

Cite this: *Digital Discovery*, 2026, 5, 1447

Reaction optimization through mechanistic insight and predictive modelling

Roger Monreal-Corona, * Anna Pla-Quintana * and Albert Poater *

The control and optimization of chemical reactions lie at the heart of modern synthetic chemistry, driving progress in efficiency, selectivity, and sustainability. This review highlights the evolution of reaction optimization strategies, from empirical one-factor-at-a-time (OFAT) approaches to statistically robust methodologies based on design of experiments (DoE). These frameworks enable a systematic exploration of reaction space, providing quantitative models that accelerate process development and mechanistic understanding. The synergy between experimental and computational chemistry is discussed as a transformative paradigm for elucidating catalytic mechanisms and rationalizing selectivity in complex systems. Advances in density functional theory (DFT) and related electronic-structure analyses have enabled detailed characterization of intermediates and transition states, supporting predictive mechanistic models. Finally, the integration of machine learning (ML) into synthetic and mechanistic chemistry is outlined as a key frontier for predictive catalysis, offering new tools for reaction deployment, development, and discovery. By uniting experimental design, theoretical modeling, and data science, this multidisciplinary framework paves the way toward automatic, data-driven reaction optimization and rational catalyst design.

Received 7th December 2025
Accepted 17th March 2026

DOI: 10.1039/d5dd00543d

rsc.li/digitaldiscovery

1. Introduction

After a long quest of more than 200 years, organic chemistry has emerged as one of the most important subjects of research, and the field has witnessed tremendous advancement along the lines of developing novel methods for efficiently synthesizing small to highly complex molecules.^{1,2} In fact, the impact of these developments could be well-observed in our day-to-day life by simply analysing the importance and need of products from various chemical-based industries like pharmaceuticals, fertilizers, pesticides, rubbers, plastics, cosmetics, *etc.*³ Applications of organic synthesis in these areas have expanded remarkably, leading to an exponential growth in the demand of synthetically challenging and complex molecular scaffolds. In order to cater to this exponentially increasing demand of efficiently accessible chemical space, the need to develop superior synthetic methodologies that are not only greener but also cost-, time-, yield-, and labour-effective is constantly intensifying.⁴

The creation of organic molecules, no matter how complex, necessitates the use of effective methods for performing specific transformations. However, developing a reaction that results in only the desired product poses a formidable challenge. Indeed, the control of selectivity is of utmost importance to transition towards green chemistry and to avoid the economic and

environmental costs associated with tedious purification and the disposal of undesired products.⁵

The control of the selectivity of a reaction, especially as molecular complexity increases, is very challenging. The conceptual approach to obtaining selectivity is to make the desired reaction channel overwhelmingly kinetically favourable compared to the rest of the transformations. One of the most efficient ways to achieve this is through the use of catalysts.⁶ Selectivity comes in various forms, such as chemo-, regio- and stereoselectivity, each presenting different grades of difficulty in control.^{7,8} Catalyst control over stereoselectivity has been a significant milestone in chemistry and intense work on this topic is leading to an increasingly better-documented knowledge of the stereoelectronic factors governing the formation of different stereoisomers. In contrast, the ability of different metal catalysts to direct the reaction of one set of reactants to the formation of different chemoselective products is meaningful but infrequent.⁹ This approach to product selection has clear advantages over the commonly employed strategy of altering reactants or reaction conditions to redirect a reaction pathway toward different products. Thus, tuning the catalyst to trigger different reaction pathways from the same set of starting materials is a highly appealing strategy for the development of new chemical transformations.

It is fundamental to improve predictability and control over selectivity, as well as to extend the typology of scaffolds that can be accessed. This is not only important for the synthesis of a targeted compound but also highly relevant for the rapid generation of libraries of complex molecules.

Institut de Química Computacional i Catàlisi, Departament de Química, Universitat de Girona, C/Maria Aurèlia Capmany 69, Girona, Catalonia 17003, Spain. E-mail: roger.monreal@udg.edu; anna.plaq@udg.edu; albert.poater@udg.edu



2. Reaction optimization

Chemical reaction optimization encompasses a range of meanings that vary significantly depending on the chemist's perspective, leading to diverse expectations regarding optimization capabilities and proficiency.¹⁰ Notably, reaction optimization remains largely underexplored in undergraduate chemistry curricula,¹¹ which may leave many research chemists unaware of established optimization techniques. As a result, chemists may rely on intuition-driven approaches throughout their careers, particularly in academic settings, where empirical trial-and-error remains common practice.

However, in industrial environments, especially in process development and scale-up, optimization is not merely a tool but a necessity. The ability to efficiently transfer a reaction from lab scale to manufacturing scale hinges on rigorous optimization protocols that ensure reproducibility, safety, and cost-effectiveness.¹² This contrasts with academic workflows, where optimization is often less formal and primarily aimed at demonstrating feasibility or mechanistic understanding rather than large-scale implementation. The divergence in approaches reflects the different objectives of each context: industry prioritizes robustness, efficiency, and regulatory compliance, while academia often values innovation, proof-of-concept, and fundamental insights. Growing evidence suggests that data-driven techniques such as reaction modelling and algorithmic optimization can significantly improve efficiency, reducing both time and material consumption. These advanced methods are routinely applied in industrial research and development, particularly within process laboratories, where multidisciplinary teams, including statisticians, data scientists, and process engineers collaborate to refine reaction conditions with precision.

2.1. One factor at a time (OFAT)

"Intuition-based optimization" primarily refers to the practice of enhancing reaction metrics based on trends and anecdotal insights from experienced chemists. In conjunction with this intuitive approach, the one factor at a time (OFAT) methodology frequently serves as an alternative strategy for chemical process optimization and understanding.^{13–16} This technique is predominantly employed in academic settings, often within the framework of a structured yet straightforward procedural approach, which renders it both effective and accessible. The OFAT methodology necessitates a degree of scientific intuition, wherein experiments are iteratively conducted by holding all process variables constant except for a single factor. Upon identifying the optimal value for this factor, it is fixed while subsequent experiments are carried out to optimize another variable. This process continues until each factor is optimized to the researcher's satisfaction, leading to the belief that optimal reaction conditions have been achieved.¹⁷ These factors may encompass a variety of experimental conditions such as temperature, catalyst, ligands, additives, stoichiometry, and reaction time, which, when considered collectively, create a multidimensional parameter space. This space contains numerous possible combinations of factors for each

experimental run and is bounded by the upper and lower limits of each variable (*e.g.*, maximum and minimum temperatures).

Although the OFAT methodology provides a fast and straightforward approach to optimization, it is commonly regarded as inaccurate and inefficient compared to more advanced methods. This approach often leads to a misinterpretation of the chemical processes under investigation, as it fails to account for potential synergistic interactions between the variables involved.¹⁸ The method overlooks any interdependencies between experimental factors, as it applies a linear framework to chemical reactions that exhibit inherently nonlinear responses.¹⁹ Such nonlinearity can be effectively addressed through statistical or physical modelling, yet it remains unexplored within the OFAT framework, which can result in the incorrect identification of optimal reaction conditions.^{20,21}

A schematic representation of an exemplary OFAT optimization strategy for a chemical reaction is depicted in Fig. 1, where reagent equivalents and temperature are the variables under consideration.¹⁰ The optimization proceeds by initially fixing the temperature while performing seven experiments to determine the optimal reagent equivalents. Once experiment 5 identifies the optimal reagent equivalents, experiments 8–14 are conducted to optimize the temperature. Since only two parameters are considered in this example, the optimization is deemed complete, assuming the identified conditions are optimal. However, because the response surface is unknown prior to experimentation, it remains challenging to assess the proximity of the identified optimal parameters to the true global optimum of the system.

Numerous examples in the literature demonstrate the application of OFAT optimization across various fields of chemistry. However, with advancements in laboratory equipment and the emergence of technologies,^{22,23} chemists must adapt and expand their skill sets to fully leverage these innovations. Modern optimization strategies are increasingly robust

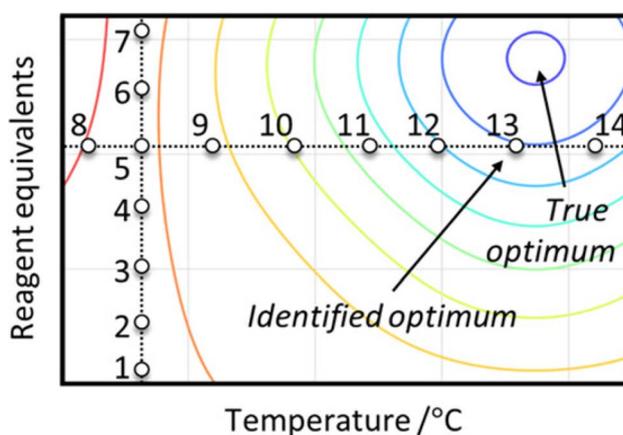


Fig. 1 Schematic representation of an OFAT experimental optimization process, in which reagent equivalents and temperature are varied. The response surface is color-coded, ranging from red (indicating a low response) to blue (indicating a high response). Reprinted with permission from the American Chemical Society.¹⁰



and efficient, often outperforming traditional OFAT approaches in terms of accuracy and reliability.^{24–26}

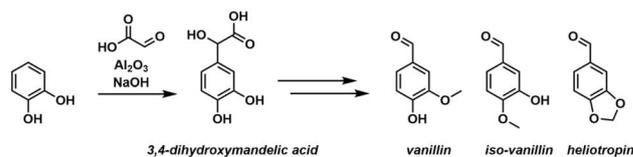
2.2. Design of experiments (DoE)

Design of experiments (DoE) is a robust and widely adopted optimization technique, particularly in the pharmaceutical and fine chemical industries. DoE encompasses a set of statistical methods aimed at constructing a mathematical model that describes the relationship between experimental inputs (*e.g.*, temperature, reaction time) and reaction outcomes (*e.g.*, yield, purity). By systematically exploring the influence of multiple factors simultaneously, DoE enables a more efficient and comprehensive understanding of reaction behaviour compared to traditional optimization approaches. There are three main objectives for DoE: screening, optimization, and robustness.^{27–29} Screening aims to identify the factors that have a significant impact on reaction outcomes while defining their upper and lower bounds. Optimization then determines the optimal levels of these factors, such as temperature and reagent equivalents, to maximize the desired reaction output. Finally, robustness testing assesses the sensitivity of the reaction to small variations in experimental conditions, ensuring reproducibility and reliability under real-world laboratory settings.

In practice, conducting a DoE campaign involves executing a set of predefined experiments following a structured experimental framework. These frameworks serve as systematic guides for exploring the selected factors and their defined ranges, ensuring efficient coverage of the parameter space. By organizing experiments in this manner, DoE facilitates the collection of data in a structured format, enabling the development of robust statistical models to analyse and optimize reaction conditions.³⁰ The structure of experimental data is crucial, as it can be challenging to analyse or interpret solely through human intuition. To address this, DoE software such as MODDE,³¹ JMP,³² Design-Expert,³³ or toolboxes in programming languages like R, MATLAB, and Python are commonly used. These tools facilitate data analysis, model fitting, and parameter optimization. Once a statistical model is constructed, optimal reaction conditions can be identified, and response surfaces are often visualized to illustrate the influence of experimental factors on reaction outcomes.

An example of reaction optimization in the frame of DoE is the synthesis of vanillin, iso-vanillin, and heliotropin, reported by Minisci and co-workers.³⁴ Several DoE studies were conducted to identify the optimal process parameters for each synthetic step, including the initial addition of glyoxylic acid to catechol to form the desired 3,4-dihydroxymandelic acid intermediate, as depicted in Scheme 1. Initial attempts to replicate published literature protocols resulted in poor selectivity and yields (<20%), prompting the authors to systematically employ DoE designs. These designs were used to identify key experimental factors and to estimate the effects of main factors and their interactions, ultimately leading to the development of an accurate statistical model that optimized product yield.

The first DoE study investigated the following variables: the amount of glyoxylic acid, aluminium oxide, reaction



Scheme 1 Synthesis of the desired 3,4-dihydroxymandelic acid intermediate.

temperature, and sodium hydroxide, while keeping the amounts of catechol, water volume, and reaction time at fixed values. This design resulted in 18 experimental runs. It was determined that excess sodium hydroxide significantly increased impurity formation rates, prompting a subsequent study under less basic conditions. The factors for this second study included glyoxylic acid, catechol-to-aluminium oxide ratio, and temperature, and this study employed a design with nine experiments. The responses measured were the recovery of catechol, selectivity for the desired product, and product yield. A statistical model was generated for each response, and the response surface for the selectivity of the desired intermediate is shown in Fig. 2. The analysis revealed that to achieve optimal product output, the glyoxylic acid amount should be increased, the catechol-to-aluminium oxide ratio should be maintained between 2.17 and 2.28, and the reaction temperature should be elevated. Further experimentation, based on these findings at higher factor bounds, resulted in an improved selectivity of 90.5% and a conversion of 78.4%, with unreacted catechol being easily recovered and recycled.

Another example of DoE is the recent work by Iudanov *et al.*,³⁵ who investigated the selective cross-metathesis (CM) of terpenes, a transformation often overshadowed by the competing ring-closing metathesis (RCM) pathway (see Scheme 2). In reactions involving prenylated 1,6-dienes, RCM is typically favoured due to entropic considerations and the formation of stable cyclic products. To overcome this inherent bias, the authors employed a sterically demanding diiodo ruthenium precatalyst, which selectively engages terminal olefins while

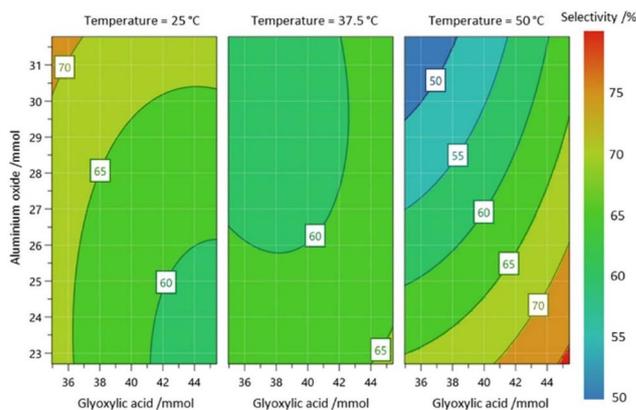


Fig. 2 Contour plot for the selectivity of the reaction forming the desired 3,4-dihydroxymandelic acid intermediate. The data, originally reported by Minisci and co-workers, were utilized to refit the model and generate the response surface, which was plotted using MODDE Pro. Reprinted with permission from the American Chemical Society.¹⁰

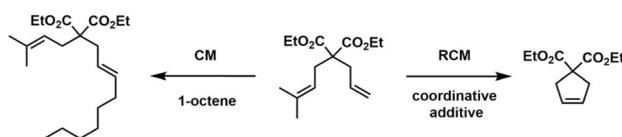


disfavouring internal alkene coordination. This catalyst design provided a promising foundation for cross-metathesis selectivity, but the reaction's complexity, driven by multiple interdependent variables, necessitated a more systematic approach to optimization.

The authors implemented a multivariate DoE strategy using MODDE software. This allowed them to simultaneously evaluate the effects and interactions of three key parameters: catalyst loading (1–3 mol%), substrate concentration (0.1–0.4 M), and equivalents of Estragole (2–15). The use of DoE was instrumental in revealing not only the individual contributions of each variable but also their synergistic effects on reaction conversion, as shown in Fig. 3. The statistical model identified substrate concentration on and catalyst loading as the most influential factors, while the number of equivalents of Estragole substrate exhibited a more nuanced, context-dependent impact. Under the predicted optimal conditions (0.31 M substrate concentration, 8.8 equivalents of Estragole, and 3 mol% catalyst) the reaction yielded 68% CM product, a significant improvement over conditions derived from OFAT methods.

Beyond the initial model system, the optimized protocol was successfully extended to other terpenes, including ocimene, *trans*- β -farnesene, and β -citronellene, demonstrating the robustness and generalizability of the DoE-guided approach. In each case, selective CM was achieved with minimal RCM byproduct formation, validating both the catalyst design and the statistical framework. This study exemplifies how DoE can serve as a strategic asset in synthetic chemistry, particularly in systems where subtle steric and electronic factors can dramatically alter reaction outcomes.

Optimization of reaction conditions utilizing DoE offers significant advantages over traditional approaches. By employing predefined, space-filling experimental designs, DoE eliminates the need for heuristic, intuition-based optimization strategies and has repeatedly proven to be a more efficient methodology. As illustrated in Fig. 4, this approach, in contrast to conventional OFAT studies, enables the development of robust statistical models that comprehensively characterize the chemical process across the entire experimental domain. This methodology is particularly advantageous for reaction prediction, as it allows for the generation of response surface contours, providing a detailed representation of parameter interactions and facilitating process optimization. A key limitation of DoE is the challenge of incorporating categorical variables, as these experimental designs are inherently suited for continuous parameters. One approach to address this issue is to represent categorical variables, such as solvent or catalyst choice, using appropriate continuous descriptors.³⁶ For



Scheme 2 Possible olefin metathesis reaction pathways (CM and RCM) for the selected substrate.

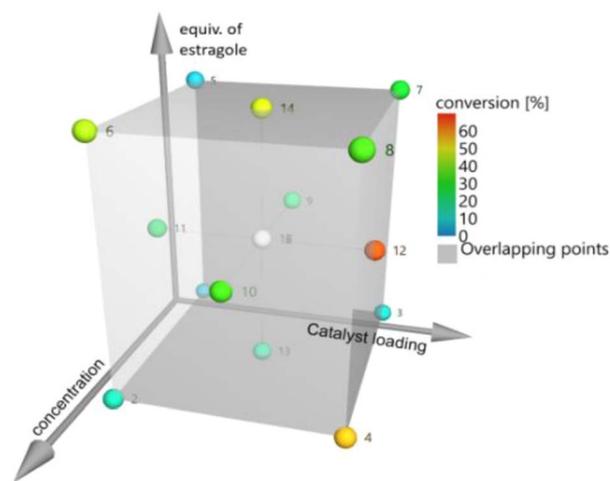


Fig. 3 Exploration of the reaction parameter space for the cross-metathesis of β -myrcene with Estragole substrate, where each sphere corresponds to an individual experiment and is color-coded according to its conversion yield. Reprinted with permission from Wiley-VCH.³⁵

instance, solvents can be encoded through physicochemical properties like polarity index or dielectric constant. These descriptors can then be mapped onto real-world categorical selections, enabling their integration into the DoE framework.

DoE has been widely employed in the optimization of chemical processes, particularly within the pharmaceutical and fine chemical industries. While DoE is commonly applied to improve product yield^{37–48} and purity,^{49–53} its utility extends to drug formulation^{54–59} and delivery,^{60–62} analytical method development,^{63–66} and various other applications.^{67–70} The increasing availability of user-friendly software and growing recognition within the chemical community have contributed to a significant rise in the adoption of this statistical approach in recent years.

3. Synergy of experiments and computational chemistry

One of the most transformative developments in the field of reaction optimization has been the emergence of

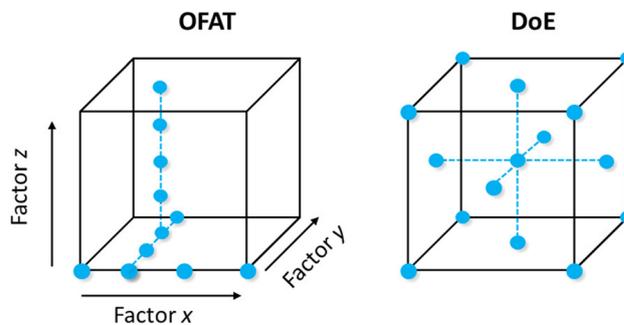


Fig. 4 Comparison of parameter space exploration between a conventional OFAT optimization and a DoE approach, where each point represents an individual experimental run. Reprinted with permission from the American Chemical Society.¹⁰



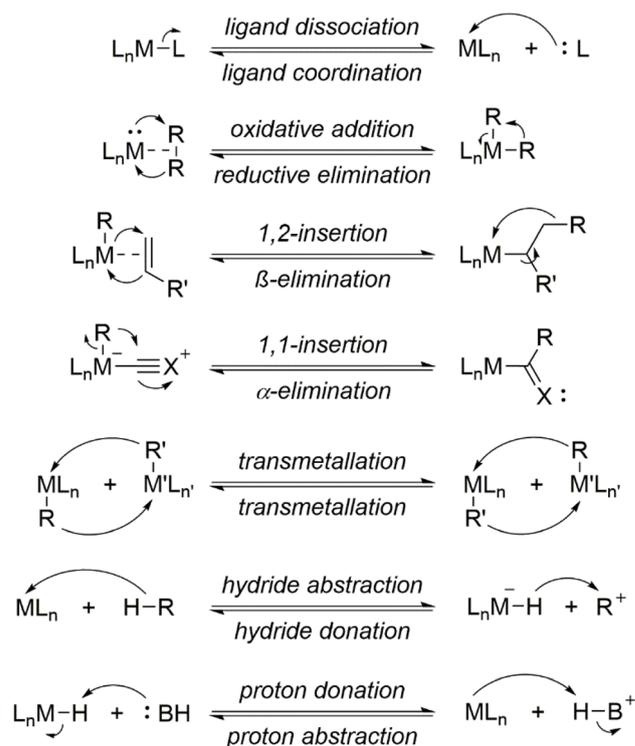
computational chemistry. This discipline leverages advanced methodologies to extract fundamental insights into chemical systems by employing models derived from both empirical data and first-principles quantum mechanics. These computational models encompass a range of theoretical approaches, each tailored to the specific system under investigation and the properties of interest, thereby offering varying degrees of accuracy and complexity.⁷¹

Computational chemistry methodologies have become indispensable for understanding catalytic processes, as they enable the detailed characterization of reaction mechanisms composed of consecutive elementary steps.^{72,73} Such approaches have been widely applied in different contexts, including organocatalysis as well as acid- and base-catalysed reactions. Organometallic catalysis poses unique challenges: the complexity of metal–substrate interactions and the diversity of accessible pathways make it especially well-suited to computational analysis.⁷⁴

In organometallic catalytic cycles, reactive intermediates generated in one step usually undergo rapid transformation in the subsequent one, which hinders their experimental isolation and characterization. Elementary reactions represent the most fundamental level of a chemical process, proceeding without the formation of stable intermediates between reactants and products.⁷⁵ In these transformations, all molecular modifications including bond formation and cleavage, as well as ligand coordination and dissociation occur within a single mechanistic step. As a result, the characterization of elementary reactions is relatively straightforward, as they can be elucidated by directly comparing the structures of the reactants and products. Scheme 3 illustrates several key elementary reactions commonly encountered in organometallic complexes.

While elementary reactions are inherently simple, their combination can give rise to a wide range of complex transformations. Thus, elucidating a catalytic cycle requires identifying the precise sequence of elementary steps that convert substrates into the final product. Given the inherent versatility of organometallic complexes, multiple reaction pathways are often feasible for the same set of intermediates, leading to competing mechanisms that may converge to the same product or diverge into different outcomes. In many cases, multiple plausible pathways can rationalize the observed reactivity, complicating the determination of the operative mechanism. Therefore, a comprehensive analysis of each species involved in the catalytic cycle is essential for achieving a detailed understanding of the system's reactivity.⁷⁶

Computational methodologies, such as DFT, are used to investigate reaction pathways by analysing both intermediates and transition states. These calculations provide molecular-level insights into the energetics and structural evolution of key steps, complementing experimental observations and helping to validate proposed mechanisms.⁷⁷ Intermediates represent stable species that form and evolve throughout the reaction, while transition states correspond to the highest-energy molecular configurations along the transformation from reactants to products. Due to their inherently unstable nature, transition states are nearly impossible to characterize



Scheme 3 Representative examples of fundamental elementary reactions in organometallic chemistry.

experimentally. However, recent advances in spectroscopic techniques such as double resonance spectroscopy and time-resolved pump-probe methods, have begun to challenge this notion. As demonstrated by Kim *et al.*,^{78,79} it is now possible to extract structural and dynamic information near the transition state region, even in relatively complex systems, by exploiting vibrational coupling and selective excitation schemes. These developments mark a turning point in our ability to probe the elusive topography of the reaction coordinate.

Moreover, the experimental identification of all intermediates involved in a catalytic reaction, as well as the transformations they undergo, is often challenging and, in many cases, unfeasible. This difficulty arises because most catalytic intermediates do not accumulate in detectable quantities at any stage of the reaction; instead, they are rapidly consumed immediately after their formation. An exception occurs when an intermediate precedes a reaction step with a sufficiently high effective energy barrier. In such cases, the transformation of this species progresses at a slower rate than its formation, leading to its accumulation in concentrations sufficient for detection. Conversely, an intermediate situated before a rate-limiting step may undergo alternative, lower-energy processes, such as the reverse reaction or side pathways that do not lead to product formation. The species interconverted through these faster processes typically exist in quasi-equilibrium, with the most thermodynamically stable among them constituting the predominant form of the free catalyst. These species, known as resting states, can often be experimentally detected. However, while resting states provide valuable insights into the overall



reactivity of the system, they do not necessarily reveal the rate-determining step, as multiple faster reactions may separate the experimentally observed species from the actual bottleneck of the catalytic cycle.⁸⁰

In the energetic span model introduced by Kozuch and Shaik,^{81,82} the concept of a rate-determining state replaces the traditional notion of a rate-determining step. Rather than focusing on a single slow step, the model identifies two key states that govern the turnover frequency (TOF) of a catalytic cycle: the TOF-determining transition state (TDTS) and the TOF-determining intermediate (TDI). These are not necessarily the highest or lowest energy points in the cycle, nor are they always adjacent (see Fig. 5). The energetic span (δE), defined by the energy difference between TDTS and TDI (with a correction for reaction free energy if needed), serves as the apparent activation energy of the entire cycle. This paradigm shift emphasizes that catalytic efficiency is shaped by the interplay of specific states rather than isolated steps, offering a more accurate and predictive framework for analysing and designing catalysts.

Moreover, certain reactions that do not directly influence the overall catalytic rate can still play a critical role in determining catalyst activity, particularly those governing selectivity. Within the complex network of interconnected mechanistic pathways, key branching points emerge, each leading to distinct products. The stereoelectronic properties of the intermediates at these junctions dictate the preferred reaction pathway, with significant kinetic differences between competing routes manifesting as measurable selectivity. Importantly, the detection of a branching intermediate does not in itself provide information on selectivity, since selectivity depends on the competition between subsequent reaction paths rather than on the mere existence of the common precursor. Furthermore, the experimental observability of intermediates is governed by their thermodynamic stability and their position relative to the rate-determining step, meaning that crucial species at selectivity-determining junctions may not always be experimentally accessible.

Experimental investigations of catalytic cycles often extend beyond the direct study of the reaction itself, incorporating complementary experiments to gain mechanistic insight. One commonly employed approach involves the stepwise, stoichiometric addition of individual substrates to the catalyst, followed by characterization of the resulting intermediates. This strategy

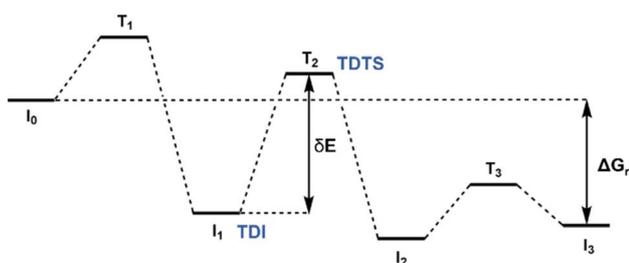


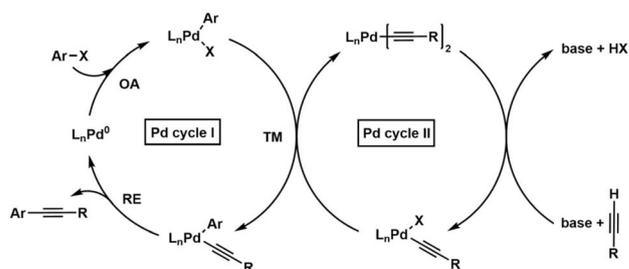
Fig. 5 Energy profile of a model catalytic cycle which shows that combination of I1 and T2 maximize δE and these states result in the TDI and TDTS.

aims to elucidate the coordination and activation modes of these species.

A compelling recent example of mechanistic dissection in catalysis is provided by Ivančić *et al.*,⁸³ who investigated the copper-free Sonogashira reaction, a palladium-catalysed cross-coupling between aryl halides and terminal alkynes, by breaking down the proposed bimetallic Pd–Pd catalytic cycle into elementary steps and studying them independently under synthetically relevant conditions (see Scheme 4). Traditionally, the Sonogashira reaction involves a dual Pd–Cu system, but the copper-free variant, often referred to as the Heck–Cassar alkylation, offers notable advantages such as reduced side reactions (*e.g.*, alkyne homocoupling), improved environmental compatibility, and operational simplicity. However, its mechanism remains controversial, particularly regarding the nature of the alkyne activation step: whether it proceeds *via* direct ligand exchange or a true transmetalation between two palladium species.

Ivančić *et al.* provide compelling evidence in favour of a bimetallic Pd–Pd mechanism involving transmetalation. They synthesized and characterized key organometallic intermediates including palladium monoacetylides, bisacetylides, and oxidative addition complexes, and compared the kinetics of individual steps such as oxidative addition, transmetalation, and reductive elimination. Their kinetic analysis revealed that transmetalation is likely the rate-determining step, challenging previous assumptions of a monometallic pathway. This work not only clarifies the mechanistic landscape of copper-free Sonogashira reactions but also highlights the utility of isolating and studying intermediates to resolve longstanding mechanistic ambiguities.

Another powerful tool for probing catalytic mechanisms is substrate variation. The use of substrates with distinct steric and electronic properties such as variations in bulkiness, nucleophilicity, or the presence of functional groups allows researchers to assess their impact on catalytic rate and selectivity.⁸⁴ Likewise, isotopically labelled substrates facilitate the measurement of kinetic isotope effects (KIE), offering crucial insights into the nature of transformations occurring during the rate-limiting step.⁸⁵ For instance, in 2008, Fristrup *et al.*⁸⁶ investigated the mechanism of a rhodium-catalysed decarbonylation of aldehydes through a combined computational and experimental approach. The authors proposed a catalytic



Scheme 4 Dual synergistic mechanistic proposal. *Cis–trans* isomerization steps are omitted for clarity. OA oxidative addition, TM transmetalation, RE reductive elimination.



cycle comprising three key steps: (I) oxidative addition of the aldehyde C–H bond to the rhodium centre, forming a rhodium-acyl intermediate; (II) migratory extrusion of carbon monoxide; and (III) reductive elimination to release the hydrocarbon product and regenerate the active catalyst (see Scheme 5).

To probe the selectivity-determining step, they performed competition experiments between benzaldehyde and its deuterated analogue (benzaldehyde- d_1), yielding a KIE of 1.77. This value indicates that the C–H bond is at least partially broken in the transition state of the selectivity-determining step, but not necessarily in the rate-limiting step. DFT calculations complemented these findings by modelling the full catalytic cycle and evaluating the energy profiles of each elementary step. The computational data revealed that the migratory extrusion of CO is the highest-energy transition state, suggesting it is the rate-determining step. Importantly, the calculated KIE values based on this mechanistic model closely matched the experimental results, reinforcing the validity of the proposed pathway.

Another frequently employed diagnostic approach involves examining the reactivity of complexes that mimic proposed intermediates, providing direct evidence for their involvement in the catalytic mechanism. In their study of the copper-free Sonogashira reaction, Gazvoda *et al.*⁸⁷ synthesized and isolated key palladium complexes believed to participate in the catalytic cycle (see Scheme 4). By mixing these species under controlled conditions and observing the formation of the cross-coupled product, they demonstrated that transmetalation between two palladium complexes is not only feasible but likely operative under catalytic conditions. This finding strongly supports their proposed Pd/Pd dual synergistic mechanism and challenges the previously accepted monometallic pathway, which does not involve transmetalation.

This study complements the work of Ivančić *et al.*⁸³ who also dissected the copper-free Sonogashira reaction into elementary steps and investigated them independently. While Ivančić *et al.* focused on kinetic analysis of each step, Gazvoda *et al.* emphasized the direct reactivity of isolated intermediates. Both studies converge on the same mechanistic conclusion: that a dual synergistic Pd–Pd pathway involving transmetalation is

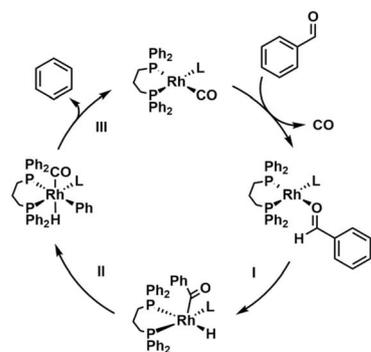
operative. However, their approaches differ in emphasis, kinetic dissection *versus* stoichiometric reactivity testing, providing orthogonal lines of evidence that reinforce the plausibility of the dual catalytic cycle.

Despite the breadth of available experimental methodologies, mechanistic studies often yield incomplete information, making it difficult to unambiguously determine all intermediates and their respective roles within the catalytic cycle. In fact, reaction mechanisms are inherently postulated models, constructed to rationalize all available evidence, but rarely validated with absolute certainty. The proposed pathway must align with kinetic data, spectroscopic observations, and reactivity profiles, yet remains a best-fit hypothesis rather than an unequivocal truth. Furthermore, precise quantification of molecular properties and reaction energetics remains inherently challenging for experimental techniques. As a result, mechanistic investigations frequently provide only partial reaction pathways and primarily qualitative insights into the effects of various factors on catalytic activity and selectivity. This limitation can hinder the formulation of well-supported mechanistic conclusions, particularly in complex systems where multiple interrelated variables influence reactivity. Consequently, achieving a comprehensive understanding of catalytic processes benefits from an integrated approach that combines experimental findings with computational and theoretical analyses. While each perspective can independently yield valuable insights, their combination often provides a more robust and nuanced picture, helping to bridge knowledge gaps and validate mechanistic hypotheses from complementary angles.

In this context, computational chemistry methodologies enable a comprehensive theoretical analysis of organometallic catalysts. These approaches allow the determination of geometries, relative energies, and other key properties of all species potentially involved in a catalytic cycle. By comparing the energetics of intermediates and transition states, computational studies not only identify the most favourable reaction pathway but also uncover alternative routes, including those leading to competing products. This provides quantitative insight into both reactivity and selectivity, with mechanistic interpretations grounded in molecular structures, bond lengths, and stereoelectronic effects. However, the systematic construction and evaluation of such complex reaction networks, often involving numerous intermediates and transition states, remains a challenging and time-consuming task.

An illustrative example of a reaction mechanism unveiled by means of computational and experimental techniques is shown in Fig. 6. Species in blue are experimentally detectable, including the catalyst, the intermediate obtained when only alkene is added to the reaction mixture, the intermediate before the highest activation energy in the whole reaction mechanism, and the product obtained. The reaction path in black is computationally obtained derived from observations, and reaction paths in grey are computationally determined but are not based on experiments.

Computational studies facilitate the identification of species that govern reaction rates and selectivity, as well as the energy differences between them, providing deeper insights into



Scheme 5 Proposed catalytic cycle for the rhodium-catalysed decarbonylation of aldehydes composed of the following steps: (I) oxidative addition, (II) migratory extrusion, and (III) reductive elimination.



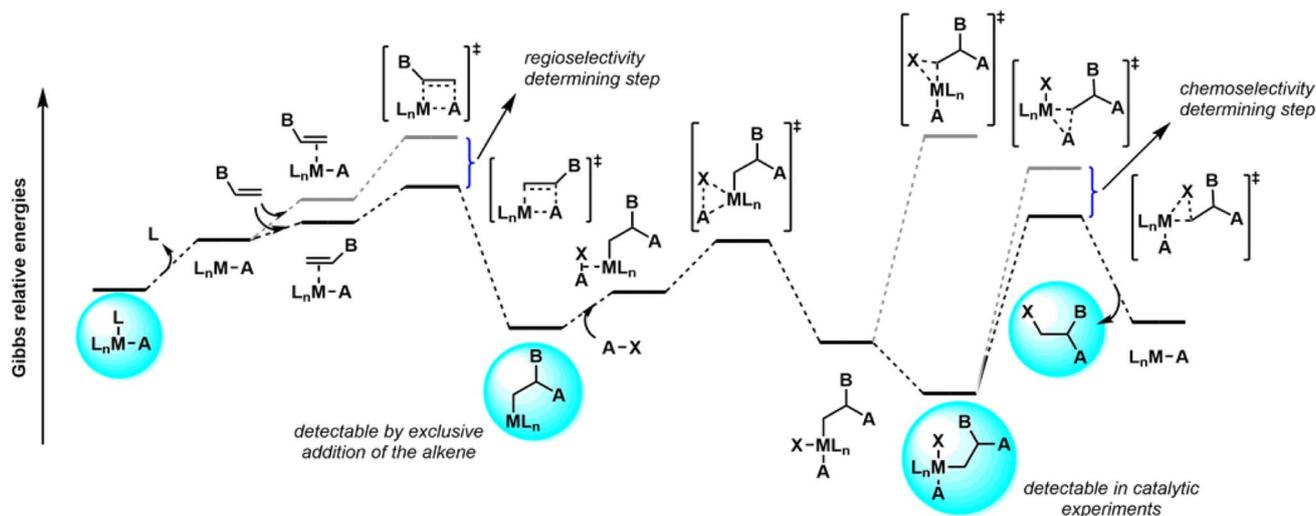


Fig. 6 A model example of a computationally and experimentally unveiled reaction mechanism.

catalytic activity. Moreover, these methodologies can be applied to various substrates, allowing for the comparative assessment of catalytic efficiency. The quantitative nature of the results, coupled with the characterization of key reactive species, enhances our understanding of the factors influencing catalysis. Advanced computational techniques, such as DFT, natural bond orbital (NBO),⁸⁸ quantum theory of atoms in molecules (QTAIM),⁸⁹ buried volume ($\%V_{\text{Bur}}$),⁹⁰ non-covalent interactions (NCI) plots,⁹¹ and Mayer bond order (MBO) analysis,⁹² provide a detailed description of the electronic structure and interactions within these species, aiding in the identification of features that define catalytic behaviour.⁹³ Each method offers complementary insights: DFT yields optimized geometries and energy profiles; NBO analysis reveals donor-acceptor interactions and charge delocalization; QTAIM characterizes bond critical points and electron density topology; $\%V_{\text{Bur}}$ quantifies the steric environment around the metal centre;⁹⁴ NCI plots visualize weak interactions such as hydrogen bonding or van der Waals forces; and MBO provides a numerical estimate of bond strength. Together, these tools enable a multifaceted understanding of the factors governing reactivity and selectivity in catalytic systems.

As a result, computational approaches have become indispensable for elucidating catalytic mechanisms. Their application not only aids in interpreting experimental observations but also informs the rational design of new catalysts by highlighting structural and electronic factors that promote or hinder specific transformations.^{95,96} However, a key challenge in computational mechanistic studies is the necessity of thoroughly exploring all possible reaction pathways to ensure reliable conclusions. Unlike experimental studies, where the observed species is typically the most thermodynamically or kinetically favoured, computational analyses risk overlooking lower energy species or alternative pathways, potentially leading to inaccurate interpretations. Given the complex reactivity of organometallic systems and the diverse molecular fragments involved,

comprehensive exploration of reaction networks can be computationally demanding.⁹⁷

To address this limitation, computational predictions must be validated and guided by experimental data. Experimental benchmarks such as activation energies, selectivity trends, and the identification of key intermediates provide essential constraints for theoretical models. Exploratory experiments, including stoichiometric reactivity tests, can further support the relevance of computed intermediates and transition states.⁹⁸

Rather than treating computational and experimental approaches as separate domains, integrating them into a unified framework enhances the reliability and depth of mechanistic insights. Computational results can suggest new experiments, while experimental findings can refine theoretical models. This reciprocal feedback loop fosters a more accurate and comprehensive understanding of catalytic processes, enabling the prediction of reactivity under diverse conditions and the optimization of catalyst performance through informed structural modifications. This synergistic strategy where theory and experiment guide and validate one another is at the heart of predictive catalysis,⁹⁹ a paradigm that seeks not only to explain observed behaviour but to anticipate and design new reactivity with precision.¹⁰⁰ Actually, despite the importance of machine learning (ML), predictive catalysis will remain a key vehicle for generating new insights, as demonstrated in iron olefin metathesis,^{101–103} particularly when datasets cannot be sufficiently expanded for robust ML applications, as recent findings from some of us indicate.^{104,105} Thus, by extension, predictive chemistry remains a valuable tool.^{106,107}

4. Machine learning

In recent years, the convergence of chemistry and data science, particularly through the integration of ML has begun to reshape synthetic and mechanistic chemistry.^{108,109} This emerging field of predictive chemistry leverages ML to accelerate the deployment, development, and discovery of chemical reactions. While



the potential is immense, it is important to recognize that many of these approaches remain in their infancy. Their success often hinges on the availability of large, high-quality datasets, and their applicability is not yet universal across reaction types or chemical domains.

Coley and coworkers classify ML applications in synthetic chemistry into three key areas: reaction deployment, reaction development, and reaction discovery.¹¹⁰

Regarding reaction deployment, ML has transformed retrosynthetic planning. Traditional expert systems, which rely on hand-coded rules, are increasingly complemented or even replaced by data-driven models. Template-based approaches use predefined reaction patterns, such as SMARTS rules, to suggest disconnections in target molecules. In contrast, template-free models treat retrosynthesis as either a graph transformation problem or a natural language translation task. A notable example is the Molecular Transformer, which uses sequence-to-sequence learning on SMILES strings to predict reaction outcomes and retrosynthetic steps.¹¹¹

In reaction development, ML models assist in selecting reaction conditions such as solvents, catalysts, and ligands, tailored to specific substrates.^{112,113} These models can operate globally across diverse reaction types or locally within a single transformation class. One promising direction is substrate scope prediction, which is becoming more systematic through ML. By training on high-throughput experimental data, models can estimate whether a new substrate will succeed under given conditions. This capability moves the field closer to algorithmically generating scope tables, traditionally built through manual trial and error. A compelling example is the use of active learning in Ni/photoredox cross-coupling reactions, where ML models iteratively suggest the most informative substrates to test next, thereby optimizing the experimental design process.¹¹⁴

As a piece of example of the applications and the size of datasets, Corminboeuf and coworkers developed a broadly applicable strategy to derive predictive linear models for enantioselectivity using minimal experimental screening data in catalytic systems featuring bidentate ligands.¹¹² To assess the robustness of the methodology, datasets were assembled and analysed spanning four distinct reaction families, encompassing 100 bidentate ligands distributed across seven structural classes. These were further complemented by the BDL-Cu-2023 dataset to enable extended ligand optimization. The protocol identified the most informative linear relationships by integrating electronic and steric descriptors together with topological parameters extracted through Bayesian ridge regression (BRR). When combined with Bayesian optimization (BO), this framework permits efficient ligand exploration even under data-scarce conditions. This integration facilitates the identification of mechanistic trends and supports the rational proposal of new ligand candidates, as demonstrated for the oxy-alkynylation transformation.¹¹² This methodology provided a practical and scalable tool for ligand selection and refinement throughout the experimental process, thereby accelerating and streamlining the development of enantioselective catalytic reactions. On the other hand, in 2026, Martínez and coworkers through the

CatalySeed database gathered up to 768 curated Ru-catalyzed ethenolysis reactions to enable reproducible benchmarking and machine learning analysis. Using ROBERT,¹¹⁵ predictive models for TON and selectivity showed robust performance, highlighting Ru atomic charge and d-orbital character as key descriptors.

In the domain of reaction discovery and mechanistic insight, ML is increasingly capable of generating new chemical knowledge.^{116–118} Mechanistic understanding has traditionally relied on foundational kinetic frameworks, such as reaction progress kinetic analysis (RPKA), which utilizes *in situ* monitoring to extract substantial information from a minimal number of experiments,¹¹⁹ and variable time normalization analysis (VTNA), which allows for the visual determination of reaction orders through simple graphical overlays.¹²⁰

Building upon these mechanistic foundations, modern approaches now link experimental kinetics with data-driven workflows. For instance, Sigman and colleagues demonstrated how multidimensional correlations of physical organic descriptors can link mechanism directly to selectivity, providing a predictive framework for catalyst optimization.¹²¹ More recently, these efforts have culminated in fully integrated computational and data-enabled platforms, such as those used to elucidate complex organometallic mechanisms by combining DFT-derived data with ML-driven analysis.¹²²

Mechanistic inference models aim to identify plausible reaction pathways or intermediates,¹²³ while selectivity prediction models address regio-, stereo-, and site-selectivity in complex transformations.¹²⁴ ML also plays a role in catalyst and ligand design,¹²⁵ often using descriptor-based models or deep generative approaches such as variational autoencoders and genetic algorithms. For example, the ELECTRO model infers mechanistic steps by modelling electron flow in polar reactions, offering a data-driven alternative to traditional mechanistic reasoning.¹²⁶

Clustering algorithms enhance machine learning applications in chemistry by enabling efficient navigation of vast chemical spaces. These methods group similar molecules or reactions based on structural or property descriptors, reducing computational demands and prioritizing promising candidates for synthesis or screening. For instance, hierarchical clustering has been applied to identify diverse ligand sets for C–H functionalization catalysts, streamlining exploration of reactivity patterns.¹²⁷ In chemical reaction optimization, clustering facilitates scalable analysis of reaction outcomes across large datasets. Recent work demonstrates density-based clustering to partition reaction spaces by yield and selectivity, guiding efficient iterative experimentation.¹²⁸

Despite these advances, it is crucial to emphasize that predictive chemistry *via* ML is not yet a turnkey solution. Many models are limited by the scope and quality of available data, and their generalizability across chemical space remains a challenge. Continued progress will depend on improved data curation, model interpretability, and integration with experimental workflows.



5 Conclusions

The optimization of chemical reactions has evolved from intuition-driven experimentation to data-rich, predictive science. Traditional approaches such as OFAT, while historically valuable, are increasingly supplanted by statistically grounded methodologies like DoE that maximize information gain and reproducibility. When coupled with computational chemistry, these approaches extend beyond empirical optimization to mechanistic understanding, enabling the rationalization of selectivity and catalytic efficiency at the molecular level.

The integration of ML further expands this paradigm by providing models capable of predicting reactivity, selectivity, and scope across diverse chemical spaces. However, the effectiveness of these models depends critically on the quality and diversity of training data, as well as their interpretability and connection to chemical theory. The future of reaction optimization lies in the synergistic convergence of experiments, computation, and ML, forming a closed feedback loop where theory guides experimentation, data refine models, and models accelerate discovery.

Achieving this vision will require closer collaboration across disciplines, from chemists and data scientists to engineers and theoreticians, and continued development of standardized, open-access datasets. As predictive chemistry matures, it will enable not only more efficient and sustainable processes but also a deeper, more unified understanding of chemical reactivity itself.

Author contributions

All authors contributed to the conceptualization and preparation of the manuscript.

Conflicts of interest

There are no conflicts to declare.

Data availability

No new data was generated.

Acknowledgements

We are grateful for the financial support from the Spanish Ministerio de Ciencia, Innovación y Universidades (Projects PID2024-155989NB-I00, PID2023-146849NB-I00/MICIU/AEI/10.13039/501100011033/FEDER, UE, RED2022-134939 T and PID-2021-127423NB-I00, predoctoral FPU20/00707 grant to R. M.-C., and the Generalitat de Catalunya (Project 2021-SGR-623).

References

- 1 E. J. Corey and X.-M. Cheng, *The Logic of Chemical Synthesis*, Wiley, New York, 1989.
- 2 K. C. Nicolaou, H. J. Vourloumis, N. Winssinger and P. S. Baran, *Angew. Chem., Int. Ed.*, 2000, **39**, 44–122.
- 3 S. D. Roughley and A. M. Jordan, *J. Med. Chem.*, 2011, **54**, 3451–3479.
- 4 O. V. Kharissova, B. I. Kharisov, C. M. Oliva González, Y. Peña Méndez and I. López, *R. Soc. Open Sci.*, 2019, **6**, 191378.
- 5 S. Posada-Pérez, A. Vidal-López, A. Comas-Vives and A. Poater, *npj Mater. Sustain.*, 2026, DOI: [10.1038/s44296-026-00100-3](https://doi.org/10.1038/s44296-026-00100-3).
- 6 (a) Y. Wang, H.-Y. Lu and P. F. Xu, *Acc. Chem. Res.*, 2015, **48**, 1832–1844; (b) N. A. Afagh and A. K. Yudin, *Angew. Chem., Int. Ed.*, 2010, **49**, 262–310; (c) R. A. Shenvi, D. P. O'Malley and P. S. Baran, *Acc. Chem. Res.*, 2009, **42**, 530–541.
- 7 B. M. Trost, The Atom Economy—A Search for Synthetic Efficiency, *Science*, 1991, **254**, 1471–1477.
- 8 R. H. Grubbs, Olefin-Metathesis Catalysts for the Preparation of Molecules and Materials, *Angew. Chem., Int. Ed.*, 2006, **45**, 3760–3765.
- 9 J. F. Hartwig, *Chem. Soc. Rev.*, 2011, **40**, 1992–2002.
- 10 C. J. Taylor, A. Pomberger, K. C. Felton, R. Grainger, M. Barecka, T. W. Chamberlain, R. A. Bourne, C. N. Johnson and A. A. Lapkin, *Chem. Rev.*, 2023, **123**, 3089–3126.
- 11 J. Andraos, Designing a green organic chemistry lecture course, in *Green Organic Chemistry in Lecture and Laboratory*, CRC Press, Boca Raton, FL, 1st edn, 2012, pp. 29–68.
- 12 D. Florez-Orrego, S. Sharma and S. Navabi, *Front. Chem. Eng.*, 2022, **4**, 961022.
- 13 D. W. Lendrem, B. C. Lendrem, D. Woods, R. Rowland-Jones, M. Burke, M. Chatfield, J. D. Isaacs and M. R. Owen, *Drug Discov. Today*, 2015, **20**, 1365–1371.
- 14 J. Boström, D. G. Brown, R. J. Young and G. M. Keserü, *Nat. Rev. Drug Discovery*, 2018, **17**, 709–727.
- 15 P. M. Murray, F. Bellany, L. Benhamou, D.-K. Bučar, A. B. Tabor and T. D. Sheppard, *Org. Biomol. Chem.*, 2016, **14**, 2373–2384.
- 16 D. Lendrem, M. Owen and S. Godbert, *Org. Process Res. Dev.*, 2001, **5**, 324–327.
- 17 V. K. Aggarwal, A. C. Staubitz and M. Owen, *Org. Process Res. Dev.*, 2016, **10**, 64–69.
- 18 M. D. Peris-Diaz, M. A. Sentandreu and E. Sentandreu, *Anal. Bioanal. Chem.*, 2018, **410**, 4275–4284.
- 19 F. Sagués and I. R. Epstein, *Dalton Trans.*, 2003, 1201–1217.
- 20 M. R. Owen, C. Luscombe, L.-W. Lai, S. Godbert, D. L. Crookes and D. Emiabata-Smith, *Org. Process Res. Dev.*, 2001, **5**, 308–323.
- 21 Z. Wahid and N. Nadir, *World Appl. Sci. J.*, 2013, **21**, 56–61.
- 22 J. Boström, D. G. Brown, R. J. Young and G. M. Keserü, *Nat. Rev. Drug Discovery*, 2018, **17**, 709–727.
- 23 P. M. Murray, F. Bellany, L. Benhamou, D.-K. Bučar, A. B. Tabor and T. D. Sheppard, *Org. Biomol. Chem.*, 2016, **14**, 2373–2384.
- 24 S. A. Weissman and N. G. Anderson, *Org. Process Res. Dev.*, 2015, **19**, 1605–1633.



- 25 P. Elliott, S. Billingham, J. Bi and H. Zhang, *Pharm. Bioprocess.*, 2013, **1**, 105–122.
- 26 H. Tye, *Drug Discov. Today*, 2004, **9**, 485–491.
- 27 A. S. Dhoot, G. J. Fernandes, A. Naha, M. Rathnanand and L. Kumar, *Pharm. Chem. J.*, 2019, **53**, 730–735.
- 28 J. J. Douglas, *Nat. Chem.*, 2019, **11**, 683–684.
- 29 M. L. Schrader, F. R. Schäfer, F. Schäfers and F. Glorius, *Nature*, 2024, **16**, 471–498.
- 30 S. E. Denmark and C. R. Butler, *J. Am. Chem. Soc.*, 2008, **130**, 3690–3704.
- 31 MODDE – Explore. Improve. Advance!, Sartorius AG, 2022, <https://www.sartorius.com/en/products/process-analytical-technology/data-analytics-software/dae-software/modde>, accessed 2025-03-10.
- 32 Statistical software for insight-driven improvement: Seeing is understanding., JMP Statistical Discovery LLC, 2023, https://www.jmp.com/en_us/home.html, accessed 2025-03-10.
- 33 Design Expert, version 13., STAT-EASE, 2023, <https://www.statease.com/software/design-expert/>, accessed 2025-03-10.
- 34 H.-R. Bjørsvik, L. Liguori and F. Minisci, *Org. Process Res. Dev.*, 2000, **4**, 534–543.
- 35 K. Iudanov, N. B. Nechmad, A. Poater and N. G. Lemcoff, *Angew. Chem., Int. Ed.*, 2024, **63**, e202412430.
- 36 M. Tomasini, L. Caporaso, M. Gimferrer and A. Poater, *Coord. Chem. Rev.*, 2026, **550**, 217383.
- 37 N. Holmes, G. R. Akien, R. J. Savage, C. Stanetty, I. R. Baxendale, A. J. Blacker, B. A. Taylor, R. L. Woodward, R. E. Meadows and R. A. Bourne, *React. Chem. Eng.*, 2016, **1**, 96–100.
- 38 N. Holmes, G. R. Akien, A. J. Blacker, R. L. Woodward, R. E. Meadows and R. A. Bourne, *React. Chem. Eng.*, 2016, **1**, 366–371.
- 39 S. J. Mear and T. F. Jamison, *J. Org. Chem.*, 2019, **84**, 15001–15007.
- 40 B. Taylor, C. Fernandez Barrat, R. L. Woodward and P. A. Inglesby, *Org. Process Res. Dev.*, 2017, **21**, 1404–1412.
- 41 S. A. Raw, B. A. Taylor and S. Tomasi, *Org. Process Res. Dev.*, 2011, **15**, 688–692.
- 42 N. L. Bell, M. Kupper and L. Cronin, *Chem. Mater.*, 2021, **33**, 7263–7271.
- 43 H. Tye and M. Whittaker, *Org. Biomol. Chem.*, 2004, **2**, 813–815.
- 44 L. A. Hobson, O. Akiti, S. S. Deshmukh, S. Harper, K. Katipally, C. J. Lai, R. C. Livingston, E. Lo, M. M. Miller, S. Ramakrishnan, S. B. Collins, P. J. Hutchinson, J. D. Smith, M. A. Green and A. R. Wilson, *Org. Process Res. Dev.*, 2010, **14**, 441–458.
- 45 H. Hajmowicz, J. Wisialski and L. Synoradzki, *Org. Process Res. Dev.*, 2011, **15**, 427–434.
- 46 G. Guercio, A. Perboni, F. Tinazzi, L. Rovatti and S. Provera, *Org. Process Res. Dev.*, 2010, **14**, 840–848.
- 47 R. C. Roemmele and M. A. Christie, *Org. Process Res. Dev.*, 2013, **17**, 422–426.
- 48 J. D. Moseley, W. O. Moss, M. J. Welham, C. L. Ancell, J. Banister, S. A. Bowden, G. Norton and M. J. Young, *Org. Process Res. Dev.*, 2003, **7**, 58–66.
- 49 V. Klöpfer, R. Eckl, J. Floß, P. M. Roth, O. Reiser and J. P. Barham, *Green Chem.*, 2021, **23**, 6366–6372.
- 50 N. K. Adlington, L. R. Agnew, A. D. Campbell, R. J. Cox, A. Dobson, C. F. Barrat, M. A. Gall, W. Hicks, G. P. Howell, A. Jawor-Baczynska, J. T. Jones, L. K. Smith, R. P. Taylor and S. M. Williams, *J. Org. Chem.*, 2019, **84**, 4735–4747.
- 51 J. ten Dam, F. Kapteijn, K. Djanashvili and U. Hanefeld, *Catal. Commun.*, 2011, **13**, 1–5.
- 52 V. Arunajatesan, S. L. McIntosh, M. Cruz, R. F. Renneke and B. Chen, in *Catalysis of Organic Reactions*, CRC Press, 2005, pp. 219–224.
- 53 F. J. Weiberth, H. S. Gill, Y. Jiang, G. E. Lee, P. Lienard, C. Pemberton, M. R. Powers, W. Subotkowski, W. Tomasik, B. J. Vanasse and Y. Yu, *Org. Process Res. Dev.*, 2010, **14**, 623–631.
- 54 N. A. Charoo, A. A. Shamsher, A. S. Zidan and Z. Rahman, *Int. J. Pharm.*, 2012, **423**, 167–178.
- 55 M. Naeem, N. U. Rahman, J. A. Khan, A. Sehti and Z. Nawaz, *Lat. Am. J. Pharm.*, 2013, **32**, 1196–1204.
- 56 P. M. Kumar and A. Ghosh, *Eur. J. Pharm. Sci.*, 2017, **96**, 243–254.
- 57 M. Malladi and R. Jukanti, *J. Drug Delivery Sci. Technol.*, 2016, **35**, 134–145.
- 58 P.-F. Chavez, P. Lebrun, P.-Y. Sacré, C. De Bleye, L. Netchacovitch, S. Cuyppers, J. Mantanus, H. Motte, M. Schubert, B. Evrard, P. Hubert and E. Ziemons, *Int. J. Pharm.*, 2015, **486**, 13–20.
- 59 R. Y. Rahali, P. Saulnier, J.-P. Benoit and Y. Bensouda, *J. Drug Deliv. Sci. Technol.*, 2010, **20**, 425–429.
- 60 A. Karatzas, S. Politis and D. Rekkas, *Drug Dev. Ind. Pharm.*, 2017, **43**, 770–779.
- 61 P. M. Kumar and A. Ghosh, *J. Drug Deliv. Sci. Technol.*, 2015, **30**, 15–29.
- 62 E. Maretti, C. Rustichelli, M. Romagnoli, A. G. Balducci, F. Buttini, F. Sacchetti, E. Leo and V. Iannuccelli, *Int. J. Pharm.*, 2016, **511**, 669–679.
- 63 L. Kumar, M. S. Reddy, R. S. Managuli and K. G. Pai, *Saudi Pharm. J.*, 2015, **23**, 549–555.
- 64 M. A. Abdelgawad, A. I. Owis and N. S. Abdelwahab, *Sep. Sci. Plus.*, 2018, **1**, 244–252.
- 65 S. Naik, P. Mullick, S. P. Mutalik, A. R. Hegde, S. A. Lewis, K. Bhat, B. S. S. Rao and S. Mutalik, *J. Chromatogr. Sci.*, 2022, **60**, 584–594.
- 66 S. Dai, B. Xu, G. Luo, J. Li, Z. Xue, X. Shi and Y. Qiao, *Open Chem. Eng. J.*, 2015, **9**, 47.
- 67 N. Patel, S. Jain, P. Madan and S. Lin, *Drug Dev. Ind. Pharm.*, 2016, **42**, 1894–1902.
- 68 M. A. Badawi and L. K. El-Khordagui, *Eur. J. Pharm. Sci.*, 2014, **58**, 44–54.
- 69 M. Shariare, M. De Matas, P. York and Q. Shao, *Int. J. Pharm.*, 2011, **408**, 58–66.



- 70 A. Al-Gheethi, E. Noman, R. M. S. Radin Mohamed, N. Ismail, A. H. Bin Abdullah and A. H. Mohd Kassim, *J. Hazard. Mater.*, 2019, **365**, 883–894.
- 71 K. N. Houk and F. Liu, *Acc. Chem. Res.*, 2017, **50**, 539–543.
- 72 J. N. Harvey, F. Himo, F. Maseras and L. Perrin, *ACS Catal.*, 2019, **9**, 6803–6813.
- 73 N. Fey and J. M. Lynam, *WIREs Comput Mol Sci.*, 2022, **12**, e1590.
- 74 A. Lledós, *Eur. J. Inorg. Chem.*, 2021, **2021**, 2547–2555.
- 75 Y. Lan, *Elementary Reactions in Organometallic Chemistry*, 2021, pp. 51–123.
- 76 G. N. Simm, A. C. Vaucher and M. Reiher, *J. Phys. Chem. A*, 2019, **123**, 385–399.
- 77 A. Brotons Rufes, S. Posada-Pérez and A. Poater, *Digit. Chem. Eng.*, 2026, **18**, 100285.
- 78 J. Kim, M. Kang, J.-H. Yoon and S. K. Kim, *Nat. Commun.*, 2025, **16**, 210.
- 79 J. H. Baraban and R. W. Field, *Nat. Commun.*, 2025, **16**, 76.
- 80 R. Poli, *Comments Inorg. Chem.*, 2009, **30**, 177–228.
- 81 S. Kozuch and J. M. L. Martin, *ACS Catal.*, 2012, **2**, 2787–2794.
- 82 S. Kozuch and S. Shaik, *Acc. Chem. Res.*, 2011, **44**, 101–110.
- 83 A. Ivančić, J. Košmrlj and M. Gazvoda, *Commun. Chem.*, 2023, **6**, 51.
- 84 B. Sawatlon, M. D. Wodrich and C. Corminboeuf, *Org. Lett.*, 2020, **22**, 7936–7941.
- 85 D. Lupp, N. J. Christensen and P. Fristrup, *Dalton Trans.*, 2014, **43**, 11093–11105.
- 86 P. Fristrup, M. Kreis, A. Palmelund, P.-O. Norrby and R. Madsen, *J. Am. Chem. Soc.*, 2008, **130**, 5206.
- 87 M. Gazvoda, M. Virant, B. Pinter and J. Košmrlj, *Nat. Commun.*, 2018, **9**, 4814.
- 88 F. Weinhold and C. R. Landis, *Chem. Educ. Res. Pract.*, 2001, **2**, 91–104.
- 89 R. F. W. Bader, *Chem. Rev.*, 1991, **91**, 893–928.
- 90 L. Falivene, Z. Cao, A. Petta, L. Serra, A. Poater, R. Oliva, V. Scarano and L. Cavallo, *Nat. Chem.*, 2019, **11**, 872–879.
- 91 J. Contreras-García, E. R. Johnson, S. Keinan, R. Chaudret, J.-P. Piquemal, D. N. Beratan and W. Yang, *J. Chem. Theory Comput.*, 2011, **7**, 625–632.
- 92 I. Mayer, *Chem. Phys. Lett.*, 1983, **97**, 270–274.
- 93 J. Munarriz, E. Velez, M. A. Casado and V. Polo, *Phys. Chem. Chem. Phys.*, 2018, **20**, 1105–1113.
- 94 S. Escayola, N. Bahri-Laleh and A. Poater, *Chem. Soc. Rev.*, 2024, **53**, 853–882.
- 95 S. Santoro, M. Kalek, G. Huang and F. Himo, *Acc. Chem. Res.*, 2016, **49**, 1006–1018.
- 96 S. Ahn, M. Hong, M. Sundararajan, D. H. Ess and M. H. Baik, *Chem. Rev.*, 2019, **119**, 6509–6560.
- 97 S. Ahn, M. Hong, M. Sundararajan, D. H. Ess and M. H. Baik, *Chem. Rev.*, 2019, **119**, 6509–6560.
- 98 R. D. Baxter, D. Sale, K. M. Engle, J. Q. Yu and D. G. Blackmond, *J. Am. Chem. Soc.*, 2012, **134**, 4600–4606.
- 99 R. Monreal-Corona, A. Pla-Quintana and A. Poater, *Trends Chem.*, 2023, **5**, 935–946.
- 100 T. Sperger, I. A. Sanhueza and F. Schoenebeck, *Acc. Chem. Res.*, 2016, **49**, 1311–1319.
- 101 A. Poater, S. V. C. Vummaleti, E. Pump and L. Cavallo, *Dalton Trans.*, 2014, **43**, 11216–11220.
- 102 A. Brotons Rufes, J. P. Martínez, S. P. Nolan, S. Posada-Pérez and A. Poater, Challenges in olefin metathesis: past, present and future, *Coord. Chem. Rev.*, 2025, **542**, 216827.
- 103 A. Brotons Rufes, J. P. Martínez, N. Joly, S. Gaillard, J.-L. Renaud, S. Posada Pérez and A. Poater, *ChemCatChem*, 2025, **17**, e00570.
- 104 R. Monreal-Corona, À. Díaz-Jiménez, A. Roglans, A. Poater and A. Pla-Quintana, *Adv. Synth. Catal.*, 2023, **365**, 760–766.
- 105 S. Dehghani, S. Sadjadi, N. Bahri-Laleh, M. Nekoomanesh-Haghighi and A. Poater, *Appl. Organomet. Chem.*, 2019, **33**, e4891.
- 106 M. Tomasini, M. Voccia, L. Caporaso, M. Szostak and A. Poater, Tuning the steric hindrance of alkylamines: a predictive model of steric editing of planar amines, *Chem. Sci.*, 2024, **15**, 13405–13414.
- 107 M. Tomasini, J. Zhang, H. Zhao, E. Besalú, L. Falivene, L. Caporaso, M. Szostak and A. Poater, A predictive journey towards trans-thioamides/amides, *Chem. Commun.*, 2022, **58**, 9950–9953.
- 108 T. Toyao, Z. Maeno, S. Takakusagi, T. Kamachi, I. Takigawa and K. Shimizu, *ACS Catal.*, 2020, **10**, 2260–2297.
- 109 E. Casillo, T. Scattolin and S. P. Nolan, *Chem. Commun.*, 2025, **61**, 18247–18272.
- 110 Z. Tu, T. Stuyver and C. W. Coley, *Chem. Sci.*, 2023, **14**, 226–244.
- 111 P. Schwaller, T. Laino, T. Gaudin, P. Bolgar, C. A. Hunter, C. Bekas and A. A. Lee, *ACS Cent. Sci.*, 2019, **5**, 1572–1583.
- 112 A. A. Schoepfer, R. Laplaza, M. D. Wodrich, J. Waser and C. Corminboeuf, *ACS Catal.*, 2024, **14**, 9302–9312.
- 113 A. Poater, S. P. García-Abellán, J. V. Alegre-Requena, B. Trzaskowski and J. P. Martínez, *ACS Catal.*, 2026, **16**, 2160–2170.
- 114 D. T. Ahneman, J. G. Estrada, S. Lin, S. D. Dreher and A. G. Doyle, *Science*, 2018, **360**, 186–190.
- 115 D. Dalmau and J. V. Alegre-Requena, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2024, **14**, e1733.
- 116 F. Strieth-Kalthoff, S. Szymkuć, K. Molga, A. Aspuru-Guzik, F. Glorius and B. A. Grzybowski, *J. Am. Chem. Soc.*, 2024, **146**, 11005–11017.
- 117 R. J. Hickman, G. Tom, Y. Zou, M. Aldeghi and A. Aspuru-Guzik, *Digit. Discov.*, 2025, **2025**, 2104–2122.
- 118 B. Sanchez-Lengeling and A. Aspuru-Guzik, *Science*, 2018, **361**, 360–365.
- 119 D. G. Blackmond, *Angew. Chem., Int. Ed.*, 2005, **44**, 4032–4320.
- 120 J. Burés, *Angew. Chem., Int. Ed.*, 2016, **55**, 16084–16087.
- 121 A. Milo, A. J. Neel, F. D. Toste and M. S. Sigman, *Science*, 2015, **347**, 737–743.
- 122 S.-W. Li, L.-C. Xu, C. Zhang, S.-Q. Zhang and Z. Hong, *Nat. Commun.*, 2023, **14**, 3569.
- 123 J. V. Alegre-Requena, S. Sowndarya SV, R. Pérez-Soto, T. M. Alturaifi and R. S. Paton, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2023, **13**, e1663.



- 124 Z. Cao, L. Falivene, A. Poater, B. Maity, Z. Zhang, G. Takasao, S. B. Sayed, A. Petta, G. Talarico, R. Oliva and L. Cavallo, *Cell. Rep. Phys. Sci.*, 2024, **6**, 102348.
- 125 L. Morán-González, A. L. Burnage, A. Nova and D. Balcells, *ACS Catal.*, 2025, **15**, 9089–9105.
- 126 C. A. Grambow, L. Pattanaik and W. H. Green, *Sci. Data*, 2020, **7**, 137.
- 127 D. Dalmau, S. García-Abellán and J. V. Alegre-Requena, *ACS Catal.*, 2026, **16**, 1–11.
- 128 D. Rana, P. M. Pflüger, N. P. Hölter, G. Tan and F. Glorius, *ACS Cent. Sci.*, 2024, **10**, 899–906.

