Sigma-Aldrich®
Lab & Production Materials

Merck KGaA
Frankfurter Strasse 250
64293 Darmstadt, Germany

SigmaAldrich.com

Lab Manager®

MERCK

BIG PICTURE

Reducing Risk and Improving Synthesis Outcomes in Drug Discovery with SYNTHIA™ Retrosynthesis Software

Read it here
www.labmanager.com/big-picture/reducing-risk-in-drug-development

Pioneer Your Pathway at
SynthiaOnline.com

© 2023 Merck KGaA, Darmstadt, Germany and/or its affiliates. All Rights Reserved. Merck, the vibrant M, Sigma-Aldrich and SYNTHIA are trademarks of Merck KGaA, Darmstadt, Germany or its affiliates. All other trademarks are the property of their respective owners. Detailed information on trademarks is available via publicly accessible resources.

MK_FL12659EN Ver. 1.0 08/2023

The Life Science business of Merck operates as MilliporeSigma in the U.S. and Canada.

Sigma-Aldrich®
Lab & Production Materials

Lab Manager®

MERCK

Table of Contents

Today's synthesis planning for drug discovery relies on connecting the wisdom and practical expertise of an experienced organic chemist with the automation and advanced algorithms found in modern retrosynthesis software. This five-part series, sponsored by Merck, will enable you to pioneer a pathway to more efficient, sustainable drug discovery. The articles featured here emphasize the importance of retrosynthesis software to help mitigate risk and augment your lab staff's expertise, offer tips to implement ‚greener' chemistry, outline the steps to scale-up a synthesis reaction, and more.

Sigma-Aldrich®
Lab & Production Materials

Overcoming Key Challenges in Drug Discovery

HOW TO REDUCE RISK, SUPPORT SUSTAINABILITY, AND ENABLE INGENUITY WITH SYNTHIA™ RETROSYNTHESIS SOFTWARE **by Merck**

Computer-aided Synthetic Design (CASD) may have been born out of the work of E.J. Corey et al. in the 1960s, but only in the twenty-first century has it begun to transform chemical synthesis. SYNTHIA™ Retrosynthesis Software is leading this effort by combining powerful artificial intelligence (AI) with expert-coded chemical synthesis rules.

Drug discovery is a risky endeavor

It can take up to 12 years or more to bring a new pharmaceutical to market with costs that can exceed $2.6 billion. Developing pharmaceutical compounds is a lengthy, iterative process that requires sifting through thousands of molecular targets to find the right candidate. This arduous process requires countless experiments across multiple research and development departments.

Researchers usually begin by screening libraries that contain tens of thousands of compounds. With the assistance of high-throughput assays and computer software, the team may uncover a small number of initial hits. Medicinal chemists then begin modifying the core chemical structure to alter both physical & chemical properties, adding or subtracting molecular components to tease out effects that may improve drug characteristics.

A battery of tests is used to determine the suitability of each candidate as a bio-active pharmaceutical. Only the compound with the best therapeutic efficacy and safety profile will make it to market. During this process, the feasibility of large-scale synthesis and, ultimately, manufacturing of the compound is also explored. Only then is the drug ready for testing in pre-clinical studies before ultimately being introduced into clinical trials.

Unfortunately, only a few compounds reach clinical trials.

Is drug discovery sustainable?

Central to drug discovery is small molecule chemical synthesis, involving a complex, step-by-step process whereby teams of chemists produce purpose-built novel molecules. This remains to be a significant bottleneck for advancing new medicines to clinical trials despite over a century of advancements in synthetic chemistry technology.

Historically, synthetic chemistry has relied on the knowledge and expertise of highly trained synthetic organic or medicinal chemists, who were responsible for designing both novel molecules and the complex routes to synthesize them without advanced decision support tools. Recent advances in AI may now improve outcomes and accelerate drug discovery projects. With CASD, the modern, sophisticated researcher can bring more targets to market more sustainably.

However, it's no longer enough for a chemical synthesis to achieve sustainability only from an economic standpoint. Greener, safer, and more eco-conscious chemistry has been a topic of global importance for the last 50 years. Yet, only in the last 25 years have chemists had access to a clear set of guidelines, known as the 12 Principles of Green Chemistry, developed by Anastas & Warner. Still, there have been few tools available to the research chemist that provide decision support to enable achieving these noble initiatives.

Merck's head of Cheminformatics, Ken Karapetyan, PhD, weighs in on the role that computer-aided synthetic design has in corporate sustainability:

"With society growing more aware of corporate sustainability, corporations are reevaluating their social responsibility. This is driving organizations to switch the consumption of electricity, for example, to renewable alternatives, or arranging smart sustainable offices. Many chemical research organizations introduce sustainability in procuring building blocks to address poor atom economy. This is common in pharmaceutical research, where the goal is to synthesize a wide range of diverse compounds quickly and reliably. SYNTHIA™ can greatly help in reevaluating an organization's synthetic chemistry footprint (e.g. manufacturing and drug design) by introducing atom efficiency for scoring the alternative pathways, thus making chemical manufacturing more efficient."

Retrosynthesis is a creative process that is hard to engineer

Retrosynthetic analysis has been taught as a bond-disconnection strategy in advanced organic chemistry courses for nearly 50 years. For many chemists, it is a creative problem-solving technique that is a "badge of honor" when an imagined synthetic route is tested and validated. This process relied only on the chemist's personal knowledge and the simple tools of a pencil and paper. However, the demands of modern drug discovery have pushed the research chemist to harness the power of cheminformatics to positively augment their chemical intuition and expand the realm of possible disconnections using software.

Today, researchers find they can super-charge their brainstorming process and spark more creative and imaginative solutions using computer-aided synthetic planning. Whereas the "old way" of synthetic planning was limited by the individual background expertise within a group, the "new way" empowers individual researchers to be more creative while still allowing for group collaboration in a virtual space.

The flexibility of a retrosynthesis tool like SYNTHIA™ appeals to both the creative and practical aspects of synthesis design. This duality of features is what makes the power of computer-aided synthetic planning so appealing and beneficial to a broad range of chemical researchers.

"We see more and more scientists desire to use CASD to get insights on possible alternative retrosynthetic pathways," explains Karapetyan. "Those scientists are mostly within the R&D drug discovery organization as synthetic chemists and medicinal chemists. Process chemists, who are focused on optimizing and scaling up the existing synthetic routes, have also found SYNTHIA™ useful to provide insights on alternative retrosynthetic pathways."

Coded by chemists and computer scientists for more than 15 years, SYNTHIA™ Retrosynthesis Software is powered by sophisticated algorithms that can help experts easily access decades of chemical research data.

SYNTHIA™ Retrosynthesis Software reduces risk, supports sustainability, and enables ingenuity

Computer-aided synthesis planning has become more than an academic curiosity. Implementing this technology into drug discovery workflows is accelerating development of therapeutic targets and clearly increasing in usage. Since 1995, more than 1,000 scientific publications include "computer-aided drug design" in the title or abstract and of those, more than 40 percent were from the last five years.

According to Karapetyan, "There are a growing number of computational drug discovery platforms where vendors create an end-to-end experience to produce virtual drug candidates for specific targets. SYNTHIATM then can integrate with those platforms via API and do a tremendous job of filtering out virtual compounds that are not easily synthesizable."

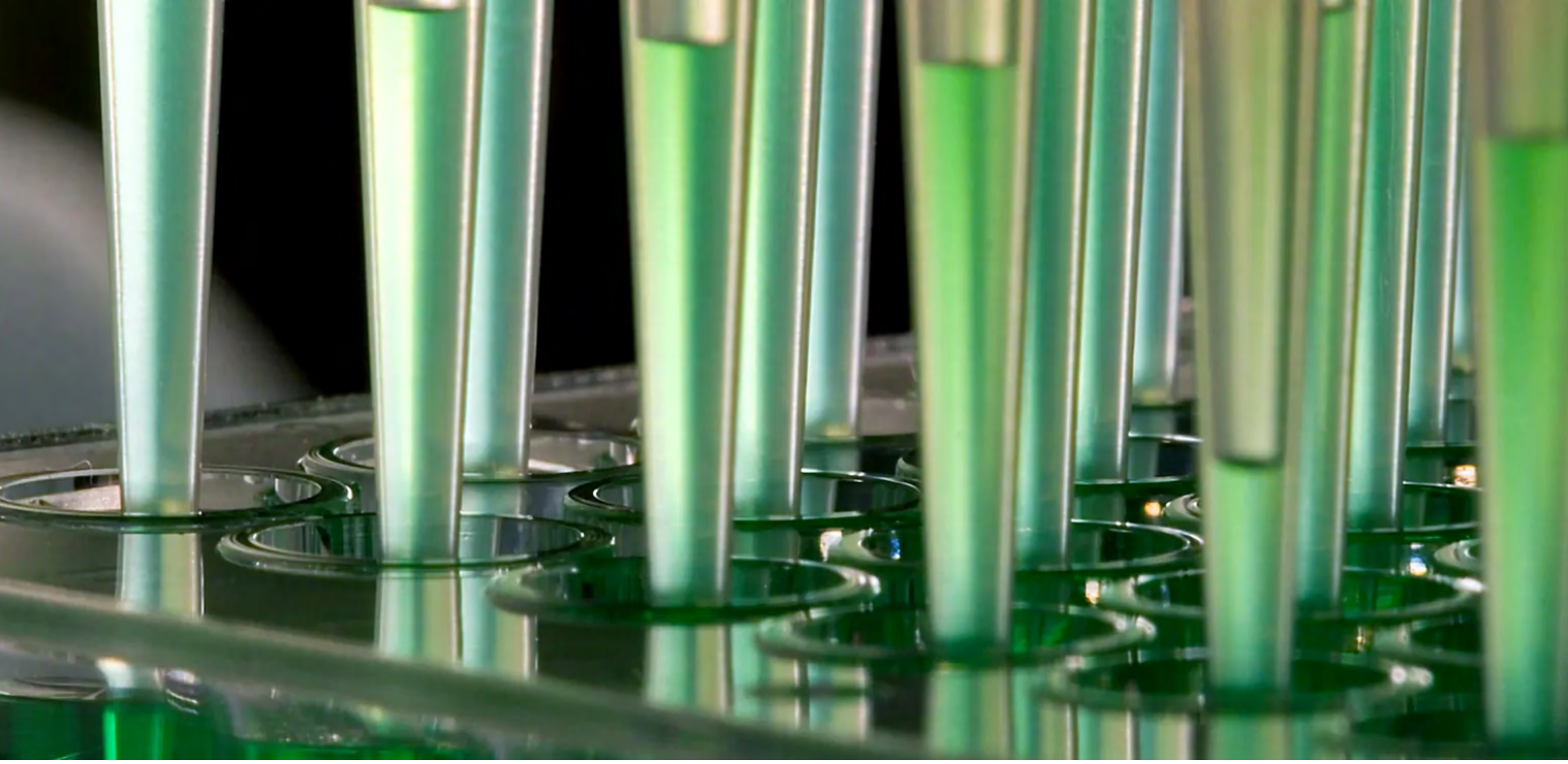
SYNTHIA™ provides valuable insights by leveraging advanced algorithms combined with more than 100,000 expert-coded organic chemistry reaction rules.

"It is very hard to keep all relevant synthesis knowledge operational in your mind and then be able to apply it on-

demand. By using SYNTHIA™, researchers can easily gain insights on alternative routes. However, we do believe that chemists must know the alternatives before designing a synthesis as it lowers the risk of missing an important, cost-effective pathway that could have a big cost implication for the organization," says Karapetyan.

Most small molecule drugs are the end-product of painstaking laboratory work —the result of an intricate process of hypothesis, design, and synthesis of individual molecules and testing their effects in biological systems. SYNTHIA™ can significantly speed up this process, accelerating drug discovery so that people can benefit from new life-changing medicines sooner.

To contact SYNTHIA™ Support, find out more, or request a demo please click HERE.



Using High-Throughput Screening to Rapidly Identify Targets

HIGH-THROUGHPUT SCREENING HAS BECOME AN IMPORTANT TOOL IN DRUG DISCOVERY **by John F. Conway, Graeme Dennis, Amandeep Ratte**

High-throughput screening (HTS) is commonly used in the drug discovery process to identify candidates or "hits" that have the desired effect on the target. These hits are then used as a starting point for a more rigorous drug discovery pipeline[17]. Historically, lead generation for drug discovery was based on empirical observations rooted in theoretical and physiological models. These drug candidates would then be painstakingly evaluated in low-throughput experimental regimes. Over time, advances in recombinant protein technology, assays, automation platforms, robotics, and fluorescence chemistry have converged to enable the screening of 100,000+ compounds per day against a single target.[4]

Steps in the HTS process

Once the therapeutic target has been identified, the HTS process can be broken down into the following phases: assay preparation, pilot screen, primary screen, secondary screen, and lead selection[8,17,18]. In the assay preparation phase the samples are prepared and an assay size, type, and detection strategy are selected. Then, a pilot screen is run using a relatively small sample size to validate the robotic, chemical, statistical, and data storage aspects of the process. The primary screen is run against the selected compound library to generate hits. Secondary screening confirms the accuracy of these hits and helps to reliably transform them into leads.

The primary screen is designed to quickly identify hits from expansive compound libraries. For this reason, it is important that test results can be quickly read via a sensor. In practice, detection methods generally rely on fluorescence measurement[11,18]. Results are presented as a measure of activity relative to positive and negative controls[11]. The quality of screen results can be assessed by evaluating the signal to noise ratio and signal variation within the assay.

Data organization

Due to the widespread adoption and continued evolution of HTS technologies over the past few decades, there has been a rapid increase in the volume of HTS trial data available. Compound libraries have also increased in volume exponentially[5]. To be able to leverage this data, an

IT and informatics infrastructure that complies with FAIR (findable, accessible, interoperable, reusable) data principles is essential[14]. A variety of software tools to capture, process, analyze, visualize, store, compare, and optimize HTS data exist. Platform-esque tools that offer users the ability to integrate these tools and systems into a common interface also exist[11]. Having a unified data environment is vital to running an effective HTS drug discovery program as it prevents the formation of data silos, tribal knowledge, manual data entry, and manual processing that put data integrity and reproducibility at risk.

AI/ML innovations

Widespread adoption of FAIR principles and developments in data standardization, organization, and labeling have expanded the usable size of datasets for analysis. The availability of these large, well-labeled datasets has enabled the evolution of complementary computational techniques. Artificial Intelligence (AI) and machine learning (ML) applications in HTS processes can consume existing data libraries to train predictive models[12,13,3]. Virtual high-throughput screening (vHTS) is an entirely in silico process that integrates AI computational techniques and existing in silico compound libraries to identify "hits"[15]. When used to complement HTS, vHTS has been shown to increase lead quality[11]. There are ongoing industry-wide efforts to improve vHTS generalizability to the point where no prior data from an existing screen needs to exist before making accurate predictions[1,2,3].

AI-driven iterative screening is another methodology that integrates AI into traditional HTS processes[2]. In traditional HTS, the hit rate is usually around one percent per assay[2]. It follows that using larger and larger compound libraries will generate a greater number of leads. Indeed, this belief dominated the early years of widespread HTS adoption in the pharmaceutical industry. However, larger libraries and increasingly complex screening technologies and chemistry lead to greater costs and lead times[2]. In recent times, information driven screening strategies are favored to balance costs, time, and hit rate. In iterative screening, results from a small screen are input into a machine learning model, which then selects the compounds to use for the next screen.[2] This process can be repeated as many times as desired to further train the model and refine hit rate in the subsequent iteration. Such an approach requires a greater degree of assay automation than traditional HTS since assays must be assembled on demand based on model outputs[18]. Hit rates from iterative screening processes can be in excess of 35 percent while using a smaller portion of the compound library[2].

Availability

As HTS continues to grow in importance within the drug discovery pipeline, many initiatives to reduce costs and improve accessibility within private, public, and academic spaces have been fruitful. "Cloud labs" are facilities that make HTS and other advanced techniques available in pay-per-use models[11]. Public-private partnerships make huge data sets available to researchers worldwide. There are also state-of-the-art HTS screening facilities operated by public-private partnerships (i.e. European Lead Factory) that make high-quality screening available to non-profit entities[10,11]. The advances in cost reduction, availability, and collaboration have increased the availability of HTS; which in turn has accelerated drug discovery in general. Advances in data management, AI, and ML will enable researchers to improve decision making and leverage ever growing datasets[14]. Multidisciplinary advances in automation, computing, data contextualization, and modeling will drive both the capability and importance of HTS in the years to come.

References:

1. David, L. et al. Applications of deep-learning in exploiting large-scale and heterogeneous compound data in industrial pharmaceutical research. Front. Pharmacol. 10, 1303 (2019).
2. Dreiman, G. H. S., Bictash, M., Fish, P. V., Griffin, L. & Svensson, F. Changing the HTS Paradigm: AI-Driven Iterative Screening for Hit Finding. SLAS Discov. Adv. life Sci. R D 26, 257–262 (2021).
3. Ekins, S. et al. Data Mining and Computational Modeling of High Throughput Screening Datasets. Methods Mol. Biol. 1755, 197 (2018).
4. MacArron, R. et al. Impact of high-throughput screening in biomedical research. Nat. Rev. Drug Discov. 2011 103 10, 188–195 (2011).
5. Mayr, L. M. & Bojanic, D. Novel trends in high-throughput screening. Curr. Opin. Pharmacol. 9, 580–588 (2009).
6. Buterez, D., Janet, J. P., Kiddle, S. & Liò, P. Multi-fidelity machine learning models for improved high-throughput screening predictions. (2022) doi:10.26434/CHEMRXIV-2022-DSBM5.
7. Roy, A., R. McDonald, P., Sittampalam, S. & Chaguturu, R. Open access high throughput drug discovery in the public domain: a Mount Everest in the making. Curr. Pharm. Biotechnol. 11, 764–778 (2010).
8. High-throughput Screening Steps · Small Molecule Discovery Center (SMDC) · UCSF. UCSF https://pharm.ucsf.edu/smdc/tech-services/hts-steps.

9. Pusterla, T. High-throughput screening (HTS) | BMG LABTECH. BMG Labtech https://www.bmglabtech.com/en/blog/high-throughput-screening/ (2019).
10. Helping to create new medicines | European Lead Factory. https://www.europeanleadfactory.eu/.
11. TRENDS IN HIGH THROUGHPUT SCREENING. 20/15 Visioneers https://www.20visioneers15.com/post/throughput-screening.
12. Advancing Drug Discovery with Artificial Intelligence. Kantify https://kantify.com/use-cases/drug-discovery.
13. Artificial Intelligence (AI) in drug discovery. Kantify https://www.kantify.com/insights/artificial-intelligence-ai-in-drug-discovery.
14. Hanton, S. D. High Throughput Experimentation Drives Better Outcomes | Big Picture | Lab Manager. Lab Manager https://www.labmanager.com/big-picture/lab-automation-benefits/high-throughput-experimentation-drives-better-outcomes-24503 (2020).
15. Virtual Screening - Creative Biolabs. Creative Biolabs https://www.creative-biolabs.com/drug-discovery/therapeutics/virtual-screening.htm.
16. Szyma?ski, P., Markowicz, M. & Mikiciuk-Olasik, E. Adaptation of High-Throughput Screening in Drug Discovery—Toxicological Screening Tests. Int. J. Mol. Sci. 13, 427 (2012).
17. Shukla, A. A. High Throughput Screening of Small Molecule Library: Procedure, Challenges and Future. J. Cancer Prev. Curr. Res. 5, (2016).
18. Bokhari, F. F. & Albukhari, A. Design and Implementation of High Throughput Screening Assays for Drug Discoveries. High-Throughput Screen. Drug Discov. (2021) doi:10.5772/INTECHOPEN.98733.
19. Brubacher, M. G. High-throughput Technologies in Drug Discovery | Technology Networks. Technology Networks: Drug Discovery https://www.technologynetworks.com/drug-discovery/articles/high-throughput-technologies-in-drug-discovery-330193 (2020).
20. Subramaniam, S., Mehrotra, M. & Gupta, D. Virtual high throughput screening (vHTS) - A perspective. Bioinformation 3, 14 (2008).
21. Wang, Y., Cheng, T. & Bryant, S. H. PubChem BioAssay: A Decade's Development toward Open High-Throughput Screening Data Sharing. SLAS Discov. 22, 655–666 (2017).

John F. Conway *has 30 years of experience in R&D on all sides of the fence; industry, software and services, and consulting. Most recently, he was global head of R&D IT at AstraZeneca.*

Graeme Dennis *has held various roles in biopharma informatics since 2004, and is currently a consultant with 20/15 Visioneers. He studied chemistry at Vanderbilt University and lives outside Nashville, Tennessee.*

Amandeep Ratte *has a background in chemical engineering, data science, and agricultural technology. He has architected and designed systems to improve automation, accessibility, and reliability in indoor farming. He is currently the head of data science at a leading ag-tech company.*

Three Ways to Improve the Sustainability of Your Lab's Chemical Synthesis

SIMPLE STEPS TO IMPLEMENT GREEN CHEMISTRY PRACTICES

by Jonathon Moir, Phd, Natalie O'Neil, Phd

Laboratories are some of the most energy intensive and wasteful spaces on university campuses and industrial sites alike[1]. In the chemical sciences, reactions are often performed with maximum yield (and/or profit) in mind, and while a lot of labs have moved toward becoming more sustainable through efforts in bringing in energy efficient equipment or recycling initiatives, many labs still perform chemistry in a wasteful way and use reagents and solvents that are hazardous to human and environmental health.

Green chemistry offers another tool in the toolbox of scientists and lab managers to reduce their environmental footprint, decrease waste, decrease exposure to hazardous chemicals, improve efficiency, and more. Green chemistry is defined by Warner and Anastas as, "the utilization of a set of principles that reduces or eliminates the use or generation of hazardous substances in the design, manufacture, and application of chemical products." It is not a field in and of itself but rather a holistic approach to practicing chemistry.

The 12 Guiding Principles of Green Chemistry (Figure 1) provide the framework through which scientists can design chemical reactions and products in a more sustainable way.

Chemists may often ask themselves, "This is all well and good, but where do I start? How do I know the best approach to take?" While it can sometimes feel overwhelming when first starting on a "greener" journey, here are some simple tips and examples to help get started.

Tip #1 – Prevent waste wherever possible

There are many simple ways to prevent waste in labs, which can ultimately mitigate hazards and exposures, too. In education and teaching labs, scaling down is one of the "lowest hanging fruit" when it comes to simple changes you can make. However, this is often not possible in industrial settings (where scale-up is usually one of the key goals). In those cases, consider solvent recycling. Many solvent recycling systems exist that can allow solvents to

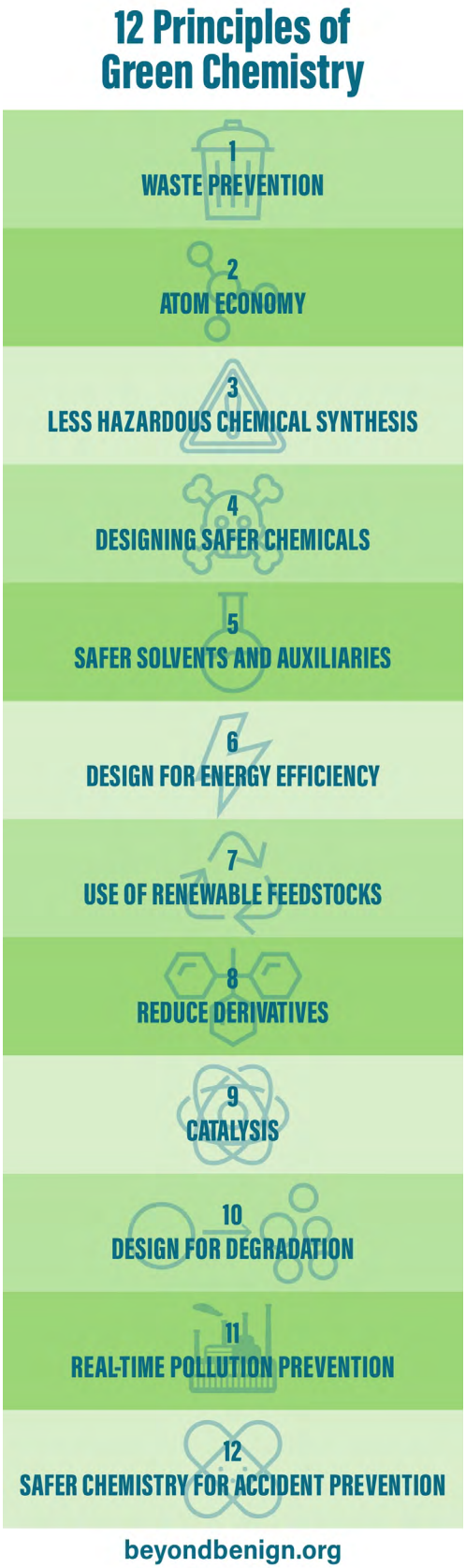


Figure 1: BEYOND BENIGN

be reused for future reactions. One-pot reactions can also help by avoiding additional workup steps between subsequent reactions including separations (columns, filtrations, liquid-liquid extractions, etc.). Of course, this is all dependent on the circumstances and reaction at hand; if the one-pot synthesis prevents kilograms of additional waste but requires the use of benzene as the solvent, a different reaction route should be considered. This is often the task of chemists looking at greener syntheses: what are the trade-offs and how can undesirable consequences be eliminated or minimized.

Tip #2 – Use safer, alternative solvents

Over the past 20 years, a number of major pharmaceutical and industrial chemical companies have released solvent selection guides showing desirable alternatives to traditional solvents[2-6]. Beyond Benign, a US-based not-for-profit dedicated to transforming chemistry education for a sustainable future, has aggregated many of these guides into a simple infographic (Figure 2). Wherever possible, alternative, safer solvents should be used in place of high hazard solvents. In general, halogenated and aromatic/short-chain aliphatic hydrocarbon solvents (such as dichloromethane (DCM), dichloroethane, chloroform, benzene, toluene, hexanes, and pentane) as well as nitrogen-containing solvents (such as pyridine, N-methyl-2-pyrrolidone (NMP), propionitrile, dimethylacetamide, and dimethylformamide (DMF)) should be avoided and replaced due to their hazardous nature.

Making these substitutions can range from being simple to more complicated. For example, if the reaction is to be done at reflux, different boiling points will impact the refluxing temperature and therefore may impact the yield or time needed for a reaction. In some cases, the solvent may participate in the reaction (such as amine-based solvents acting as bases). Again, trade-offs must be assessed carefully by the attending chemist to understand if there is greater benefit, and some testing may be required, but the rewards can be substantial. A wonderful example case study in industry is the early synthesis of Viagra, which saw a reduction in waste solvent from 1300 L per kg active pharmaceutical ingredient (API) in 1990 down to 7 L/kg API in the late 1990s[7]. The team was able to not only eliminate the use of chlorinated solvents in the process but also reduce the use of highly volatile solvents like methanol and diethyl ether.

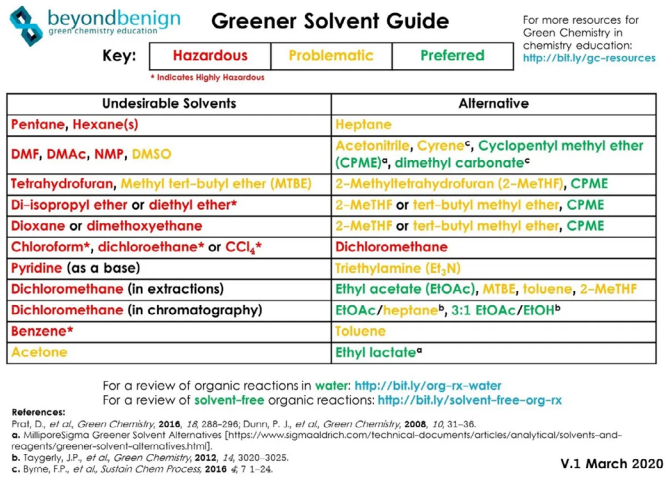
Tip #3 – Measure your impact and get a baseline

Scientists love metrics. They feed off quantifiable data. And who wouldn't? Having the numbers to rationalize a result, understand deeper meaning behind an observation, or argue for more funding are the "bread and butter" of a scientist's daily work. Green chemistry is no exception, but there are many different metrics and tools for measuring "greenness" and choosing the right one can

sometimes be daunting. There is no single metric that can evaluate all aspects of the "greenness" of a chemical product or process. However, a few quick descriptions of some of the more common metrics are available below to help give a sense of when one metric might be more appropriate than another. In fact, they are often best used in combination.

- Yield – The quintessential measure of chemical success, yield only considers the amount of product formed but does not consider waste or other factors in the reaction.
- Atom Economy – The percent mass of atoms from the reagents incorporated into a product. An example of a reaction with a high atom economy is a Diels-Alder reaction, while an example of a reaction with low atom economy is the production of hydrogen from benzenethiol and lithium aluminum hydride. Under the right conditions, both reactions can progress readily and produce high yields, but the second reaction incorporates only a small mass fraction of the atoms from the starting materials into the product.
- Process Mass Intensity (PMI) – The ratio of total mass of all components used in a reaction to the mass of isolated product. The PMI can never be lower than one, whereas the E-factor (which is similar) can theoretically be as low as zero.
- E-Factor – The ratio of total waste mass (anything that is not incorporated into the product) to the mass of isolated product. An ideal E-factor is zero, where no waste is produced.
- Life Cycle Analysis (LCA) – An LCA is a much more complex analysis of a reaction product or process and must take into account metrics across a wide range of impact categories (the selection of which can vary greatly depending on the circumstance). Some examples of categories include global warming potential, acidification, terrestrial eutrophication, human toxicity, ionizing radiation, and more.
- DOZN – A tool developed by Merck, this is a quantitative online green chemistry evaluator that can assess the relative greenness of chemicals and chemical processes across all twelve principles of green chemistry.

This is just a sample and there are many more metrics available. Ultimately, using even just a few of these metrics to begin to understand how green your reaction is can lead to small changes and improvements that, over time, produce significant impact and long-term lab sustainability.



beyondbenign green chemistry education

Greener Solvent Guide

For more resources for Green Chemistry in chemistry education: http://bit.ly/gc-resources

Key: Hazardous | Problematic | Preferred

* Indicates Highly Hazardous

Undesirable Solvents	Alternative
Pentane, Hexane(s)	Heptane
DMF, DMAc, NMP, DMSO	Acetonitrile, Cyrene[c], Cyclopentyl methyl ether (CPME)[a], dimethyl carbonate[c]
Tetrahydrofuran, Methyl tert-butyl ether (MTBE)	2-Methyltetrahydrofuran (2-MeTHF), CPME
Di-isopropyl ether or diethyl ether*	2-MeTHF or tert-butyl methyl ether, CPME
Dioxane or dimethoxyethane	2-MeTHF or tert-butyl methyl ether, CPME
Chloroform*, dichloroethane* or CCl_4*	Dichloromethane
Pyridine (as a base)	Triethylamine (Et_3N)
Dichloromethane (in extractions)	Ethyl acetate (EtOAc), MTBE, toluene, 2-MeTHF
Dichloromethane (in chromatography)	EtOAc/heptane[b], 3:1 EtOAc/EtOH[b]
Benzene*	Toluene
Acetone	Ethyl lactate[a]

For a review of organic reactions in water: http://bit.ly/org-rx-water
For a review of solvent-free organic reactions: http://bit.ly/solvent-free-org-rx

References:

Prat, D., *et al*, *Green Chemistry*, **2016**, *18*, 288–296; Dunn, P. J., *et al*, *Green Chemistry*, **2008**, *10*, 31–36.

a. MilliporeSigma Greener Solvent Alternatives [https://www.sigmaaldrich.com/technical-documents/articles/analytical/solvents-and-reagents/greener-solvent-alternatives.html].

b. Taygerly, J.P., *et al*, *Green Chemistry*, **2012**, *14*, 3020–3025.

c. Byrne, F.P., *et al*, *Sustain Chem Process*, **2016** *4*, 7 1–24.

V.1 March 2020

Figure 2: BEYOND BENIGN

References

1. Jain, N. Integrating Sustainability into Scientific Research. Nat. Rev. Methods Primer 2022, 2 (1), 35. https://doi.org/10.1038/s43586-022-00126-6.
2. Byrne, F. P.; Jin, S.; Paggiola, G.; Petchey, T. H. M.; Clark, J. H.; Farmer, T. J.; Hunt, A. J.; Robert McElroy, C.; Sherwood, J. Tools and Techniques for Solvent Selection: Green Solvent Selection Guides. Sustain. Chem. Process. 2016, 4 (1), 7. https://doi.org/10.1186/s40508-016-0051-z.
3. Prat, D.; Pardigon, O.; Flemming, H.-W.; Letestu, S.; Ducandas, V.; Isnard, P.; Guntrum, E.; Senac, T.; Ruisseau, S.; Cruciani, P.; Hosek, P. Sanofi's Solvent Selection Guide: A Step Toward More Sustainable Processes. Org. Process Res. Dev. 2013, 17 (12), 1517–1525. https://doi.org/10.1021/op4002565.
4. Henderson, R. K.; Jiménez-González, C.; Constable, D. J. C.; Alston, S. R.; Inglis, G. G. a.; Fisher, G.; Sherwood, J.; Binks, S. P.; Curzons, A. D. Expanding GSK's Solvent Selection Guide – Embedding Sustainability into Solvent Selection Starting at Medicinal Chemistry. Green Chem. 2011, 13 (4), 854–862. https://doi.org/10.1039/c0gc00918k.
5. Adams, J. P.; Alder, C. M.; Andrews, I.; Bullion, A. M.; Campbell-Crawford, M.; Darcy, M. G.; Hayler, J. D.; Henderson, R. K.; Oare, C. A.; Pendrak, I.; Redman, A. M.; Shuster, L. E.; Sneddon, H. F.; Walker, M. D. Development of GSK's Reagent Guides – Embedding Sustainability into Reagent Selection. Green Chem. 2013, 15 (6), 1542. https://doi.org/10.1039/c3gc40225h.
6. Taygerly, J. P.; Miller, L. M.; Yee, A.; Peterson, E. A. A Convenient Guide to Help Select Replacement Solvents for Dichloromethane in Chromatography. Green Chem. 2012, 14 (11), 3020. https://doi.org/10.1039/c2gc36064k.
7. Dunn, P. J.; Galvin, S.; Hettenbach, K. The Development of an Environmentally Benign Synthesis of Sildenafil Citrate

(ViagraTM) and Its Assessment by Green Chemistry Metrics. Green Chem 2004, 6 (1), 43–48. https://doi.org/10.1039/B312329D.

Jonathon Moir *is program manager at Beyond Benign for the upcoming Green Chemistry Teaching and Learning Community (GCTLC)—an online collaboration and networking platform set to launch in 2023 that will include green chemistry education materials, collaboration spaces, events, job postings, and more. Jonathon received his B.Sc.H. from Queen's University in 2010 and his PhD from the University of Toronto in 2016 in Inorganic Chemistry with a focus on nanomaterials and electrochemistry. Following graduate school, Jonathon transitioned to the not-for-profit sector, where he has helped manage and support the development of numerous multimillion-dollar international, interdisciplinary research programs, scientific projects, workshops, symposia and national scholarship programs across Canada and the world.*

Natalie O'Neil *is director of higher education at Beyond Benign. Natalie earned her PhD in chemistry from the University at Albany in 2017. During her graduate studies, she felt that the topics of sustainability, toxicology, and environmental hazards were missing from the traditional graduate chemistry curriculum. Therefore, she pursued a one-year certification in Green Chemistry and Chemical Stewardship and attended the American Chemical Society (ACS) Green Chemistry & Sustainable Energy Summer School. She now uses her teaching experience and her passion to empower the next generation of scientists to use sustainable approaches through green chemistry through the Beyond Benign Green Chemistry Commitment (GCC) program.*

How to Scale Up a New Synthesis Reaction

A SUCCESSFUL SCALE-UP REQUIRES INPUT FROM A VARIETY OF TEAM MEMBERS TO PRODUCE A SAFE AND EFFECTIVE PRODUCT **by Michael J. Williams**

The first step to successfully scaling up from bench research to production requires building the right team. A wide variety of skills is needed, including chemists, chemical engineers, analytical chemists, environmental health and safety (EHS) experts, supply chain managers, and even sales and marketing specialists. The team will set the tone for the entire scale-up process.

As products are introduced into the production environment, the questions that need to be answers have evolved. The supply chain team is no longer simply tasked with finding the raw materials on the market. They also consider the distance from the supplier to the production site to reduce the carbon footprint. Raw materials can be obtained from both a bio-renewable source and a synthetic source and have the same material specifications. However, the materials may not perform identically due to "unspecified specifications." Seasonal fluctuations of bio-renewable materials may affect supply chain distribution and impact the production process.

Steps for the chemist

The chemist's role starts with understanding the basic chemistry and working with the product development team or application team to know how the product must perform in a given application. The chemist needs sufficient experience with the reaction in the lab to know what caused an upset in the plant, and most importantly, how to fix the problem. This means making material that is at the edge of and beyond the product's specifications. As part of the runs, the chemist needs to learn what influences the product and test the robustness of the synthesis. For example, if a reaction requires a temperature of 150C, what happens if the plant's control system actually maintains a temperature range of 145C to 155C? Will the product have the desired conversion? When running through these trial lab batches, the chemist must consider all of the product specifications, which are reflected in the product's Certificate of Analysis.

To have a successful scale up, a product needs to have clearly defined analytical targets and the appropriate ranges for those targets. As a product is scaled up to the plant, it needs to be analyzed on equipment the plant has, using the methods that the plant utilizes. The specifications that

the team sets must fall within the accuracy limits of the analytical tests being used. The team needs to ensure that the analytical methods exist in the production environment before scale-up.

After the chemist has thoroughly vetted the process, they will develop the bill of materials. Unfortunately, this typically is not just scaling up the reaction quantities because plants prefer to consume whole containers of materials. When developing batch sizes, the chemist must consider what material is hardest to handle in the process. The batches are then scaled around consuming whole containers of this material. If the material is a solid, is it available as a pellet or hot melt? If it is a hot melt, does the plant have the right equipment to charge it? If it is only available as a solid and the plant cannot handle the material, then this may mean the process has to start again and the chemist must develop a new process that can work within the plant's limitations.

Factors that influence a scale-up

When a product moves from benchtop scale to production, it is not as simple as multiplying the reactant quantities by 200x. Most lab scale reactions are sized so that their reaction vessels can quickly dissipate any heat that is evolved during the reaction. Heat-up and cool-down cycles are much longer in production scale vessels. This additional reaction time may impact a product's properties. If the reaction is performed in glassware, it is often vented to atmosphere or flows through a condenser, so that pressure is not a concern. If the lab scale work is done in pressure vessels, the vessels are usually rated for well beyond any expected pressure. For this reason, it is important to investigate and document the vapor pressure of a reaction mixture through the entire process. As the process experiences the complete temperature cycle, the viscosity of the material may change. While it is hard to measure the viscosity during the entire process, the chemist can measure the viscosity of the starting materials and final products. The chemical engineers need to know the viscosity profile over the course of the reaction.

For example, a raw material with the viscosity of water may work well in the reactor's pump at room temperature; however, as the temperature rises, the viscosity may be too low to properly lubricate the pump seals on the reactors. If the viscosity of the product rises over the course of the reaction, the product may flow without issue through the pump at elevated temperatures but become too viscous for the pump to adequately circulate the material as it cools.

Safety considerations

When a product is developed in the lab, all the work is typically done inside a fume hood. In a production environment, the ventilation is usually natural ventilation; the wind and air moving around the vessels. As the process is prepared for scale up, the EHS group will look at the materials and processes. It is important to identify if the product and all of its isolated intermediates are on the US Toxic Substances Control Act (TSCA). If the material is on TSCA, but has never been produced at production scale before, has the preliminary manufacturing notification (PMN) been completed? If the material has a PMN, does it have a significant new use rule (SNUR). Another concern is staying within the production facilities' air permits. Many common chemicals in use in the lab will exceed air permits if they are vented to the atmosphere during a reaction. Simply heating water has the potential to exceed a production vessel's pressure rating if the water is present in high enough concentration. Lower boiling solvents can exacerbate this effect even more.

Topics such as maximum allowable working pressure (MAWP) and vessel pressure ratings are just two of the many topics that need to be discussed when the safety team member leads the process hazards analysis (PHA) of a new product. The PHA team will take into account material compatibility, vessel pressure and temperature ratings, control systems, plant interlocks, relief device settings, and many other risk factors. The technical team of the product introduction team normally participates. The goal is to ensure that the new process fits within the safety capabilities of the existing plant equipment. If the product or process falls outside of the existing systems, recommendations are made to improve the equipment. Protocols differ among production facilities, but all of them undergo some type of 'management of change' process when new materials are introduced to the plant.

A successful first production run can take many shapes. It may simply be completing the first large scale run of a product in the plant or it may be making millions of pounds of product for a customer. No matter what the definition of success is, it is only possible when the organization works as a team to get it finished. This team includes many departments, such as chemists, chemical engineers, EHS, supply chain, and more. The chemist plays a significant role in kickstarting the project by defining targets and vetting the process. Once the process is under way, other departments can assists with evaluating all the factors and safety considerations that need to be prioritized to ensure a successful scale-up run.

Michael J. Williams *received his MSc in organic chemistry from the University of Oregon in 1995. Immediately after school he worked in the pharma industry doing process development and scale-up. In 1998, Mike became a production chemist at a small surfactant company called Tomah Products in Milton, WI. Since then, he has been a synthetic research chemist, group leader and senior group leader. After several corporate name changes, Mike still sits at the same desk, but he works for the Evonik Corporation where he now manages two synthetic and one analytical R&D groups. He is additionally responsible for scale-up and product transfer activities at five production sites in the Americas.*

Computer-Aided Synthesis Reduces Complexity and Accelerates Novel Chemical Discoveries

MODERN SYNTHESIS PLANNING SOFTWARE CAN BE BEST USED AS A DECISION SUPPORT TOOL TO REDUCE RISK OF FAILURE AND COST OF RESEARCH **by Merck**

We've come a long way from the earliest days of synthesis planning. Though ambitions of alchemical transmutation from base elements such as lead to precious metals such as gold have long since passed[1], the modern chemist still labors over a different kind of bubbling cauldron—one of cheminformatics. With hundreds of millions of known compounds and documented reactions, today's synthesis planning relies on connecting the wisdom and practical expertise of an experienced organic chemist with the automation and advanced algorithms found in modern retrosynthesis software.

Chemical Production—Extraction to synthesis

The discovery of dyes produced by labor-intensive extraction from natural sources, such as insects like Kermes Vermilio for Vermillian[2] or Tyrian Purple (6,6'-dibromoindigo) extracted from predatory rock snails[3], accelerated the demand for novel approaches to the production of synthetic dyes in the 1800s[4] In the early twentieth century, the development of complex, medicinal compounds by executing multi-step organic synthesis led to a formalized process of "reverse engineering" such targets, paving the way for modern synthesis planning, using what is now referred to as "retrosynthetic analysis[5]."

The benefits of synthetic planning were obvious, however, as with chemical extraction, organic synthesis is not without its own inherent risks. While a carefully planned synthesis centralized the production of a target at scale, a new set of challenges remained that threatened any successful operation. Scale-up was and still is easier planned than executed. A chemical reaction perceived to be "'clean"' at a small scale (generating undetectable amounts of side-products) could have tremendous challenges when scaled up[6]. Planned starting materials and reagents, that were later found to threaten the environment, may have needed to be replaced due to ever-changing regulatory constraints. Supply chain disruptions due to natural disasters, geo- political events, and other causes were, and remain, a significant threat to production[7].

Retrosynthetic analysis and mitigating risk

Addressing the many intrinsic and external challenges to chemical synthesis has led to a highly complex system of regulations that involve numerous academic and governmental organizations in modern times. While the simplicity and elegance of E.J. Corey's retrosynthetic analysis allows it to remain a pillar for the student of organic chemistry, it has been supplemented by powerful computer software with advanced algorithms to support the synthetic chemist in the last few decades. This new age of enhanced computer-aided retrosynthesis is still in its infancy, and like all new tools, has been both lauded and criticized. While the potential for these tools to further mitigate risks to industrial synthesis has already become clear, it would be an oversight to omit the impact on human capital, ingenuity, or resources when discussing risk in any economic endeavor—and nothing evokes this more strongly than tools claiming to have artificial intelligence.

Nevertheless, a more powerful calculator has yet to fully replace the mathematician. Like all tools, there is an art to being productive with each one—and art vis-a-vis creativity, is a domain that will forever belong to human intellect. Moreover, we must remain cognizant of the specific use-cases that have been successfully managed with automation and artificial intelligence. Many grudgingly mundane physical tasks have been replaced by machines, allowing human operators the freedom and creativity to direct their energy towards decision-making and other project-related goals.

Improving synthesis outcomes

It is with this in mind that we review several recent publications featuring the usage of SYNTHIA™ Retrosynthesis Software. Perhaps the most obvious benefit to computer-aided synthetic planning is in the efficient identification of a low-cost, yet robust synthetic pathway. Chemistry has long since drawn culinary metaphors, referring to a chemical synthesis procedure as a "'recipe."' In this manner, one might draw analogy of a computer-aided retrosynthesis plan as a catering guide for a successful 10-course meal. As demonstrated by Klucznik et al., such a "'buffet"' of synthetic targets could be readily produced in a chemical laboratory after using SYNTHIA™ Retrosynthesis Software (formerly Chematica) to devise efficient pathways[8].

"The Chematica program was used to autonomously design synthetic pathways to eight structurally diverse targets, including seven commercially valuable bioactive substances and one natural product. All of these computer-planned routes were successfully executed in the laboratory and offer significant yield improvements and cost savings over previous approaches, provide alternatives to patented routes, or produce targets that were not synthesized previously[8]."

The utility of a computer-aided synthesis planning tool is made crystal clear from the work of Klucznik et al., demonstrating more robust route planning with better yields and reduced costs that steers around patents and generates valuable intellectual property. What may be less clear is whether this success is "cherry-picked" or where limitations of scope exist in the tool. However, this is an easy criticism for any tool, the boundaries of which may only be found by the skillful creativity of the user. One such example of creativity was demonstrated by Gajewska et al. who explored the landscape of tactical combinations (TCs) in organic synthesis and discovered a computer-aided algorithm that revealed 4.85 million combinations within 46,000 reaction classes[9]. Gajewska asserts that although difficult to identify, and with only 500 TCs previously catalogued, tactical combinations are uniquely useful in the synthetic planning of complex organic targets, opening up possibilities for pronounced structural simplification in subsequent, downstream steps[9].

Though it seems such creative research projects may belong exclusively to the field of cheminformatics, the potential for identifying and capitalizing on new intellectual property is clear from such endeavors, surely extending to drug discovery, process chemistry, and beyond. Computer-aided synthesis planning promises to reduce complexity in this space and accelerate novel chemical discoveries. Molga et al. showed that navigating the whitespace for a synthetic target to avoid routes with patents pending was more easily accomplished with SYNTHIA™ Retrosynthesis Software[10].

All of these computer-planned routes were successfully executed in the laboratory and offer significant yield improvements and cost savings over previous approaches, provide alternatives to patented routes, or produce targets that were not synthesized previously.

"By keeping track of lists of specific bonds one wishes to preserve, a computer program is able to identify the key disconnections used in the patented syntheses and design synthetic routes that circumvent these approaches10."

The researcher notes the complexity in the syntheses of blockbuster drugs which can be protected by tens to hundreds of patents claiming hundreds of synthetic steps and altogether forming quite complex reaction networks.[10]

Moreover, they explored the state-of-the-art with all relevant and available software tools which allowed them to draw a stark comparison in efficiency and ease of use:

"Reactions from patented syntheses are easily available by querying repositories such as Reaxys or SciFinder; however, they come as individual steps rather than complete synthetic plans, and the atoms in them are not numbered. Therefore, our first step is to assign atom mappings to each of these reactions. We do so by using Chematica's SMARTS reaction templates and our in-house atom- mapping codes (commercial mappers in ChemDraw or Marvin can also be used) to match both substrates and products and also unambiguously assign the reaction type[10]."

A new age in computer aided synthesis planning

Navigating the murky waters of intellectual property is a complex problem. The savvy researcher may need to use numerous tools to find an optimal path, especially since many chemical patents far exceed 100 pages of complex language designed to protect what seemingly is all potential uses of a particular chemical entity. The complementary nature of similar tools is well understood by various trades. It is no wonder that research is being conducted to more precisely define the scope and limits of each synthesis planning tool.

If the ocean of intellectual property is murky, then supply chain disruptions are like icebergs. Obviously, there is tremendous value in tactical agility around such unfortunate and often unforeseen excursions. Therefore, it should seem crucial to build contingencies into a synthesis plan at the very beginning stages. Moreover, multi-route synthesis planning can dramatically benefit the development of quick- to-market drugs required to combat a future pandemic. Such an approach was published by Szymkuć et al. in early 2020 a few months after the worldwide outbreak of Sars-CoV-2, the virus that causes COVID-19. The team used SYNTHIA™ Retrosynthesis Software to identify multiple viable routes to synthesize two potential COVID-19 therapeutics[11].

"A computer program for retrosynthetic planning helps develop multiple 'synthetic contingency' plans for hydroxychloroquine and also routes leading to remdesivir, both promising but yet unproven medications against COVID-19. These plans are designed to navigate, as much as possible, around known and patented routes and to commence from inexpensive and diverse starting materials, so as to ensure supply in case of anticipated market shortages of commonly used substrates[11]."

The authors acknowledge that such a strategy is merely one example of risk reduction to loss-of-life from a global health crisis. They further note that the "development of similar contingency syntheses is advocated for other already-approved medications, in case such medications become urgently needed in mass quantities to face other public- health emergencies[11]." Climate scientists largely agree that the frequency of pandemics is likely to increase as microorganisms have the ability to bio-diversify faster than other organisms and climate stressors such as food scarcity will affect population dynamics creating more occurrences for interspecies interactions and the transfer of infectious disease[12].

Like all technologies, computer-aided synthesis planning continues to evolve. Today's modern synthesis planning software combines human ingenuity with artificial intelligence to quickly propose multi-step synthetic pathways. However, the proposed routes still require careful assessment and bench execution by a skilled synthetic organic chemist. It is at the discretion of the end user to determine which route is the best fit for the purpose at hand. Therefore, the state of the art is best considered a decision support tool that can reduce risk of failure in the lab and plant, thereby reducing the cost of research and increasing profitability for commercial organizations. The game hasn't yet changed, but this new age of synthesis planning feels more revolutionary than evolutionary.

Please visit SigmaAldrich.com/SYNTHIA for more information on SYNTHIA™ Retrosynthesis Software.

References:

1. Greenberg, A. A chemical history tour: Picturing chemistry from alchemy to modern molecular science. John Wiley & Sons, 2000.
2. Eastaugh, N. Pigment Compendium: A Dictionary of Historical Pigments. Butterworth-Heinemann, 2004.
3. Kassinger, R. G. Dyes: From Sea Snails to Synthetics. 21st century, 2003.
4. Hagan, E., Poulin, J. Statistics of the early synthetic dye industry. Herit. Sci. [Online], 2021, 33. https:// doi.org/10.1186/s40494-021-00493-5 (accessed 17 Mar 2022).
5. Corey, E. J.; Jorgensen, William L. Computer-assisted synthetic analysis. Synthetic strategies based on appendages and the use of reconnective transforms. J. Am. Chem. Soc. [Online], 1976, 1, 189–203. https://doi.org/10.1021/ja00417a030 (accessed 17 Mar 2022).
6. Bisio, A., Scaleup of Chemical Processes. Wiley Interscience, 1985.
7. Jüttner, U. et. al. Supply chain risk management: outlining an

agenda for future research. International Journal of Logistics Research and Applications [Online], 2003, 4, 197-210. https://doi.org/10.1080/1367556031000162701 6 (access 17 Mar 2022).

8. Klucznik, T., et. al. Efficient syntheses of diverse, medicinally relevant targets planned by computer and executed in the Laboratory. Chem [Online], 2018, 3, 522–532. https:// doi.org/10.1016/j.chempr.2018.02.002
9. Ewa P. Gajewska, et. al. Algorithmic Discovery of Tactical Combinations for Advanced Organic Syntheses. Chem. [Online], 2020, 1, pp. 280-293. https://doi.org/10.1016/j.chempr.2019.11.016 (accessed 17 Mar 2022).
10. Molga, Karol, et. al. Navigating around Patented Routes by Preserving Specific Motifs along Computer-Planned Retrosynthetic Pathways. Chem. [Online], 2019, 5. https://doi.org/10.1016/j.chempr.2018.12.004 (accessed 17 Mar 2022).
11. Szymkuć, S., et. al. Computer-generated "synthetic contingency" plans at times of logistics and supply problems: scenarios for hydroxychloroquine and remdesivir. Chem. Sci. [Online], 2020, 11, 6736-6744. https://doi.org/10.1039/D0SC01799J (accessed 17 Mar 2022)
12. Gallana, M., et. al. Climate change and infectious diseases of wildlife: Altered interactions between pathogens, vectors and hosts. Current Zoology. [Online], 2013, 3, pp. 427–437, https://doi.org/10.1093/czoolo/59.3.427 (accessed 17 Mar 2022).