continued evolution of this trend depends on the development of new reaction **methods** and new technologies that facilitate modularity and efficiency. This review is therefore split into three parts that highlight: (i) new modular synthetic routes to natural products; (ii) the development of new methods that enable quick modular access to structurally complex motifs related to natural compounds; and (iii) new technologies that facilitate accelerated reaction discovery and modular total synthesis.

Modular Total Syntheses

In this section, we highlight recent representative modular syntheses in which analogs have been reported. The research group of Andrew Myers is well known for pioneering modern modular total syntheses, with one prominent early example embodied by their 2005 tetracycline campaign [19,20]. The tetracycline class of antibiotics are essential medicines that have been subject to intensive synthetic studies. Noting that the invariant AB ring of the tetracycle is primarily responsible for target binding, the Myers group developed a synthetic strategy facilitating modification of the D ring that is believed to be responsive to structural variation (Figure 2A). In stark contrast to existing literature that detailed semisynthetic approaches only or constructed the antibiotic scaffold linearly beginning from the D or the CD ring, Myers chose to join the AB ring with a modular D ring using a diastereoselective Michael–Dieckmann reaction sequence to form the C ring. This modular strategy enabled the synthesis of deoxytetracycline analogs in as few as 14 linear steps and with up to 7.0% overall yield while also enabling sweeping SAR studies on the D ring. A few of the over 50 reported fully synthetic deoxytetracycline analogs are displayed in Figure 2A. Notably, as a result of this research, a start-up company called Tetraphase was launched that utilizes this modular strategy to develop novel tetracycline antibiotics. As antibiotic resistance is an emerging crisis, novel antibiotics effective against resistant strains are direly needed.

Glossary

Convergency (in total synthesis): convergent total syntheses typically divide a target molecule into two or more building blocks of relatively equal molecular complexity. Convergent synthesis is most often contrasted with linear total synthesis, which adds increasing complexity to one primary building block *en route* to the target. Often, convergent routes are more concise and higher yielding than linear ones.

Divergent synthesis: a strategy designed to obtain a library of compounds derived from a common synthetic intermediate that serves as a point of structural divergence.

Diversity-oriented synthesis (DOS): characterized by a series of modular building blocks and convergent coupling reactions, DOS is a strategy to generate a library of small molecules that exhibit high structural variation and that do not necessarily map onto natural product scaffolds.

Methods (in organic chemistry): the development of new bond-forming reactions between atoms to access chemical motifs. New methods may lead to higher-yielding, more efficient routes to existing chemical structures or may enable access to completely new chemical space.

Modularity (in total synthesis): based on the design principle of modularity, which involves dividing a system into smaller 'modules' that can be interchanged, modular total synthesis can be defined as a convergent approach to a target molecule using interchangeable building blocks. By varying the individual building blocks, diverse analogs of the target can be synthesized. Modular syntheses can be challenging to develop as they require robust reactions with high chemocompatibility to accommodate a range of different building blocks. Nevertheless, modular approaches to natural products are desirable because they enable SAR interrogations.

Retrosynthesis: an analysis technique used to design conceptual routes toward a natural product. Retrosynthetic analysis involves a process of working backward, deconstructing a complex target step by step into increasingly simpler precursors.

Semisynthesis: a strategy to synthesize novel target molecules through modifying complex compounds

In 2013, Myers reported a modular total synthesis of the antitumor natural product trioxacarcin A, DC-45-A1, and various analogs (Figure 2B) [21,22]. The trioxacarcins are polycyclic bacterial metabolites that display potent cytotoxicity with subnanomolar growth inhibition (GI_{70}) values, making them attractive potential chemotherapeutics. Their bioactivity is thought to be due to the ability of a reactive epoxide to *N*-alkylate guanine residues in DNA. Myers took advantage of the two strategic glycosidic appendages by deconstructing the molecule into five modular components: two glycosides, **1** and **2**, and three core building blocks, **3**–**5**, which could be synthesized from commercial reagents in six to eight steps (Figure 2B). In the forward direction, a classical Kraus–Sugimoto annulation forged the quinone tricycle **6** from two modular fragments, **3** and **5** (Figure 2B). The tricycle was then advanced using a carbonyl-ylide cycloaddition with diazo **4** to forge the aglycon core of the molecule. From there, the authors developed glycosylation conditions to selectively append the two differing carbohydrates, yielding trioxacarcin A. The modularity of this route was exploited by varying core building blocks **3** and **5** and the two sugars **1** and **2** to synthesize over 30 structurally diverse analogs, some of which displayed potency comparable with that of trioxacarcin A (Figure 2B).

Two other notable modular syntheses in 2013 were pioneered by Micalizio and coworkers. The first utilized three relatively simple building blocks in a stereoselective synthesis of three macrocyclic lactams based on a polyketide natural product scaffold, the anticancer benzoquinone ansamycins [23]. At nearly the same time in 2013, Micalizio reported another modular synthesis, of the marine natural product lehualide B and 25 analogs [24]. Their efforts enabled SAR studies of the activity of lehualide B against multiple myeloma, helping to establish its biological mechanism of action.

In 2017, He reported a modular synthesis of various members of the dictyodendrin class of natural products (Figure 2C) [25]. The dictyodendrins have been subject to numerous synthetic efforts, as they display potent inhibitory activity against telomerase and β-amyloid cleaving enzyme 1 (BACE1). While modular routes to these molecules utilizing C–H functionalization were reported by Davies and Gaunt [26,27], only the He group demonstrated the modularity of their route through analog synthesis. Using four cyclic building blocks, **8**–**11**, He synthesized four of these fascinating natural products and six of their analogs, all in 11 steps with overall yields of 8–16%. Key to their modular efforts was their development of a formal [3+2] cycloaddition reaction [28–30] between aryne **13** and the core 2-aminoquinone intermediate **12** to forge the carbazolequinone core of the final product.

In a *tour de force* of modern modular synthesis, Myers reported in 2016 a convergent approach toward the synthesis of hundreds of macrolide antibiotics, which are indispensable broad-spectrum bacteriostatics used to treat a wide range of infections (Figure 3) [31]. Prior to this report, new macrolides such as azithromycin and clarithromycin were developed exclusively through semisynthetic modification of the fermented natural product erythromycin, resulting in limitations to the structural variations of these analogs. In Myers' approach, a highly convergent route is employed utilizing eight modular building blocks, **14**–**21** (Figure 3B), each of which can be modified to result in unprecedented access to novel macrolide analogs. Through variation of each building block, hundreds of structural modifications were made on varying locations on the ring (Figure 3B). Important to the success of this synthesis were robust stereoselective coupling reactions (Figure 3A). For instance, the initial step coupling ketone **18** and α-aminoamide **14** was facilitated by the development of a pseudoephenamine auxiliary-mediated stereoselective aldol reaction, reported in 2014 [32]. In another recent asymmetric methodology developed in 2007, a stereoselective palladium-catalyzed ene reaction [33] was utilized to couple enol ether **16** and dicarbonyl **19**. Later, in the penultimate macrocyclization step, Myers applied a ketene

isolated from natural sources such as plants or bacteria. In contrast to total synthesis, semisynthesis requires few steps to achieve a target compound, as the chemical synthesis begins with an already advanced substrate in which molecular complexity has been biosynthetically constructed by nature's enzymes. Representative drugs made through semisynthesis include tetracycline, paclitaxel, and artemether.

Structure–activity relationship (SAR): identifies patterns relating the structure of a molecule to its biological activity, including but not limited to properties such as solubility, potency, and metabolic stability. Often, modifying certain portions of a molecule will affect these biological properties.

Synthon: a formalism for a synthetic equivalent of a molecule used in a total synthesis. For example, C2 synthons include acetaldehyde and acetylene. A carbanion synthon includes organolithium reagents.

Total synthesis: the chemical synthesis of small-molecule natural products from simpler, commercially available building blocks. Total synthesis is distinguished from semisynthesis, in which natural products are synthesized by modifying advanced intermediates isolated through biological means such as fermentation. Theoretically, total synthesis enables access to wider chemical space than semisynthesis, which is constrained to the modification of an advanced intermediate that is isolated from a biological source.

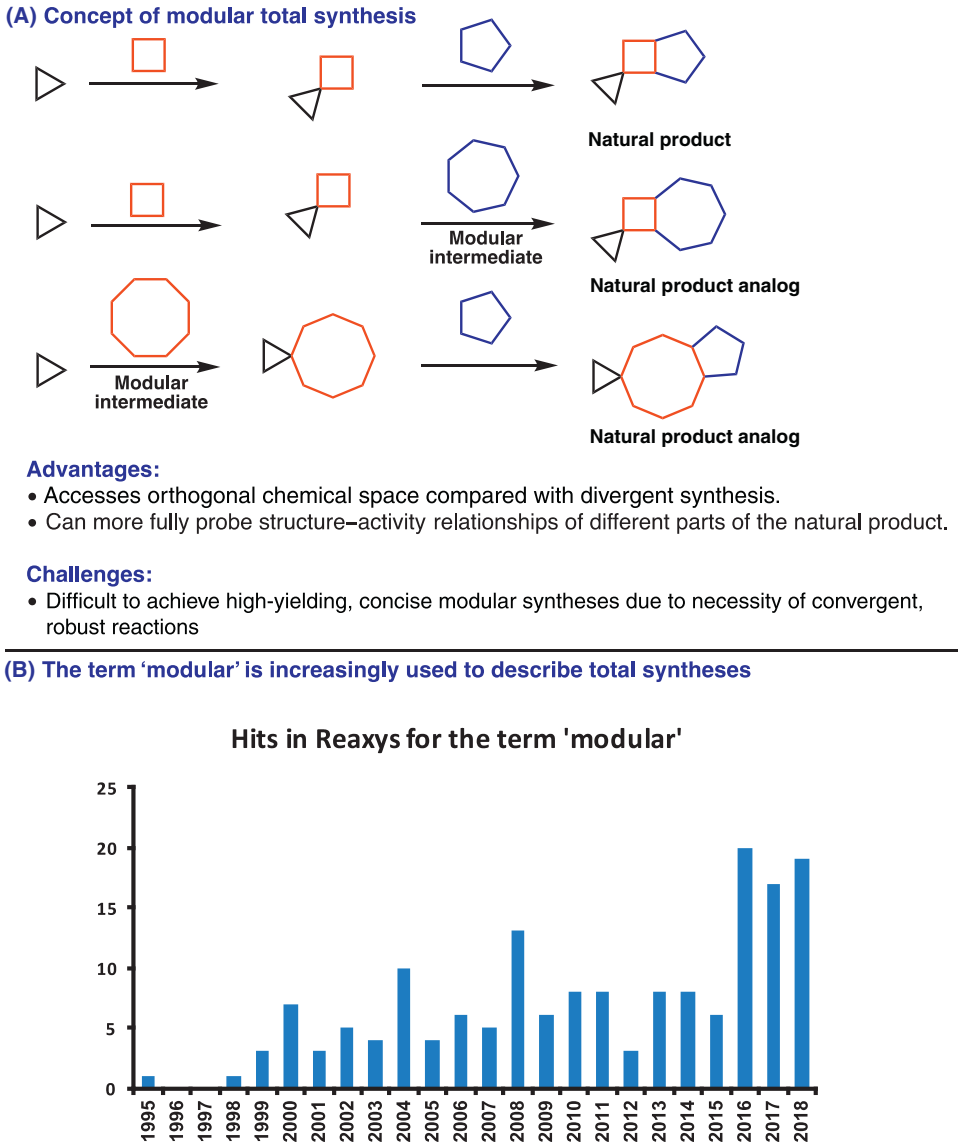

Figure 1. Convergent Modular Synthesis, an Emerging Trend.

cyclization previously developed by Boeckman to generate cyclized intermediate **29**; this impressive cyclization proved to be both high yielding and general [34]. The robust, enantioselective, and convergent coupling reactions utilized in this modular synthesis highlight the importance of new reaction development as an enabling tool for more efficient and concise syntheses of natural products. Notably, this synthetic platform formed the technology basis for Macrolide Pharmaceuticals, a start-up company focusing on the production of novel macrolide antibiotics.

Complex multicyclic natural products have also been synthesized using modular routes. In 2017, Magauer reported a modular synthesis of six related tetracyclic cytotoxic meroterpenoid natural

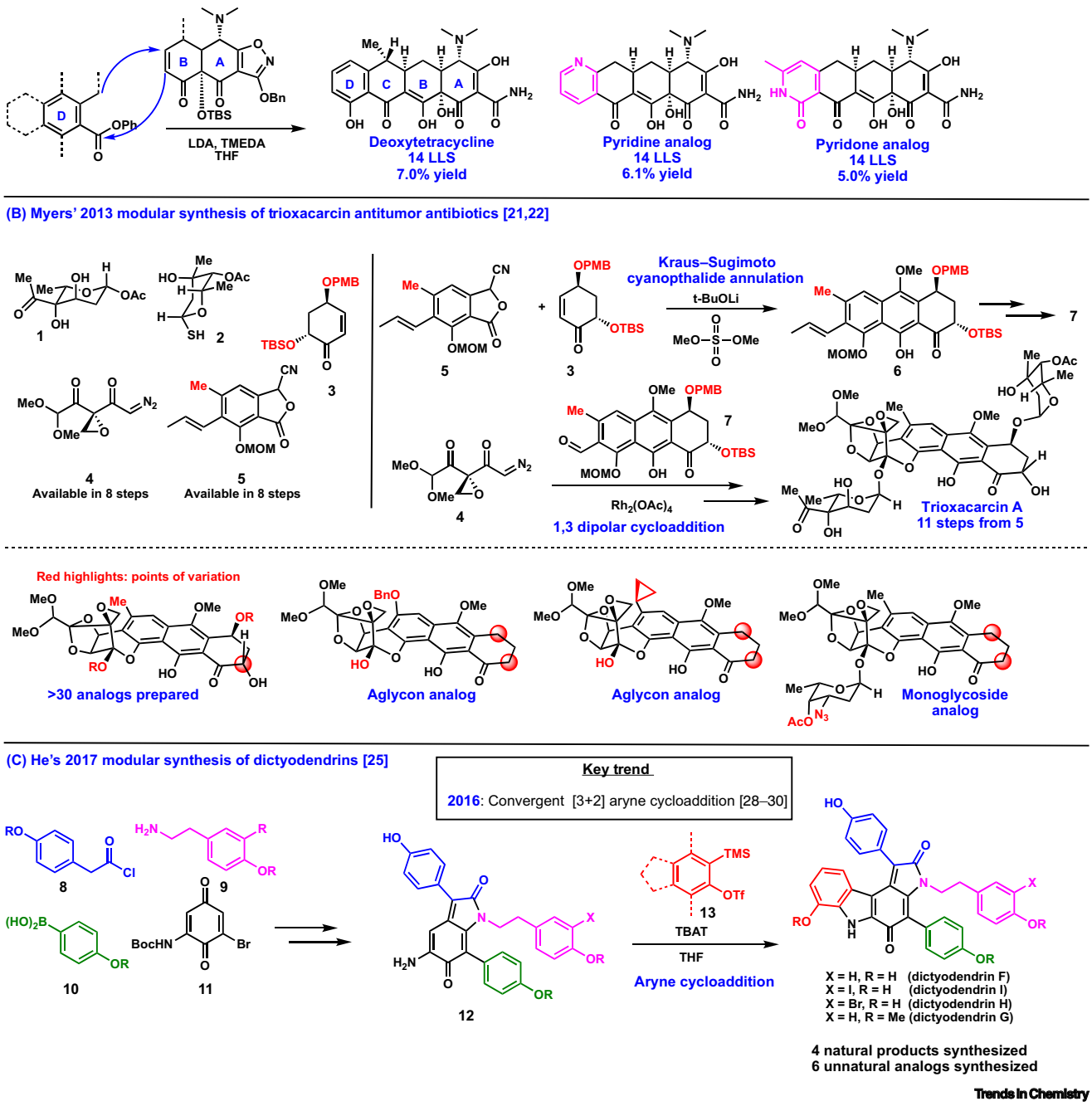

Figure 2. Recent Representative Modular Syntheses in Which Analogs Have Been Reported. (A) Modular total synthesis of tetracycline antibiotics [19,20]. (B) Modular total synthesis of cytotoxic trioxacarcin antitumor antibiotics [21,22]. (C) Modular total synthesis of dictyodendrin natural products [25]. See also [28–30]. Abbreviations: LDA, lithium diisopropylamide, a strong base; LLS, longest linear sequence, the largest number of reactions needed to transform any of several starting materials to the desired product; THF, tetrahydrofuran, a polar aprotic cyclic ether often used as a solvent; TMEDA, tetramethylethylenediamine, often used as a ligand for metal ions.

products: (+)-stachflin, (+)-aureol, (+)-smenoqualone, (+)-strongylin A, (–)-cyclosmenospongine, and (–)-mamanuthaquinone (Figure 4A) [35]. By varying three modular pieces, **30**–**32**, they were able to synthesize 15 non-natural analogs, even discovering an analog that displayed

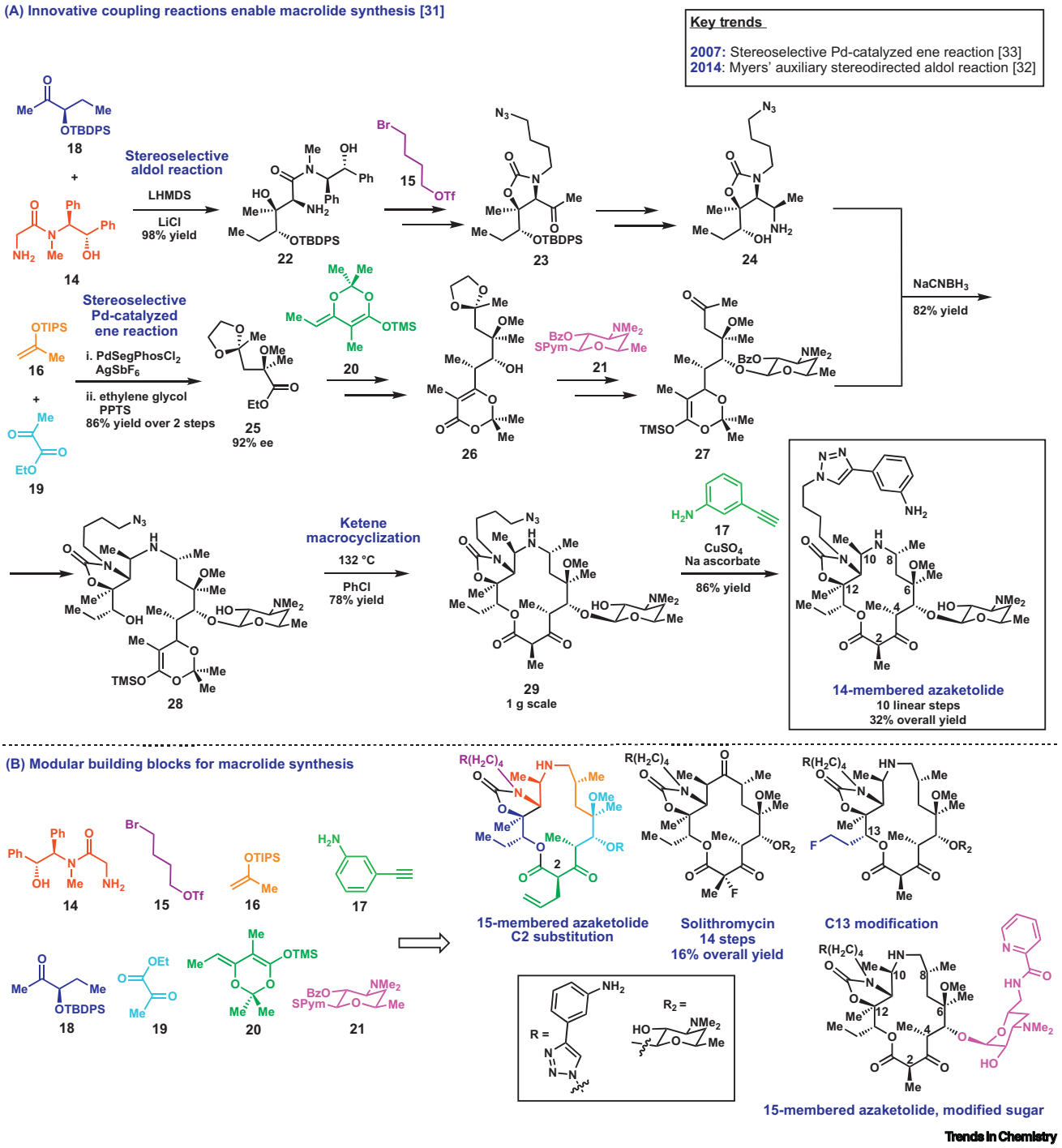

Figure 3. Modular Synthesis of Macrolide Antibiotics [31]. See also [32,33]. Abbreviations: LHMDS, lithium bis(trimethylsilyl)amide, a strong non-nucleophilic base; PPTS, pyridinium *p*-toluenesulfonate, an acid catalyst; OTBDPS, *O*-*tert*-butyldiphenylsilyl, a silyl ether resulting from the protection of a free alcohol with *tert*-butyldiphenylsilyl, a common protecting group.

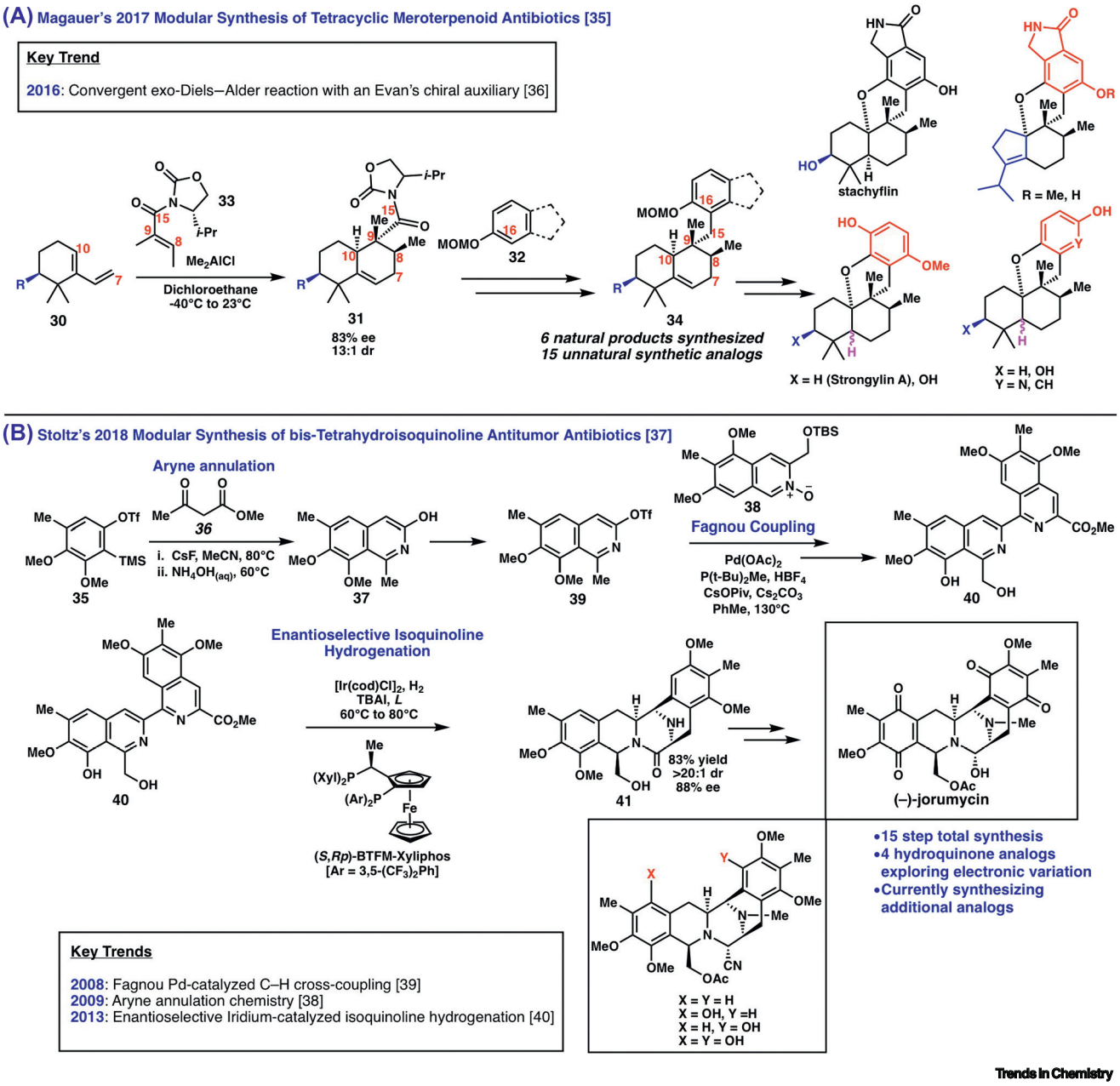

Figure 4. Modular Synthesis of Complex Multicyclic Natural Products. (A) Modular synthesis of meroterpenoid antibiotics [35]. (B) Modular synthesis of jorumycin [37]. See also [36–40]. Abbreviations: dr, diastereomeric ratio; ee, enantiomeric excess; OTBS, *O-tert*-butyldimethylsilyl, a silyl ether resulting from the protection of a free alcohol with *tert*-butyldimethylsilyl, a common protecting group; TMS, trimethylsilyl, a functional group often used to protect free alcohols or as a precursor to generate benzyne intermediates.

greater potency against methicillin-resistant *Staphylococcus aureus* (MRSA) than the reference natural product strongylin A. The key methodological advance that enabled their modular route was a novel Evans' auxiliary-mediated Diels–Alder reaction that unified diene **30** and dienophile **33** to assemble the 4,5-dehydrodecalin core **34** with a 13:1 diastereomeric ratio (dr) and 83% enantiomeric excess (ee) [36].

Recently, Stoltz and colleagues reported a modular synthesis of the highly cytotoxic bis-tetrahydroisoquinoline (bis-THIQ) natural product jorumycin and several analogs (Figure 4B) [37]. Owing to their extraordinary potency (IC_{50} 0.24 nM vs A549 lung cancer, 0.49 nM vs DU145 prostate cancer, and 0.57 nM vs HCT116 colon cancer) and unique chemical structures, bis-THIQ natural products have been extensively studied over the past 40 years, with ecteinascidin 743 (Yondelis) even receiving FDA approval for the treatment of various solid tumors. In diverging from conventional biomimetic cyclization strategies, their highly convergent route beginning from arene **35** enables for the first time the preparation of electronically varied hydroquinone analogs for biological studies; as the quinone ring of jorumycin is rapidly reduced *in vivo* to the hydroquinone, such analogs should more readily enable SAR investigations (Figure 4B).Three modern reactions were needed to accomplish this synthesis: first, an aryne annulation using β-ketoester **36** developed by their laboratory [38]; second, a key 2009 Fagnou C–H activating C–C cross-coupling to merge the two isoquinoline fragments **38** and **39** [39]; and third, a highly diastereo- and enantioselective iridium-catalyzed hydrogenation using an electron-deficient Xyliphos ligand to form the bis-THIQ core **41** [40]. The asymmetric hydrogenation reaction is especially remarkable, delivering four equivalents of hydrogen to convert the essentially flat bis-isoquinoline intermediate **40** into an sp^3-rich bis-THIQ core. This modular synthesis will enable biological investigations of unprecedented jorumycin analogs for chemotherapeutic purposes. Synthetic efforts in their laboratory are currently under way to design such analogs.

We believe that these aforementioned modular syntheses are heralds for the beginning of a rich era of convergent analog-oriented synthetic strategies. To facilitate the growth of modularity in total synthesis, the development of robust new reactions, especially stereoselective reactions, will be of great importance. For example, whereas acyclic compounds may be more easily deconstructed into modular components, bioactive complex polycyclic natural products present a special challenge to modular synthesis. For these polycyclic compounds, the development of convergent modular approaches amenable to variation is significantly complicated by the challenge of identifying modular fragments that can be efficiently coupled in sterically hindered environments. Future efforts toward the modular synthesis of such polycyclic targets will be greatly enabled by the development of robust reactions that can quickly generate complexity.

Novel Synthetic Methods That Facilitate Modular Synthesis

By necessity, modular synthesis must be convergent, and thus the development of highly stereoselective and convergent methods that quickly generate molecular complexity will be crucial in designing modular routes to natural products. With the goal of building complexity in a fast and straightforward way, multiple methods have emerged in the past few decades [41]. Such strategies include domino/cascade reactions (including recent examples from the Nagorny group) [42–45], multicomponent reactions [46,47], and direct C–H functionalization [48,49]. A recent example of such a novel methodology is a cycloaddition–cross-coupling sequence developed by the Baran group [50]. Through this sequence, various relevant natural and pharmaceutical compounds and key intermediates can be obtained in significantly fewer steps than existing state-of-the-art routes (Figure 5A,B). Additionally, to make full use of this new methodology, new 2π **synthons** were explored in this work, which concisely shows a vital interaction of methodology and synthesis: applying methods in a new way for more efficient and modular syntheses; realizing a need to extend the available synthons to access a wider range of relevant intermediates; and then, based on targets, investigating the required extensions of methodology.

Other such new methodologies include an enantioselective version of the venerable multicomponent Ugi reaction developed by the Tan group [51] and an oxidative enolate coupling/ring-closing metathesis sequence developed by the Thomson group [52], each giving easy access to a large variety of

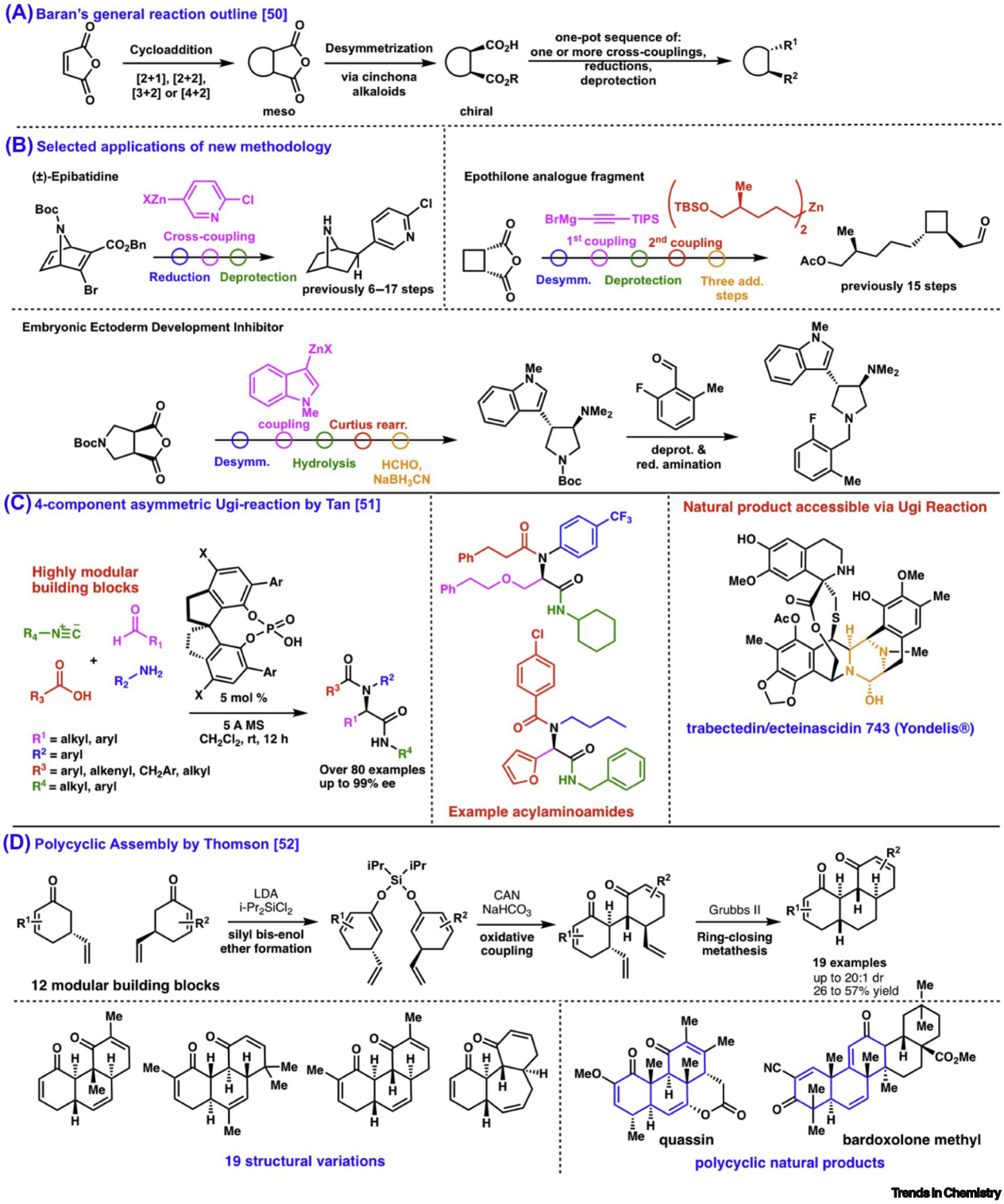

Figure 5. Novel Synthetic Methods That Facilitate Modular Synthesis. (A,B) Cycloaddition–cross-coupling sequence for facile access to (sp^3)-rich motifs [48]. (C) Four-component Ugi reaction to access acylaminoamides [51]. (D) Oxidative coupling of silyl bis-enol ethers to access complex polycyclic motifs [52]. See also [50]. Abbreviations: dr, diastereomeric ratio; ee, enantiomeric excess.

useful building blocks related to natural compounds (Figure 5C,D). Tan's enantioselective Ugi reaction gives biologically relevant α-acylaminoamides in good yields and enantiomeric excess (ee) while allowing a variety of substituents. Using two similar SPINOL-derived chiral phosphoric acid catalysts, they provided over 80 examples of chiral acylaminoamide products with impressive molecular diversity. The reaction enables a variety of alkyl-, aryl-, and alkenyl-substituted products to be synthesized in a single step starting from fairly simple and available substrates. The **convergency** and modularity of this stereoselective approach holds great potential for the modular synthesis of peptide-related natural products; for example, the FDA-approved chemotherapeutic trabectidin may be accessed through this modular Ugi reaction (Figure 5C) [53]. Similarly, starting from readily available, easily silylated enones, the Thomson group's sequence features a sequential diastereoselective oxidative enolate coupling/ring-closing methathesis strategy to forge polycyclic compounds containing a variety of readily modifiable enone and olefin functional handles (Figure 5D). Using 12 diverse modular enone building blocks, they could assembly a variety of substituted carbocyclic compounds that readily map onto a variety of natural products. Through this simple sequence, one can potentially access a large variety of terpenoid- and steroid-related scaffolds with ease. The above highlighted reactions represent a few of the imaginative strategies that researchers have developed to quickly generate complexity in a manner that may be adaptable for natural product synthesis.

New Tools to Facilitate Modular Total Syntheses

Historic technologies such as infrared (IR) spectroscopy, X-ray crystallography, and NMR have greatly enabled all aspects of organic synthesis [54], from theory and reaction development to total synthesis. In recent years, newly developed technologies promise to further enhance the speed of discovery in the field of organic synthesis, and by extension will greatly facilitate modular total synthesis (Figure 6). In this section, we briefly describe a few such emerging research tools, with automated technologies as the central theme [55].

Automation is an encompassing theme of the 21st century. In the realm of chemistry, flow reactors represent one example of the potential of automation to change our practice of organic synthesis: flow chemistry involves performing a chemical reaction in a reactor system involving a flowing stream (gas or liquid) of reactants [56–60]. This setup contrasts with a traditional batch system such as a round-bottom flask. A typical flow setup involves pumps that are required to move the fluid stream in a continuous system. Flow chemistry offers many advantages: atypical reaction conditions such as extremely high pressures or temperatures exceeding a solvent's boiling point may be achieved. Additionally, the high surface area results in optimal heat transfer and mixing of reactants, which can accelerate reaction times and improve yields. We envision that flow chemistry may accelerate total synthesis as scaling up intermediates is a time-consuming process. As flow setups become increasingly accessible to academic research laboratories, they may be used to automate scale-up for certain synthetic intermediates, freeing time for the researcher to continue the pioneering aspects of the total synthesis. The synthesis machine developed by the Burke laboratory is a notable evolution of flow reactor design to include chromatographic purification in series: using modular *N*-methyliminodiacetic acid (MIDA) boronates in sequential Suzuki cross-couplings, Burke assembled complex polyene intermediates that could then undergo biomimetic cyclizations to rapidly generate the cores of complex polycyclic caged natural products (Figure 6B) [61,62]. For natural products amenable to biomimetic cyclizations from linear polyene precursors, the Burke flow reactor represents a significant synthetic advance.

Automation has also been extended toward techniques to enhance reaction discovery and analysis (Figure 6A). For instance, high-throughput robotic arrays have been reported that allow massively parallel reaction screening of dozens of parameters such as reactants, ligand, and solvent, to identify new reactivity and to optimize existing reactions [63–67]. Such high-throughput approaches when coupled with HPLC-mass spectrometry (MS) readouts allow quantification, optimization, and

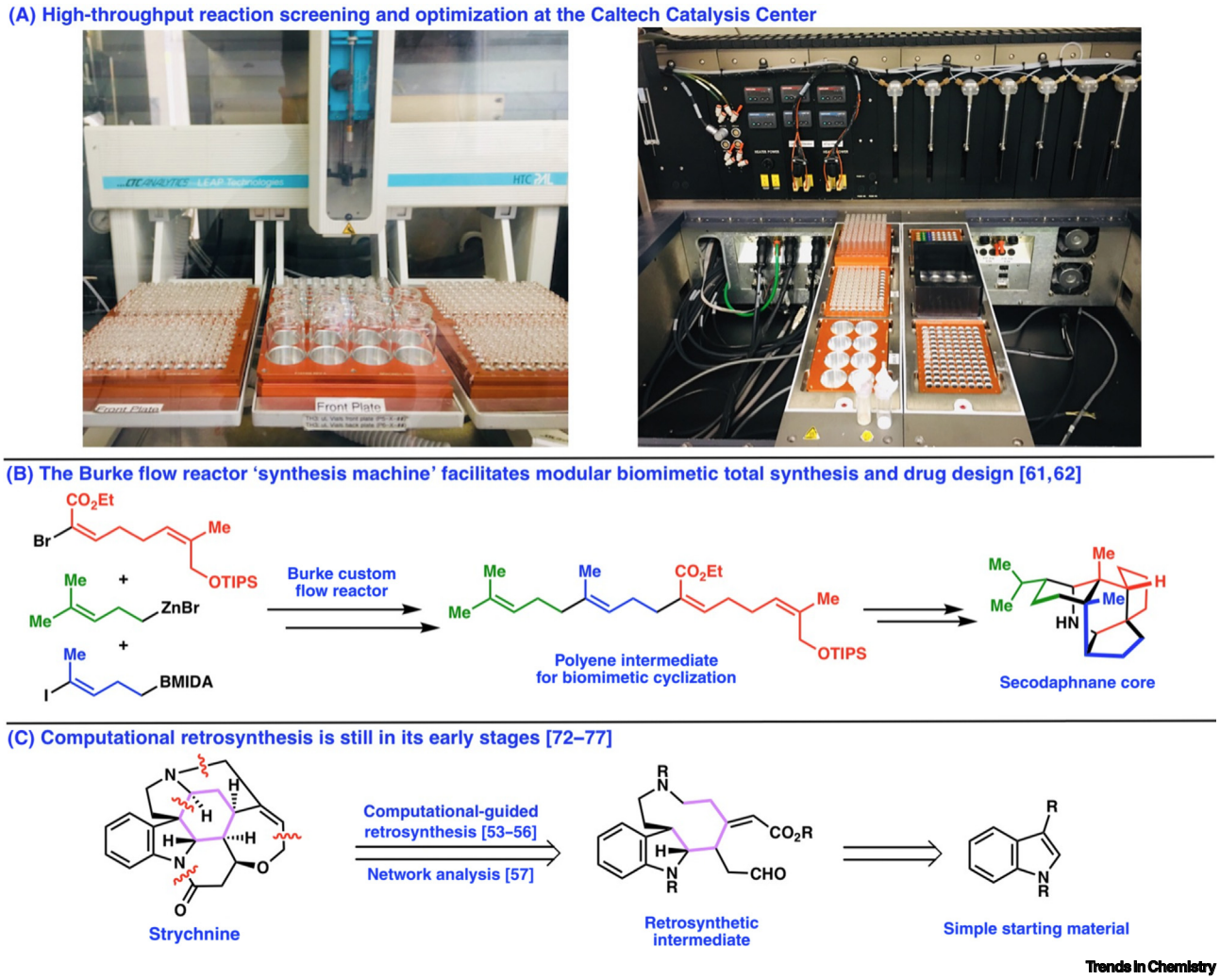

Figure 6. Automation and Computation-Assisted **Retrosynthesis** Will Facilitate Modular Total Synthesis. See also [53–57,61,62,72–77].

identification of new products. Recently, Cronin reported a machine-learning approach to reaction discovery involving a robotic apparatus linked to flow NMR and an IR spectrometer [64]. After establishing a training set of various combinations of known reactive and nonreactive compounds, the researchers used new reactants to identify new modes of reactivity. Ultimately, they reported four new reactions, including a three-component reaction to build a new dihydrofuran ring system. This type of machine-learning approach, coupled with high-throughput screening systems, has the potential to accelerate the discovery of reactions that will facilitate modular total synthesis. However, machine learning as applied to organic chemistry is still in its infancy, and protocol design must be carefully considered to obtain data that can be extrapolated to new chemical systems [68,69]. In addition, various high-throughput technologies with expensive hardware may not be broadly accessible to academic research groups. A less hardware-intensive reaction-screening approach was reported by Hartwig using simpler arrayed wells to conduct smaller-scale screens [65]. These technologies are increasingly becoming more automated. For instance, an automated reaction optimization protocol was reported using a machine-learning algorithm to analyze reaction

outputs and optimize various reactions [57,63]. The optimization system is capable of performing up to 32 reactions per day, exceeding the capabilities of an average bench chemist. We envision that researchers may adapt such reaction optimization protocols to more quickly improve yields on difficult steps of a modular total synthesis.

Finally, we highlight the application of cryoelectron microscopy (CryoEM) toward small-molecule structural elucidation [70,71]. CryoEM originally fell under the purview of molecular biologists as a tool to view the atomic structure of cells and proteins, but was recently extended to enable the astounding determination of the molecular structure of various microcrystalline small molecules. The primary advantage of this method, termed microcrystalline electron diffraction (MicroED), is that X-ray-quality crystals are unnecessary; samples possessing microcrystallinity only 1/6000th of the volume of a typical X-ray-quality crystal are needed to produce a diffraction pattern. MicroED thus promises to greatly facilitate structural determination in total synthesis. Often in the course of reactions, unidentified products are isolated, sometimes in minute quantities; MicroED presents a potentially rapid way to determine structure, and through continued development it may rival NMR analysis and X-ray crystallography as the gold standards for structural determination. Thus, MicroED promises to accelerate natural product total synthesis, reaction discovery, and drug development.

Outstanding Questions

How can computer scientists and organic chemists best collaborate to design computational retrosynthesis programs capable of designing routes toward complex 3D natural products?

Can a universal automated synthesis machine be developed and what is the timeline for it?

What kinds of new highly enantioselective bond-forming reactions will researchers develop to enable complex natural product synthesis?

What is the best way to synergistically implement tools of automation such as flow chemistry, reaction optimization, and new-reaction screening, toward modular total synthesis?

Concluding Remarks

As total synthesis continues to advance, pioneering researchers in the field are searching for new, exciting research avenues, one of which lies in identifying the biological and medicinal applications of their synthetic work. In this review, we have highlighted an emerging trend of modular total synthesis as a strategy to access structural analog libraries of bioactive natural products. Additionally, we explored the development of automated technologies such as flow chemistry, high-throughput reaction screening, and machine-learning-driven reaction optimization, which promise to facilitate modular total synthesis in several ways: through the discovery of new robust stereoselective and/or convergent reactions that can rapidly build molecular complexity for total synthesis; through the widening of total synthesis bottlenecks such as scale-up procedures; and through enhancing reaction yield and efficacy.

Currently, the computational design of retrosynthetic routes toward complex 3D targets is one of the greatest challenges in automation [72]. Recent software successfully identifies competent synthetic routes to simpler flat molecules [73–75]. However, the retrosynthetic analysis of complex natural products remains mostly limited to the purview of highly competent human chemists. Nevertheless, as computer scientists and organic chemists increasingly advance the software, such automation programs may eventually identify feasible modular routes to complex targets on a routine basis. This field is progressing, as exemplified by a recent computational route to the complex natural product epicolactone [72]. Additionally, to aid in the retrosynthetic deconstruction of complex polycyclic natural products, the Sarpong group has developed an algorithmic form of network analysis to identify the ring with most bridgehead atoms (Figure 6C, maximally bridged ring highlighted in lavender) [76]. In the forward direction, chemists are already using high-level density functional theory calculations to predict substrate variants with the highest likelihood of success in total synthesis routes [77]. These advances should be welcome, as they would enable the chemist to contemplate previously unconsidered modular synthetic routes. Ultimately, the human chemist must still be the final arbitrator in determining and executing the most promising of such computationally designed routes, at least for the time being.

These new technologies, combined with an ever-increasing synthetic compendium of reactions as well as human synthetic ingenuity, will lead to more frequent modular approaches toward natural products. Modular approaches represent a shift in the ideals of total synthesis toward not only achieving a synthetic target but also exploring its SAR through analog synthesis.

This strategy comes with its own challenges (see Outstanding Questions), however, as there is a high bar for conciseness and convergence in modular synthesis: the target and its analogs must be achieved in a relatively short amount of time and in sufficient quantity to enable streamlined biological studies. Thus, the practice of modular synthesis seeks to push the field to new heights through utilizing chemical ingenuity to identify and execute modular routes to natural products. Ultimately, one goal of modular synthesis is to identify lead-like compounds that may be developed into new medicines. Natural products have always represented a rich source of and inspiration for new drugs and thus this trend is poised to address some of the most pressing medical challenges in the 21st century: antibiotic resistance, cancer, HIV, malaria, metabolic syndrome, and neurological disorders such as Alzheimer's, Parkinson's, and depression.

Acknowledgments

The NIH-NIGMS (R01GM080269) and Caltech are thanked for support of our research program. A.W.S. thanks the NIH-NIGMS for a predoctoral fellowship (Ruth L. Kirschstein Institutional National Research Service Award F30GM120836) and a UCLA-Caltech Medical Scientist Training Program Fellowship (T32GM008042). S.L. acknowledges the Deutsche Forschungsgemeinschaft (DFG) for a postdoctoral fellowship (LA 4220/1-1).

References

1. Lovering, F. *et al.* (2009) Escape from Flatland: increasing saturation as an approach to improving clinical success. *J. Med. Chem.* 52, 6752–6756
2. Lovering, F. (2013) Escape from Flatland 2: complexity and promiscuity. *Med. Chem. Commun.* 4, 515–519
3. Expósito, O. *et al.* (2009) Biotechnological production of Taxol and related taxoids: current state and prospects. *Anticancer Agents Med. Chem.* 9, 109–121
4. Paddon, C.J. *et al.* (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496, 528–532
5. Cuevas, C. and Francesch, A. (2009) Development of Yondelis (trabectedin, ET-743). A semisynthetic process solves the supply problem. *Nat. Prod. Rep.* 26, 322–337
6. Burke, M.D. and Schreiber, S.L. (2004) A planning strategy for diversity-oriented synthesis. *Angew. Chem. Int. Ed. Engl.* 43, 46–58
7. Connor, C.J.O. *et al.* (2012) Diversity-oriented synthesis: producing chemical tools for dissecting biology. *Chem. Soc. Rev.* 41, 4444–4456
8. Galloway, W.R.J.D. *et al.* (2010) Diversity-oriented synthesis as a tool for the discovery of novel biologically active small molecules. *Nat. Commun.* 1, 80
9. Kato, N. *et al.* (2016) Diversity-oriented synthesis yields novel multistage antimalarial inhibitors. *Nature* 538, 344–349
10. Schreiber, S.L. (2000) Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* 287, 1964–1969
11. Li, L. *et al.* (2018) Divergent strategy in natural product total synthesis. *Chem. Rev.* 118, 3752–3832
12. Hugelshofer, C.L. and Magauer, T. (2015) Total synthesis of the leucosceptroid family of natural products. *J. Am. Chem. Soc.* 137, 3807–3810
13. Jones, S.B. *et al.* (2011) Collective synthesis of natural products by means of organocascade catalysis. *Nature* 475, 183–188
14. Mizoguchi, H. *et al.* (2014) Biogenetically inspired synthesis and skeletal diversification of indole alkaloids. *Nat. Chem.* 6, 57–64
15. Wender, P.A. *et al.* (2017) Scalable synthesis of bryostatin 1 and analogs, adjuvant leads against latent HIV. *Science* 358, 218–223
16. Abbasov, M.E. *et al.* (2019) Simplified immunosuppressive and neuroprotective agents based on gracilin A. *Nat. Chem.* 11, 342
17. Charaschanya, M. and Aubé, J. (2018) Reagent-controlled regiodivergent ring expansions of steroids. *Nat. Commun.* 9, 934
18. McLeod, M.C. *et al.* (2014) Probing chemical space with alkaloid-inspired libraries. *Nat. Chem.* 6, 133–140
19. Charest, M.G. *et al.* (2005) A convergent enantioselective route to structurally diverse 6-deoxytetracycline antibiotics. *Science* 308, 395–398
20. Sun, C. *et al.* (2008) A robust platform for the synthesis of new tetracycline antibiotics. *J. Am. Chem. Soc.* 130, 17913–17927
21. Švenda, J. *et al.* (2011) A multiply convergent platform for the synthesis of trioxacarcins. *Proc. Natl Acad. Sci. U. S. A.* 108, 6709–6714
22. Magauer, T. *et al.* (2013) Component-based syntheses of trioxacarcin A, DC-45-A1 and structural analogues. *Nat. Chem.* 5, 886–893
23. Jeso, V. *et al.* (2013) Synthesis of benzoquinone ansamycin-inspired macrocyclic lactams from shikimic acid. *Angew. Chem. Int. Ed. Engl.* 52, 4800–4804
24. Jeso, V. *et al.* (2013) Synthesis and SAR of lehualide B: a marine-derived natural product with potent anti-multiple myeloma activity. *ACS Chem. Biol.* 8, 1241–1252
25. Banne, S. *et al.* (2017) A unified modular synthetic strategy for dictyodendrins F, H, I, and G. *Org. Lett.* 19, 4996–4999
26. Yamaguchi, A.D. *et al.* (2015) Concise syntheses of dictyodendrins A and F by a sequential C–H functionalization strategy. *J. Am. Chem. Soc.* 137, 644–647
27. Pitts, A.K. *et al.* (2015) A concise and scalable strategy for the total synthesis of dictyodendrin B based on sequential C–H functionalization. *Angew. Chem. Int. Ed. Engl.* 54, 5451–5455
28. Guo, J. *et al.* (2016) Synthesis of carbazolequinones by formal [3 + 2] cycloaddition of arynes and 2-aminoquinones. *Org. Lett.* 18, 2499–2502
29. Guo, J. *et al.* (2016) Total synthesis of calothrixins and their analogues via formal [3+2] cycloaddition of arynes and 2-aminophenanthridinedione. *Tetrahedron Lett.* 57, 3481–3484
30. Gilmore, C.D. *et al.* (2008) Orthogonal synthesis of indolines and isoquinolines via aryne annulation. *J. Am. Chem. Soc.* 130, 1558–1559
31. Seiple, I.B. *et al.* (2016) A platform for the discovery of new macrolide antibiotics. *Nature* 533, 338–345
32. Seiple, I.B. *et al.* (2014) Stereocontrolled synthesis of *syn*-β-hydroxy-α-amino acids by direct aldolization of pseudoephenamine glycinamide. *Angew. Chem. Int. Ed. Engl.* 53, 4642–4647
33. Mikami, K. *et al.* (2007) Enantioselective catalysis of ketoester-ene reaction of silyl enol ether to construct quaternary carbons by chiral dicationic palladium(II) complexes. *J. Am. Chem. Soc.* 129, 12950–12951
34. Boeckman, R.K. and Pruitt, J.R. (1989) A new, highly efficient, selective methodology for formation of medium-ring and macrocyclic lactones via intramolecular ketene trapping: an application to a convergent synthesis of (–)-kromycin. *J. Am. Chem. Soc.* 111, 8286–8288
35. Wildermuth, R. *et al.* (2017) A modular synthesis of tetracyclic meroterpenoid antibiotics. *Nat. Commun.* 8, 2083

36. Buter, J. *et al.* (2016) Stereoselective synthesis of 1-tuberculosinyl adenosine; a virulence factor of *Mycobacterium tuberculosis*. *J. Org. Chem.* 81, 6686–6696
37. Welin, E.R. *et al.* (2019) Concise total syntheses of (–)-jorunnamycin A and (–)-jorumycin enabled by asymmetric catalysis. *Science* 363, 270–275
38. Allan, K.M. *et al.* (2009) Expedient synthesis of 3-hydroxyisoquinolines and 2-hydroxy-1,4-naphthoquinones via one-pot aryne acyl-alkylation/condensation. *Org. Biomol. Chem.* 7, 4960–4964
39. Campeau, L-C. *et al.* (2008) Site-selective sp^2 and benzylic sp^3 palladium-catalyzed direct arylation. *J. Am. Chem. Soc.* 130, 3266–3267
40. Iimuro, A. *et al.* (2013) Asymmetric hydrogenation of isoquinolinium salts catalyzed by chiral iridium complexes: direct synthesis for optically active 1,2,3,4-tetrahydroisoquinolines. *Angew. Chem. Int. Ed. Engl.* 52, 2046–2050
41. Jones, A.C. *et al.* (2014) Toward a symphony of reactivity: cascades involving catalysis and sigmatropic rearrangements. *Angew. Chem. Int. Ed. Engl.* 53, 2556–2591
42. Grondal, C. *et al.* (2010) Organocatalytic cascade reactions as a new tool in total synthesis. *Nat. Chem.* 2, 167–178
43. Poulin, J. *et al.* (2009) Pericyclic domino reactions: concise approaches to natural carbocyclic frameworks. *Chem. Soc. Rev.* 38, 3092–3101
44. Cichowicz, N.R. *et al.* (2015) Concise enantioselective synthesis of oxygenated steroids via sequential copper(II)-catalyzed Michael addition/intramolecular aldol cyclization reactions. *J. Am. Chem. Soc.* 137, 14341–14348
45. Khatri, H.R. *et al.* (2019) Modular total synthesis and cell-based anticancer activity evaluation of ouabagenin and other cardiotonic steroids with varying degrees of oxygenation. *J. Am. Chem. Soc.* 141, 4849–4860
46. de Graaff, C. *et al.* (2012) Recent developments in asymmetric multicomponent reactions. *Chem. Soc. Rev.* 41, 3969
47. Dömling, A. *et al.* (2012) Chemistry and biology of multicomponent reactions. *Chem. Rev.* 112, 3083–3135
48. Gutekunst, W.R. and Baran, P.S. (2011) C–H functionalization logic in total synthesis. *Chem. Soc. Rev.* 40, 1976–1991
49. Sinha, S.K. *et al.* (2018) Natural product synthesis by C–H activation. *Asian J. Org. Chem.* 7, 1178–1192
50. Chen, T-G. *et al.* (2018) Building $C(sp^3)$-rich complexity by combining cycloaddition and C–C cross-coupling reactions. *Nature* 560, 350
51. Zhang, J. *et al.* (2018) Asymmetric phosphoric acid-catalyzed four-component Ugi reaction. *Science* 361, eaas8707
52. Robinson, E.E. and Thomson, R.J. (2018) A strategy for the convergent and stereoselective assembly of polycyclic molecules. *J. Am. Chem. Soc.* 140, 1956–1965
53. Endo, A. *et al.* (2002) Total synthesis of ecteinascidin 743. *J. Am. Chem. Soc.* 124, 6552–6554
54. Whitesides, G.M. (2018) Complex organic synthesis: structure, properties, and/or function? *Isr. J. Chem.* 58, 142–150
55. Boström, J. *et al.* (2018) Expanding the medicinal chemistry synthetic toolbox. *Nat. Rev. Drug Discov.* 17, 709–727
56. Adamo, A. *et al.* (2016) On-demand continuous-flow production of pharmaceuticals in a compact, reconfigurable system. *Science* 352, 61–67
57. Perera, D. *et al.* (2018) A platform for automated nanomole-scale reaction screening and micromole-scale synthesis in flow. *Science* 359, 429–434
58. Plutschack, M.B. *et al.* (2017) The hitchhiker's guide to flow chemistry. *Chem. Rev.* 117, 11796–11893
59. Newton, S. *et al.* (2014) Accelerating spirocyclic polyketide synthesis using flow chemistry. *Angew. Chem. Int. Ed. Engl.* 53, 4915–4920
60. Ziegler, R.E. *et al.* (2018) 7-Step flow synthesis of the HIV integrase inhibitor dolutegravir. *Angew. Chem. Int. Ed. Engl.* 57, 7181–7185
61. Li, J. *et al.* (2015) Synthesis of many different types of organic small molecules using one automated process. *Science* 347, 1221–1226
62. Lehmann, J.W. *et al.* (2018) Towards the generalized iterative synthesis of small molecules. *Nat. Rev. Chem.* 2, 0115
63. Bédard, A-C. *et al.* (2018) Reconfigurable system for automated optimization of diverse chemical reactions. *Science* 361, 1220–1225
64. Granda, J.M. *et al.* (2018) Controlling an organic synthesis robot with machine learning to search for new reactivity. *Nature* 559, 377
65. Robbins, D.W. and Hartwig, J.F. (2011) A simple, multidimensional approach to high-throughput discovery of catalytic reactions. *Science* 333, 1423–1427
66. McNally, A. *et al.* (2011) Discovery of an α-amino C–H arylation reaction using the strategy of accelerated serendipity. *Science* 334, 1114–1117
67. Collins, K.D. *et al.* (2014) Contemporary screening approaches to reaction discovery and development. *Nat. Chem.* 6, 859–871
68. Ahneman, D.T. *et al.* (2018) Predicting reaction performance in C–N cross-coupling using machine learning. *Science* 360, 186–190
69. Chuang, K.V. and Keiser, M.J. (2018) Comment on "Predicting reaction performance in C–N cross-coupling using machine learning". *Science* 362, eaat8603
70. Jones, C.G. *et al.* (2018) The cryoEM method MicroED as a powerful tool for small molecule structure determination. *ACS Cent. Sci.* 4, 1587–1592
71. Gruene, T. *et al.* (2018) Rapid structure determination of microcrystalline molecular compounds using electron diffraction. *Angew. Chem. Int. Ed. Engl.* 57, 16313–16317
72. Szymkuć, S. *et al.* (2016) Computer-assisted synthetic planning: the end of the beginning. *Angew. Chem. Int. Ed. Engl.* 55, 5904–5937
73. Klucznik, T. *et al.* (2018) Efficient syntheses of diverse, medicinally relevant targets planned by computer and executed in the laboratory. *Chem* 4, 522–532
74. Segler, M.H.S. *et al.* (2018) Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* 555, 604–610
75. Coley, C.W. *et al.* (2017) Computer-assisted retrosynthesis based on molecular similarity. *ACS Cent. Sci.* 3, 1237–1245
76. Marth, C.J. *et al.* (2015) Network-analysis-guided synthesis of weisaconitine D and liljestrandinine. *Nature* 528, 493–498
77. Kim, D.E. *et al.* (2019) Total synthesis of paspaline A and emindole PB enabled by computational augmentation of a transform-guided retrosynthetic strategy. *J. Am. Chem. Soc.* 141, 1479–1483