Check for updates

Mapping Chemical Reactivity

Utilizing Similarity Measures to Map Chemical Reactivity

Robert Rauschen, Dean Thomas, and Leroy Cronin*

Abstract: The development of fully autonomous chemical synthesis platforms requires robust, real-time assessment of reactivity that does not rely on prior mechanistic knowledge. Existing methods often depend on predefined reaction models or chemical intuition, limiting their generalizability and adaptability. To address this challenge, we introduce a chemically agnostic approach that quantifies reactivity dynamics by applying similarity metrics to the full informational content of in-situ spectroscopic data. Using NMR, UV/Vis, IR, and EPR spectroscopy, we demonstrate that spectral similarity trajectories can reliably indicate reaction progress, detect kinetic features such as autocatalysis, and resolve complex behaviours including oscillations. Across diverse reaction classes, this method enabled estimated end-point detection and kinetic profiling without reaction-specific tuning. For example, the formation of lophine was followed by NMR at different reaction temperatures revealing Arrhenius-type kinetics. In a different experiment, the Belousov Zhabotinsky reaction was monitored by UV/Vis and the chemical oscillation with a periodicity of 7.25 s was captured. These results establish a generalizable framework for real-time, data-driven reactivity monitoring, representing a critical step toward autonomous synthesis guided by multidimensional spectroscopic feedback.

Introduction

Scientists have long sought to understand the complex processes occurring within a chemical reaction flask. Traditionally, thin-layer chromatography (TLC) serves a cheap and relatively universal technique to monitor reaction progress and determine when a mixture is ready for further processing.[1–3] Only then is a comprehensive characterization of the reaction outcome performed after the reaction has concluded. Fundamentally, TLC does not offer deeper structural insights and is limited to *human-in-the-loop* workflows owing to the heavily manual procedure. With advancements in process analytical technology (PAT) however, a growing array of automated techniques allows for real-time, in situ analysis of reactions, providing critical insights into the formation of intermediates, products and subsequent degradation pathways.[4–13] Due to the complex nature and vast variety of chemical reactions however, chemically-relevant, automated monitoring is non-trivial and usually specific to each use case.[14] Consequently, the technological solutions to gather reaction data are similarly diverse.[15–20] With advancing technological capabilities – particularly benchtop analytics – there is a growing repertoire of spectroscopic techniques that afford in situ structural analysis of products formed as well as overall reaction progression. This includes on-line spectrometers for NMR,[21–25] IR,[22,26,27] UV/Vis,[28–32] and EPR spectroscopy[33] but also on-line HPLC[34–36] and mass spectrometry setups.[37–39]

Despite this impressive suite of ever-improving technologies, the automated analysis of typically low-resolution, low-sensitivity, crude spectra presents a series of challenges. First, not all chemical reactions exhibit diagnostic peaks for tracking reaction progress as signals from reagents, products, and impurities often extensively overlap making quantification case-dependant.[40] Instead of relying on pre-known regions, one can automatically track the evolution of all distinguishable signals in a spectrum[41] however this approach requires powerful peak picking and deconvolution techniques and despite significant spectrum processing development, even machine learning–assisted algorithms still struggle to reliably analyse many datasets, reflecting the high complexity and information density of chemical signals.[42] Second, transient species may not yield quantitative signals in viable time frames and/or might not exhibit strong enough responses to external stimuli. Balancing sampling time with signal-to-noise is complex as longer analysis times (in terms of spectroscopic scans) may hinder kinetic investigations into rapid reactions.[43] Additionally, no single analytical technique is applicable to every chemical problem as some compounds are not UV-active whilst others do not contain characteristic spin-active species. Finally, and logistically frustrating, often analytical workflows are accompanied by bespoke, licensed software which can limit interoperability and cross-compatibility. All of these factors complicate automated analysis, reducing generalised approaches to a case-by-case basis.[14]

An agnostic tool for reaction monitoring is needed that complements multiple spectroscopic techniques to simplify

[*] Dr. R. Rauschen, Dr. D. Thomas, Prof. Dr. L. Cronin
School of Chemistry, University of Glasgow, University Avenue, Glasgow G12 8QQ, UK
E-mail: lee.cronin@glasgow.ac.uk

Additional supporting information can be found online in the Supporting Information section
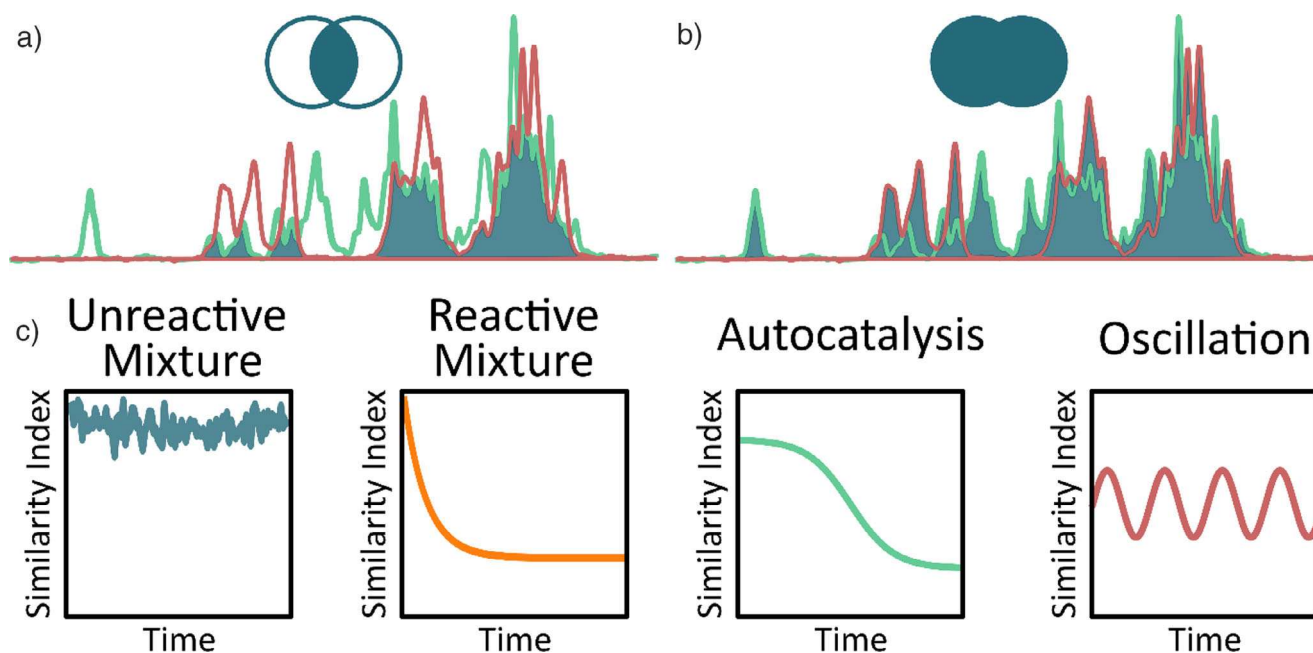
Figure 1. Illustration of the Jaccard similarity index calculation. a) The area under the intersection is divided by b) the area under the union of two spectra to obtain their Jaccard similarity index. c) Exemplar plots of similarity index against time for typical reaction profiles.

and harmonise the analysis of data for rapid fingerprinting.[22] Herein, we propose that similarity metrics like cross correlation or intersection over union (Jaccard index),[44,45] which are widely used in statistics for comparing the similarity of series and sample sets, respectively, can serve as a proxy for extracting comprehensive information about reactivity from spectroscopic data (Figure 1). In the context of spectral analysis, we adopt a geometric interpretation of the similarity index as follows: the intersection of two spectra can be calculated by choosing the minimum intensity data point at each chemical shift/wavelength and integrating the resulting curve (Figure 1a). The union is derived from the maximum intensity, accordingly (Figure 1b). Instead of disentangling the peaks in a spectrum, obtaining a similarity metric is sufficient to draw information about the reaction progression by comparing how "dissimilar" the spectra become over time in comparison to the first spectrum in the sample set. At the beginning of the experiment, the first spectrum is compared to itself, giving a Jaccard index with a maximum value of 1.0. Thereafter, it varies as a function of the reaction profile (Figure 1c).

The Jaccard index itself is then obtained by dividing the absolute value of the intersection area ($|A \cap B|$) by the union area ($|A \cup B|$) (Equation 1).

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \qquad (1)$$

When it comes to the comparison of two time-dependant data series $f(t)$ and $g(t)$ where the geometrical integration is not chemically plausible, we use the maximum cross-correlation and normalise it with the mean autocorrelation to measure the similarity $S(f, g)$ (Equation 2).

$$S(f, g) = \frac{\max_{\tau \in [0, t]} (f * g)(\tau)}{\sqrt{\sum_{\tau=0}^{t} f(\tau)^2 \cdot \sum_{\tau=0}^{t} g(\tau)^2}} \qquad (2)$$

Of which, the cross-correlation is a function of the shift τ (Equation 3) where $\overline{f(t)}$ denotes the complex conjugate of $f(t)$.

$$(f * g)(\tau) \triangleq \sum_{t=-\infty}^{\infty} \overline{f(t)} \cdot g(t + \tau) \qquad (3)$$

This cross-correlation similarity is particularly useful when comparing raw free induction decay (FID) data from NMR experiments (vide infra Figure 6c) as well as in cases where analysis with the Jaccard index is inappropriate as negative areas would lead to cancellation effects.

Assuming that the initial state of a reactive mixture is known, i.e., a spectrum was acquired before the reaction started, the similarity between the initial spectrum and the spectrum at a given timepoint can serve as a proxy for the reactivity of that mixture. This is particularly valuable for automated platforms requiring autonomous decision-making capabilities as traditionally, case-specific reaction monitoring required either predefined characteristic properties or *human-in-the-loop* intervention during execution.

As many chemical transformations are performed in the solution phase, a solution-based automation platform equipped with a wide array of in situ and on-line monitoring technologies was required to ensure real-time feedback and high reproducibility.[14] The modular Chemputer[46–48] platform served as the operational backbone owing to the flexible

liquid-handling architecture alongside the broad suite of integrated sensing modalities,[49] making it ideally suited for method development across diverse chemistries. Crucially, the Chemputer supports both standard and user-defined experimental workflows, enabling direct integration of spectroscopic inputs for intelligent reaction control.[31] These characteristics are particularly well-suited for evaluating analysis strategies aimed at general applicability, where minimal assumptions are made about the underlying reaction mechanisms or reagent identities.

Beyond liquid handling and analysis, frameworks such as the Chemputer embody the foundational principles of digital chemistry,[50] i.e., capturing all procedural, physical, and analytical steps as codified instructions. This formalization enables automated error detection and correction through reaction feedback. For instance, deviations in real-time analytical signals, such as stalled reactivity or unexpected side-product formation, can trigger workflow modifications, such as reagent re-dosing or extended reaction times,[51] thereby closing the loop between observation and execution. This approach to feedback-driven adaptivity is essential for achieving truly autonomous chemical synthesis.

Importantly, benchtop NMR, IR, UV/Vis, and EPR spectrometers have become increasingly viable for integration into automated synthesis platforms due to their reduced footprint, cost, and user-friendly design. These instruments now offer sufficient sensitivity and resolution for reaction monitoring in many cases [52] and require only minor hardware modifications (e.g., flow cell attachments) and software extensions to interface with existing infrastructure.[31] Their compatibility with open-source control frameworks and standard data communication protocols makes them ideal candidates for closed-loop automation, where spectroscopic feedback can be used not just for data logging but for active decision-making. (For detailed implementation strategies, including hardware schematics and software integration pathways, see Supporting Information, Sections 1.2 and 2).

Results and Discussion

To investigate the versatility of this reaction monitoring approach, the algorithm was applied to a range of different chemistries and spectroscopic techniques (Figure 2). A variety of different chemical reactions including imine formations,[51] formose reactions,[53] Buchwald-Hartwig aminations,[54] Maillard reactions,[55–60] heterocycle formation,[61–65] photocatalytic debromination,[66] and others were executed on the Chemputer platform and monitored using the on-line sampling sequences. The scope of reactions included those with known and unknown outcomes to examine the performance of the monitoring algorithm with minimal bias in chemical space. All reactions were subjected to flow-cell sampling in regular intervals, depending on the number of scans acquired. When the plateau detection indicated that the reaction had finished, the reactions were quenched, the product phase was isolated via liquid-liquid extraction and purified when a single product was expected or subjected to chromatographic analysis if multiple undefined products were expected.

Considering NMR first, whenever possible, a solvent suppression pulse sequence was used to allow for the accurate tracking of reagents in non-deuterated solvents. Preliminary test acquisitions were concluded in advance to optimise the suppression parameters via Bayesian optimisation (see Figure S1). When diagnostic signals for the product and the starting material were available, their integrals were measured, resulting in a benchmark conversion for comparison over the course of the reaction (see Figure S3 for an example of conversion time series). Satisfyingly, for the synthesis of lophine (Figure 3a), the time point at which the Jaccard plot reaches a stable plateau (Figure 3c) compares well with the plateau time index obtained from a traditional conversion plot derived from analysing diagnostic signals only (Figure 3b). The plateau was defined to be stable when the slope from a linear regression of five consecutive data points falls below a threshold value of 1×10^{-3}.

To demonstrate that the similarity of a sequence of spectra provides a robust measure of the reaction progress and is not instead analysing the evolution of noise or the shift of instrument parameters, the reaction kinetics of lophine's formation was investigated at different temperatures. To this end it has been recently demonstrated[67] how signal averaging during post-processing can be exploited to simultaneously obtain adequate signal-to-noise ratio and temporal resolution which we extensively utilised whenever temporal resolution is critical. The reaction vessel was held at a constant temperature and aliquots from the reaction were analysed with NMR at regular intervals. An exponential decay function was fitted to the course of the similarity index to extract a time constant, which was subsequently plotted against the reaction temperature. As can be seen (Figure 3d), a linear relationship emerges, which indicates that the performed analysis is a good measure for the physical kinetic properties of the reaction.

Having established that a plateau in the similarity metric is insightful for reaction endpoint detection, the next logical step is to understand how the Jaccard index categorises unreactive or unchanging systems. When monitoring a blank sample, containing just solvent, the Jaccard index fluctuated randomly leading to erratic behaviour of the plateau-detection algorithm. To avoid this, noisy datasets can be detected by calculating their Spearman correlation coefficient (see Supporting Information, Section 5) and excluded from monitoring. Although this serves as a practical workaround it does not explain the erratic behaviour of the Jaccard index when the underlying spectra are very similar. To understand this, a deeper mathematical investigation of the Jaccard algorithm was carried out (see Supporting Information, Section 6). The numerical analysis revealed that the application of the minimum (or maximum) function on the intensity values in the experimental spectrum is mathematically "ill-conditioned" meaning that it can arbitrarily amplify the experimental error of the intensity values in certain cases. This phenomenon is known in the context of subtraction in floating point arithmetic where it leads to so-called "catastrophic cancellation".[68] Having identified the ill-conditioned minimum function as the problematic part
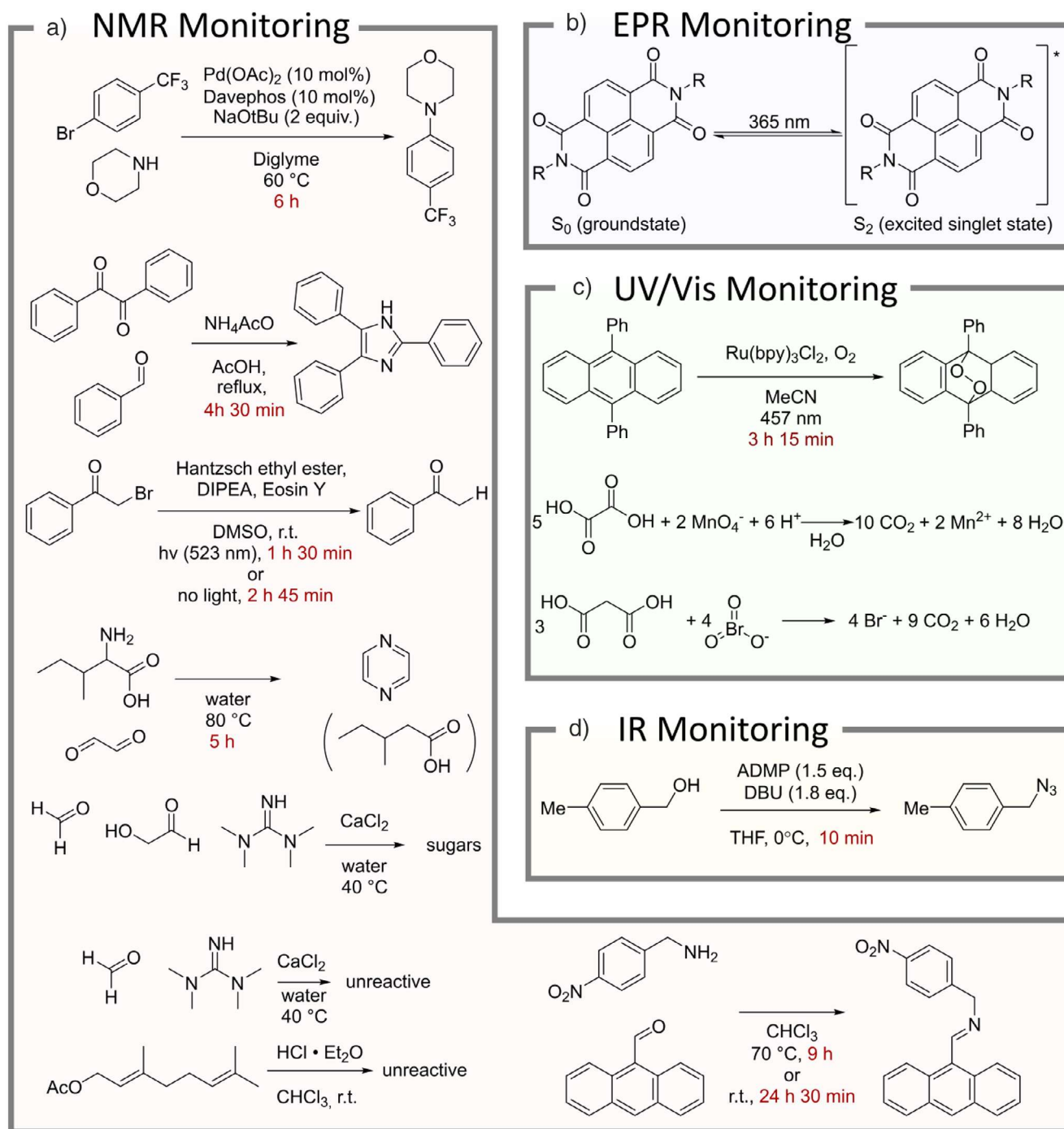
Figure 2. Overview of Reactions Subjected to Reaction Monitoring. Where applicable, reaction times in red refer to the time after which the similarity index of the spectra reached a plateau. Unreactive, oscillating, and autocatalytic mixtures were purposefully included in this scope to demonstrate how the similarity-based monitoring approach can distinguish different kinetic behaviours.

of the algorithm, a mitigation strategy is clear. Prior to the calculation of the Jaccard similarity index, the instrument error on the intensity values is estimated and whenever the difference between two intensity values during the comparison of two spectra is smaller than the instrument error, the values are treated as equal. This modified procedure reliably avoids catastrophic error amplification and correctly identifies unchanging reactions.

The similarity score was then validated over a diverse set of chemical reactions (see Supporting Information, Section 3), demonstrating that the agreement between the plateau time of the conversion plot and the Jaccard plot is reproducible. Similar plateau-detection approaches with alternate reaction properties could also have been taken, e.g., heat flux for exothermic/endothermic reactions, pressure change for reactions which evolve gas, or colour change if
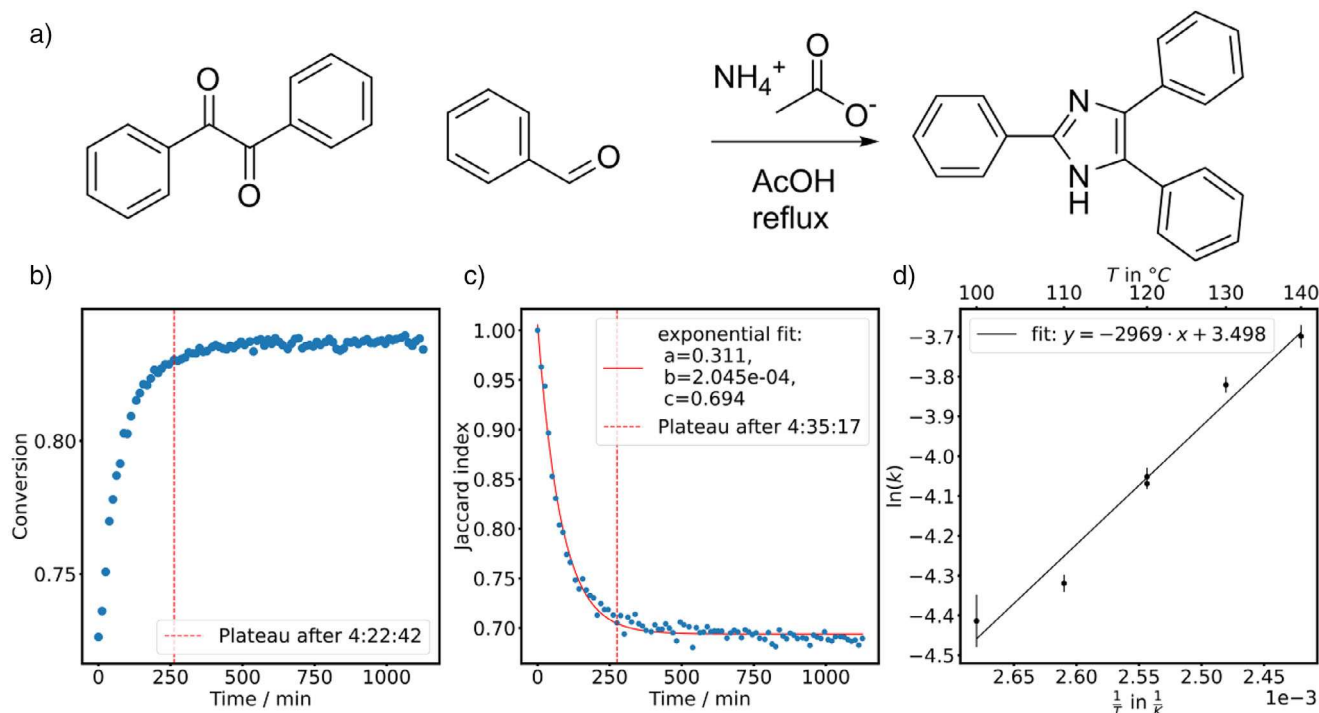
Figure 3. Exemplar monitoring data from Lophine formation. a) Reaction scheme. b) Tracking of the conversion over time by comparing diagnostic NMR signals of the starting material and the product in the aromatic region against each other. c) Similarity index of the NMR spectra over time compared against the initial spectrum following an exponential decay that reaches a plateau around the same time as the conversion time series. d) Arrhenius analysis of the Jaccard index at different temperatures: The time series of the Jaccard index can be fitted well with an exponential decay function and the time decay constant from that fit function was plotted against the reaction temperature. Data obtained with an 80 MHz bench-top NMR spectrometer.

UV/Vis-active compounds were part of the reaction (See Supporting Information, Section 3.8 for the photochemical oxidation of an anthracene derivative). It should be noted however that plateau-detection approaches are not suitable for transient intermediates or the characterisation of oscillatory behaviour. Only systems that reach a long-term stable state are amenable to endpoint detection analysis.

Inspired by the robust results in NMR, the similarity-based analysis approach was applied to spectra obtained with electron paramagnetic resonance (EPR) spectroscopy as well. As the paramagnetic nature of species with unpaired electrons interferes with the signal obtained in NMR experiments, EPR spectroscopy serves as a complementary analytical technique, often providing critical insights into transition-metal containing complexes, thereby opening another fraction of chemical space to automated reaction monitoring. The detected signal in an EPR experiment is normally proportional to the first derivative of the absorption due to the lock-in detection method,[69] which improves sensitivity and noise suppression. Consequently, the data was integrated as part of the pre-processing before applying the similarity measures to obtain a traditional absorption-mode spectrum. Initially, the time-resolved monitoring of the photochemical excitation of a naphthalene diamide (NDI) gel was chosen to explore the applicability of the Jaccard similarity index to integrated EPR data. Naturally, the first spectrum in the dataset will appear as a flat baseline because no EPR signal is expected at the beginning of the excitation period. This means the similarity index will rapidly decay to zero when taking the first spectrum as a reference (see Supporting Information, Section 3.11), which holds limited information in this case. In fact, it is more insightful to compare against the last spectrum in the dataset to find out if the excitation process reached an equilibrium. The resulting Jaccard indices clearly show that the excitation process has not reached an equilibrium within the 50 min of monitoring.

Next, the same approach was demonstrated with on-line IR spectroscopy (Figure 4). IR spectra are infamously hard to deconvolute, especially in the fingerprint region because of the vast number of overlapping signals and the richness in chemical information. This makes IR spectra particularly interesting for a meta-analysis with similarity metrics. The in-situ formation of azides from alcohols was monitored over a total period of 50 minutes in a spectral window between 1500 cm^{-1} and 2500 cm^{-1} and it was found that there are two peaks in the spectrum that are amenable to isolated analysis. The product azide peak at 2097 cm^{-1} grows over time, following an exponential trend that is characteristic for first-order kinetics. Moreover, there is a transient intermediate, observed at 2173 cm^{-1} which follows an exponential "rise-and-fall" curve.[70,71] Analysing the similarity of the whole spectrum over time leads to the conclusion that the reaction nears completion after roughly 10 minutes.
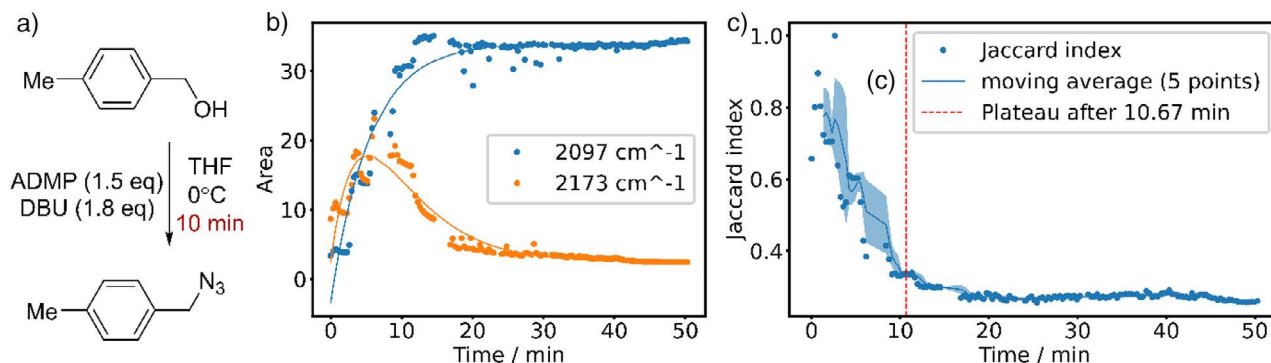
Figure 4. IR Monitoring of Azide Formation. a) Reaction scheme. b) Isolated, diagnostic peaks can be analysed for kinetic information about intermediates. c) The overall Jaccard index suffices to detect when the reaction is mostly completed.
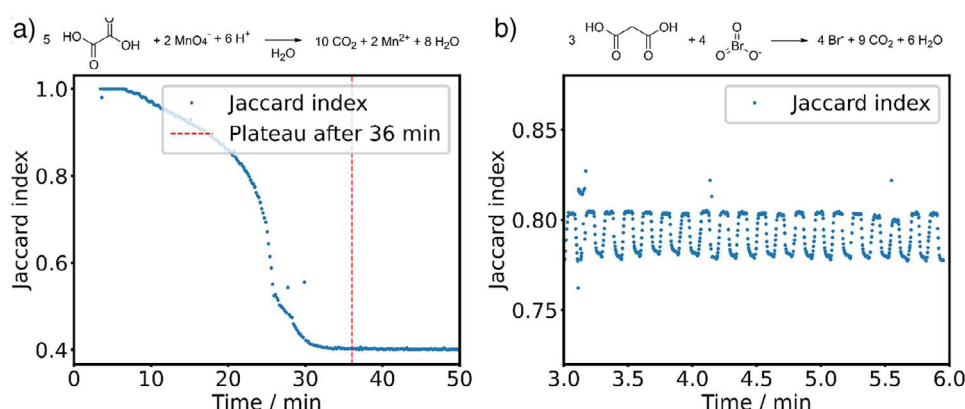
Figure 5. Complex kinetic behaviour like oscillation and autocatalysis. Similarity metrics can also be used to track reactions that have kinetic profiles more complex than apparent first-order exponential time series. a) The autocatalytic behaviour of the permanganate-oxalate-reaction is easily detected by the sigmoidal shape of the time series. b) The chemical oscillation frequency of the Belousov-Zhabotinsky reaction can be obtained by Fourier transformation of the time series of the similarity index of the UV/vis spectra.

Overall, the reactions where the similarity index reaches a stable plateau were found to follow apparent first-order kinetics, meaning that monitoring it through spectroscopic techniques will most likely yield a data series that can be fitted with a simple exponential function. The similarity-based analysis is not restricted to these cases though. This approach is also valuable for complex systems, where dynamic behaviours such as switching and oscillations are known to propagate through complex networks to drive material change.
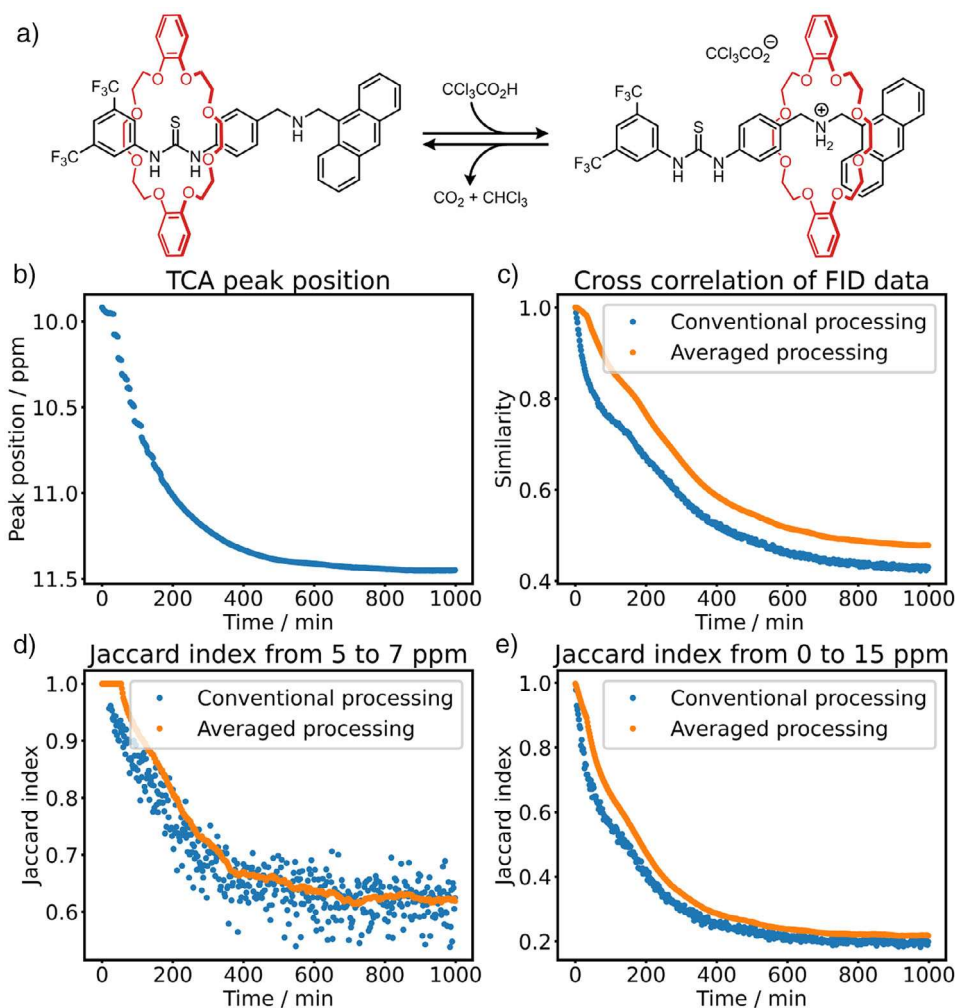
A model example for non-standard kinetics is the reaction between potassium permanganate and oxalic acid,[72] owing to the autocatalytic behaviour as soon as critical amounts of Mn(II) are present in solution. A variety of kinetic models[72–74] have been developed to describe the kinetics of this reaction and a numerical simulation of the time course of the concentration of relevant species was carried out based on a recent kinetic model to find out that the species with the most prominent UV/Vis signal (permanganate) is expected to follow a sigmoidal decay function (see Supporting Information, Section 3.9). Monitoring in short intervals and similarity-based analysis of the resulting spectra confirmed that the data is, indeed, following a sigmoidal function (Figure 5a).

Next, a reaction was chosen that does not reach a stable plateau initially because it displays oscillatory behaviour. The

Belousov-Zhabotinsky reaction is the most famous example of such oscillatory behaviour and can be readily monitoring through UV/Vis spectroscopy (Figure 5b). The analysis of the similarity of the resulting spectra revealed the chemical oscillation with a period of 7.25 s (See Figure S10).

Finally, functional assemblies can undergo controlled switching between distinct states in response to chemical fuels or environmental stimuli. NMR spectroscopy was employed to investigate the switching of a [2]rotaxane using trichloroacetic acid (TCA), providing an automated platform for observing the response over time (Figure 6). The macrocycle shuttles between two distinct positions depending on the protonation state of the system. A single pulse of trichloroacetic acid (CCl_3CO_2H), transiently protonates the rotaxane and induces shuttling before undergoing base-catalysed decarboxylation and decomposing into carbon dioxide (CO_2) and chloroform ($CHCl_3$) as the only waste products. During this process, basic conditions in the medium are restored, reversing the macrocycle's position and completing one full operational cycle (Figure 6a). Owing to the dynamic strengths of the similarity approach, the cycle can be monitored along multiple reaction coordinates – ranging from chemical conversion, as seen in the changing $CCl_3CO_2\underline{H}$ chemical shift values (Figure 6b), to spectral pattern evolution

Angew. Chem. Int. Ed. **2025**, e21519 (6 of 10)

© 2025 The Author(s). Angewandte Chemie International Edition published by Wiley-VCH GmbH

Figure 6. Molecular Switching Insights. a) The programmable shuttling of the macrocycle between binding sites of a [2]rotaxane can be monitored using various analytical approaches, selected according to the specific experimental objective and the level of prior knowledge available. For instance: b) Tracking the CCl_3CO_2H (TCA) chemical shift, which is sensitive to pH, provides a direct readout of fuel decomposition over time. c) The averaged free induction decay (FID) similarity over time offers a global, unsupervised view of spectral evolution. d) The Jaccard index calculated within the 5–7 ppm region highlights local spectral changes associated with protonation-state-dependent switching. e) The Jaccard index computed across the full spectrum captures overall compositional or dynamic change throughout the switching cycle.

measured by local (5–7 ppm) or global Jaccard index values (Figure 6d,e, respectively), as well as overall signal similarity through time-resolved FID comparison (Figure 6c). Each approach offers distinct insights into the [2]rotaxane's operation and can be tuned depending on the analytical focus. Such systems are ideally suited for optimisation with autonomous workflows, which can iteratively adjust reaction conditions in response to feedback, accelerating the development of molecular machines capable of performing complex, adaptive functions.

Conclusion

Whether utilising bespoke and expensive analytical platforms or simpler, commercially available sensors, the pressing need to analyse and understand reactions is ubiquitous across academic and industrial sciences. This workflow however relies

heavily on the skills and intuition of trained experts, which is challenging to fully capture in autonomous workflows. To address this, we have developed a simple yet powerful, chemically agnostic algorithm that analyses of the similarity of spectroscopic data over time. The broad applicability of this approach was demonstrated across the fundamental pillars of analytical spectroscopy (NMR, EPR, IR and UV/Vis) to gain a plethora of insights into reactivity for various chemical scenarios. Different kinetic behaviours, including apparent first-order kinetics, oscillations, and autocatalysis are identified in the analysis of the spectroscopic similarity, depending on the spectroscopic technique and the obtained signal type. Furthermore, the series of similarity values can be subjected to a plateau-detection algorithm to determine the endpoint of reactions. This information can then be used to autonomously trigger more complex sub-routines including quenching or work-up procedures upon reaction completion. The applications of this approach open the

door to fully autonomous reaction monitoring and control, enabling real-time decision making in chemical synthesis without the need for prior mechanistic knowledge or expert intervention.

Supporting Information

The python notebooks to reproduce the figures presented throughout the manuscript and the supplementary information, along with the entire dataset, were uploaded to Zenodo.org, doi: 10.5281/zenodo.17241207. The necessary software for reproducing the plots can be obtained from GitHub: github.com/croningp/spectroscopic_similarity. The software framework for operating the Chemputer platform is subject to change and the active development version of the software can be made available to collaborators upon request to the corresponding author. A referenced version of the software stack[49] is publicly available on Zenodo, doi: 10.5281/zenodo.6534009.

Author Contributions

R.R.: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing — original draft, Writing — review & editing. **D.T.**: Supervision, Visualization, Writing — review & editing. **L.C.**: Conceptualization, Funding Acquisition, Resources, Supervision, Writing — review & editing

Conflict of Interests

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are openly available in zenodo at [10.5281/zenodo.1724120], reference number 17241207.

Keywords: Automation · Cross-correlation similarity · Digital chemistry · Jaccard index · Reaction Monitoring

[1] Y.-J. Tseng, Z.-R. Hong, H. Lin, *J. Chem. Educ.* **2024**, *101*, 3635–3642, https://doi.org/10.1021/acs.jchemed.4c00084.

[2] J. Silver, *J. Chem. Educ.* **2020**, *97*, 4217–4219, https://doi.org/10.1021/acs.jchemed.0c00437.

[3] H. Dickson, K. W. Kittredge, A. M. Sarquis, *J. Chem. Educ.* **2004**, *81*, 1023, https://doi.org/10.1021/ed081p1023.

[4] L. Chen, X. Ding, Z. Wang, S. Xu, Q. Jiang, C. Dun, J. J. Urban, *Surface Science and Technology* **2024**, *2*, 9.

[5] X. Chen, Y. Cheng, M. Matsuba, X. Wang, S. Han, M. Jordan, Q. Zhu, *Appl. Spectrosc.* **2019**, *73*, 1299–1307.

[6] L. Jin, A. Seifitokaldani, *Catalysts* **2020**, *10*, https://doi.org/10.3390/catal10050481.

[7] A. A. L. Michalchuk, F. Emmerling, *Angew. Chem. Int. Ed. Engl* **2022**, *61*, e202117270, https://doi.org/10.1002/anie.202117270.

[8] C. Mu, C. Lv, X. Meng, J. Sun, Z. Tong, K. Huang, *Adv. Mater. Interfaces* **2023**, *10*, 2201842.

[9] C. Prüfert, R. D. Urban, T. G. Fischer, J. Villatoro, D. Riebe, T. Beitz, D. Belder, K. Zeitler, H.-G. Löhmannsröben, *Anal. Bioanal. Chem.* **2020**, *412*, 7899–7911, https://doi.org/10.1007/s00216-020-02923-y.

[10] K. S. Reddy, B. Siva, S. D. Reddy, N. R. Naresh, T. V. Pratap, B. V. Rao, Y.-A. Hong, B. V. Kumar, A. K. Raju, P. M. Reddy, A. Hu, *Molecules* **2020**, *25*, https://doi.org/10.3390/molecules25092199.

[11] M. Rößler, P. U. Huth, M. A. Liauw, *Reaction Chemistry & Engineering* **2020**, *5*, 1992–2002.

[12] I. Telegeiev, O. Thili, A. Lanel, P. Bazin, Y. Levaque, C. Fernandez, M. El-Roz, *Anal. Chem.* **2018**, *90*, 14586–14592, https://doi.org/10.1021/acs.analchem.8b04754.

[13] M. A. Morin, W. (Peter) Zhang, D. Mallik, M. G. Organ, *Angew. Chem. Int. Ed.* **2021**, *60*, 20606–20626, https://doi.org/10.1002/anie.202102009.

[14] M. Christensen, L. P. E. Yunker, P. Shiri, T. Zepel, P. L. Prieto, S. Grunert, F. Bork, J. E. Hein, *Chem. Sci.* **2021**, *12*, 15473–15490, https://doi.org/10.1039/D1SC04588A.

[15] C. J. Taylor, A. Pomberger, K. C. Felton, R. Grainger, M. Barecka, T. W. Chamberlain, R. A. Bourne, C. N. Johnson, A. A. Lapkin, *Chem. Rev.* **2023**, *123*, 3089–3126, https://doi.org/10.1021/acs.chemrev.2c00798.

[16] M. Abolhasani, E. Kumacheva, *Nat. Synth* **2023**, *2*, 483–492, https://doi.org/10.1038/s44160-022-00231-0.

[17] S. Bonciolini, A. Pulcinella, T. Noël, *J. Am. Chem. Soc.* **2025**, *147*, 28523–28545, https://doi.org/10.1021/jacs.5c10303.

[18] A. F. De Almeida, R. Moreira, T. Rodrigues, *Nat. Rev. Chem.* **2019**, *3*, 589–604, https://doi.org/10.1038/s41570-019-0124-0.

[19] A.-C. Bédard, A. Adamo, K. C. Aroh, M. G. Russell, A. A. Bedermann, J. Torosian, B. Yue, K. F. Jensen, T. F. Jamison, *Science* **2018**, *361*, 1220–1225.

Angew. Chem. Int. Ed. **2025**, e21519 (8 of 10)

© 2025 The Author(s). Angewandte Chemie International Edition published by Wiley-VCH GmbH

[20] O. Bayley, E. Savino, A. Slattery, T. Noël, *Matter* **2024**, *7*, 2382–2398, https://doi.org/10.1016/j.matt.2024.06.003.

[21] V. Sans, L. Porwol, V. Dragone, L. Cronin, *Chem. Sci.* **2015**, *6*, 1258–1264, https://doi.org/10.1039/C4SC03075C.

[22] J. M. Granda, L. Donina, V. Dragone, D.-L. Long, L. Cronin, *Nature* **2018**, *559*, 377–381, https://doi.org/10.1038/s41586-018-0307-8.

[23] T. Maschmeyer, L. P. E. Yunker, J. E. Hein, *React. Chem. Eng.* **2022**, *7*, 1061–1072, https://doi.org/10.1039/D2RE00048B.

[24] T. Dai, S. Vijayakrishnan, F. T. Szczypiński, J.-F. Ayme, E. Simaei, T. Fellowes, R. Clowes, L. Kotopanov, C. E. Shields, Z. Zhou, J. W. Ward, A. I. Cooper, *Nature* **2024**, *635*, 890–897, https://doi.org/10.1038/s41586-024-08173-7.

[25] A. Slattery, Z. Wen, P. Tenblad, J. Sanjosé-Orduna, D. Pintossi, T. Den Hartog, T. Noël, *Science* **2024**, *383*, eadj1817, https://doi.org/10.1126/science.adj1817.

[26] C. F. Carter, H. Lange, S. V. Ley, I. R. Baxendale, B. Wittkamp, J. G. Goode, N. L. Gaunt, *Org. Process Res. Dev.* **2010**, *14*, 393–404.

[27] R. Bavisotto, Q. Owen, N. Hopper, W. T. Tysoe, *ChemCatChem.* **2024**, *16*, e202301620, https://doi.org/10.1002/cctc.202301620.

[28] Y. Jiang, D. Salley, A. Sharma, G. Keenan, M. Mullin, L. Cronin, *Sci. Adv.* **2022**, *8*, eabo2626, https://doi.org/10.1126/sciadv.abo2626.

[29] A. Madani, B. Pieber, *ChemCatChem* **2023**, *15*, e202201583, https://doi.org/10.1002/cctc.202201583.

[30] M. L. Bols, J. Ma, F. Rammal, D. Plessers, X. Wu, S. Navarro-Jaén, A. J. Heyer, B. F. Sels, E. I. Solomon, R. A. Schoonheydt, *Chem. Rev.* **2024**, *124*, 2352–2418.

[31] B. M. Matysiak, D. Thomas, L. Cronin, *Angew. Chem. Int. Ed.* **2024**, *63*, e202315207, https://doi.org/10.1002/anie.202315207.

[32] H. Barrington, T. J. D. McCabe, K. Donnachie, C. Fyfe, A. McFall, M. Gladkikh, J. McGuire, C. Yan, M. Reid, *Angew. Chem. Int. Ed.* **2025**, *64*, e202413395, https://doi.org/10.1002/anie.202413395.

[33] V. V. Ptushenko, V. N. Linev, *Molecules* **2022**, *27*, https://doi.org/10.3390/molecules27185996.

[34] J. Liu, Y. Sato, F. Yang, A. J. Kukor, J. E. Hein, *Chemistry–Methods* **2022**, *2*, e202200009, https://doi.org/10.1002/cmtd.202200009.

[35] T. Ha, D. Lee, Y. Kwon, M. S. Park, S. Lee, J. Jang, B. Choi, H. Jeon, J. Kim, H. Choi, H.-T. Seo, W. Choi, W. Hong, Y. J. Park, J. Jang, J. Cho, B. Kim, H. Kwon, G. Kim, W. S. Oh, J. W. Kim, J. Choi, M. Min, A. Jeon, Y. Jung, E. Kim, H. Lee, Y.-S. Choi, *Sci. Adv.* **2023**, *9*, eadj0461, https://doi.org/10.1126/sciadv.adj0461.

[36] A. I. Leonov, A. J. S. Hammer, S. Lach, S. H. M. Mehr, D. Caramelli, D. Angelone, A. Khan, S. O'Sullivan, M. Craven, L. Wilbraham, L. Cronin, *Nat. Commun.* **2024**, *15*, 1240, https://doi.org/10.1038/s41467-024-45444-3.

[37] D. E. Fitzpatrick, C. Battilocchio, S. V. Ley, *Org. Process Res. Dev.* **2016**, *20*, 386–394, https://doi.org/10.1021/acs.oprd.5b00313.

[38] D. Caramelli, J. M. Granda, S. H. M. Mehr, D. Cambié, A. B. Henson, L. Cronin, *ACS Cent. Sci.* **2021**, *7*, 1821–1830, https://doi.org/10.1021/acscentsci.1c00435.

[39] S. H. M. Mehr, D. Caramelli, L. Cronin, *Proc. Natl. Acad. Sci. U.S.A.* **2023**, *120*, e2220045120, https://doi.org/10.1073/pnas.2220045120.

[40] P. Klukowski, M. Augoff, M. Zieba, M. Drwal, A. Gonczarek, M. J. Walczak, *Bioinformatics* **2018**, *34*, 2590–2597, https://doi.org/10.1093/bioinformatics/bty134.

[41] D. W. Li, A. L. Hansen, C. Yuan, L. Bruschweiler-Li, R. Brüschweiler, *Nat. Commun.* **2021**, *12*, 5229, https://doi.org/10.1038/s41467-021-25496-5.

[42] S. B. Federico Paruzzo, *Automatic Signal Region Detection in 1H NMR Spectra Using Deep Learning*, Bruker, **2024**.

[43] R. Chung, J. E. Hein, *Top. Catal.* **2017**, *60*, 594–608, https://doi.org/10.1007/s11244-017-0737-9.

[44] P. Jaccard, *New Phytologist* **1912**, *11*, 37–50, https://doi.org/10.1111/j.1469-8137.1912.tb05611.x.

[45] A. H. Murphy, *Weather and Forecasting* **1996**, *11*, 3–20.

[46] D. Angelone, A. J. S. Hammer, S. Rohrbach, S. Krambeck, J. M. Granda, J. Wolf, S. Zalesskiy, G. Chisholm, L. Cronin, *Nat. Chem.* **2021**, *13*, 63–69, https://doi.org/10.1038/s41557-020-00596-9.

[47] P. S. Gromski, J. M. Granda, L. Cronin, *Trends in Chemistry* **2020**, *2*, 4–12, https://doi.org/10.1016/j.trechm.2019.07.004.

[48] R. Rauschen, M. Guy, J. E. Hein, L. Cronin, *Nat. Synth* **2024**, *3*, 488–496, https://doi.org/10.1038/s44160-023-00473-6.

[49] S. Rohrbach, M. Šiaučiulis, G. Chisholm, P.-A. Pirvan, M. Saleeb, S. H. M. Mehr, E. Trushina, A. I. Leonov, G. Keenan, A. Khan, A. Hammer, L. Cronin, *Science* **2022**, *377*, 172–180, https://doi.org/10.1126/science.abo0058.

[50] M. Guillén-Soler, L. Cronin, *Chem.* **2024**, *10*, 2621–2623, https://doi.org/10.1016/j.chempr.2024.07.032.

[51] R. Rauschen, J.-F. Ayme, B. M. Matysiak, D. Thomas, L. Cronin, *Chem* **2025**, *11*, 102504, https://doi.org/10.1016/j.chempr.2025.102504.

[52] J. Phuong, B. Salgado, J. Heiß, E. Steimers, P. Nickolaus, L. Keller, U. Fischer, E. von Harbou, D. J. Holland, F. Jirasek, H. Hasse, K. Münnemann, *Food Research International* **2025**, *203*, 115741, https://doi.org/10.1016/j.foodres.2025.115741.

[53] S. M. Sutton, S. Pulletikurti, H. Lin, R. Krishnamurthy, C. L. Liotta, *Chem* **2025**, 102553.

[54] M. Kim, S. Chang, *Org. Lett.* **2010**, *12*, 1640–1643, https://doi.org/10.1021/ol100437j.

[55] J. M. Ames, *Trends Food Sci. Technol.* **1990**, *1*, 150–154, https://doi.org/10.1016/0924-2244(90)90113-D.

[56] V. M. Burin, S. Marchand, G. de Revel, M. T. Bordignon-Luiz, *Talanta* **2013**, *117*, 87–93, https://doi.org/10.1016/j.talanta.2013.08.037.

[57] C. Cerny, *Ann. N. Y. Acad. Sci.* **2008**, *1126*, 66–71, https://doi.org/10.1196/annals.1433.011.

[58] H. M. Charnock, G. J. Pickering, B. S. Kemp, *Frontiers in Microbiology* **2022**, *13*, https://doi.org/10.3389/fmicb.2022.979866.

[59] P. Jeandet, S. S. Heinzmann, C. Roullier-Gall, C. Cilindre, A. Aron, M. A. Deville, F. Moritz, T. Karbowiak, D. Demarville, C. Brun, F. Moreau, B. Michalke, G. Liger-Belair, M. Witting, M. Lucio, D. Steyer, R. D. Gougeon, P. Schmitt-Kopplin, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 5893–5898, https://doi.org/10.1073/pnas.1500783112.

[60] O. W. Moore, L. Curti, C. Woulds, J. A. Bradley, P. Babakhani, B. J. W. Mills, W. B. Homoky, K.-Q. Xiao, A. W. Bray, B. J. Fisher, M. Kazemian, B. Kaulich, A. W. Dale, C. L. Peacock, *Nature* **2023**, *621*, 312–317, https://doi.org/10.1038/s41586-023-06325-9.

[61] L. Erdey, I. Buzás, *Anal. Chim. Acta* **1956**, *15*, 322–324, https://doi.org/10.1016/0003-2670(56)80059-7.

[62] K. Maeda, H. Ojima, T. Hayashi, *Bull. Chem. Soc. Jpn.* **1965**, *38*, 76–80, https://doi.org/10.1246/bcsj.38.76.

[63] D. F. Marino, F. Wolff, J. D. Ingle, *Anal. Chem.* **1979**, *51*, 2051–2053, https://doi.org/10.1021/ac50048a041.

[64] E. H. White, M. J. C. Harding, *J. Am. Chem. Soc.* **1964**, *86*, 5686–5687, https://doi.org/10.1021/ja01078a062.

[65] D. Yanover, M. Kaftory, *Acta Crystallographica Section E* **2009**, *65*, o711, https://doi.org/10.1107/S1600536809006552.

[66] J. Jung, J. Kim, G. Park, Y. You, E. J. Cho, *Adv. Synth. Catal.* **2016**, *358*, 74–80, https://doi.org/10.1002/adsc.201500734.

[67] A. Flook, G. C. Lloyd-Jones, *J. Org. Chem.* **2024**, *89*, 16586–16593, https://doi.org/10.1021/acs.joc.4c01882.

[68] D. Goldberg, *ACM Comput. Surv.* **1991**, *23*, 5–48, https://doi.org/10.1145/103162.103163.

[69] M. Tseitlin, S. S. Eaton, G. R. Eaton, *J. Magn. Reson.* **2011**, *209*, 277–281, https://doi.org/10.1016/j.jmr.2011.01.027.

[70] M. Kitamura, S. Kato, M. Yano, N. Tashiro, Y. Shiratake, M. Sando, T. Okauchi, *Org. Biomol. Chem.* **2014**, *12*, 4397, https://doi.org/10.1039/c4ob00515e.

[71] C. Lee, H. Son, S. Park, *Phys. Chem. Chem. Phys.* **2015**, *17*, 17557–17561, https://doi.org/10.1039/C5CP02368H.

[72] A. Skrabal, *Zeitschrift für anorganische Chemie* **1904**, *42*, 1–59, https://doi.org/10.1002/zaac.19040420103.

[73] H. F. Launer, D. M. Yost, *J. Am. Chem. Soc.* **1934**, *56*, 2571–2577, https://doi.org/10.1021/ja01327a013.

[74] S. J. Adler, R. M. Noyes, *J. Am. Chem. Soc.* **1955**, *77*, 2036–2042, https://doi.org/10.1021/ja01613a002.