

ANNALES

PROCEEDINGS OF THE ACADEMY OF SCIENCES OF BOLOGNA

CLASS OF PHYSICAL SCIENCES



ANNALES

PROCEEDINGS OF THE ACADEMY OF SCIENCES OF BOLOGNA
CLASS OF PHYSICAL SCIENCES

1



Board of Governors of the Academy of Sciences of Bologna

President: Prof. Luigi Bolondi

Vice-President: Prof.ssa Paola Monari

Secretary of the Class of Physical Sciences: Prof. Lucio Cocco

Vice-Secretary of the Class of Physical Science: Prof. Aldo Roda

Secretary of the Class of Moral Sciences: Prof. Giuseppe Sassatelli

Vice-Secretary of the Class of Moral Sciences Prof. Riccardo Caporali

Treasurer: Prof. Pierluigi Contucci

Annales. Proceedings of the Academy of Sciences of Bologna Class of Physical Sciences

Editor in Chief

Pierluigi Contucci

Editorial Board

Daniele Bonacorsi (Physics)

Luca Ciotti (Astronomy)

Matteo Guidotti (Chemistry)

Pier Luigi Martelli (Biology)

Alberto Parmeggiani (Mathematics)

Susi Pelotti (Medicine)

Nicola Rizzo (Medicine)

Marco Rocchetti (Computer science)

Cesare Saccani (Engineering)

Editorial Consultant of the Academy of Sciences of Bologna

Angela Oleandri

Fondazione Bologna University Press

Via Saragozza 10, 40123 Bologna

tel. (+39) 051 232 882 – fax (+39) 051 221 019

ISSN: 2975-2302

ISBN: 979-12-5477-324-6

ISBN online: 979-12-5477-325-3

DOI: 10.30682/annalesps2301

www.buponline.com

info@buponline.com

Copyright © the Authors 2023

The articles are licensed under a Creative Commons Attribution CC BY-NC-SA 4.0

Cover: Pellegrino Tibaldi, *Odysseus and Ino-Leocothea*, 1550-1551,
detail (Bologna, Academy of Sciences)

Layout: Gianluca Bollina-DoppioClickArt (Bologna)

First edition: December 2023

Table of contents

| | |
|---|----|
| Prefazione , <i>Luigi Bolondi</i> | 1 |
| Introduzione/Introduction , <i>Pierluigi Contucci</i> | 3 |
| Scienza e Pace <i>Giorgio Parisi</i> | 7 |
| A biomedical tale linked to the Academy of Sciences of Bologna Institute <i>Lucio Ildebrando Maria Cocco</i> | 15 |
| Some specific historical elements on the evolution of “Mathematics Education” as a research discipline <i>Bruno D’Amore</i> | 23 |
| Stream-of-consciousness thoughts on language and AI <i>Marina Frasca-Spada</i> | 35 |
| “Dal cervello impariamo cosa è dolce e cosa è amaro”. Sapori, odori e neuroscienze <i>Andrea Stracciari</i> | 51 |
| Sensi chimici. La scienza degli odori e dei sapori <i>Silvano Fusco</i> | 65 |
| How the first two decades of the twenty-first century are reshaping the science world. The perspective of synthetic organic chemistry <i>Claudio Trombini</i> | 75 |
| The endless war: the long-term impact on health and environment of armed conflicts <i>Matteo Guidotti, Andrea Leisewitz, Massimo C. Ranghieri</i> | 91 |

| | |
|---|-----|
| Evoluzione della Tavola Periodica: un ciclo di conferenze supportato da ParliamoneOra e ospitato dall'Accademia delle Scienze dell'Istituto di Bologna | 105 |
| <i>Margherita Venturi</i> | |
| Neutralità climatica e transizione energetica: il ruolo delle città intelligenti e delle comunità energetiche | 119 |
| <i>Carlo Alberto Nucci</i> | |
| Galaxy Clusters and the <i>Euclid</i> mission | 141 |
| <i>Gianluca Castignani, Giulia Despali, Carlo Giocoli, and Lauro Moscardini, on behalf of the SWG Galaxy Clusters of the Euclid Consortium</i> | |

How the first two decades of the twenty-first century are reshaping the science world

The perspective of synthetic organic chemistry

Claudio Trombini

Dipartimento di Chimica “Giacomo Ciamician”, Alma Mater Studiorum - Università di Bologna; Benedictine Academician

Abstract

The relevant role of organic synthesis in providing the society with essential new products for everyday life applications, is emphasized. The first two decades of this century witness the transition from the classic human intelligence-centred design and synthesis of new molecular architectures to the new digitalized practices. Thus, most of the ancillary and repetitive activities of organic synthesis will be committed to artificial intelligence-machine learning tools (AI/ML), while the chemist’s role can focus on the most intellectual challenges underlying a chemical structure and its synthetic design. Improved computing power, data availability and algorithms, are overturning the concept that AI is a competitor for bench chemists, but, rather, a fantastic aid to boost productivity, with potential to streamline and automate chemical synthesis. How AI/ML can empower tomorrow’s chemists is briefly discussed, together with the change chemists are going to adopt in their professional life. For example, the new publishing and referee policies, the sharing of chemical information using digital repositories, and, as members of the wide family of scientists, the fight against unethical behaviours and the spread of pseudoscience.

Keywords

Organic synthesis, Retrosynthetic analysis, AI/ML-assisted retrosynthetic analysis, Robotics and automation, Publishing policies.

1. Introduction

In 2000, the prediction of which major changes will have characterized the world of Science in the new century was quite approximate. For example, the 2023 unprecedented situation, defined by the World Economic Forum as the “*Polycrisis*” [1], namely the contemporary interaction of people with pandemics, brutal wars, energy, cost-of-living and climate crises, was not an even conceivable nightmare [2]. A major debate was about fundamental research, namely curiosity-driven ideas, and risk-taking visionary thinking. Major breakthroughs in technology originate from fundamental research, not from simple incremental improvements of processes by stepwise re-engineering or adaptation of operational conditions. So, the question was: must scientific research be directed towards private economic objectives, aiming at short-term payoff, to the detriment of open and unfocused research in the general interest? The idea that fundamental research must be funded exactly as applied projects for short-term practical goals found a general consensus.

Of course, in 2000 we were aware that the biotechnological sector (the *Human Genome Project* launched in 1990 and completed in 2003, one of the greatest feats in the history of science) [3], and *Artificial Intelligence* (AI) as well, (born in the 1940s and 1950s, with the landmark paper by Alan Turing on thinking machines) [4], should have impressively shaped the scientific development. However, at the turn of the century nobody could imagine that DNA sequencing, providing the exact order of the bases of a human DNA, could go from a cost of about US\$ 2 billion in a 15 years effort, to about US\$ 500 in less than 12 hours, thanks to the fundamental contribution of Craig Venter [5]. New-generation DNA sequencing methods arrived on the scene, allow high-accuracy sequencing of variation across the human genome, driving future discovery in human health and disease [6].

Similarly, it was impossible in 2000 to imagine the terrific and ever-expanding power of AI in transforming how we live and work. Technologies based on new information handling and computation, robotics, machine learning, etc. are ensuring unprecedented opportunities to the scientific endeavour. At the same time, AI represents a real, unspoken concern over communication, advertising, commercial management, social organisation, and political control¹ [7]. Indeed, a major risk is fostering misinformation, and, in front of new digital data, we could wonder someone has spoofed it. In Science there are not good or bad discoveries, but good or bad use of them.

2. The power of synthesis in molecular sciences

Historically, molecular sciences emerge as an interdisciplinary research field that crosses over chemistry, physics, and more recently biology. Biology was the last discipline which entered the club of scientists who study the properties of substances at the molecular scale thanks to the unbelievable progress of biochemistry and genetic engineering in the last century. Chemistry, among hard and life sciences, has had a unique feature. Chemistry studies not only the existing

¹ On May 2023, Geoffrey Hinton, known as the godfather of AI, quitted Google. The reason was that conversational chatbots can be dangerous if exploited by “bad actors”. Chatbots know much more than any person and apply predictive intelligence to enable personalization based on user’s profile. They can be extraordinary tools in the hands of authoritarian leaders to manipulate the electorate [7].

world, as all the sciences do, but it is also capable to build up new objects of study. This ability lies within the domain of chemical synthesis which everyday puts at disposal of the society new compounds for the most varied applications² [8]. Synthesis will continue to contribute molecular solutions to the most daunting challenges of our time, until life sciences, material sciences, environmental sciences require new more effective and valuable compounds, considering the vastness of molecular space, which is estimated to contain 10^{63} possible stable organic compounds with molecular weights under 500 Da [9].

When in the early 1970s genetic engineering was born [10], biologists learnt how to create new genetically modified organisms (GMO), operating at the molecular level. For example, first generation human recombinant insulin, a peptide hormone containing two chains, insulin A and B cross-linked by disulfide bridges, was produced in *Escherichia coli* [11]. That day biology, too, acquired the ability to create new objects of study. This research field is today named as synthetic biology.

Figure 1 shows the role of chemical synthesis in the discovery of new products. Figure 1 is adapted from the *Design–Build–Test–Learn* cycle, as formulated first by synthetic biology to define the optimization of metabolic engineered strains as catalysts for the production of new compounds.

The design–synthesize–test–learn cycle highlights the logic sectors the optimization of a new product is divided into. The design stage identifies the problem and suggests the first target molecule by correlating structure and function. This stage is favoured by the progress in data mining and computational autonomous design able to explore the chemical space and predict the properties of molecular architectures [12]. The Synthesis stage designs a synthetic plan and execute it, the Test sector verifies the physical properties and the product performance, eventually, the

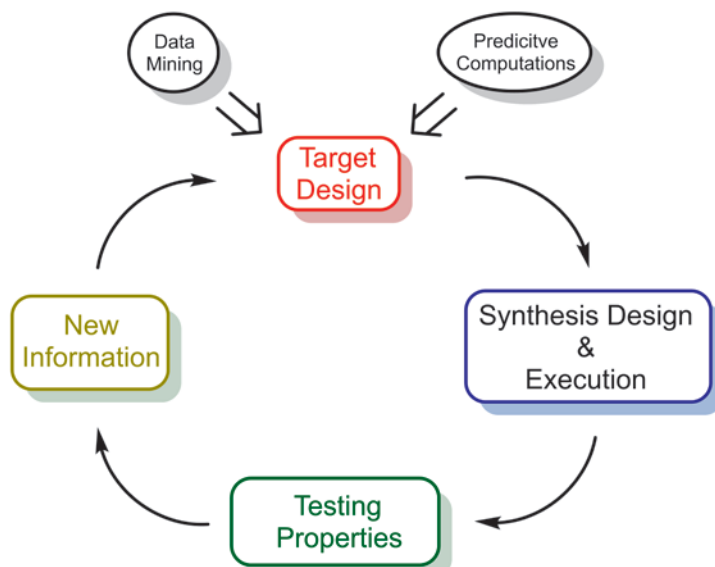


Fig. 1. The design-synthesize-test-learn paradigm in the discovery of new products.

² The French chemist Marcellin Berthelot wrote in 1876: “Chemistry creates its objects. This creative faculty, akin to that of art, forms an essential distinction between chemistry and the other natural or historical sciences” [8].

Learn stage analyses data and proposes modifications to the original target molecule, fostering iterations of the cycle until a new product is identified with the desired functional optimization.

In the past century organic synthesis has served as the testing ground for new methodologies and strategies verified in dexterous total syntheses of complex molecules. However, total syntheses looked like academic exercises that ended when enough product was obtained to characterize it. Organic synthesis maintains its importance, but it is pursuing more ambitious goals, boosting efficiency, and striving to compete with nature in terms of synthetic efficiency and scale. In short, the new guidelines are scalability and practicality, two terms that require new innovative synthetic strategies, while continuing to shed new light into fundamental reactivity.

3. The rational approach to organic synthesis

Elias J. Corey in 1990 won the Nobel Prize in Chemistry “for his development of the theory and methodology of organic synthesis”. In the 1970s Corey was a forerunner in the application of AI to the problem of how to plan a total synthesis. He adopted a mental process called retrosynthetic analysis, and developed an algorithm translated into lines of code which could sequentially run by a computing machine. The corresponding software was named *LHASA (Logic and Heuristics Applied to Synthetic Analysis)*. The *LHASA* program was exceedingly complex in the context of mid 1970s, with about 400 subroutines, 30,000 lines of *FORTRAN* code and a data base of over 600 common chemical reactions. But *LHASA* and later similar software did not take off, perhaps because the gap in computing power between the 1970s and now.³ The great power of retrosynthetic analysis was in teaching and learning how to plan a total synthesis, since once the synthesis plan had been decoded in a sequence of elementary actions, it could be taught to students.

Retrosynthesis means to design a synthetic plan of the target molecule by working backward compared to the synthesis in the lab.

I collected the main strategies of retrosynthetic analysis in a recent open-access book [13]. We can distinguish the following steps:

1. *Strategic bonds identification*. It means looking for bonds that can be constructed in the lab using known reliable reactions (*transforms*). Using ionic transforms, bond dissection first generates formal ions (*formal synthons*) that must be translated into true reagents (the *synthetic equivalents* of *synthons*). This step depends on the chemist's expertise to recognize structural and functional group patterns that suggest a specific reaction. In this step fantasy and creativity may play a fundamental role in devising nonobvious new routes.
2. Each strategic bond is mentally disconnected into structurally simpler precursors which can be combined in the lab by exploiting the known transform. This step requires a careful information retrieval process. Repositories of published reactions such as *Reaxys*, *Scifinder*, *Scopus*, *Web of Science*, are consulted first since they offer the possibility to search by structure

³ The number of transistors on integrated circuit chips was around 50,000 in late 1970s and close to 40 billion in 2018, with a growth that followed Moore's law (the number of transistors on integrated circuits approximately doubles every two years), and now technology is approaching the physical limit of the transistor size of 5 nm (the minimal radius of a sphere that could contain a 500 kDa protein).

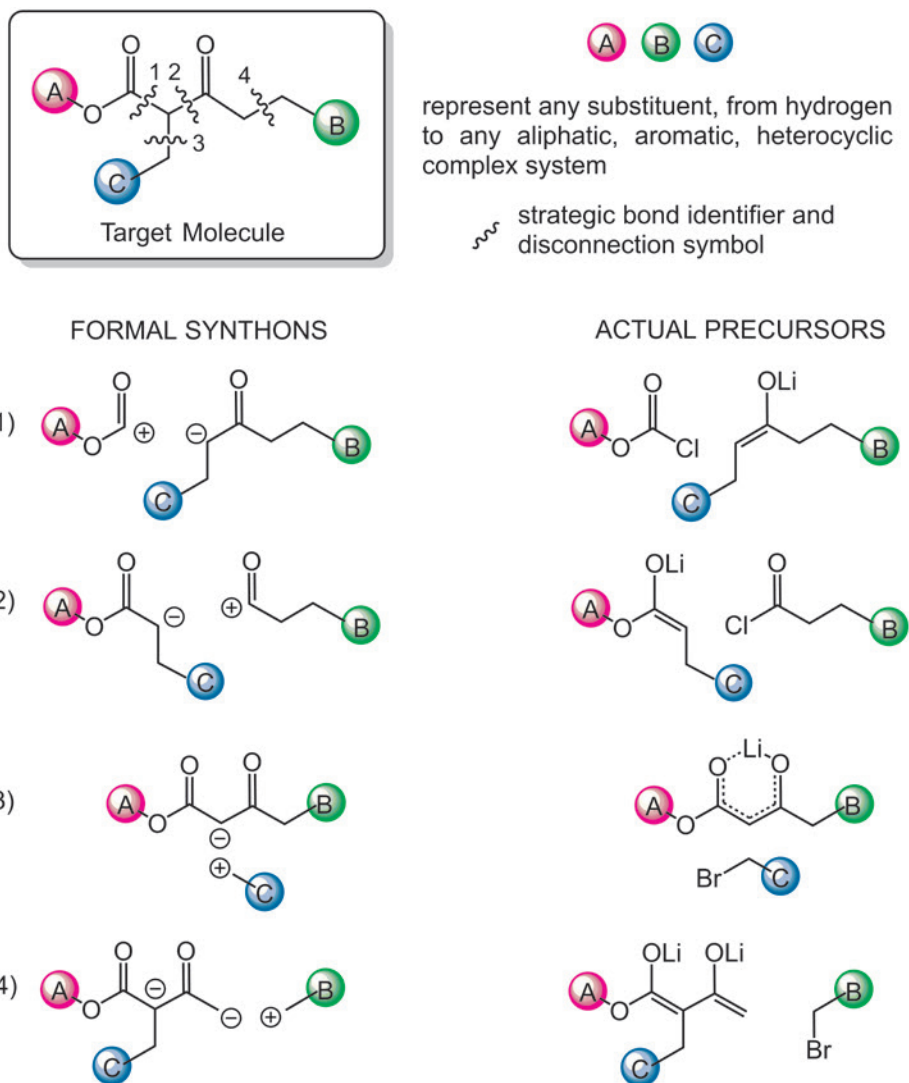


Fig. 2. Example of retrosynthetic analysis. Each bond disconnected is transformed first in formal ionic precursors (synthons), then translated into their corresponding synthetic equivalents.

and/or find information on specific reactions by searching by reactants and/or products. Once the query structure is matched to a known reaction product, information about the reaction gets available. Databases link to primary sources that contain the needed experimental details. Since several strategic bonds may be identified, and each disconnection can be related to different transforms, chemist can devise a vast number of solutions. Figure 2 shows an example of reasonable alternative disconnections of a 1,3-ketoester.

- Each precursor identified in the first disconnection analysis then undergoes the same retrosynthetic analysis, until precursors are identified which correspond to commercially available starting materials.

4. The global result is that several alternative synthetic strategies to the target molecule are devised, then chemists rank them on a hierarchical scale based on their expertise, biases, preferences, and ways of thinking. Criteria for the choice of the first plan to be tested in the lab are based on its simplifying power (how much simpler are the precursors structures compared to the target molecule), efficiency (step counting, estimated yields), economy, and sustainability standards.

4. The chemist's role in a synthesis lab

Once an attractive synthetic plan is devised, chemists must put it into action. The organic synthesis lab is a perfect environment to educate students while scouting out talents. At any stage of a reaction, unexpected difficulties require and develop critical thinking skills. The execution of a chemical reaction consists of several sequential actions: i) setting up the reaction apparatus, ii) deciding what is the best addition order of the reactants, solvents, catalysts, etc., iii) establishing which reaction conditions are to be adopted (stirring, heating/cooling, etc.), iv) following the reaction progress, v) extracting products when the reaction has completed, and finally, vi) purifying and vii) analysing them. At each stage, a careless mistake may affect the final yield or lead to complete failure. Making choices requires exploiting the practitioner's ability, skills, hands-on competence, education, and human qualities like intuition, creativity, passion, and guts, including emotions and imagination. There is an instinctive ability to face lab problems or grasp information from the subtle changes of a reaction mixture, which are almost impossible to formalize, but they are certainly related to the capacity of a fellow to concentrate, carefully observe, and critically analyse what she/he is doing. It is a matter of an inborn aptitude, or the relationship between explicit *vs.* tacit knowledge.⁴ What trainees must avoid is an emotional approach to lab work, particularly if it generates fear and self-doubt, or they are biased against the final success, or want to see only positive proofs in line with expectations and neglect single negative evidence.

Work in a synthesis lab is never annoying and always stimulates the student's wisdom by merging logic, heuristics, and creativity. For this reason, a synthesis lab is an ideal gym to train students. In short, traditional synthesis involves significant mental and physical work, rewarded by the pleasure when crossing the finish line after a demanding steeplechase. A direct consequence of the stimulating environment represented by a chemical synthesis lab is that well-trained practitioners are the most sought after by the private sector. Their excellent problem-solving skills, indeed, can be easily repositioned for countless translational applications in all sectors of applied chemistry, from life sciences (drugs, agrochemicals, etc.) to materials and nanoscience, and other cutting-edge transdisciplinary fields (for example in neuroscience, environmental science, omics, etc.).

⁴ While *explicit knowledge* (knowing-that) is codified in documents (books, reports, memos, etc.), *tacit knowledge* (knowing-how) is embedded in the mind through experience. Tacit knowledge includes insights. The impact of *tacit knowledge* on the activity in a chemical lab is critical, particularly on chemical yields, that integrate the mix of dexterity and clumsiness characterizing each person. Cfr. Polanyi, M. *The tacit dimension*. Chicago: University of Chicago Press, 1966.

The digital revolution of the last 20 years (also referred to as the fourth industrial revolution) also entered the life of synthetic chemists. Chemical databases that span the entire chemical space of known molecules reduce the time required for information retrieval to a few seconds. Software working in the backward direction according to the retrosynthetic strategy instantly generates good solutions, reducing, or removing the time-consuming aspects of synthesis planning [14-16]. A tool like *Reaxys Synthesis Planner* [17] allows the user to produce up to ten steps pathways in a bunch of seconds, and thus generate multiple, alternative synthesis pathways for the target molecule. An alternative tool is the *SciFinder's SciPlanner* [18]. Eventually, the best options of each step are arranged into a synthetic plan.

Since planning a chemical synthesis is an open problem, virtually infinite solutions exist. There is not the best plan, there is the best among a set of solutions devised. The choice of which plan execute in the lab is in the hands of chemists, their expertise, ways of thinking, preferences. When we apply the best-at-each-step strategy, only one option is selected at each step, while the total number of options to explore and evaluate to identify the “best” pathway of “*n*” steps should be astronomically high for a medium-complexity target structure. Computers assist chemists to design, sometimes replace them for simple processes, but intuition, insights and creativity cannot be encoded, yet, in appropriate formats and taught to a learning machine. This will always make the difference between a talented scientist and the tremendous aid of AI/ML, which empowers chemical synthesis work by enabling scientists to structure, analyse, and evaluate an unlimited amount of data [19].

Where will we end up is always a critical guess. However, we must remember that algorithms are the result of human intelligence, even though, once created, they run independently. The ever-increasing power of computing machines allows the development of ever more sophisticated chemistry software. *Chematica*, an organic retrosynthesis software [19, 20] now marketed under the brand name *Synthia*, offers a new way to plan a total synthesis. The program designs and evaluates many more routes than a human could possibly evaluate and scores them according to criteria established by the user, for example cost, circumventing patented technologies, originality, the time span, no use of toxic or regulated substance, and sustainability issues [21]. The strength of *Chematica* is to evaluate extraordinarily large numbers of possible synthetic pathways while applying multiple optimization criteria/constraints simultaneously. This work is clearly beyond human cognition if performed manually for the huge number of possibilities to consider and for the complex logical combinations of cross-referencing multiple constraints.

A second strength of the *Chematica* platform refers to the execution phase of a synthesis. Indeed, it provides for each reaction suggestions for reaction conditions, literature citations, the chemistry involved, functional groups that need to be protected etc., with the indirect effect to empower less experienced chemists to perform synthetic work typically carried out in classic synthetic laboratories. These digital aids contribute to forge talented chemists with both the written and tacit knowledge that will be accumulated over the years. However, when using these platforms, we must remember we are not blindly in the hands of a magic black-box which does our job quickly and efficiently, instead, we must remember the essential role played by expert chemists in producing scientific results in a preliminary lab investigation that will be coded later in appropriate algorithms.

At the industrial scale, statistics and automation warrant a high degree of reproducibility. Reproducibility is the essential tenet of a scientific study and discovery. It is the proof that an established and documented work can be verified, repeated, and reproduced, and it is at the basis of the trust of scientists in their colleagues. The great advantage of automated robotic systems also at the fundamental research level, is the creation of reliable, reproducible results that do not depend on the operator hands. Thus, research reproducibility is implemented while minimizing human errors [22, 23]. The chemist' community already masters the spectacular tool of automated solid-phase peptide synthesis (SPPS), a fundamental technology to produce chemically engineered peptides by the Fmoc/t-Bu protecting-group strategy,⁵ the most used one. Analogous solid-phase synthesis of oligonucleotides technologies fully integrated into an automated setup is exploited daily for use in research laboratories, hospitals, and industry using the phosphoramidite strategy [24]. Today it is possible to apply robotization virtually to all the operations in the lab, particularly in the top public and private research centres, but with a lower presence in resource-limited university labs. Robotics, whose functioning must be configured by expert chemists, will not cut jobs, but requires new ways to train future chemists capable to manage these new tools, share ideas on how best to harness automation to benefit academia, industry, and society. Automation is useful for performing simple processes repeatedly, but (new chemists are necessary for teaching robots how to perform new processes. New curricula must prepare scientists to design automated inexpensive instruments easy to deploy in smaller laboratories, and, at the same time, to adapt advanced equipment in a centralized laboratory setting to their specific quests. Recent examples within the domain of synthetic biology of highly automated central facilities are biofoundries [25]. A biofoundry is a technological infrastructure which enables the rapid design, construction, and testing of genetically reprogrammed organisms for biotechnological applications and research. Instead of developing automation in house, research groups can have access to biofoundries through proper scientific agreements, exploiting the cutting-edge technology available. A *Global Biofoundry Alliance* has recently been established to coordinate activities worldwide [26]. They have a beneficial role in improving data reproducibility and quality of protocol reporting, a today weakness in life science publications, owing to the inherently complex matter managed.

5. Lifestyle changes in chemists' professional life

5.1 New priorities in synthetic chemistry

In the 20th century the total synthesis of complex natural products defined the frontiers of organic chemistry by achieving fundamental insights into reactivity and selectivity concepts. However, the virtuosity of the scientists was not only measured by the ability to produce a few milligrams of a complex natural product, rather by the discovery and development of new

⁵ In the coupling of α -aminoacids to give peptides, the selective and orthogonal protection of the amino and the carboxylic acid groups is mandatory to prevent polymerization of the amino acids and to minimize undesirable side reactions. Fluorylmethoxycarbonyl (Fmoc) is the most common protective group of the amino functionality removable with bases; the acid labile tBu protective group of the carboxylic function is commonly associated with the Fmoc group.

strategies, methodologies and reagents expressly devised for that total synthesis. Baran underlined that even if we collected all the taxol samples published in the 1990s, its amount should be less than 30 mg [27]. In contrast, the current plant cell fermentation output of taxol in the process developed by *Bristol Myers Squibb* [28], is on the scale of 100 g, clearly stressing the gap between chemical and biological scales and efficiency. Synthetic chemists must commit themselves to increase efficiency, trying to close the gap between biosynthesis and chemical synthesis. This goal may be achieved by making total syntheses scalable, at least at the gram level. Baran says “The ultimate expression of a natural product synthesis is to make it in gram quantities and invent methods that have wider uses”, underlying the double effort of synthetic chemists of building up complex architectures while inventing new chemistry [29]. Simplicity will be a perfect benchmark for ideality [30].

A successful scaling-up requires streamlining a synthetic plan by reducing the number of steps. To reach this objective, a relentless effort is necessary to develop new chemoselective transformations (no need for protective groups) and new strategies, requiring an ever-lower amount of labour and material expenses. So, the message to new generations is that organic synthesis, after a glorious history, has an exciting and challenging future that will ensure joy and satisfaction to people who will invest their passion and stamina in this intellectually rewarding job. A synthesis lab will be never a simple service provider. A signal of how organic synthesis stays relevant today comes from the portfolio of *Nature Journals*, enriched in 2022 with *Nature Synthesis*. The Editorial board defines the new Journal as “a home for new and important syntheses of molecules and materials that can make the world a better place”.

5.2 Editorial and publishing policies

The past two decades witnessed the arrival on the chemistry stage of a vast number of scientists, mostly new young people from Far East Countries. Mainland China was almost absent from the chemistry scene in the XX century, but in 2018 the Chinese Academy of Sciences has been the world’s top-ranking Institution in Chemistry [31]. In the list of affiliations of the most cited researchers, the Chinese Academy of Sciences come third in 2019, after Harvard and Stanford. The welcomed new communities of chemists greatly contribute to the development of science.

While the world population growth has reached an unsustainable level (it increased from 5.3 billion in 1990 to 8 billion on Jan. 2023, independently of the fixed constraints of the biosphere, that cannot provide food and water to so many human beings), the growth of science is a terrific opportunity, if we will be able to exploit the good face of scientific discovery. A long-term commitment of all public and private stakeholders must focus on the great ambition of stopping environmental destruction, finding science-based solutions to achieve carbon neutrality, and improving general well-being [32]. Unfortunately, in 2023 this aspirational goal seems to be visionary, since in most countries people continue to damage the environment and other human beings (local, regional & global pollution issues, climate change, drought, starvation, new cruel wars, weapons business, poisons, recreational drugs, etc.).

A consequence of the growing number of chemists is the increasing number of chemistry-related journals (2325 journals are listed in the 2022 Edition of the *Journal Citation Report* in

the Category Chemistry) [33]. Since 1990, the number of papers holding the words “*Organic Synthesis*” in the title, abstract, and keywords has seen a 14-fold increase (*Scopus*). If more science is welcomed, a “con” is the level of accuracy of the peer review process, which can be hardly preserved when the number of papers to be examined on our desk increases too fast.⁶ Publishing scientific reports is also a business. In 2010, a top-level publisher reported profits of almost US\$1 billion [34]. The pressure to publish (the “publish or perish” refrain) due to university hiring procedures, also reflects the increase of scientific unethical practices. Duplicate submission, falsification or fabrication of data, plagiarism, and fraud⁷ pass more easily through the net of an approximate referee’s analysis [35]. The reproducibility of results is the first victim of a poor-quality study. A 2016 *Nature* survey reported that 80% of respondents failed to reproduce a published chemical experiment from a different research group, pointing out the great problem of reproducibility in recent literature [36]. Under the pressure of fast publication rates necessary to get better positions, chemists are increasingly adopting marketing rules in their publication policy. They invest plenty of energy and time in the emotional appeal of their manuscripts or research proposals, but scientific reputation does not depend on aesthetics. We all have experience of a well-advertised research (creative artworks, high-sounding titles, covers, and impressive figures), that hides weaknesses. The worst consequence of these behaviors is that they contribute to lower the trust in science.

The existence of a deceitful publishing business model that involves charging publication fees to authors without checking articles for quality and legitimacy, was first noticed by Jeffrey Beall in 2015 who defined “predatory publishers” those who adopt these unethical behaviors [37-38]. To oppose this risk, a few journals [39], for example *Nature Communication* [40], are promoting transparency in science by adopting an open peer review, where referees reports, author rebuttals or letters from the editor, as well as reviewers’ identities, are published alongside the articles [41]. Of course, such a policy discourages careless reports, independently of whether they are due to lack of time or conflict of interest.

5.3 AI-ML empowers tomorrow’s chemists

Automation, data-enabled machine-learning (ML) methods have been shown to provide results sometimes superior to humans [24]. A breakthrough has been the prediction of protein folding based on the bioinformatic analysis of the protein primary structure in a very short time with a high level of accuracy thanks to the AI-based open program *AlphaFold 2*, which also provides reliable information about protein-protein interactions, so important in enzymatic machineries [42].

AI/ML tools offer us the opportunity to learn complex patterns of chemical reactivity from organic reaction data, data-driven ML models have been designed and applied to planning

⁶ It looks bizarre the triple payment system within public research in which the State pays 3 times the production of scientific literature. A State agency pays for the research, then scientists, paid by the State, offer their results for free to publishers or must pay fees for open access publications. Peer review is done by scientists on a voluntary basis, and eventually publishers sell the product back to government-funded institutional and university libraries.

⁷ For a “theatrical” example of scientific fraud in chemistry in India, see Killugudi Jayaraman, *Chemistry’s ‘colossal’ fraud 2008* [35].

synthetic pathways, recommending reaction conditions for each reaction, and even predicting what can be the major products of a yet-untested reaction. Jensen at MIT [43] is developing an automated platform integrating generative ML models and property prediction ML models. He can use generative models to propose new molecules not previously reported and having desired properties, then exploiting automation to synthesize,⁸ purify, and characterize them, and eventually update the predictive ML models for a subsequent cycle of explorations, according to the workflow shown in Figure 1.

In all ML applications, the performance of models for a given task heavily relies on the quality and scope of the training data that cover this task. Data quality require large, well-curated, datasets. Thus, chemists must change their research data management practices by embracing an accepted standardized machine-actionable platform that integrates data collection, processing, and capture, while removing barriers to access and reuse [44]. A second change of mindset required is the urgency to report negative results and zero-yields [45]. These data are necessary to build up a sufficiently large dataset from which reliable predictions can be extracted about kinetics and thermodynamics, in turn related to reaction yield, stereoselectivity, and all the experimental outcomes chemists are interested in. For instance, an *Open Reaction Database* consisting of a structured data format for chemical reaction data, has the goal to allow data sharing, by providing an interface for browsing/downloading that data [46]. A further example of open science is offered by the *Chemotion Repository* for the storage and provision of chemical reactions and processes, including chromatography or spectroscopy data, that can provide full and open access to results data [47]. At the same time efforts are necessary to migrate the existing datasets from non-digital to digital formats (data curation). The larger the databases, more reliable are the synthetic strategies or reaction conditions that can be extracted.

Independently of authors claims about AI/ML potency, their applications to process optimization and synthetic design are still in their infancy [48]. However, there are no doubts about the need to provide next-generation chemists with new digital competences. As teachers, we also are confident that, AI and immersive reality will offer students unbelievable efficient advantages of understanding the 3D molecular world by visual learning, and speeding up concept learning thanks to immediately comprehensible visual aids.

5.4 Fighting pseudoscience

Let us discuss now the call for scientists to fight pseudoscience. Too many people fall for deniers of climate change, anti-vaccine activism, theories about racial classifications, etc. A major consequence of improper use of social media is the spreading of a serious loss of confidence in scientific activities. The flooding on the internet of rubbish and fake news, makes it difficult for public opinion to acknowledge competence and form their own opinion based on true facts, not affected by misinformation campaigns. Unfortunately, misinformation is in the position to dictate the agenda of the public debate, as it happened in the early COVID era. The risk is that

⁸ Automated assembly lines can outperform humans in reaction processes that require accuracy, reproducibility, and reliability. It is not a case that the earlier examples have been developed for the synthesis of short-lived ¹⁸F radiolabelled compounds, used in positron-emission tomography.

when truth and a well-tailored pseudo-truth are proposed together with an ingenious mix of true information and made-up fake news, the emotional impact is destructive. The younger component of our society, who participates in public affairs less frequently, with less knowledge and enthusiasm, is more vulnerable to commentaries on podcasts, YouTube videos, and other social media.⁹ The market considers young people just as consumers, instilling them making money right away as the only standard of achievement, not their long-term job perspectives, satisfaction, professional growth, generosity, etc.

However, the demarcation between people who promote science and pseudoscience is recognizable through their attitude to deal with a scientific problem. Science debates about probability that a problem has been understood and solved. Science is never based on tenets. When a theory is accepted as true, this is done provisionally. Nothing is ever proven, even if we have a lot of data to support it. What today is truth, tomorrow new evidence can change or dismantle it. Every discovery leads to more questions, something else that needs explaining. On the other hand, pseudoscience supporters disparage values such as memory of the past, study, and work, consider competence a self-referential non-value of a discredited elite, and does not recognize authority, promoting distrust of scientific institutions, endorsing irrational antiscientific movements. Assumptions in pseudoscience spreaders are absolute, indisputable, even though their proofs are based on secondhand heard or read self-complacent truths that appeal to emotion and personal irrational belief. Science is the provisional acceptance of good data, pseudoscience is an unshakable certainty based on invented data.

6. Conclusions

Future generations of chemists will manage what we can do now in a much faster and more efficient way. The transition to digital chemistry, namely the AI/ML age, is underway [48]. This will force chemists to modify their mindset and raise their professional profile by minimizing routine activities. Computer-assisted retrosynthetic analysis will solve the most time-wasting steps of chemical research, from the identification of alternative synthetic plans to the execution stage, finding good reaction conditions. It is a common feeling that the hard work and commitment in the lab of total synthesis, often repetitive and time-consuming, will be committed to smart machines. But we must not forget that chemical databases, computer-assisted synthetic software, lab automation and all the digitalized new services are powered by the work of expert chemists, and that any software may be improved in an endless optimization effort, a duty that still is in the hands of chemists. The experimental chemist must first evaluate each process in the lab, find all the critical parameters, optimize them, and eventually provide the programmer with all the details necessary for encoding algorithms and, eventually, enjoying the benefits of automation.

⁹ Teens suffered a lot of loneliness and social isolation during the COVID-19 pandemic, becoming more vulnerable to psychological disorders. Social media may help by increasing social interactions, self-expression, and making the access to information easier, but they can also heighten stress, up to trigger psychiatric issues. The defence of privacy in the networked information age is replaced by the need of visibility. Ashwin A., Cherukuri S.D., Rammohan A. *Glob. Health* 2022, 12: 03009.

Discussing now about science in general terms, not limited to chemistry, scientists failed to build-up a social consensus about the environmental risks in the era of “*polycrisis*” [1]. Democratic regimes had proved unable to promote the necessary and often unpopular long-term economic and social reforms under the fear of frequent short-term electoral risks. Even direr is the scientist’s failure in forming an environmental awareness among people of third-world countries, in part ruled by autocrats and theocrats who often are deniers of the climate change. In poorest countries people do not want to renounce to the hope to approach the welfare standards that the western countries, reached thanks to their past predatory misuse of natural resources, a practice that is now endured by new geopolitical players. Unfortunately, exactly these poorest countries in Africa and Asia are paying the highest price to climate change because of droughts, extreme floods, and accelerated spread of infectious diseases. My personal view is pessimistic in the short term. Indeed, even though science can evaluate long-term effects of human actions, most people of any social class in all geographical areas are living on the moment, considering the environment something to exploit, and pursuing their short-term self-interests. Fighting this reality is the leading challenge of science, today and tomorrow.

References

1. The World Economic Forum. *The Global Risks Report 2023*, 18th ed. https://www3.weforum.org/docs/WEF_Global_Risks_Report_2023.pdf (accessed on 4 July 2023).
2. Gomollón-Bel, F.; García-Martínez, J. *Angew. Chem. Int. Ed.* 2023, e202218975.
3. Human genome project. Available online: <https://www.genome.gov/about-genomics/educational-resources/fact-sheets/human-genome-project> (accessed on 20 June 2023).
4. Turing, A. *Mind* 1950, 59, 433-460. <https://doi.org/10.1093/mind/LIX.236.433>.
5. Venter, J.C.; *et al.* *Science* 2001, 291, 1304-1351.
6. Nurk, S.; *et al.* *Science* 2022, 376, 44-53.
7. <https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html> (accessed on 15 June 2023).
8. Berthelot, M. *La Synthèse Chimique*. G. Baillière: Paris, 1876, 275. Available online: <https://archive.org/details/lasynthsechimi00bert> (accessed on 25 May 2023).
9. Lipkus, A.H.; Watkins, S.P.; Gengras, K.; McBride, M.J.; Wills, T.J. *J. Org. Chem.* 2019, 84, 13948-13956.
10. Tamura, R.; Toda, M. *Neurol. Med. Chir. (Tokyo)* 2020, 60, 483-491.
11. Walsh, G. *Appl. Microbiol. Biotechnol.* 2005, 67, 151-159.
12. Joshi, R.P.; Kumar, N. *Molecules* 2021, 26: 6761. doi: 10.3390/molecules26226761.
13. Trombini, C. *The Construction of Molecular Architectures. Guidelines to Organic Synthesis Design*. BUP: Bologna, 2021.
14. Szymkuć, S.; Gajewska, E.P.; Klucznik, T.; Molga, K.; Dittwald, P.; Startek, M.; Bajczyk, M.; Grzybowski, B. *Angew. Chem. Int. Ed.* 2016, 55, 5904-5937.
15. Mikulak-Klucznik, B.; *et al.* *Nature* 2020, 588, 83-88.

16. Ishida, S.; Terayama, K.; Kojima, R.; Takasu, K.; Okuno, Y. *J. Chem. Inf. Model.* 2022, 62, 1357-1367.
17. <https://www.elsevier.com/en-xm/solutions/reaxys/how-reaxys-works/synthesis-planner> (accessed on 22 June 2023).
18. <https://www.cas.org/solutions/cas-scifinder-discovery-platform/cas-scifinder/synthesis-planning> (accessed on 22 June 2023).
19. Oliveira, J.C.A.; Frey, J.; Shuo-Qing Zhang, S-Q.; Xu, L-C.; Li, X.; Li, S-W.; Hong, X.; Ackermann, L. *Trends in Chem.* 2022, 4, 863-885.
20. Klucznik, T.; Mikulak-Klucznik, B.; McCormack, M.P.; Mrksich, M.; Trice, S.L.J.; Grzybowski, B.A. *Chem* 2018, 4, 522-532.
21. Mikolajczyk, A.; *et al.* *Green Chem.* 2023, 25, 2971-2991.
22. *Empowering Tomorrow's Chemist: Laboratory Automation and Accelerated Synthesis: Proceedings of a Workshop in Brief (2022)*. National Academies Press. Washington, 2022. Available online: <http://nap.nationalacademies.org/26497> (accessed on 26 Jun 2023).
23. Gromski, P.S.; Granda, J.M.; Cronin, L. *Trends in Chem.* 2020, 2. doi: 10.1016/j.trechm.2019.07.004.
24. Sandahl, A F.; Nguyen, T.J.D.; Hansen, R.A.; Johansen, M.; Skrydstrup, T.; Gothelf, K.V. *Nat. Commun.* 2021, 12 (2760). doi: 10.1038/s41467-021-22945-z.
25. Jessop-Fabre, M.M.; Sonnenschein, N. *Front. Bioeng. Biotechnol.* 2019, 7 (18). doi: 10.3389/fbioe.2019.00018.
26. Hillson, N.; *et al.* *Nat. Commun.* 2019, 10, 2040. doi: 10.1038/s41467-019-10079-2.
27. Mendoza, A.; Ishihara, Y.; Baran, P.S. *Nat. Chem.* 2012, 4, 21-25.
28. Mountford, P.G. The Taxol® Story – Development of a Green Synthesis via Plant Cell Fermentation. In *Green Chemistry in the Pharmaceutical Industry*; Dunn, P.J., Wells, A.S., Williams, M.Y. Eds.; Wiley-VCH, 2010; Chapter 7.
29. M. Peplow. *Chemistry World* 2014, Apr. 11.
30. Peters, D.S; Pitts, C.R.; McClymont, K.S.; Stratton, T.P.; Bi, C.; Baran, P.S. *Acc. Chem. Res.* 2021, 54, 605-617.
31. <https://www.natureindex.com/annual-tables/2020/institution/all/chemistry/global> (accessed on 03 August 2023).
32. Wang, F.; *et al.* *The Innovation* 2021, 2, 1-22.
33. <https://jcr.clarivate.com/jcr/browse-categories>.
34. Creus, G.J. *University World News*, 15 October 2022. <https://www.universityworldnews.com/post.php?story=20221014071213444> (accessed on 30 June 2023).
35. Jayaraman, K. *Chem. World* 2008, 25 March. <https://www.chemistryworld.com/news/chemistrys-colossal-fraud/3000886.article> (accessed on 03 July 2023).
36. Baker, M. *Nat.* 2016, 533, 452-454.
37. <https://beallslist.net/> (accessed on 25 June 2023).
38. Hvistendahl, M. *Science* 2013, 342, 1035-1039.
39. Polka, J.K.; Kiley, R.; Konforti, B.; Stern, B.; Vale, R.D. *Nat.* 2018, 560, 545-547.
40. Editorial. *Nat. Commun.* 2022, 13 (6173). doi: 10.1038/s41467-022-33056-8.
41. Wolfram, D.; Wang, P.; Hembree, A.; Park, H. *Scientometrics* 2020, 125, 1033-1051.

42. Bryant, P.; Pozzati, G.; Elofsson, A. *Nat. Commun.* 2022, 13 (1265). doi: 10.1038/s41467-022-28865-w.
43. <https://jensenlab.mit.edu/publications-recent/> (accessed on 27 June 2023).
44. Herres-Pawlis, S.; *et al.* *Angew. Chem. Int. Ed.* 2022, 61, e202203038.
45. Editorial. *Org. Lett.* 2023, 25, 2945-2947.
46. Kearnes, S.M.; *et al.* *J. Am. Chem. Soc.* 2021, 143, 18820-18826.
47. Tremouilhac, P.; Lin, C.-L.; Huang, P.-C.; Huang, Y.-C.; Nguyen, A.; Jung, N.; Bach, F.; Ulrich, R.; Neumair, B.; Streit, A.; Bräse, S. *Angew. Chem. Int. Ed.* 2020, 59, 22771-22778.
48. Editorial. *Nat.* 2023, 617, 438.

