

Robot-assisted mapping of chemical reaction hyperspaces and networks

<https://doi.org/10.1038/s41586-025-09490-1>

Received: 19 March 2025

Accepted: 4 August 2025

Published online: 24 September 2025

Open access

 Check for updates

Yankai Jia^{1,2,3}, Rafał Frydrych^{1,3}, Yaroslav I. Sobolev¹, Wai-Shing Wong¹, Bibek Prajapati¹, Daniel Matuszczyk¹, Yasemin Bilgi¹, Louis Gadina¹, Juan Carlos Ahumada¹, Galymzhan Moldagulov^{1,2}, Namhun Kim¹, Eric S. Larsen¹, Maxence Deschamps¹, Yanqiu Jiang¹ & Bartosz A. Grzybowski^{1,2}

Despite decades of investigation, it remains unclear (and hard to predict^{1–4}) how the outcomes of chemical reactions change over multidimensional ‘hyperspaces’ defined by reaction conditions⁵. Whereas human chemists can explore only a limited subset of these manifolds, automated platforms^{6–12} can generate thousands of reactions in parallel. Yet, purification and yield quantification remain bottlenecks, constrained by time-consuming and resource-intensive analytical techniques. As a result, our understanding of reaction hyperspaces remains fragmentary^{7,9,13–16}. Are yield distributions smooth or corrugated? Do they conceal mechanistically new reactions? Can major products vary across different regions? Here, to address these questions, we developed a low-cost robotic platform using primarily optical detection to quantify yields of products and by-products at unprecedented throughput and minimal cost per condition. Scanning hyperspaces across thousands of conditions, we find and prove mathematically that, for continuous variables (concentrations, temperatures), individual yield distributions are generally slow-varying. At the same time, we uncover hyperspace regions of unexpected reactivity as well as switchovers between major products. Moreover, by systematically surveying substrate proportions, we reconstruct underlying reaction networks and expose hidden intermediates and products—even in reactions studied for well over a century. This hyperspace-scanning approach provides a versatile and scalable framework for reaction optimization and discovery. Crucially, it can help identify conditions under which complex mixtures can be driven cleanly towards different major products, thereby expanding synthetic diversity while reducing chemical input requirements.

Owing to rapid advances in chemical automation, there has been a growing number of high-throughput campaigns to discover new modes of reactivity^{6,7,12,16,17}, maximize reaction yields (often under the guidance of various optimization algorithms^{9,15,18}) and produce standardized data sought urgently¹⁹ by chemical artificial intelligence. Our goals here were different, as we wished to reconstruct complete portraits of chemical reactions over multidimensional parameter spaces (concentrations, temperatures), with the quantification of yields not only for the major products but also for as many by-products as could be identified. Anticipating thousands of crude mixtures to analyse (here >9,000 in total), we wished to minimize the use of techniques such as nuclear magnetic resonance (NMR) or liquid chromatography–mass spectrometry (LC-MS), which, at their maximum throughput of only a few samples per hour, become costly and require prolonged or even dedicated access to equipment, which may be prohibitive for most academic groups. Accordingly, we relied heavily on inexpensive (cents per sample) and rapid (about 100 samples per hour) ultraviolet–visible (UV-Vis) detection augmented with algorithms to decompose complex

spectra (that is, not just those featuring distinct and easy-to-interpret spectral features^{20,21}) and with autocorrelation metrics to quantify missing information and detect anomalous outcomes.

Robotic set-up and analysis of reaction outcomes

Unless otherwise stated, experiments were performed on a robotic platform (Fig. 1a and Supplementary Video 1) house-built to support various organic solvents and harsh reagents and allowing us to execute and characterize up to roughly 1,000 reactions per day (for design blueprints and computer codes, see Methods and Supplementary Information Sections 1 and 2).

For a given reaction under study, this robot examines the hyperspace of conditions at points of some *N*-dimensional grid (for example, uniform in Fig. 1b), setting up reactions and acquiring UV-Vis absorption spectra at each point and at desired time(s). The acquisition of an individual spectrum takes, on average, roughly 8 s, whereas the accompanying pipetting and spectrometer washing–drying take about 50 s in total

¹Center for Algorithmic and Robotized Synthesis (CARS), Institute for Basic Science (IBS), Ulsan, Republic of Korea. ²Department of Chemistry, Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea. ³These authors contributed equally: Yankai Jia, Rafał Frydrych. [✉]e-mail: jyqinception@gmail.com; nanogrzybowski@gmail.com

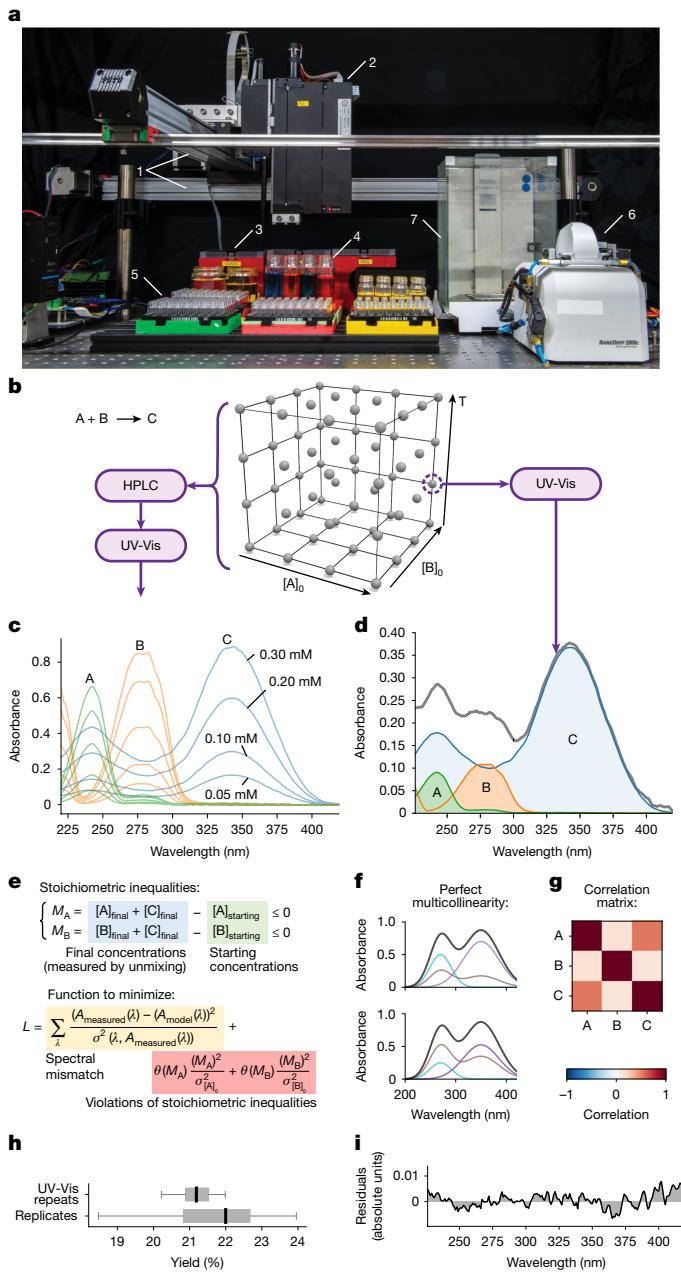


Fig. 1 | Automated reaction platform and optical yield determination.

a, Parts of the house-built, inexpensive (about \$25K) system: (1) horizontal gantry; (2) liquid-handling module; (3) racks of pipette tips; (4) stock solutions; (5) 54-vial plates; (6) UV-Vis spectrophotometer; (7) balance for weighing of solutions (see Supplementary Videos 1 and 2). **b**, Hypothetical A + B → C reaction space defined by initial concentrations $[A]_0$ and $[B]_0$ and temperature T . **c**, UV-Vis spectra of the HPLC-isolated species are measured at different concentrations (here for Claisen–Schmidt condensation from Fig. 2a). **d**, At each point of the conditions' space, the UV-Vis spectrum (grey) is decomposed—here nearly perfectly—into the component spectra from **c**, quantifying the concentrations/yields of these components in the crude mixture. **e**, Stoichiometry of reaction places constraints on possible concentrations (blue) of components in the crude mixture (see Methods). **f**, Example of a perfect linear dependence ('multicollinearity'), in which the brown spectrum in the top subplot happens to be a linear sum of the cyan and purple ones (scaled by, respectively, 0.50 and 0.25). Consequently, unmixing is an ill-conditioned problem, as the spectrum of the entire mixture (black) can be decomposed into infinitely many combinations of the brown, cyan and purple components (1:1:1 in the top and 0.5:0.75:2 in the bottom subplots). **g**, Correlation matrix for components A, B and C of the Claisen–Schmidt condensation from **d**. Low off-diagonal elements indicate no strong multicollinearity. **h**, Statistical distributions of yields with respect to the limiting reagent. Relative standard deviations are 2% between repeated optical detection performed on the same crude mixture ($n = 54$) and 5% between experimental repeats ($n = 27$) of the entire workflow for the same conditions (data from Fig. 5). **i**, Differences between experimental and modelled spectra (from **d**) as a function of wavelength. The fit is adequate, as this trace closely resembles uncorrelated ('white') noise.

(Supplementary Video 2). In a subsequent (and perhaps counterintuitive) step, the crudes from all hyperspace points are combined (Fig. 1b, left). The resultant mixture is separated by chromatography and the isolated fractions are identified by traditional spectroscopic (NMR, MS) analyses. This 'bulk' analysis is to identify the 'basis set' of reaction products that form in any appreciable quantities anywhere in the hyperspace. The UV-Vis absorption spectra of these purified products (as well as all substrates, solvents and reagents used) at different concentrations are taken to construct concentration–absorbance calibration curves (Fig. 1c). Then the crude and often complex UV-Vis spectra acquired at each point of the hyperspace (Fig. 1b, right) are fitted by linear combinations of the reference spectra of the basis-set components (Fig. 1d) using the so-called vector decomposition techniques (also known as spectral unmixing²²). Overall, this protocol requires conventional high-performance liquid chromatography (HPLC)/NMR/MS analysis of only one complex mixture combining all hyperspace samples (Fig. 1c). It returns on this time investment manifolds by rapidly estimating the yields and mixture compositions at thousands of individual data points within the hyperspace of the reaction.

This general scheme merits two comments. First, it is important to reject potential unmixing solutions violating reaction stoichiometry (Fig. 1e) and to ensure stability of the fit. The latter benefits from the absorption spectra of components being of similar magnitudes and not linearly dependent (Fig. 1f), which is diagnosed by looking at off-diagonal elements of the correlation matrix between the concentrations of the components (Fig. 1g). Satisfying these conditions (1) is increasingly challenging as the number of components increases but (2) is helped by extending the spectral range as much as possible into the UV, at which the absorption bands of organic species are usually more numerous and narrower (see the example in Fig. 5b). With these precautions, yield estimates are within 5% (for example, a 20% yield would have a spread of 19–21%), with optical measurement and spectral unmixing contributing 2% and the remaining 3% caused by uncertainty of pipetting, fluctuations of temperature and residual evaporation (Fig. 1h and Supplementary Information Sections 2.10 and 2.11).

Second, the fitting procedure is accompanied by an algorithm to detect anomalous outcomes in some regions of the hyperspace. This is done by tracking differences between experimental and fitted spectra, calculating the variance of residuals (Fig. 1i) and evaluating the autocorrelation (across wavelengths, not time) of these residuals by the Durbin–Watson statistic. The mismatch is considered notable if the root mean square residual exceeds 0.01 absorbance units or if the Durbin–Watson statistic at 30 nm 'lag' deviates from the value of 2 (which corresponds to the absence of autocorrelation) more strongly than the respective statistic for the baseline of a given spectrophotometer. If systematic deviation is detected, it then signifies formation of a product unexpected based on the initial scan of the hyperspace (see Fig. 3d,e). Further mathematical details of spectral unmixing and anomaly detection are discussed in Supplementary Information Section 3.

Reaction scope

We confirmed that this method is applicable to a range of reactions used widely in both academic groups and industry, including various couplings, condensations, cycloadditions and rearrangements (reactions **1–8** on the green background in Extended Data Fig. 1a), as

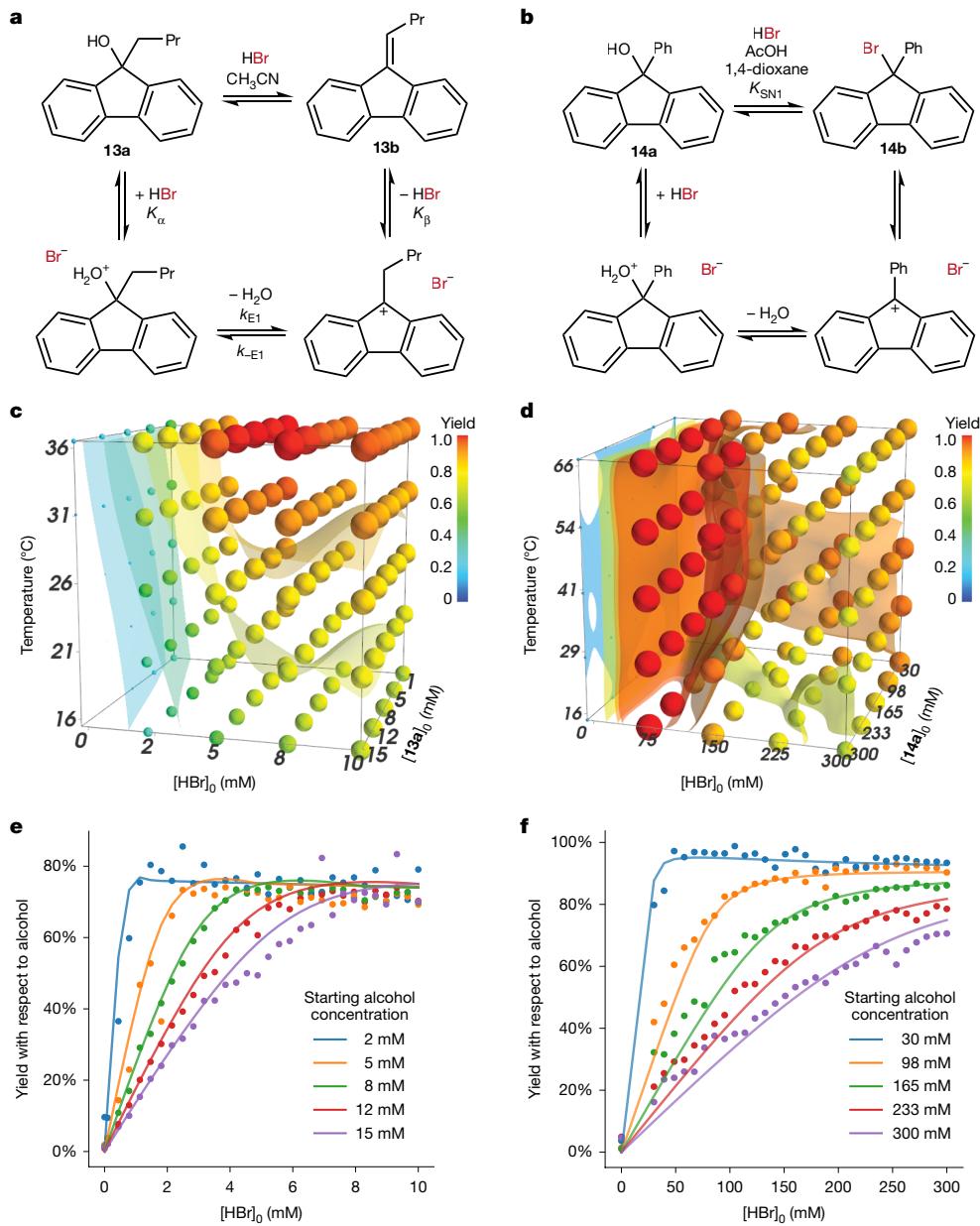


Fig. 2 | Yield distributions over the reaction spaces of E1 and S_N1 reactions. **a,b**, Schemes of E1 (**a**) and S_N1 (**b**) reactions and their mechanisms. For the former, we use 9-butyl-9H-fluoren-9-ol substrate **13a** and for the latter, to prevent elimination, 9-phenyl-9H-fluoren-9-ol **14a**. For E1, the parameters of the mechanism are the temperature dependences of equilibrium constant K_α for dynamic equilibrium between **13a** and its protonated form, of equilibrium constant K_β for dynamic equilibrium between **13b** and its protonated form, of pK_a of HBr in acetonitrile and the Eyring–Polanyi equation parameters for kinetic rate constants k_{E1} and k_{-E1} for water molecule elimination from the protonated **13a** (Supplementary Information Section 5.2). For S_N1 at a given temperature, the only parameter is the effective equilibrium constant $K_{SN1} = [14b][H_2O]/([14a][HBr])$ (Supplementary Information Section 5.1). E1 used acetonitrile as

solvent. For S_N1, the solvent was dioxane with admixture of AcOH (to homogenize the biphasic dioxane–water mixture formed in situ and complicating optical detection). **c,d**, Corresponding yield distributions of products **13b** and **14b**. For clarity, only 125 conditions per cube are shown. Marker colours and sizes are proportional to yields. Yield colour scales are the same for both cubes. Surfaces of equal yield are at 20%, 40%, 70% and 85%. **e,f**, Kinetics model (curves) fitted to experimental yields with respect to alcohol (circles) for E1 and S_N1 reactions. To better visualize the quality of fits, only two-dimensional snapshots of the entire dataset are shown, here $T = 21^\circ\text{C}$ for E1 and $T = 36^\circ\text{C}$ for S_N1 (for similar plots at all temperatures, see Supplementary Information Sections 5.1 and 5.2). Horizontal axis has the initial concentrations of HBr, with the colour corresponding to different initial concentrations of the alcohol indicated in the keys.

well as substitutions, eliminations and multicomponent reactions (MCRs), discussed in detail in Figs. 2–6. Expectedly, the approach is not suitable for, for example, reactions of aliphatic scaffolds whose products give no signal in the UV-Vis range above 220 nm and some reactions are borderline in terms of fit quality, owing to the obstruction of the product signal by reactant/solvent peaks (for example, **9–12** in the yellow and red portions of Extended Data Fig. 1a). Notably, yields quantified by the robotized approach correlate strongly, $R^2 = 0.96$, with

the yields of the same reactions performed, purified and analysed ex robo (Extended Data Fig. 1b).

Reaction hyperspaces

With these capabilities, we began to explore and analyse hyperspaces of several classic reaction types differing in mechanistic complexity. In doing so, we focused on yield distributions at reaction time of several

hours (that is, the situation relevant to synthetic organic practice) and on hyperspace regions featuring maximal degree of variability (as opposed to regions in which, for example, concentrations are so high that yields are all ‘saturated’; see Supplementary Fig. 121). For clarity, in Figs. 2–6, we plot only the subsets of data points, with all raw data available at Zenodo (<https://doi.org/10.5281/zenodo.14880579>).

Hyperspaces of basic reactions with no anomalies

Starting simple, we considered the spaces of the E1 elimination and S_N1 substitution^{23,24} using, respectively, substrates **13a** and **14a**. The space of E1 was examined at $t = 4$ h for 775 conditions and that of S_N1 at $t = 48$ h for 930 conditions. The three-dimensional yield distributions shown in Fig. 2c,d for reaction products **13b** and **14b** are approximately concave, steadily increasing towards the global maximum (in bimolecular S_N1, there is shoulder maximum in the region in which the concentration of the limiting reagent **14a** is low; see also Extended Data Fig. 2). Mathematically, the ‘steepness’ of these surfaces can be quantified by the slopes of the product concentration with respect to the initial concentration of substrate(s), $D_{ij} = \partial C_i / \partial C_{0j}$ —as shown in Supplementary Fig. 166, the absolute values of these slopes, $|D_{ij}|$, are small, at or below unity. Analysis of residuals detects no anomalies and, under all conditions, only substrates and the product are present in various proportions (for example, E1 is not accompanied by S_N1 side reaction). Of note, yield data contained in the ‘cubes’ allow for fitting approximate kinetic models (by time-integrating the underlying kinetic equations) and for deriving reasonable values of some kinetic and thermodynamic parameters, such as ΔH , ΔH^\ddagger or ΔS (see caption to Fig. 2, Methods and Supplementary Information Sections 5.1 and 5.2 for derivations). Even though the data correspond to only one time point, the abundance of yield values for various substrate concentrations imposes hundreds of constraints against which a candidate model must simultaneously fit, thus limiting the acceptable parameters of the model. In the ‘A four-dimensional hyperspace’ section, we will use this approach to interrogate more complex mechanisms.

A hyperspace with anomalous outcomes

Next, we revisited the S_N1 reaction but with **15a** used as a substrate (Fig. 3a). The main product **15b** forms through the rearrangement of the carbocation **15c**, and its yield distribution over the 1,085 conditions fits closely to the expected first-order kinetics (see caption to Fig. 3b). The algorithm also detects a minor anthraquinone by-product **15d** congruent with the Diels–Alder reaction with singlet-state oxygen^{25,26} and distributed as illustrated in Fig. 3c. Most notably, analysis of residuals (Fig. 1i) provides the first example of a systematic anomaly and an unexpected outcome in a narrow region of low HBr concentrations marked in pink in Fig. 3d. This species exhibits an intensely pink colour that gradually wanes over two days (and disappears rapidly during purification attempts). Its UV-Vis spectrum (Fig. 3e) does not agree with absorption spectra of **15c** or structurally similar carbocations^{27–29}, but extensive analyses based on MS as well as time-dependent density functional theory (TD-DFT) calculations (Supplementary Information Sections 4.13 and 6.2) support assignment as a carbocation dimer **15e** resulting from a reaction between **15c** and **15a** (with enough HBr to create **15c**, but with the amount of water introduced by this HBr insufficient to quench the intermediate into **15b**). To our best knowledge, dimerization of a substrate with its own derived carbocationic intermediate has only been observed under superacidic conditions at temperatures down to -50 °C (ref. 30) but never in the presence of quenching nucleophiles (here H₂O and Br⁻) and under ambient conditions. We emphasize, however, that even with the presence of this anomaly, the hyperspace remains simple—that is, each of its constituent species features only one yield maximum—and the yield distributions of all species are slowly varying, $|D_{ij}| \approx 1$ (Supplementary Information Fig. 166), including **15e**, which forms over a narrow range but in only approximately 1 nM concentration.

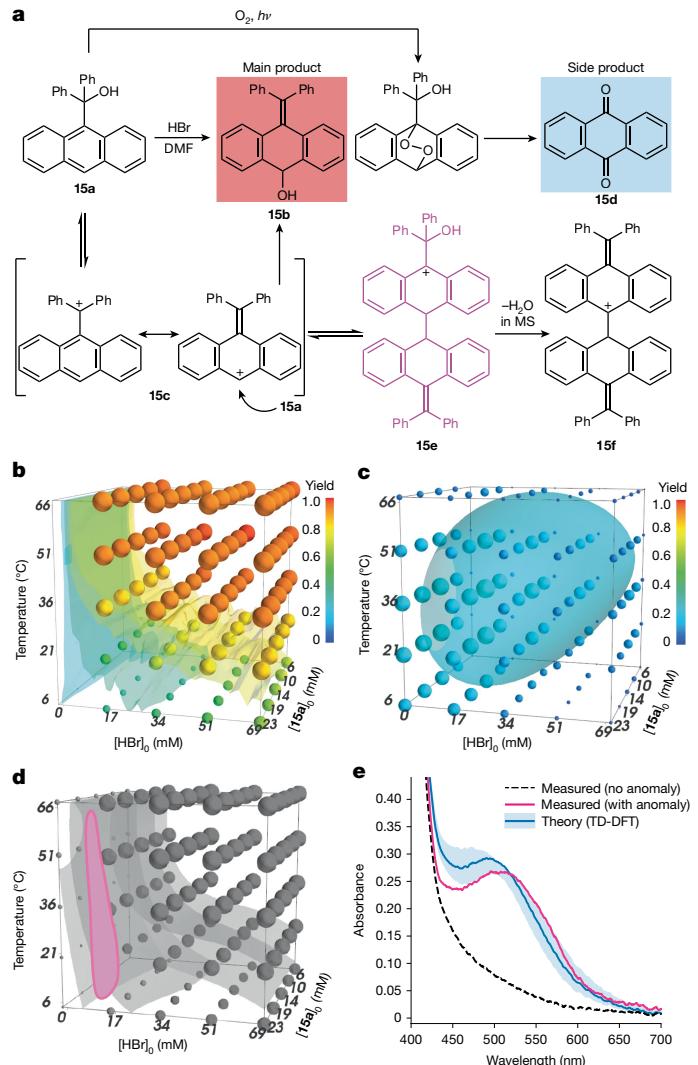


Fig. 3 | S_N1 space with an anomalous outcome. **a**, Scheme of a reaction between alcohol **15a** and HBr leading to major product **15b** through anthranyl carbocation **15c**. The minor anthraquinone product **15d** originates from the cycloaddition of substrate **15a** with singlet oxygen to give the endoperoxide, followed by subsequent fragmentations (for details, see Supplementary Fig. 32). The ‘unexpected’ carbocation dimer structure is assigned to **15e** (for details of structural assignment, see Supplementary Information Sections 4.13 and 6.2). **b**, Space of reaction yields of the main product **15b** under different conditions (reaction time: 24 h; for clarity, only 125 out of 1,085 points are shown; for all data, see Supplementary Information Section 4.15). Fitting the kinetic model to these experimental data reveals that yields of product **15b** scale as $1 - \exp\left(-tk_1[HBr]\exp\left(-\frac{E_a}{k_B T}\right)\right)$, in which t is reaction time, $k_1 \exp\left(-\frac{E_a}{k_B T}\right)$ is the kinetic rate constant at a given temperature (prefactor k_1 is a free parameter), E_a is the activation energy free parameter and $[HBr]$ is the concentration of HBr—that is, these conversions increase with $[HBr]$ and with temperature, do not depend on the initial concentration of the alcohol substrate and the rate-limiting step is formation of a carbocation **15c**, all as expected for S_N1 reaction. The isosurfaces correspond to 20%, 40%, 70% and 85% yield. **c**, Distribution of yields for the side product **15d**. The isosurface corresponds to 20% yield. **d**, Pink cloud delineates a narrow region of conditions in which an anomalous outcome—that is, carbocation dimer **15e**—is observed. **e**, Pink line is a typical UV-Vis spectrum from the region of conditions in which an anomalous outcome is observed. The dashed line is a typical experimental spectrum for which this anomaly is absent. Blue line is the ab initio theoretical absorption spectrum of 1.23 nM of **15e** after summation with the dashed ‘baseline’, blue band is the 1σ confidence interval (see Supplementary Information Section 6.2).

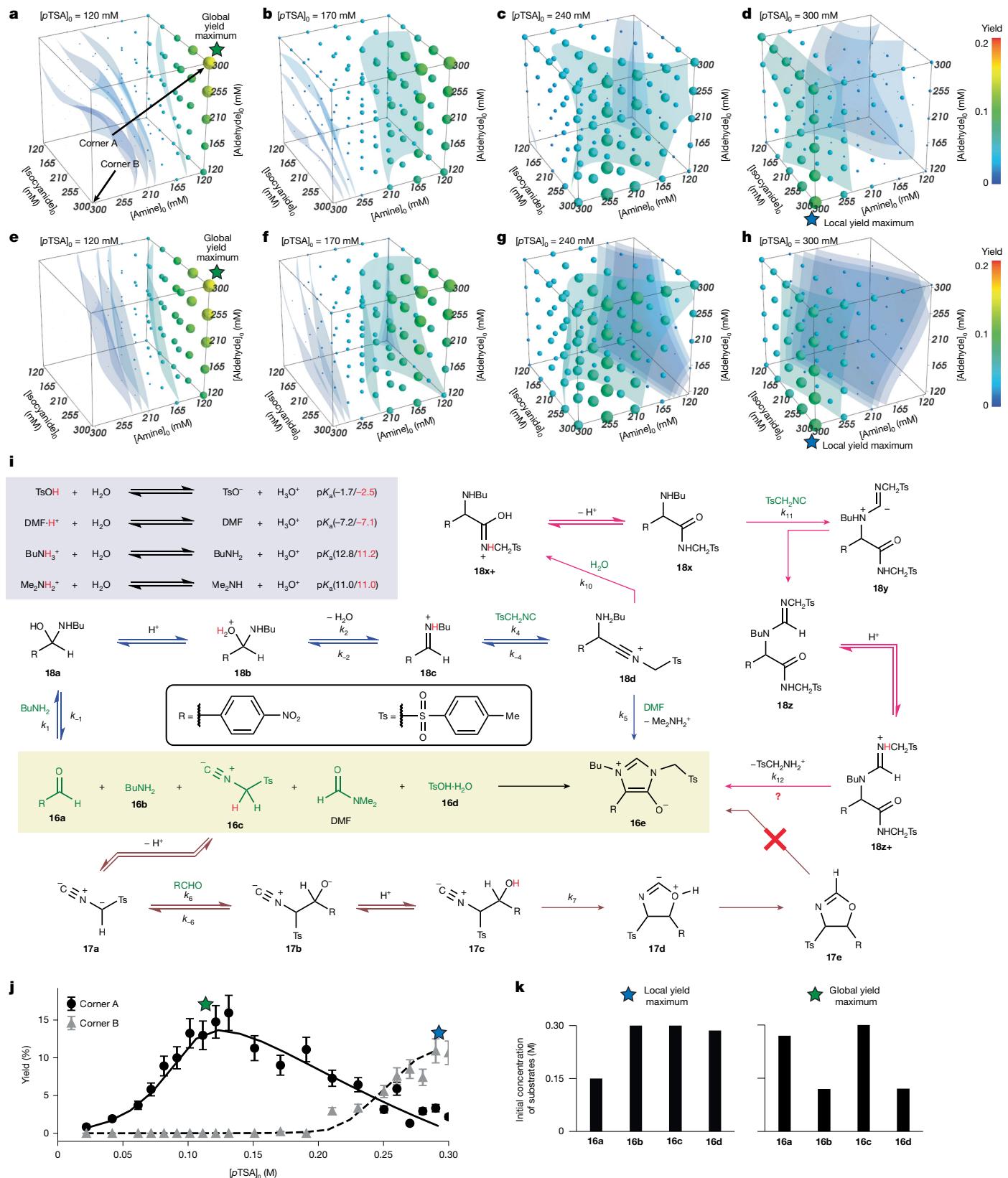


Fig. 4 | See next page for caption.

A four-dimensional hyperspace

In search for more topologically complex hyperspaces, we turned to reactions that involve more than two substrates and are based on

much more complex mechanisms. As the first case, we interrogated a four-component Ugi-type reaction from ref. 31, illustrated in Fig. 4 and chosen because of distinct UV-Vis signal of the cyclized product **16e**. For this cyclization reaction, 3,234 conditions were investigated

Fig. 4 | Experimental and modelled hyperspace of a Ugi-type, four-component reaction. **a–d**, Snapshots of the reaction's four-dimensional hyperspace. The axes of each plot are initial amine, aldehyde and isocyanide concentrations for initial concentrations of *p*TSA indicated above each plot. Experimental yields (after 16 h at 26 °C) of product **16e** are colour-coded according to the legend in **d**. Only 125 interpolated conditions per cube are shown for clarity (see Supplementary Information Section 4.15 and Supplementary Video 3). Green and blue stars mark two distinct yield maxima. **e–h**, Corresponding yield distributions of **16e** predicted by a kinetic model fitted simultaneously to all experimental data (yield isosurfaces are at 0.5%, 1.0%, 1.5% and 5.0%). This model is based on a kinetic network outlined in **i** and further side reactions (see Supplementary Information Scheme 5). The original reaction of **16a**, **16b**, **16c** and DMF using initiator **16d** and leading to product **16e** is on the yellow background. The classic Ugi mechanism is summarized by

blue arrows (for details, see Supplementary Information Scheme 6). A competing oxazoline-based path (brown) was experimentally excluded (red cross). A third mechanism (magenta), potentially active at high [pTSA]₀ and involving aldehyde, amine and isocyanide in a 1:1:2 ratio, did not improve model fit. The theoretical model (Supplementary Tables 19, 20, 22 and 23) comprised this entire mechanistic network of 12 reactions and 15 proton transfer steps. Best-fitted model followed the data closely while retaining p*K_a* values (in red font) of most species within 1–2 units from literature values (black font) used as initial guess (see Supplementary Information Section 5.3). **j**, Experimental and modelled yield profiles at the two maxima (corners A and B in **a**) along [pTSA]₀ confirm that yields at global/local maxima are higher than at any four-dimensional paths connecting them (see Supplementary Video 3 and Supplementary Fig. 161). Error bars correspond to 14.6% relative standard error. **k**, Initial concentrations at global and local maxima.

within a four-dimensional space defined by different initial concentrations of 4-nitrobenzaldehyde **16a**, *n*-butylamine **16b** and *p*-tosylmethyl isocyanide **16c** substrates (DMF substrate was also a solvent of constant concentration), as well as *p*-toluenesulfonic acid monohydrate, *p*TSA **16d**, acting as a reaction initiator.

The hyperspace illustrated in Fig. 4a–d and Supplementary Video 3 now reveals the presence of two distinct yield maxima for the heterocyclic product **16e**—a global maximum marked by a green star and a local maximum marked by a blue star. Further sampling of the conditions' grid in the region between the maxima confirms that they are separated in four-dimensional space by a region of low yield (Fig. 4j and Supplementary Figs. 160c and 161).

Notably, MS analyses (Supplementary Information Section 4.8) evidence distinct signals at these two maxima: the local maximum gives a peak *m/z* ratio corresponding to the iminium ion **18c** expected for the classical Ugi mechanism^{32,33}, whereas the global maximum also features a peak attributable to the oxazoline **17e**. This, in turn, may suggest an extra mechanism (brown arrows) through the oxazoline **17e** intermediate—alternative mechanisms for the Ugi reaction have, indeed, been postulated for decades but never proved^{34,35}. However, such a possibility is directly disqualified by separate, ex roboto experiments in which isolated **17e** fails to give product **16e** on further reaction with *n*-butylamine in the presence of *p*TSA (Supplementary Information Section 4.8.4). Although other mechanisms can also be considered (see magenta arrows in Fig. 4 and corresponding caption), fitting the experimental hyperspace data to the kinetic network of 12 reactions and 15 proton transfer steps (Fig. 4i and theoretical details in Supplementary Information Section 5.3) suggests that an excellent agreement can be achieved (Fig. 4e–h) based only on the Ugi mechanism—that is, inclusion of extra mechanisms offers no perceptible improvement. Instead, the emergence of the two maxima can be attributed to the shifts in the equilibria underlying the network (Supplementary Information Sections 5.3.4 and 5.3.5). We also note that evolution of these maxima following changes in *p*TSA concentration is very gradual (Supplementary Video 3) and the yield distributions are characterized by very small 'slopes', $|D_{ij}| \lesssim 1$ (Supplementary Fig. 166).

A hyperspace supporting a switchable reaction network

Next, we considered the classic Hantzsch pyridine synthesis (Fig. 5), studied for almost 150 years (refs. 36–38) and interesting for several known intermediates and competing pathways (coloured in blue in Fig. 5c). Notably, examining the four-dimensional hyperspace (concentrations of the three substrates, 26 °C and 80 °C temperatures; a total of 2,582 conditions) revealed the presence of many more components than previously thought. Their isolation required several HPLC repurification cycles performed in a closed-loop manner. Specifically, after each cycle, the newly identified substances were added to the spectral unmixing algorithm and the global fits to experimental spectra from the entire hypercube gradually improved (Fig. 5a,b); meanwhile, more

fractions were selected for the next round of purification if they featured yet-unassigned signals (Supplementary Information Sections 4.9 and 4.10). After eight such fitting–purification cycles, the mismatch between the fitted and experimental spectra reached instrumental noise (dashed green line in Fig. 5a), at which point our knowledge of the hyperspace composition can, for all practical reasons, be deemed nearly complete. This knowledge spans not only the seven known species^{36–38} but also nine new ones (coloured red in Fig. 5c) that have not been reported in the classical Hantzsch reaction but can be of interest in the context of biological activity³⁹.

By considering the causal relationships between these and other species discovered within the hyperspace, it is possible to establish the synthetic connectivity of the isolated species (Supplementary Information Section 4.10). This analysis, ultimately, reconstructs the complete network of the Hantzsch reaction detailed in Fig. 5c with key steps confirmed by further ex roboto reactions (marked by green arrows). Notably, analysis of yield distributions (with the help of HPLC, as spectral unmixing of all 16 components is no longer unique) revealed that, by adjusting substrate concentrations at 80 °C, the network can be switched between three different major products (**19d**, **19e**, **19k**), each forming in >60% yield and with maxima located at different corners of the conditions' cube (Fig. 5d–f). One of these switchovers is between the Hantzsch ester (**19d**) and the Petrenko-Kritschenko product (**19e**)⁴⁰, meaning that even the so-called named reactions can, in reality, be part of the same hyperspace. Another observation is that, despite the switchovers, the individual yield distributions for individual species remain smooth, as visualized by the isosurfaces in Fig. 5d–f.

Five-dimensional compositional space

Finally, we consider a hyperspace in which dimensions correspond to systematic variations in composition. Here these compositions are the contents of different metals in the Prussian blue analogues (PBAs)⁴¹ (Fig. 6), which are perovskite-type materials described by a general formula KM_B[M_A(CN)₆] and widely studied in the context of catalysis and energy storage⁴². The robot surveys a five-dimensional space defined by the contents of metals at sites M_A (two types, Fe and Co) and M_B (five types, Mn, Fe, Co, Ni and Cu) over a uniform grid with the granularity of 0.2 (for example, Mn_{0.2}Ni_{0.4}Cu_{0.4}–Fe_{0.6}Co_{0.4} PBA versus Mn_{0.2}Ni_{0.2}Cu_{0.6}–Fe_{0.4}Co_{0.6} PBA, with molar fractions at A and B sites each summing to unity). Each of the 756 PBAs is prepared *in situ* by straightforward co-precipitation method (Fig. 6a,b) and is used to catalyse reaction of styrene **20a** and *t*-BuOOH to give styrene oxide **20c** (given that styrene oxide is the key intermediate to synthesize many fine chemicals and pharmaceuticals⁴³, this reaction has been the subject of several studies using various catalysts, including PBAs^{44,45}). Figure 6c shows the distribution of reaction yields and Fig. 6d quantifies the selectivities with respect to benzaldehyde **20b**, the main by-product of the reaction. As seen, the yield hyperspace is much more corrugated than concentration–temperature hyperspaces from Figs. 2–5 and features

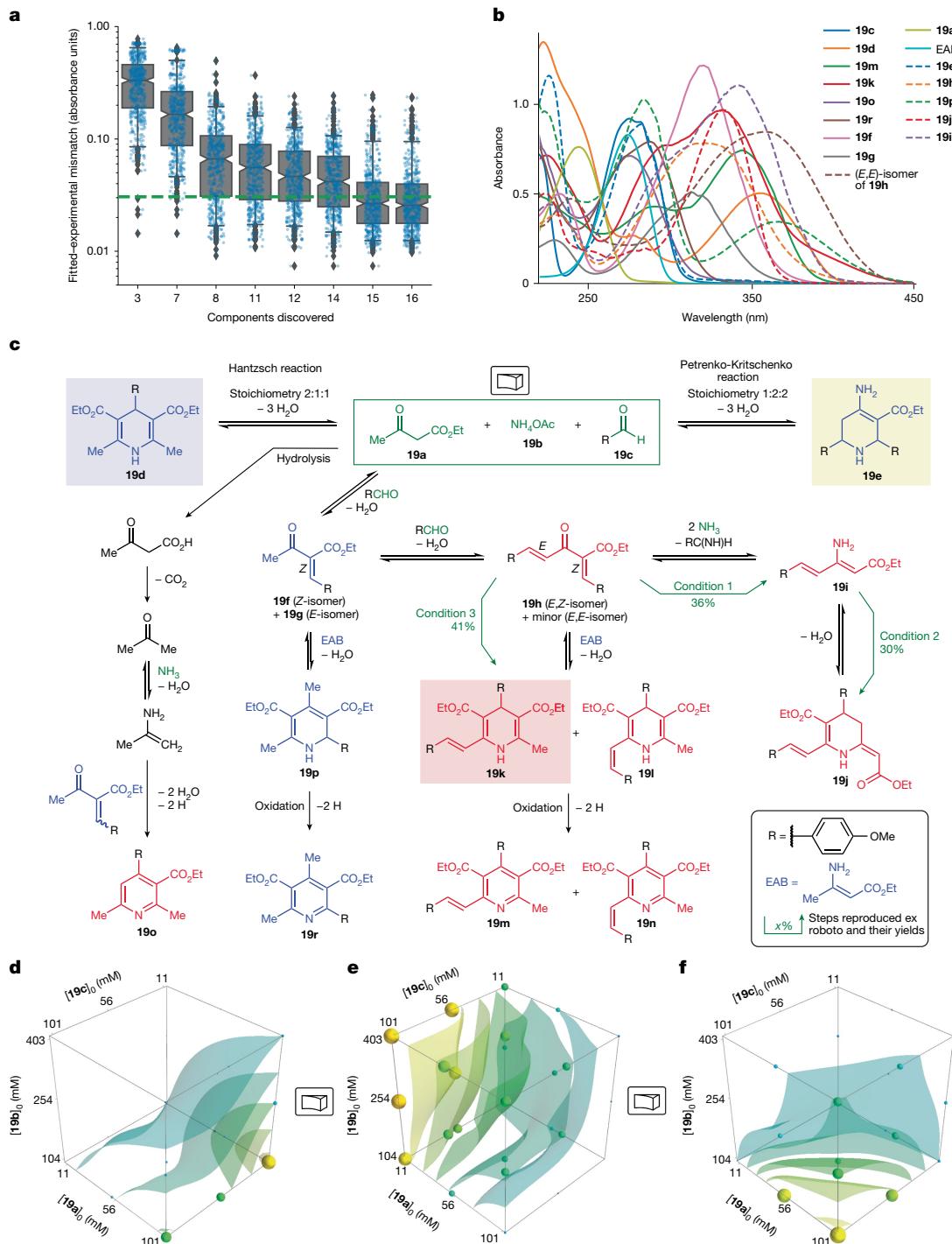


Fig. 5 | 'Switchable' hyperspace of the Hantzsch reaction network.

a, Identification of reaction products by closed-loop repurification and refitting against four-dimensional hypercube UV-Vis spectra (see Supplementary Information Section 4.15). Within eight repurification rounds, 16 components (14 products plus unreacted **19a** and **19c**) were identified. The mismatch between fitted and experimental spectra (vertical axis) decreased to the median value of instrumental uncertainty (dashed green line; see Supplementary Information Section 4.9). The vertical axis shows the highest spectral mismatch per condition. Box plot elements indicate the first quartile, median and third quartile, whiskers show the 5th and 95th percentiles, with outliers as black diamonds. Blue scatter plot shows individual data points. **b**, UV-Vis 'basis set' of the 16 components. Their discovery follows the cycles in **a**: 3 components (**19a**, **19c**, **19d**), 7 components (+**19k**, **19r**, **19o**, **19m**), 8 components (+EAB), 11 components (+**19j**, **19f**, **19g**), 12 components (+**19e**), 14 components (+**19h**, **19p**), 15 components (+**19i**), 16 components (+**19h**). **c**, Reaction network: green,

substrates; blue, isolated known products/intermediates; red, new products/intermediates. Network connectivity was verified by separate reactions (green arrows): (1) $\text{NH}_4\text{OAc}, \text{AcOH}, 80^\circ\text{C}, 18\text{ h}$; (2) ethyl acetoacetate, EtOH, $80^\circ\text{C}, 18\text{ h}$; (3) EAB, EtOH, $80^\circ\text{C}, 24\text{ h}$. Note the presence of products of two separate 'named reactions', Hantzsch (**19d**) and Petrenko-Kritschenko (**19e**) (mechanisms are shown in Supplementary Information Schemes 2 and 3). **d–f**, Within the hyperspace, the reaction network at 80°C can be switched between three main products observed in different corners of the space: Hantzsch ester **19d** (maximum yield 63.5%, <3% **19p**) (**d**); Petrenko-Kritschenko product **19e** (maximum yield about 66%) (**e**); benzylidene-extended Hantzsch derivative **19k** (maximum yield 61.7%) (**f**). These yields (after 48 h) were quantified by HPLC and are colour-coded according to the legend. Isosurfaces correspond to 20%, 30%, 40%, 50% and 60% yield. For full experimental details, see Supplementary Information Sections 4.7, 4.9 and 4.10.

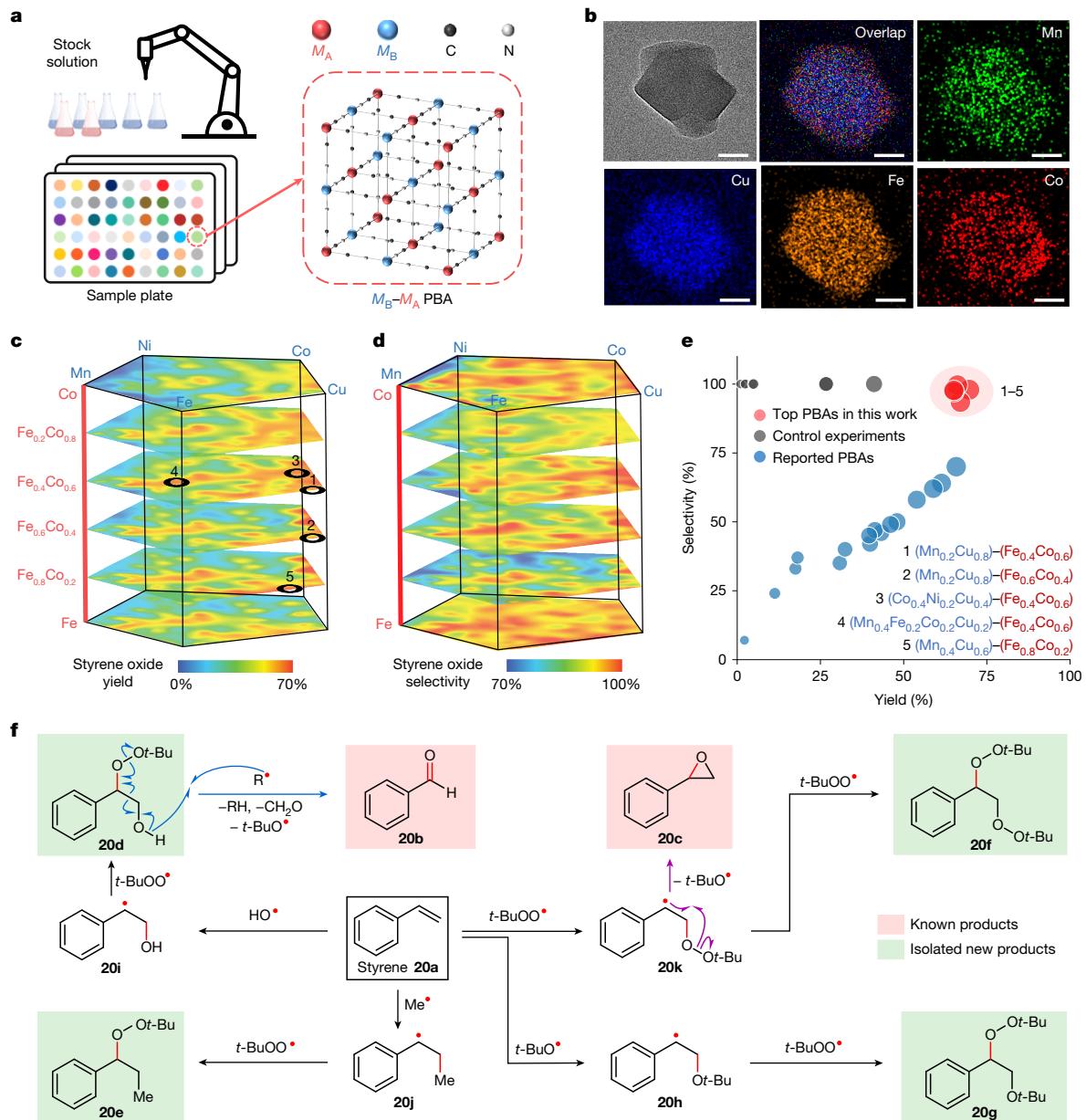


Fig. 6 | Five-dimensional space of catalyst compositions. **a**, 756 PBAs, catalysts with a general formula $KM_B[M_A(CN)_6]$, were prepared by the robot in separate vials using the co-precipitation method. Up to two aqueous solutions of M_A and up to five of M_B were pipetted automatically with specified volumes to prepare the PBA catalyst suspensions, each with a final volume of 200 μ l in individual vials, followed by 24 h of ageing at room temperature. **b**, Irrespective of the composition, these nanomaterials were particles approximately 50 nm in size. Uniformity of metal distributions was confirmed by energy-dispersive spectroscopy. For large-area scanning electron microscopy images and all experimental details, see Supplementary Information Section 8. Scale bars, 20 nm. **c**, Five-dimensional hyperspace of yields of styrene oxide reconstructed by spectral unmixing. **d**, Five-dimensional hyperspace of selectivity of styrene oxide reconstructed by spectral unmixing. **e**, At least five PBAs (red markers 1–5

here and black circles in the corresponding hyperspace in **c**) showed yield-selectivity properties better than previously studied PBAs (blue markers), as well as some other controls (grey markers: Mn^{2+} , Fe^{2+} , Co^{2+} , Ni^{2+} , Cu^{2+} , $Fe(CN)_6^{3-}$, $Co(CN)_6^{3-}$ and the last one without any catalyst; see raw data in Supplementary Table 34). **f**, The robotically reconstructed hyperspace allowed for the identification of four products unreported in previous PBA studies of this reaction (**20d**, **20e**, **20f** and **20g** shown on green backgrounds) and for the reconstruction of the reaction's mechanistic network involving four intermediates **20h**, **20i**, **20j** and **20k**. The blue fishhook arrows on species **20d** indicate the electron flow to furnish aldehyde **20b** ($R = t\text{-BuOO}^\bullet$, $t\text{-BuO}^\bullet$ or HO^\bullet), whereas the purple fishhook arrows on species **20k** indicate the electron flow to furnish epoxide **20c** (ref. 56). For the generation of radicals from *tert*-butyl hydroperoxide, see Supplementary Information Section 8.7.

several local yield maxima. Notably, it contains several PBA compositions that offer better yield-selectivity characteristics than previously reported PBAs^{44,45} or other controls (Fig. 6e). As in other examples that we considered, the hyperspace reconstruction reveals the presence of unreported intermediates and by-products and helps reconstruct a mechanistic network (Fig. 6f) that had previously remained elusive^{44–46}.

Discussion

Hyperspace structure

One of the insights from these studies is the relatively simple structure of the concentration–temperature spaces, with yield distributions of any single species featuring, at most, two yield maxima and with the slope values $|D_{ij}|$ at or below unity (Supplementary Fig. 166).

As described in Supplementary Information Sections 7.1 and 7.2, it can be rigorously proved that, for strongly connected, directed hypergraphs^{47,48} (that is, those that cannot be split into non-interacting subsets of reactions) and for first-order and pseudo-first-order reactions, $|D_{ij}| \leq 1$. For higher-order kinetics of individual reactions, the theoretical upper bound on $|D_{ij}|$ is too high to be practically relevant, but numerical studies of networks comprising first-order and second-order steps and no cycles confirm that $|D_{ij}|$ is also typically on the order of unity. A corollary to this result is that, along concentration coordinates C_{0j} , $j = 1, \dots, k$, the hypersurfaces can be examined faithfully at sparse intervals (see Supplementary Information Section 7.6). This, in turn, can help reduce the number of experiments during conditions' screening and optimization campaigns. For more discussion of hyperspace properties (in terms of topology, differential yields and Shannon information theory), see Supplementary Figs. 161 and 162 and Supplementary Information Section 4.12.

Naturally, these considerations do not apply to hyperspaces in which compositions are varied. For instance, different catalysts (as in Fig. 6) can lower the activation barriers substantially and to different extents. As a result, they can alter the network-wide kinetics more than changes over a limited range of concentrations or temperatures, resulting in 'steeper' concentration derivatives and in the hyperspace featuring many local yield maxima.

Network reconstruction

The method by which we reconstructed reaction networks is loosely analogous to the approaches used in electronics to reverse-engineer a circuit inside a 'black box' by applying inputs and analysing corresponding outputs. Here the inputs are the various substrate concentrations and temperatures and the outputs are the identities of the products found by the global analysis of the hyperspace. With the set of products identified, the 'wiring' of the network is prescribed by the general rules of chemical reactivity. In the cases studied here, the wirings could be reconstructed relatively readily by humans but, as illustrated in Supplementary Information Section 4.11, this process can also be aided by network algorithms operating either at the level of mechanistic steps^{4,49} or full reactions⁵⁰, thereby linking the hyperspace of conditions to the 'space of reaction grammars'⁵, a link argued to be key to chemical innovation⁵. In this effort, the broad investigation of the hyperspace is essential, as different inputs can serve to 'activate' or 'deactivate' different branches of the network (by modifying reaction rates and shifting equilibria), enabling the formation of as many products as possible. This also means investigating the reaction space at stoichiometries that a human chemist might not necessarily find intuitive or relevant to a particular reaction. Here this approach more than doubled the knowledge of the Hantzsch hyperspace and we have seen similar discovery enhancements in other reactions that we have studied since then (for example, Pechmann, Biginelli). From a technical point of view, our ability to interrogate entire spaces of conditions capitalizes on the use of inexpensive (cents per sample) optical detection, which reduces the need for costly (about \$45–300 per sample) HPLC/quantitative NMR analyses by a factor of several hundred (for example, one HPLC/NMR cycle per 1,085 UV-Vis experiments to reconstruct the S_N1 network in Fig. 3; eight closed-loop cycles to ensure complete knowledge of the Hantzsch space examined by UV-Vis at 2,582 conditions). As long as the equilibrium is not fully reached (Supplementary Figs. 154 and 160), sweeping the reaction rates through control of starting concentrations accesses the type of data similar to conventional kinetic experiments (that 'sweep' the time to observation instead) and estimation of kinetic parameters becomes possible.

Network control

A related aspect is the practical ability to direct these reaction networks towards desired outcomes by simply adjusting concentrations (that is, without using different reagents, as done in some recent excellent

papers, such as ref. 51). For simple reactions, such stoichiometric control is, of course, well known (for example, we can influence the ratio of monosubstituted to disubstituted product of alkyl dibromide by using one versus two equivalents of the nucleophile). However, for multicomponent reaction mixtures, it is less obvious that they can be pushed cleanly towards different major products. We have seen this capability realized for the Hantzsch reaction, in which the yields of major products **19d**, **19e** and **19k** were maximized at concentration ratios very different from those we may expect based on product stoichiometry alone. For instance, product **19e** incorporates one copy of substrate **19a**, two copies of **19b** and two copies of **19c**, but in the hyperspace in Fig. 5e, the regions of appreciable yield, >40%, start from about 1:4.5:2 ratio (and continue for even higher molar excess of **19b** at the yield maximum, which, however, becomes wasteful in terms of substrate use; see Extended Data Fig. 2 and Supplementary Information Section 4.17).

Conclusions

In conclusion, this work is an initial effort—made possible by modern reaction automation—to understand the structure of reaction hyperspaces, one of the five foundational spaces of mathematical chemistry⁵. It reinforces the view of chemical reactions as networks^{4,51,52} embedded in multidimensional spaces of conditions and, at least in some cases, switchable between different major products, which is reminiscent of some biochemical networks^{53,54} and promising in terms of diversity-oriented synthesis⁵⁵. The rapid and cost-effective approach to hyperspace reconstruction can (1) systematize and accelerate reaction discovery and optimization and (2) foster fundamental research on reaction networks as dynamic systems. This effort can be aided by hyperspace visualization and analysis tools such as those we developed and make available at Zenodo (<https://doi.org/10.5281/zenodo.14880579>; see also Supplementary Video 4). In a broader context, inclusion of our experimental yield maps to the benchmarks used for testing the general-purpose yield optimization algorithms^{9,11,15} would appreciably widen the diversity of these benchmarks, given the current scarcity of such datasets^{7,9,13,15}. In future work, we aim to extend our robotic platform to reactions that require solid dispensing and/or strictly oxygen-free conditions (this will broaden the scope of hyperspaces we can analyse), to accelerate dispensing and measuring operations (to analyse fast reactions) and also to monitor hyperspace evolution over time, which will be necessary to reconstruct highly nonlinear mechanisms (for example, oscillating reactions).

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-025-09490-1>.

- Skoraczyński, G. et al. Predicting the outcomes of organic reactions via machine learning: are current descriptors sufficient? *Sci. Rep.* **7**, 3582 (2017).
- Saebi, M. et al. On the use of real-world datasets for reaction yield prediction. *Chem. Sci.* **14**, 4997–5005 (2023).
- Liu, Z., Moroz, Y. S. & Isayev, O. The challenge of balancing model sensitivity and robustness in predicting yields: a benchmarking study of amide coupling reactions. *Chem. Sci.* **14**, 10835–10846 (2023).
- Szymkuć, S., Wotus, A., Roszak, R. & Grzybowski, B. A. Estimation of multicomponent reactions' yields from networks of mechanistic steps. *Nat. Commun.* **15**, 10286 (2024).
- Restrepo, G. Spaces of mathematical chemistry. *Theory Biosci.* **143**, 237–251 (2024).
- Granda, J. M., Donina, L., Dragone, V., Long, D. L. & Cronin, L. Controlling an organic synthesis robot with machine learning to search for new reactivity. *Nature* **559**, 377–381 (2018).
- Lin, S. et al. Mapping the dark space of chemical reactions with extended nanomole synthesis and MALDI-TOF MS. *Science* **361**, eaar6236 (2018).
- Coley, C. W. et al. A robotic platform for flow synthesis of organic compounds informed by AI planning. *Science* **365**, eaax1566 (2019).

9. Angeloo, N. H. et al. Closed-loop optimization of general reaction conditions for heteroaryl Suzuki-Miyaura coupling. *Science* **378**, 399–405 (2022).
10. Rohrbach, S. et al. Digitization and validation of a chemical synthesis literature database in the ChemPU. *Science* **377**, 172–180 (2022).
11. Slattery, A. et al. Automated self-optimization, intensification, and scale-up of photocatalysis in flow. *Science* **383**, eadj1817 (2024).
12. Dai, T. et al. Autonomous mobile robots for exploratory synthetic chemistry. *Nature* **635**, 890–897 (2024).
13. Buitrago Santanilla, A. et al. Nanomole-scale high-throughput chemistry for the synthesis of complex molecules. *Science* **347**, 49–53 (2015).
14. Davies, I. W. The digitization of organic synthesis. *Nature* **570**, 175–181 (2019).
15. Shields, B. J. et al. Bayesian reaction optimization as a tool for chemical synthesis. *Nature* **590**, 89–96 (2021).
16. Wilbraham, L., Mehr, S. H. M. & Cronin, L. Digitizing chemistry using the chemical processing unit: from synthesis to discovery. *Acc. Chem. Res.* **54**, 253–262 (2021).
17. Mahjour, B. et al. Rapid planning and analysis of high-throughput experiment arrays for reaction discovery. *Nat. Commun.* **14**, 3924 (2023).
18. Wang, J. Y. et al. Identifying general reaction conditions by bandit optimization. *Nature* **626**, 1025–1033 (2024).
19. Strieth-Kalhoff, F. et al. Artificial intelligence for retrosynthetic planning needs both data and expert knowledge. *J. Am. Chem. Soc.* **146**, 11005–11017 (2024).
20. Stadler, E. et al. A versatile method for the determination of photochemical quantum yields via online UV-Vis spectroscopy. *Photochem. Photobiol. Sci.* **17**, 660–669 (2018).
21. Lu, J.-M. et al. Robotized AI-assisted microfluidic photocatalytic synthesis and screening up to 10,000 reactions per day. *Nat. Commun.* **15**, 8826 (2024).
22. Bioucas-Dias, J. M. et al. Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **5**, 354–379 (2012).
23. Banert, K. & Kurnianto, A. Nucleophile substitution bei 4,4-dimethyl-2-adamantyl-substraten: rückseitenangriff bei 2-adamantan-derivaten. *Chem. Ber.* **119**, 3826–3841 (1986).
24. Thibblin, A. & Sidhu, H. Mechanisms of competing solvolytic elimination and substitution reactions. The role of ion-pair intermediates in aqueous solvents. *J. Chem. Soc., Perkin Trans. 2* **22**, 1423–1428 (1994).
25. Clennan, E. L. Aromatic endoperoxides. *Photochem. Photobiol.* **99**, 204–220 (2023).
26. Klaper, M., Wessig, P. & Linker, T. Base catalysed decomposition of anthracene endoperoxide. *Chem. Commun.* **52**, 1210–1213 (2016).
27. Ammer, J., Sailer, C. F., Riedle, E. & Mayr, H. Photolytic generation of benzhydryl cations and radicals from quaternary phosphonium salts: how highly reactive carbocations survive their first nanoseconds. *J. Am. Chem. Soc.* **134**, 11481–11494 (2012).
28. Nishimae, Y., Kurata, H. & Oda, M. Arylbis(9-anthryl)methyl cations: highly crowded, near infrared light absorbing hydrocarbon cations. *Angew. Chem. Int. Ed.* **43**, 4947–4950 (2004).
29. Nojima, M., Takagi, M., Morinaga, M., Nagao, G. & Tokura, N. Reaction of some triarylmethyl radicals, polyaryllkenes, and 9,10-dihydro-9,10-epidioxyanthracenes with sulphur dioxide; detection of radicals and/or cations derived from the corresponding cation radicals. *J. Chem. Soc., Perkin Trans. 1* **15**, 488–495 (1978).
30. Hollenstein, S. & Laali, K. K. Efficient conversion of 9-isopropenylphenanthrene to 4,6,6-trimethyl-6H-benz[de]anthracene in FSO₃H; 5,6-dihydro-4H-benzanthracen-4-iun ion and its charge delocalization mode. *Chem. Commun.* 2145–2146 (1997).
31. Cankafová, N., Nemec, I. & Krchňák, V. p-TSA-mediated four-component reaction: one-step access to mesoionic 1*H*-imidazol-3-iun-4-olates, direct NHC precursors. *Adv. Synth. Catal.* **364**, 2996–3003 (2022).
32. Medeiros, G. A. et al. Probing the mechanism of the Ugi four-component reaction with charge-tagged reagents by ESI-MS/(MS). *Chem. Commun.* **50**, 338–340 (2014).
33. Rocha, R. O., Rodrigues, M. O. & Neto, B. A. D. Review on the Ugi multicomponent reaction mechanism and the use of fluorescent derivatives as functional chromophores. *ACS Omega* **5**, 972–979 (2020).
34. Alvim, H. G. O., da Silva Júnior, E. N. & Neto, B. A. D. What do we know about multicomponent reactions? Mechanisms and trends for the Biginelli, Hantzsch, Mannich, Passerini and Ugi MCRs. *RSC Adv.* **4**, 54282–54299 (2014).
35. Chéron, N., Ramozzi, R., Kaim, L. E., Grimaud, L. & Fleurat-Lessard, P. Challenging 50 years of established views on Ugi reaction: a theoretical approach. *J. Org. Chem.* **77**, 1361–1366 (2012).
36. Hantzsch, A. Condensationsprodukte aus Aldehydammoniak und ketonartigen Verbindungen. *Ber. Dtsch. Chem. Ges.* **14**, 1637–1638 (1881).
37. Shen, L. et al. A revisit to the Hantzsch reaction: unexpected products beyond 1,4-dihydropyridines. *Green Chem.* **11**, 1414–1420 (2009).
38. Santos, V. G. et al. The multicomponent Hantzsch reaction: comprehensive mass spectrometry monitoring using charge-tagged reagents. *Chem. Eur. J.* **20**, 12808–12816 (2014).
39. Chang, C.-C. et al. Antagonism of 4-substituted 1,4-dihydropyridine-3,5-dicarboxylates toward voltage-dependent L-type Ca²⁺ channels Ca_v1.3 and Ca_v1.2. *Bioorg. Med. Chem.* **18**, 3147–3158 (2010).
40. Petrenko-Kritschko, P. Über die kondensation des acetondicarbonsäureesters mit aldehyden, ammoniak und aminen. *J. Prakt. Chem.* **85**, 1–37 (1912).
41. Singh, B. & Indra, A. Prussian blue- and Prussian blue analogue-derived materials: progress and prospects for electrochemical energy conversion. *Mater. Today Energy* **16**, 100404 (2020).
42. Li, W. et al. Chemical properties, structural properties, and energy storage applications of Prussian blue analogues. *Small* **15**, 1900470 (2019).
43. Choo, J. P. S. & Li, Z. Styrene oxide isomerase catalyzed Meinwald rearrangement reaction: discovery and application in single-step and one-pot cascade reactions. *Org. Process Res. Dev.* **26**, 1960–1970 (2022).
44. Guo, S. et al. Synthesis of trimetallic Prussian blue analogues and catalytic application for the epoxidation of styrene. *Ind. Eng. Chem. Res.* **59**, 13831–13840 (2020).
45. Liang, Y. et al. Prussian blue analogues as heterogeneous catalysts for epoxidation of styrene. *RSC Adv.* **5**, 17993–17999 (2015).
46. Zhang, L., Zhang, Z., He, X., Zhang, F. & Zhang, Z. Regulation of the products of styrene oxidation. *Chem. Eng. Res. Des.* **120**, 171–178 (2017).
47. Pal, A. et al. Finding thermodynamically favorable pathways in chemical reaction networks using flows in hypergraphs and mixed-integer linear programming. *J. Chem. Inf. Model.* **65**, 6772–6787 (2025).
48. Grzybowski, B. A., Bishop, K. J. M., Kowalczyk, B. & Wilmer, C. E. The ‘wired’ universe of organic chemistry. *Nat. Chem.* **1**, 31–36 (2009).
49. Krzeszewski, M. et al. Computer-generated, mechanistic networks assist in assigning the outcomes of complex multicomponent reactions. *J. Am. Chem. Soc.* **147**, 15636–15644 (2025).
50. Mikulak-Klucznik, B., Klucznik, T., Beker, W., Moskal, M. & Grzybowski, B. A. Catalyst: curtailing the scalable supply of fentanyl by using chemical AI. *Chem* **10**, 1319–1326 (2024).
51. Mahjour, B., Shen, Y., Liu, W. & Cernak, T. A map of the amine–carboxylic acid coupling system. *Nature* **580**, 71–75 (2020).
52. Baltussen, M. G., de Jong, T. J., Duez, Q., Robinson, W. E. & Huck, W. T. S. Chemical reservoir computation in a self-organizing reaction network. *Nature* **631**, 549–555 (2024).
53. Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314**, 1585–1588 (2006).
54. Daniel, R., Rubens, J. R., Sarapeshkar, R. & Lu, T. K. Synthetic analog computation in living cells. *Nature* **497**, 619–623 (2013).
55. Wołos, A. et al. Computer-designed repurposing of chemical wastes into drugs. *Nature* **604**, 668–676 (2022).
56. Halder, J. et al. Insight of solvent effect on CeO₂ catalyzed oxidation of styrene with tert-butyl hydroperoxide: a combined experimental and theoretical approach. *Catal. Commun.* **164**, 106413 (2022).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025

Methods

Automation platform

Reagent addition was carried out using a commercial pipetting module (ZEUS, Hamilton), optimized for pipetting a range of organic solvents in aliquots ranging from 10 μl to 1 ml, with volumetric errors of less than 1% for volumes greater than 50 μl and less than 2% for volumes between 10 and 50 μl . Liquids were dispensed into 2-ml glass vials arranged in custom-designed 54-well plates, with each reaction typically set to a final volume of 500 μl . Following reagent addition, vials were hermetically sealed using a flat lid, with a rubber sheet and perfluoroalkoxy film placed between the lid and the vial. Initial mixing was performed on an orbital shaker at 250 rpm for 5 min. Subsequent stirring is omitted, as separate studies (see Supplementary Information Sections 2.9 and 4.5) have shown that, for vials of these dimensions, passive mixing over the course of the reaction (hours) is as efficient as mechanical stirring, thereby eliminating the need for cumbersome stir bars.

After reactions were run for desired times, the crude mixture from each vial was, if necessary, automatically diluted to align with the detection range of the UV-Vis spectrophotometer (NanoDrop, Thermo Fisher Scientific). The operation of the spectrophotometer was also automated, including lid opening and closing, along with the flushing and drying of the measuring pedestal (see Supplementary Video 2). All of these operations were orchestrated by house-written software. The entire assembly was placed inside a hood and, if needed, constantly purged with nitrogen. The manual operations in the entire workflow were changing the 54-well plates, covering them with rubber/perfluoroalkoxy sheets and placing on the orbital shaker. For some hyperspaces, two such systems were used. Further technical details and the blueprints for system replication are provided in Supplementary Information Section 2 and at Zenodo (<https://doi.org/10.5281/zenodo.14880579>)⁵⁷. The software design is illustrated in Supplementary Fig. 2 and the source code is deposited at <https://github.com/yaroslavsvobolev/robowski-maps>.

The full cubes of 775 E1 conditions (five temperatures, 16–36 °C; five initial substrate concentrations, 1.5–15.0 mM; 31 HBr concentrations, 0.1–10.0 mM; reaction time: 4 h) and 930 S_N1 conditions (six temperatures; 16–66 °C; five 9-butyl-9H-fluoren-9-ol substrate concentrations, 0.03–0.30 M; 31 HBr concentrations; reaction time: 48 h) are provided in Supplementary Information Section 4.15, with all raw data in the Zenodo repository.

Stoichiometry constraints

Stoichiometry constraints, as shown in Fig. 1e, were included into the spectral unmixing algorithm by modifying the log-likelihood function L to be minimized: L is a sum of squared spectral mismatch (yellow) of experimental $A_{\text{measured}}(\lambda)$ and modelled $A_{\text{model}}(\lambda)$ absorbances, divided by the instrumental variance of the spectrophotometer $\sigma^2(\lambda, A_{\text{measured}}(\lambda))$, and the term representing the violation of stoichiometric inequalities (red) containing the Heaviside function $\theta(x)$ and the experimental uncertainties $\sigma_{[A]_0}$ and $\sigma_{[B]_0}$ of preparing the reaction mixture with given starting concentrations. See Supplementary Information Section 3 for more details about the spectral unmixing algorithm

Isosurface extraction

The isosurfaces in Figs. 2–5 were extracted by the marching cubes algorithm⁵⁸ operating on a 50 × 50 × 50 regular grid obtained through the radial basis functions interpolator applied to the raw data. This method was implemented in the HyperspaceViewer software, which is provided in the Supplementary Information and is open-sourced at a GitLab repository (https://gitlab.com/az-steak/hyperspace_viewer).

Kinetic fitting

The kinetic fitting in Fig. 2e,f was done by numerically integrating the equations of the mechanism and optimizing the parameters to achieve

the best-fit of the model's predicted product yields simultaneously to all experimental data. This procedure gave reasonable estimates of the reactions' thermodynamic parameters: $\Delta H = -30.7 \pm 1.4 \text{ kJ mol}^{-1}$ for S_N1 and $\Delta H = 81.98^{+3.52}_{-1.64} \text{ kJ mol}^{-1}$, $\Delta S = 273.8^{+10.8}_{-4.9} \text{ J mol}^{-1} \text{ K}^{-1}$ and $\Delta H^\ddagger = 21.10^{+0.33}_{-0.49} \text{ kJ mol}^{-1}$ (transition state enthalpy) for E1. For derivations of these values and comparisons against related literature examples, see Supplementary Information Sections 5.1 and 5.2.

Initial guess of kinetic parameters (Fig. 4e–h) was provided by the Markov chain Monte Carlo algorithm⁵⁹, kinetic equations at each condition were integrated numerically and the discrepancy against the experimental yield distribution (over the hyperspace) was minimized using the trust region reflective algorithm⁶⁰. See Supplementary Information Section 5.3 for further details of the kinetic model.

PBA-catalysed styrene epoxidation

PBAs were synthesized following a reported co-precipitation method⁶¹. Two stock solutions containing metal M_A (Fe, Co) were prepared by dissolving K₃[Fe(CN)₆] or K₃[Co(CN)₆] in deionized water to form clear solutions of type A (0.10 M). Five stock solutions containing metal M_B (Mn, Fe, Co, Ni, Cu) were prepared separately by dissolving corresponding metal nitrates/chlorides (Mn(NO₃)₂, FeCl₂, Co(NO₃)₂, Ni(NO₃)₂, Cu(NO₃)₂, each at 0.10 M) in the presence of trisodium citrate dihydrate (Na₃C₆H₅O₇·2H₂O, 0.1125 M) in deionized water to form clear solutions of type B. Afterwards, mixing of solutions of type A and type B were performed in specified volume ratios using the automated liquid-handling system according to the target composition. For instance, for a PBA with composition Mn_{0.2}Ni_{0.4}Cu_{0.4}–Fe_{0.6}Co_{0.4}, the sequence and volume of stock solution addition were as follows: 20 μl Mn(NO₃)₂, 40 μl Ni(NO₃)₂, 40 μl Cu(NO₃)₂, 60 μl K₃[Fe(CN)₆] and 40 μl K₃[Co(CN)₆]. The total volume of solutions of type A was 100 μl and the total volume of solutions of type B was also 100 μl —therefore, the total volume of each PBA solution was always 200 μl . After sealing, each 54-vial well plate housing different mixed solutions was shaken for 15 min (250 rpm) and aged at room temperature for 24 h to form PBAs.

For a typical set of styrene epoxidation reactions, the reaction stock solution was prepared as follows: first, styrene (0.50 M) and *tert*-butyl hydrogen peroxide (0.75 M) were dissolved in 100 ml acetonitrile. Then, 30 ml cetyltrimethylammonium bromide aqueous solution (0.25 M) was added to the mixture. Subsequently, the obtained mixture was sonicated for 5 min to obtain a homogeneous stock solution. Afterwards, the robotic system pipetted 1.3-ml aliquots of the thus prepared stock solution into the vials containing previously prepared PBAs (200 μl aqueous suspension). The reaction vials were then sealed and every reaction plate was placed on a thermal shaker (72 °C, 700 rpm) for 6 h. After reaction, when the PBAs sedimented, the robot acquired 40 μl of the supernatant from every vial, diluted it 1,000 times and then measured the UV-Vis spectrum with the NanoDrop spectrophotometer. The yield of styrene oxide (y_{so}) and benzaldehyde (y_{alde}) for each sample can be obtained by spectral unmixing and the selectivity of styrene, calculated by $y_{\text{so}}/(y_{\text{so}} + y_{\text{alde}}) \times 100\%$, was then obtained.

Data availability

All data in support of the findings of this study are included in the Supplementary Information; furthermore, raw spectra and their fitted results for all automated reactions (9,362 data points in total) are deposited at Zenodo (<https://doi.org/10.5281/zenodo.14880579>)⁵⁷. Deposition numbers 2418839 (**15b**) and 2418838 (**15b**·C₃H₇NO) containing the crystallographic data are available free of charge from the Cambridge Crystallographic Data Centre at www.ccdc.cam.ac.uk/data_request/cif.

Code availability

All codes and pipelines used for this study are provided at GitHub (<https://github.com/yaroslavsvobolev/robowski-maps>). HyperspaceViewer

software for the visualization and analysis of hyperspaces is deposited at GitLab (https://gitlab.com/az-steak/hyperspace_viewer) and Zenodo (<https://doi.org/10.5281/zenodo.14880579>⁵⁷), along with the source code and user manual, and as the HyperspaceViewer.zip Supplementary File. Furthermore, We release a Google Colab notebook (https://colab.research.google.com/drive/1uXaoWjkHYaPoCt8N6q4GzYY3e92T_nC?usp=sharing) with the step-by-step tutorial of the core functionality of calibration and spectral unmixing for the Claisen–Schmidt condensation case shown in Fig. 1c,d.

57. Jia, Y. et al. Code and raw data for ‘Robot-assisted mapping of chemical reaction hyperspaces and networks’. Zenodo <https://doi.org/10.5281/zenodo.14880579> (2025).
58. Schroeder, W., Martin, K. & Lorensen, B. *The Visualization Toolkit: An Object-oriented Approach to 3D Graphics* 4th edn (Kitware, 2006).
59. Goodman, J. & Weare, J. Ensemble samplers with affine invariance. *Commun. Appl. Math. Comput. Sci.* **5**, 65–80 (2010).
60. Branch, M. A., Coleman, T. F. & Li, Y. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. *SIAM J. Sci. Comput.* **21**, 1–23 (1999).
61. Du, M. et al. High-entropy Prussian blue analogues and their oxide family as sulfur hosts for lithium-sulfur batteries. *Angew. Chem. Int. Ed.* **61**, e202209350 (2022).

Acknowledgements We thank A. Shevlyakov for helpful discussions. We thank Fan Guo for assistance with the EDS analysis. This work was generously supported by the taxpayers of South Korea through the Institute for Basic Science, project code IBS-R020-D1.

Author contributions Y. Jia and Y.I.S. constructed the automation hardware, designed its software, conducted experiments with R.F. and B.P., performed data analysis and wrote the initial draft with R.F. R.F. contributed to automation design, led the chemical part of the

project, selected, performed and interpreted all reactions in roboto and performed MS and MS/MS mechanistic studies. Y.I.S. designed software for spectral unmixing and numerical models of reaction kinetics (with input from R.F. and W.-S.W.), fitted kinetics to the experimental data, performed TD-DFT calculations of absorption spectra, quantum thermochemistry calculations and theoretical study of the smoothness of yield maps. B.P. performed all Hantzsch reactions with the robotic platform and partial HPLC purifications of the Hantzsch reaction and performed one reaction for versatility test. L.G. proposed the reaction, made the proof-of-concept test and determined products structure and reaction mechanism for PBA catalysts. J.C.A. worked on synthesis, products isolation, structure determination and mechanism analysis for PBA-catalysed reactions. J.C.A. also performed and analysed the Suzuki–Miyaura coupling experiments. Y. Jiang worked on PBA synthesis, reactions on robotic platform, data analysis and wrote the initial draft for the PBA part. M.D. developed the HyperspaceViewer software. L.G. wrote the manual for the HyperspaceViewer software. W.-S.W. worked on four reactions for the versatility test, assigned structures to unexpected Hantzsch reaction products and proposed the corresponding mechanisms. Y.B. optimized S_N1 and E1 reactions and their adaptation to the robotic platform. G.M. and others conducted a reaction versatility test. G.M. and Y.B. conducted the system accuracy test. D.M. contributed to robotic synthesis (Hantzsch reaction), separation of crude mixtures, purification, NMR analysis of isolated compounds and performed one reaction for the versatility test. N.K. analysed single-crystal X-ray diffraction data. E.S.L. conducted synthesis and aided with the analysis and determination for the rearrangement reaction of anthracen-9-ylidiphenylmethanol and analysed several sets of single-crystal X-ray diffraction data to determine their three-dimensional structures. B.A.G. conceived and supervised the project.

Competing interests The authors declare no competing interests.

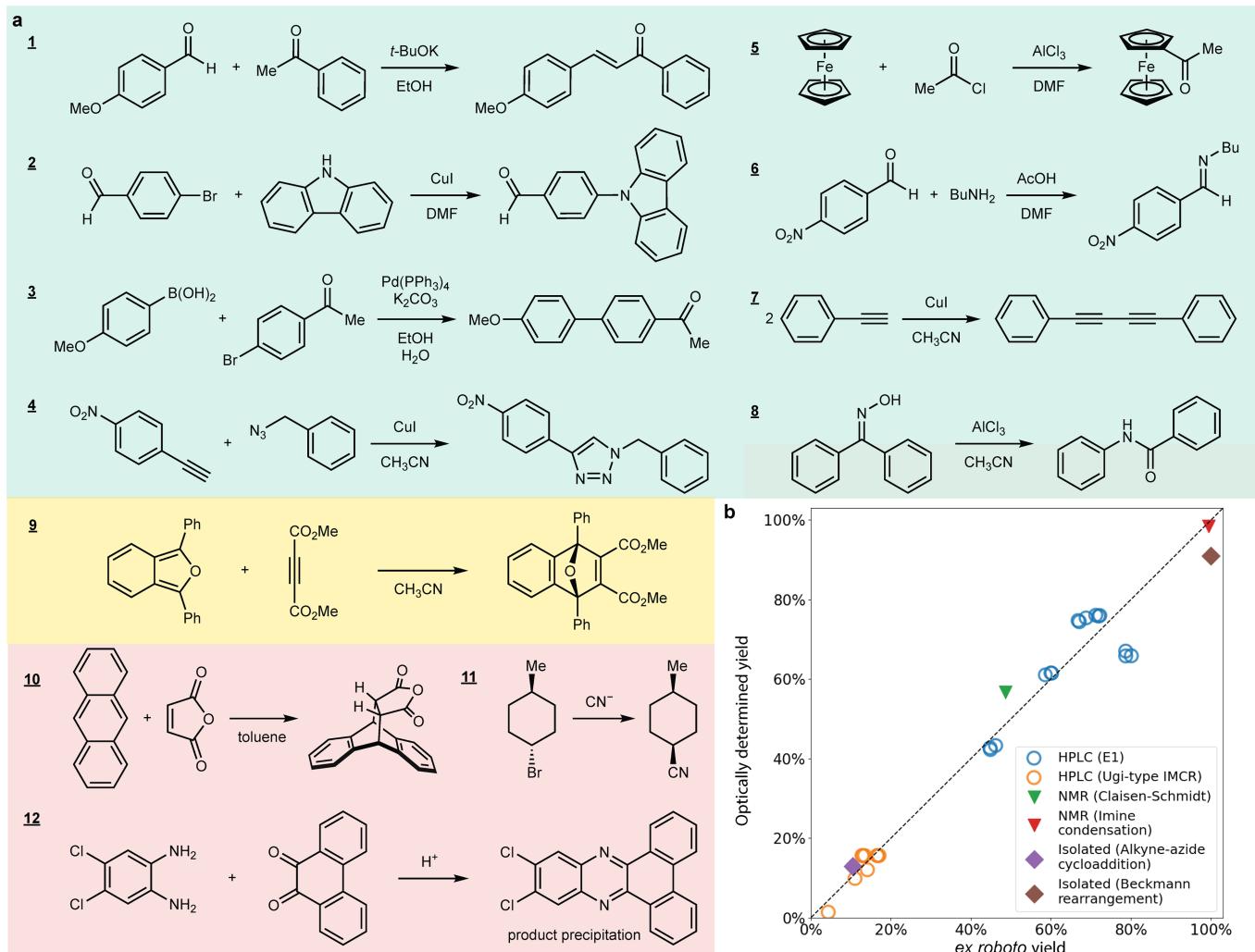
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-025-09490-1>.

Correspondence and requests for materials should be addressed to Yanqiu Jiang or Bartosz A. Grzybowski.

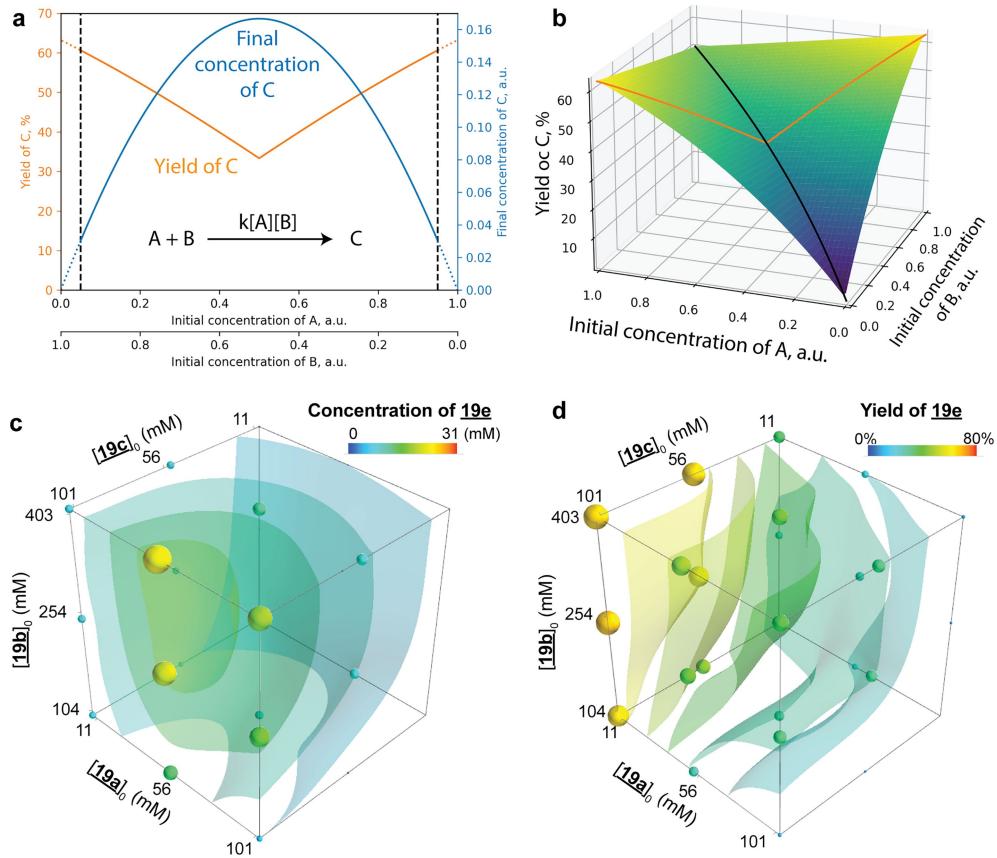
Peer review information *Nature* thanks Guillermo Restrepo and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | Partial scope and quality of optical yield estimates. **a**, Examples of reactions compatible (green background), borderline (yellow) and incompatible (pink) with yield determination using UV-Vis and spectral unmixing. The compatible reactions include but are not limited to Claisen-Schmidt condensation **1**, Ullmann-type coupling **2**, Suzuki-Miyaura cross-coupling **3**, Cu(I)-catalysed alkyne-azide cycloaddition **4**, Friedel-Crafts acylation **5**, imine condensation **6**, Glaser-type coupling **7** or Beckmann rearrangement **8**. The Diels-Alder reaction **9** shown on yellow background is borderline in terms of the quality of unmixing (the overlap of the UV-Vis signals from the product and from dimethyl acetylenedicarboxylate prevents accurate fitting, leading to an overestimation of the yield). For the Diels-Alder reaction **10** shown on red background, the band of the product ($\lambda_{\text{max}} < 286 \text{ nm}$)

is obscured by the toluene solvent ($\lambda_{\text{max}} = 286 \text{ nm}$). Other reactions unsuitable for the method are, for example, S_N2 reaction of alkyl bromide **11** (not absorbing above 220 nm) or condensation of dichlorodiaminobenzene with phenanthrenequinone **12** (poor solubility of the product). **b**, $R^2 = 0.96$ correlation between the optically determined yields versus the yields of the same reactions determined by conventional means (HPLC for E1 from Fig. 3 and Ugi-type MCR from Fig. 4; ^1H NMR for Claisen-Schmidt condensation **1** and imine condensation **6**; the weights of isolated compounds for Cu(I)-catalysed alkyne-azide cycloaddition **4** and Beckmann rearrangement **8**). For numerical values of each data point, see Supplementary Tables 5–7 in Supplementary Information Sections 4.3 and 4.4.



Extended Data Fig. 2 | Concentrations versus yields near domain boundaries and the need to sample domain interiors. **a**, Taking as an example a basic bimolecular reaction $A + B \rightarrow C$ (with kinetic rate $k[A][B]$, reaction time $t = 1$ and kinetic constant $k = 1$), the plot has the yield and concentration of C for various combinations of A and B substrates. **b**, The plot shows the yield surface explicitly (black curve is the boundary at which the limiting reactant switches from A to B). As seen, yield—defined with respect to the concentration of the limiting reagent—tends to maximize at the boundaries of the region (orange curve in **a**) but the concentration of the product tends to maximize in its interior (blue curve in **a**). There are two implications of this ‘inverse’ dependence: (1) yields at very low concentrations may be artificially high, whereas—in reality—they involve very high (wasteful) excesses of some reagents; (2) detecting and quantifying the product reliably may be compromised in such a low-concentration regime. Accordingly, in the current work—especially for the multicomponent mixtures—we have worked over concentration ranges that do not extend to very small values, thereby assuring reliable detection while

avoiding highly wasteful excess ratios. Such a ‘practical’ range is shown schematically in **a** by the dashed vertical lines. Another corollary is that discovery of unexpected products may be severely compromised when using a strategy of investigating only the edges and boundaries of an n -dimensional region while leaving the interior unexplored—a strategy that may otherwise seem intuitive to a chemist and has some merit in maximization of the yield of known products. **c,d**, Comparison of final concentration map (**c**) to yield map (**d**) for product **19e** in Hantzsch reaction. Note the difference between the conditions that maximize concentration versus ones that maximize yield: yield is maximized at the vertices of the cube, whereas concentration is maximized in the middle of the front face and top edge, with higher concentrations protruding deeper into the interior of the cube (see blue curve in **a**). The isosurfaces correspond to 5 mM, 10 mM, 15 mM, 20 mM, 25 mM and 30 mM concentrations of **19e** in **c**, 20%, 30%, 40%, 50% and 60% yields of **19e** in **d** and are calculated as explained in the caption of Fig. 2c,d.