
Supplementary information

Robot-assisted mapping of chemical reaction hyperspaces and networks

In the format provided by the
authors and unedited

Supplementary Information for manuscript entitled “*Robot-assisted mapping of chemical reaction hyperspaces and networks*”.

Table of Contents

1. Materials	1
2. Robotic platform	2
2.1. Platform Overview	2
2.2. Well plates	5
2.3. XY gantry	8
2.4. Zeus pipetting module	9
2.5. UV-Vis spectroscopy	13
2.6. Containers and 3D-printed parts	16
2.7. Vial washer	16
2.8. List of materials for the robotic system	18
2.9. Reaction mixing	19
2.10. Influence of solvent evaporation	20
2.11. Uncertainty analysis	21
3. Data processing	22
3.1. Spectral unmixing algorithm	22
3.2. PCA-based model of the spectral baseline	23
3.3. Experimental map of spectrophotometer’s heteroscedastic uncertainties	24
3.4. Stoichiometric inequalities	26
3.5. Zero-dose extrapolation to compensate for photobleaching	27
3.6. Synthetic dynamic range from measurements at two dilutions	28
3.7. Quality control of the spectral data	29
3.8. Construction of reference spectrum of a given compound	31
3.9. Analysis of residuals with autocorrelation statistics	33
3.10. Tables of calibration and spectral unmixing options used in each experiment	34
3.11. Implementation	37
3.12. Examples of unmixed spectra for specific reactions from Extended Data Fig. 1	37
3.13. Outlier filtering and smoothing of the Ugi reaction yield maps	42
4. Experimental procedures and results	43
4.1. Compounds synthesized manually	43
4.2. Compounds synthesized using the automated platform	50
4.3. Synthetic procedures for versatility test (Extended Data Fig. 1)	71
4.4. Comparison of yields	82
4.5. Stirred and unstirred reactions yield comparison	86

4.6.	¹ H NMR benchmarks for Hantzsch reaction	87
4.7.	HPLC data for Hantzsch reaction maps.....	89
4.8.	Mechanistic study of the Ugi-type Multicomponent Reaction.....	98
4.9.	Iterative discovery in Hantzsch reaction network	115
4.10.	Hantzsch reaction network studies	118
4.11.	Computer-driven analysis of Hantzsch mechanistic networks	126
4.12.	Condition hyperspace boundaries between distinct outcomes of Hantzsch reaction	133
4.13.	Anomaly identification of S _N 1 reaction by mass spectrometry	137
4.14.	Control measurements of the water content in E1 reaction mixtures	138
4.15.	Raw data of yield maps	140
4.16.	Crystallographic data.....	160
4.17.	Reaction mass efficiency for the Hantzsch reaction.....	164
5.	Kinetic models and fitting them to the data	165
5.1.	Case of S _N 1 reaction (main text Figure 2b)	165
5.2.	Case of E1 reaction (main text Figure 2a).....	168
5.3.	Case of Ugi reaction	181
6.	Quantum chemical calculations.....	207
6.1.	Quantum thermochemistry of E1 and S _N 1 reactions	207
6.2.	Theoretical light absorption spectra for identification of product 15e	208
7.	Theoretical perspective on the “slopes” of yield maps	214
7.1.	Overview	214
7.2.	Case of first-order-kinetics networks	218
7.3.	Case of higher-order-kinetics networks.....	220
7.4.	Case of constrained second-order-kinetics network.....	223
7.5.	Numerical exploration of reaction networks’ smoothness	225
7.6.	Ramifications of limited slopes for the experimental sampling strategies	233
8.	Experimental section for the PBA-catalyzed styrene epoxidation.	235
8.1.	PBA particle size distribution.....	235
8.2.	Grid generation and t-SNE for visualization	236
8.3.	Yield map of benzaldehyde	240
8.4.	Comparison of yield and selectivity	241
8.5.	NMR UV yield comparison.....	242
8.6.	Recycling of PBAs	243
8.7.	Generation of four radicals that react with styrene.....	244
8.8.	¹ H NMR data of isolated compounds	245
9.	Supplementary Videos.....	246
10.	References.....	247

11. Mass spectra	258
12. NMR spectra.....	272
13. HPLC chromatograms	339

1. Materials

All compounds and solvents used for synthesis were purchased from Sigma-Aldrich, TCI (Tokyo Chemical Industry) or Alfa Aesar. Deionized water was purified through a Milli-Q water filtration system.

For thin liquid chromatography (TLC), silica gel on TLC Al foil 60/F254 was used (Sigma-Aldrich No. 60778). For flash column chromatography, silica gel 60 Å 220–440 mesh was used (Sigma-Aldrich No. 60738). ^1H , ^{13}C and 2D NMR spectra were acquired on 400 MHz FT-NMR Bruker Avance III HD spectrometer and 600 MHz FT-NMR Bruker AVANCE NEO 600 with 5mm Cryo probe at 298 K. NMR spectra were calibrated to residual solvent peak (for ^1H : chloroform-*d* 7.26 ppm, DMSO-*d*₆ 2.50 ppm, dichloromethane-*d*₂ 5.32 ppm, toluene-*d*₈ 2.08 ppm (–CD₃ shift); for ^{13}C : chloroform-*d* 77.14 ppm, DMSO-*d*₆ 39.52 ppm, dichloromethane-*d*₂ 53.84 ppm). Mass spectra were acquired on JEOL AccuTOF 4G+ DART using ESI ionization, Xevo® G2-XS ToF StepWave™ ion optics with the XS Collision Cell using ESI ionization and Advion expression® Compact Mass Spectrometer with APCI ion source. Analytical HPLC: Shimadzu Prominence with SIL-10AP autosampler, CTO-40C oven, SPD-20AV detector equipped with column JAIGEL-ODS-AP-A, SP-120-10 (Reversed Phase Column, particle size: 10 μm, 6.0×250 mm) or JAIGEL-ODS-BP-A, SP-120-10 (Reversed Phase Column, 6.0×250 mm), 2 channels: acetonitrile and 10 mM solution of ammonium acetate in water. Recycling preparative HPLC: JAI LaboACE LC-7080 Plus equipped with a RID (RI-700 LA) and UV–Vis (UV–VIS4ch 800LA) detectors and preparative GPC columns JAIGEL-2HR Plus and JAIGEL-2.5HR Plus (size 20.0 mm ID × 600 mm each) working in pair, preparative Reverse Phase (RP) column JAIGEL-ODS-AP, SP-120-10 (size 20.0 mm ID × 500 mm) and preparative Normal Phase (NP) column JAIGEL-SIL, SP-043-10 (size 20.0 mm ID × 500 mm (size 20.0 mm ID × 500 mm)).

2. Robotic platform

2.1. Platform Overview

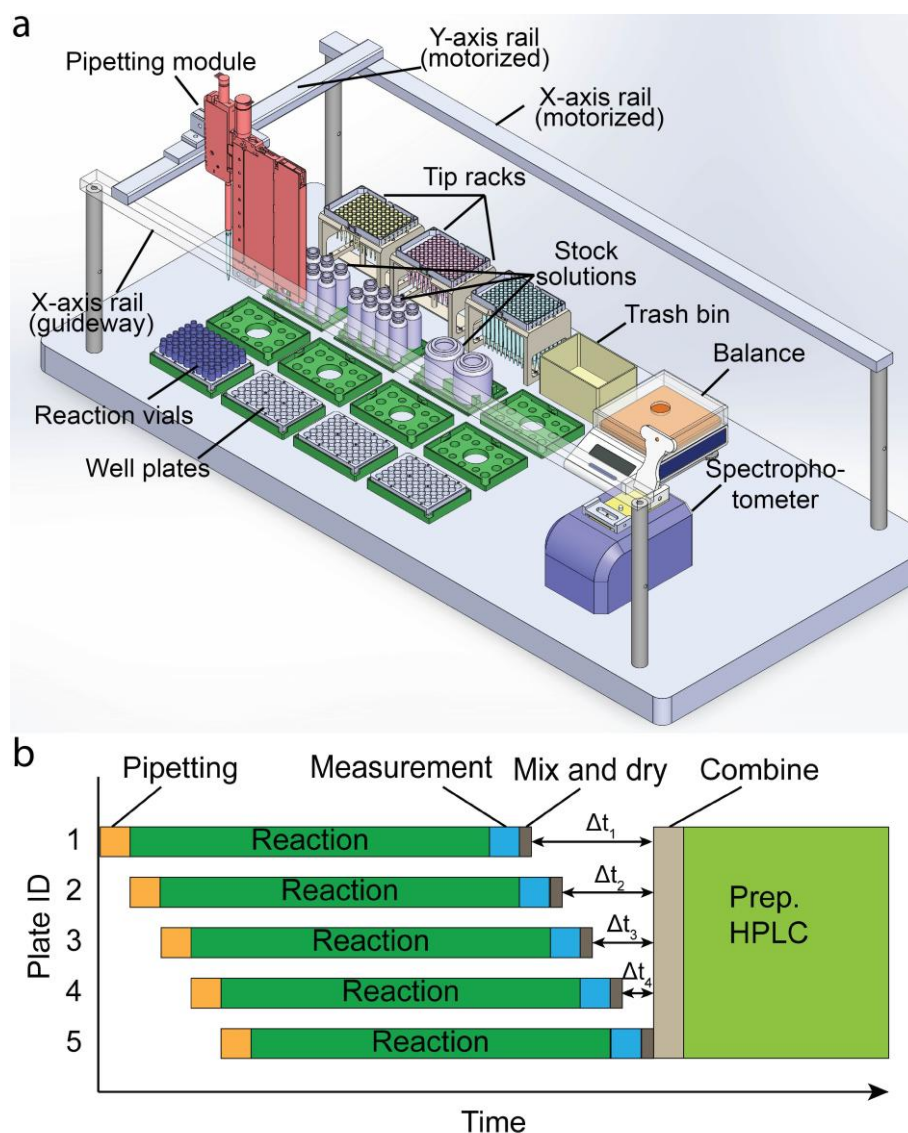


Figure S1. a. 3D model of the automated platform. To assemble the automation platform, an optical breadboard (600 mm $\times$ 1350 mm) is installed inside a fume hood for housing all parts (Figure S1). First, an XY gantry is installed onto the board, and the Hamilton pipetting channel is then mounted onto the gantry. Then, multiple 3D-printed holders are fixed on the board for accommodating the well plates, stock solution containers, and pipetting tips. After that, the Nanodrop spectrophotometer and an analytical balance used for calibration are installed. Also included on the breadboard is the electronic control box. The control box contains an Arduino controller for the XY gantry, a USB-CAN communication module for the pipetting channel, another Arduino controller for arm close/open and flushing of Nanodrop, and a USB-serial cable for the analytical balance. **b.** Gantt chart illustrating the workflow for an exemplary five-plate experiment, with each plate containing 54 reactions. For each plate, the process begins with pipetting reagents to each vial, followed by reaction progression (typically more than 12

h) with temperature control in an oven. Subsequently, each plate undergoes UV-Vis spectrophotometric measurement. After the measurement of one plate, the reaction crudes are mixed and dried using a rotary evaporator to quench the reactions. Once all plates are processed, the dried samples are re-dissolved and pooled for preparative HPLC purification. The intervals (Δt_1 – Δt_4) represent the time gaps between drying the crude reaction mixtures and the final combination of all samples. These intervals do not influence reaction progression before the samples are combined, since the reactions are effectively stopped after drying.

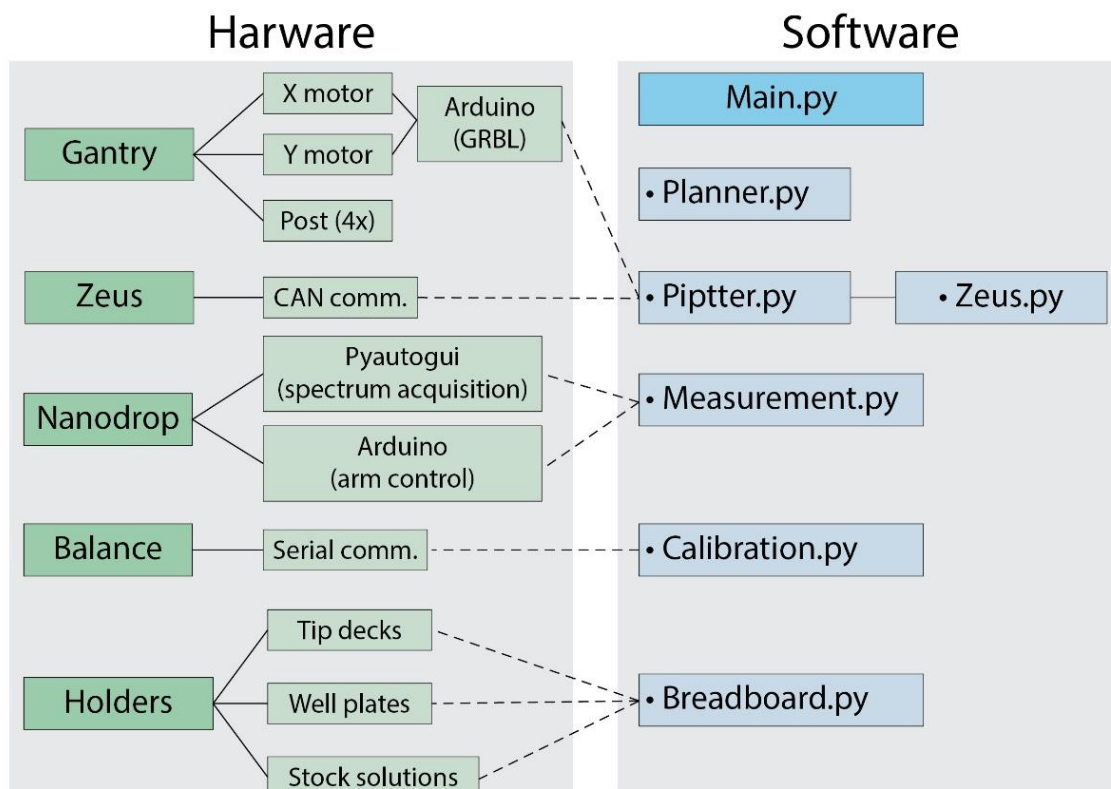


Figure S2. Hardware and software composition of the automated platform. Control of the XY gantry is managed through an Arduino board running the open-source GRBL library. The pipetting channel Zeus is communicated via CAN protocol using a commercial cable (Kvaser Leaf v3). The acquisition of spectral data from Nanodrop is automated through the PyAutoGUI package, which simulates mouse clicks within the Nanodrop software. The Nanodrop arm's opening and closing operations are governed by another Arduino board. The balance (Mettler Toledo XPE205) interfaces with the PC through a serial port connection. The 3D-printed holders are designed to secure the positioning of pipette tips, well plates, and stock solutions. The software records the status of the pipette tips and the volume of liquid in each container. These modules are detailed in the following section.

2.2. Well plates

The well plates housing the reaction vials are CNC-machined from aluminum with a customized design (**Figure S3** and **Figure S5**). The plate has 54 wells (6×9), each hosting a 2 mL vial (outer diameter 11.6 mm and height 32.0 mm). Using glass vials instead of conventional plastic 96-well plates allows for the application of harsh solvents in organic reactions, facilitating a broader range of reaction conditions. The well plate has the same footprint as a conventional 96-well plate (ANSI SLAS standards), ensuring compatibility with analytical equipment, such as plate readers.

Inspired by previous designs^{62,63}, we pair each well plate with a lid that is attached with a silicone sheet and a PFA film (**Figure S3**). Eight M4 screws are used to secure the lid tightly against the vials. The silicone sheet and the PFA film between the lid and the vials enable good sealing and allow for reactions at high temperatures (< 100 °C). Thermal-camera imaging confirms that the well plate maintains uniform solvent temperature, ensuring consistent reaction conditions across vials during heating (**Figure S3d**). During the automated experiments, while the primary operations (i.e. reaction preparation and measurements) proceed without human intervention, it is noted that the reaction plates are manually screwed and transferred between locations. For reproducible positioning of plates on the breadboard, we designed custom holders (green in **Figure S1**), which are 3D-printed from PETG plastic and screwed onto the breadboard. These holders are designed as so-called “compliant mechanisms”: they have spring-like elements that hold the inserted plate at a precise position. CAD files for these holders can be found in the Zenodo data archive associated with this paper – specifically, in the directory “mechanical_designs\holder_for_plates”.

To assess the sealing performance during heating, the fully assembled well plate, containing 1 mL ethanol in each vial, is placed in an oven at 80 °C for 48 hours. The temperature 80 °C is the highest used in this work. It was found that applying a torque of 4 N·m to each screw resulted in significant solvent evaporation in some vials, with one vial experiencing up to a 20% loss. When the torque was increased to 8 N·m, effective sealing was achieved, limiting the maximum loss to less than 1% (**Figure S4**). Accordingly, each plate is sealed during the reaction with a torque more than 8 N·m per screw to ensure secure containment.

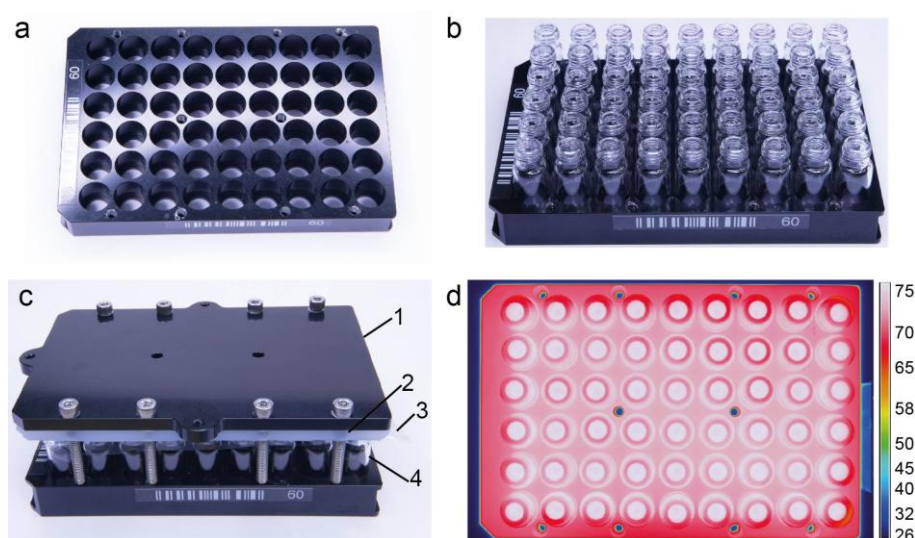


Figure S3. Customized 54-well plate with reaction vials: **a**, CNC-machined 54-well plate, **b**, the plate filled with 2 mL glass vials, **c**, assembly of the well plate with sealing for reactions: 1) plate lid, 2) silicone sheet, 3) PFA film, 4) reaction vials, **d**, Thermal camera image of the well plate immediately after removal from a 75 °C oven. Here, each vial contains 1.3 mL of.

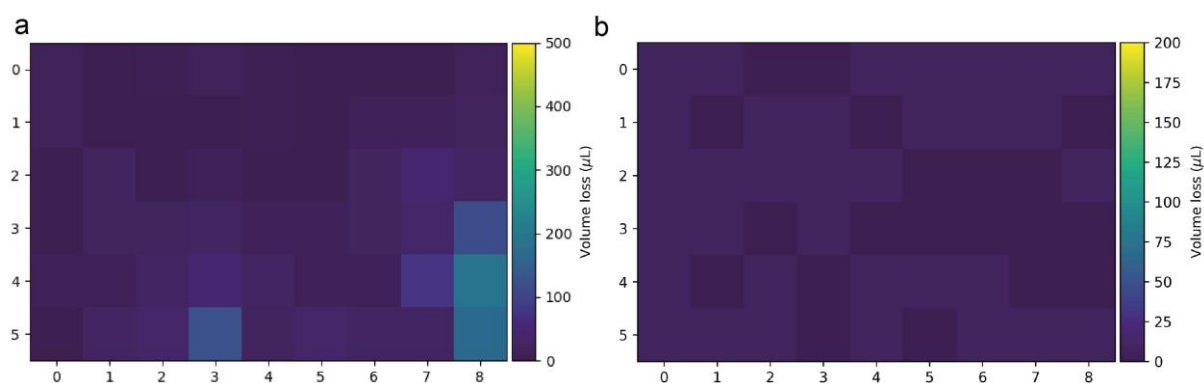


Figure S4. Validation of the well plates' sealing. The color map illustrates the solvent (ethanol) loss in a plate of 54 vials after being placed in an oven at 80 °C for 48 h, with the tightening torque on the plate set to 4 N·m (**a**) or 8 N·m (**b**). The original volume of the solvent in each vial is 1 mL.

2.3. XY gantry

The XY gantry is built with two motorized linear rails and a linear guide (parallel to the X axis), as shown in **Fig. 1a** and **Figure S1**. The gantry is lifted 300 mm above the optical breadboard by four support posts. The X- and Y-axis linear rails are 900 mm and 500 mm in length, respectively, providing a working area of 700 mm × 450 mm. A limit switch is installed on each linear rail to establish the initial (home) position of the gantry. The gantry when mounted with the pipetting module has a position accuracy of 0.1 mm.

The stepper motors on the X- and Y-axes are driven by TB6600 motor drivers, with their motion controlled by an Arduino board (**Figure S6**). The Arduino runs on an open-source motion control library: grbl (<https://github.com/gnea/grbl>). The control panel for automation includes an Arduino controller, two motor drivers, a PCB board for USB–FCC connection, a USB–CAN bus cable, 48–24 V DC converter and 48 V DC power supply (**Figure S6**).

For mounting the Zeus module on the gantry, we designed and machined a dedicated mount visible in the left of **Figure S7a**. CAD file (SolidWorks format) for this mount can be found in the Zenodo data archive associated with this paper – specifically, in the directory “mechanical_designs\Connector_for_ZEUS_YJV02”.

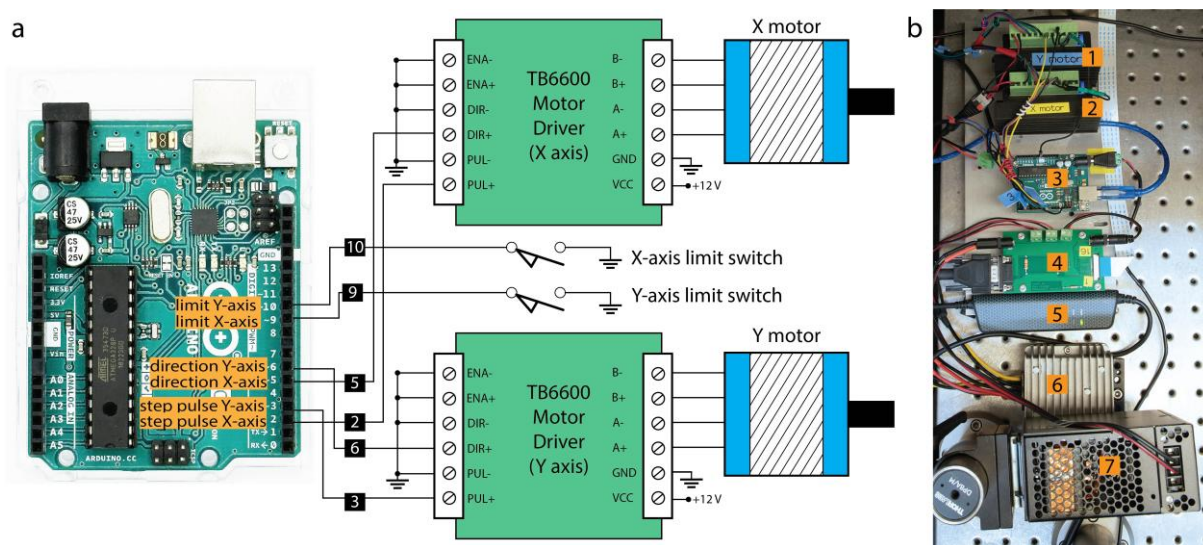


Figure S6. a, Scheme of the motion control of the XY gantry. **b**, Electronics for automation: 1. Y-axis motor driver, 2. X-axis motor driver, 3. Arduino control board, 4. PCB board for USB–FCC connection. 5. USB–CAN bus cable (Kvaser Leaf v3), 6. 48–24 V DC converter, 7. 48 V DC power supply.

2.4. Zeus pipetting module

The Zeus pipetting module, as shown in **Figure S7**, is a commercial, electronic air-displacement pipette designed primarily for biological or medical applications. It communicates with the controlling computer through CAN protocol (controller area network). Our software operates Zeus via CAN protocol using a python module named ‘zeus.py’ (See **Figure S2**), which is adapted from an open-source package available at <https://github.com/nicholasmorrow/zeus>.

With factory settings, the pipetting module works only for certain liquid types, namely, water, serum, ethanol and glycerin. For each liquid type, the pipetting module applies a set of pipetting parameters, such as aspiration and dispensing flow rates, over-aspirated volume, blowout air volume, and more (**Table S1**). The pipetting parameters are largely influenced by the physical properties of the liquids, such as density, surface tension, viscosity, boiling point, vapor pressure, and other related factors.

To expand the scope of pipetting, we began by replacing the original rubber sealing O-ring in the pipetting module with a more chemically inert Viton™ O-ring. We then replaced it monthly to remain reliable, which proved essential for accurate pipetting. Then, we carefully tested and established the pipetting parameters for the organic solvents utilized in this study, including dimethylformamide (DMF), 1,4-dioxane, 1,2-dichloroethane (DCE), and acetonitrile. The pipetting parameters for these solvents were adapted from the built-in liquid types. Specifically, DMF was treated with the same parameters as water, while dioxane, DCE, and acetonitrile were assigned the same parameters as ethanol. These parameter assignments ensure smooth liquid transfer, avoiding possible pipetting fails, including 1) liquid dripping from the tip during gantry movement, 2) residual liquid volume inside the tip after dispensing, and 3) liquid splashes during dispensing. It is important to note that, even with these precautions, we still could not accurately pipette certain highly volatile solvents, including dichloromethane, 1,2-dichloroethane and acetone.

Subsequently, to get good pipetting precision (**Figure S8**), the organic solvents were calibrated using an analytical balance (Mettler Toledo, XPE205; see Supplementary Video 1). Each organic solvent was calibrated for three types of tips: 50 μL , 300 μL , and 1000 μL . Each solvent underwent calibration across eight points per tip type, resulting in a total of 24 calibration points per solvent (**Figure S9**). Each calibration point was measured three times. The balance control and readings were managed by a customized program through serial communication (see **Figure S2**). By integrating balance control with automated pipetting, the entire calibration process was fully automated. Following calibration, pipetting errors were reduced to less than 2% for volumes under 50 μL and less than 1% for volumes greater than 50 μL for all solvents utilized in this study, as demonstrated in **Figure S8**.

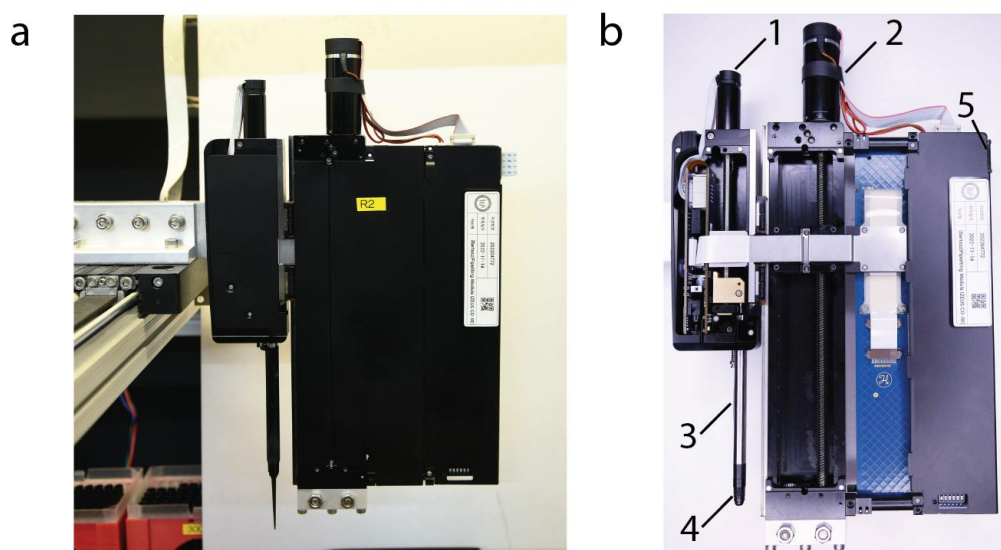


Figure S7. **a**, Zeus pipetting module. **b**, Internal components of the pipetting module. 1, pipetting drive. 2, Z-axis drive. 3, barrel with piston. 4, tip mounting adapter. 5, power and CAN connector. CAD file (SolidWorks format) the mount visible in the left of panel **a** can be found in the Zenodo data archive associated with this paper – specifically, in the directory “mechanical_designs\Connector_for_ZEUS_YJV02”.

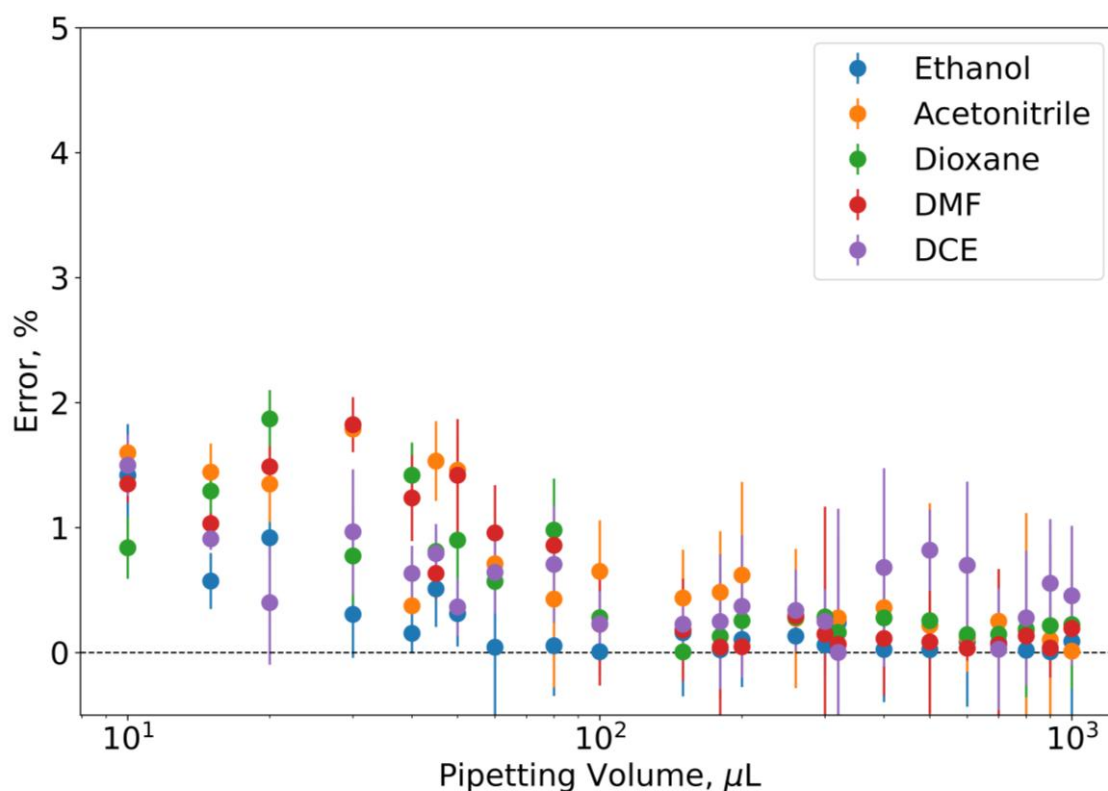


Figure S8. Pipetting precision for organic solvents: ethanol, acetonitrile, 1,4-dioxane, dimethylformamide (DMF) and 1,2-dichloroethane (DCE).

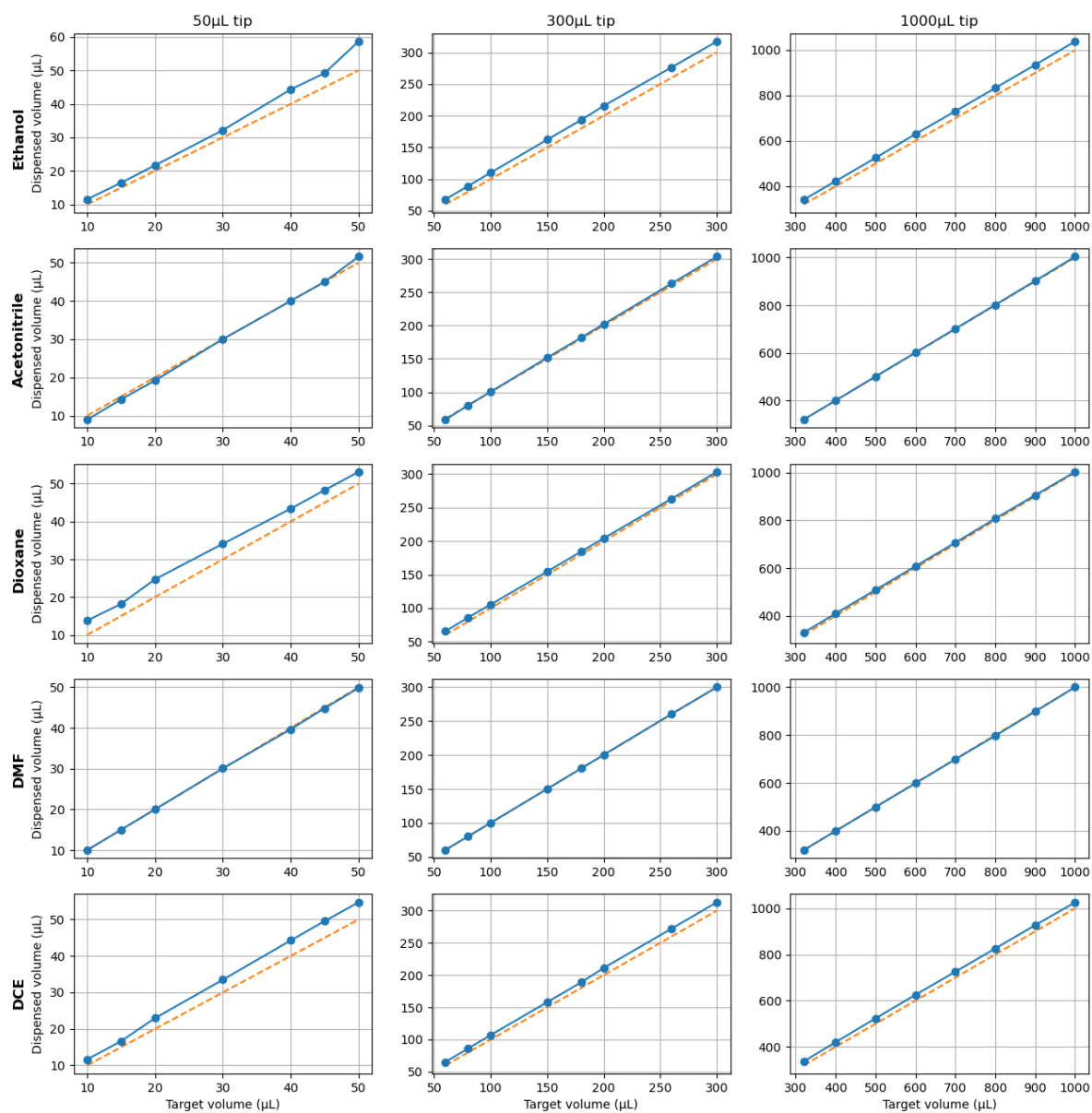


Figure S9. Calibration curves for pipetting organic solvents: ethanol, acetonitrile, 1,4-dioxane, dimethylformamide (DMF) and 1,2-dichloroethane (DCE).

Table S1. Pipetting parameters for water and ethanol with different tip type. These parameters are extracted from the Zeus pipetting module.

id	Parameters	Water (50 μ L)	Water (300 μ L)	Water (1000 μ L)	Ethanol (50 μ L)	Ethanol (300 μ L)	Ethanol (1000 μ L)
1	index	0	1	2	12	13	14
2	aspirationFlowRate	5000	5000	5000	2000	2000	2000
3	overAspiratedVolume	20	50	50	0	50	50
4	aspirationTransportVolume	30	50	80	30	50	50
5	blowoutAirVolume	300	250	400	500	250	400
6	aspirationSwapSpeed	200	200	200	200	200	200
7	aspirationSettlingTime	10	10	10	10	10	10
8	lld	0	0	0	1	1	1
9	clldSensitivity	3	3	3	3	3	3
10	plldSensitivity	3	3	3	3	3	3
11	adc	0	0	0	1	1	1
12	dispensingMode	0	0	0	0	0	0
13	dispensingFlowRate	5000	5000	5000	3000	3000	3000
14	stopFlowRate	0	0	0	2000	2000	2000
15	stopBackVolume	0	0	0	0	0	0
16	dispensingTransportVolume	30	50	50	30	50	50
17	acceleration	40	40	80	40	40	40
18	dispensingSwapSpeed	200	200	200	200	200	200
19	dispensingSettlingTime	10	10	10	10	10	10
20	flowRateTransportVolume	325	325	325	325	325	325

The description of the parameters, excerpted from the Zeus user manual, is provided below.

1, liquid class index. 2, Flow Rate in 0.1 μ L/s. 3, Over-Aspirated volume in 0.1 μ L. 4, Transport Volume in 0.1 μ L. 5, Blowout Air Volume in 0.1 μ L. 6, Swap Speed in 0.1 mm/s. 7, Settling Time in 0.1 s. 8, Liquid Level Detection (LLD) Mode (0 = cLLD, 1 = pLLD). 9, cLLD Sensitivity (1 = very high, 2 = high, 3 = medium, 4 = low). 10, pLLD Sensitivity (1 = very high, 2 = high, 3 = medium, 4 = low). 11, Anti-Droplet Control (ADC; 0 = off, 1 = on). 12, Dispensing Mode: 0 = jet empty, 1 = jet part, 2 = surface empty, 3 = surface part. 13, Flow Rate in 0.1 μ L/s. Dispensing speed of the plunger. 14, Stop Flow Rate in 0.1 μ L/s. 15, Stop Back Volume in 0.1 μ L. 16, Transport Volume in 0.1 μ L. 17, Acceleration in 0.1 μ L/s². 18, Swap Speed in 0.1 mm/s. 19, Settling Time in 0.1 s. 20, Flow Rate Transport Volume in 0.1 μ L/s.

2.5. UV-Vis spectroscopy

Light absorption spectra were measured either by CRAIC (micro-)spectrophotometer, or by NanoDrop spectrometer.

CRAIC spectrophotometer.

CRAIC spectrophotometer did not require additional hardware automation: the plate of 54 vials was placed on the automated table of microspectrometer, and vials were viewed from above through Olympus PlanApo N 1.25x/0.004 microscope objective. On the software side, we used PyAutoGUI package to interface with CRAIC software: the respective Python script can be found in the directory “robowski/uv_vis_absorption_spectroscopy/craic_automation” of the code repository associated with this article.

NanoDrop spectrophotometer

During the acquisition of the UV–Vis spectrum, a liquid bridge of the sample forms between the lower pedestal and the upper arm of the Nanodrop (**Figure S10a**). The measurement path length is fixed to 1 mm, controlled by a position-limiting post located beneath the arm. Typically, the measurement requires a sample volume of 2 μL . The pedestal’s diameter can be expanded from 3 mm to 6 mm by attaching a custom metal ring around it (**Figure S10b**), for hosting a large sample volume (up to 20 μL) on the pedestal. The mechanical design for this metal ring can be found in the CAD file (SolidWorks format) in the [Zenodo data archive](#) associated with this paper – specifically, in the directory “mechanical_designs\Nanodrop\ring_for_Nanodrop_pedestal”.

The Nanodrop spectrophotometer, originally designed for manual UV–Vis spectral analysis, operates across a wavelength range of 220 nm to 750 nm. It requires approximately 2 μL of liquid per measurement, dispensed onto a pedestal, with the measurement path length secured at 1 mm by its arm. The diameter of the pedestal has been increased from 3 mm to 6 mm by affixing a metal ring around it. This modification enables the accommodation of a larger volume (up to 10 μL) atop the pedestal and promotes the formation of a stable liquid bridge between the pedestal and the upper part. This is particularly crucial for organic solvents, which typically exhibit surface tensions much lower than that of water. The spectrophotometer arm is equipped with a servo motor for automated close/open operations. Furthermore, an aluminum chamber is mounted around the pedestal of the Nanodrop designed to contain the flow of fluids during automated flushing and drying operations .

The measurement process is automated through the addition of two functions to the spectrophotometer (**Figure S10c**). First, the Nanodrop arm’s automatic opening and closing are facilitated by a servo motor affixed to the arm’s hinge, which is controlled by an Arduino board. Second, the pedestal is automatically cleaned after each measurement via an aluminum chamber encasing it. The aluminum chamber features an inlet channel and an outlet. The outlet links to the hood’s vacuum line, while the inlet can be connected to either ethanol, for flushing, or air,

for drying purposes. Switching between flushing and drying functions is managed through solenoid valves, which are operated by the same Arduino board.

The quality of cleaning of the Nanodrop after each measurement is tested (**Figure S11**). It is shown that after cleaning, there is no sign of cross contamination. The reproducibility of the measurement is tested by acquiring multiple consecutive spectra (e.g., 54 spectra of one solution, Hantzsch ester in ethanol at a concentration of 0.1 mg/mL as shown in **Figure S12**). The standard deviation of absorbance at 280, 340 and 400 nm are 5.45%, 5.14% and 2.92%, respectively.

Downstream processing of the spectral data files produced by NanoDrop spectrophotometer has the following nuance. The output spectrum of NanoDrop spectrophotometer may contain wavelengths from 190 nm to 850 nm, but sometimes it does not: the spectral range cannot be set in software and instead varies from sample to sample. Sometimes the range is 220 nm to 750 nm, for example. Through experimentation we have established that the guaranteed wavelength range covers at least from 220 nm to 600 nm. In our data processing code, for ensuring uniform workflow, the wavelengths below 220 nm and above 600 nm are removed upon loading the data from NanoDrop files. This functionality is implemented in `“load_nanodrop_csv_for_one_plate()”` method of `“robowski/uv_vis_absorption_spectroscopy/process_wellplate_spectra.py”` module in the code repository associated with this article. The left and right limits are defined by class variables `“nanodrop_lower_cutoff_of_wavelengths”` and `“nanodrop_upper_cutoff_of_wavelengths”` of `“SpectraProcessor”` class defined in the same module.

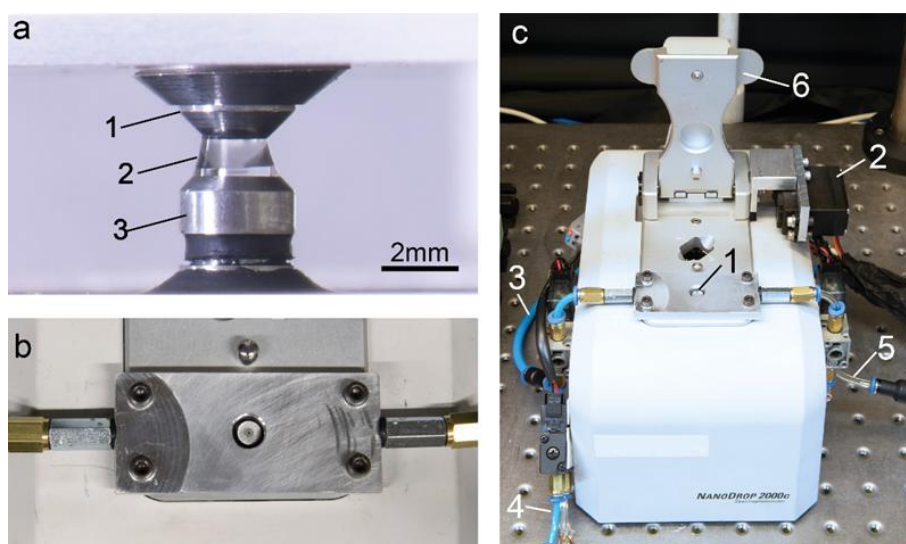


Figure S10. Photographs of Nanodrop adapted for automated spectrum acquisition. **a**, The liquid bridge (2) connecting the pedestal (3) and the upper probe (1) during measurement. **b**, The pedestal mounted with a metal ring and surrounded by a chamber for washing. **c**, The Nanodrop featuring automation cleaning and automative arm opening/closing: 1, pedestal with metal ring; 2, flushing chamber; 3, Nanodrop's arm; 4, servo motor; 5, 6, 7, solenoid valves controlling the vacuum tubing (5), flushing solvent (6) and air (7).

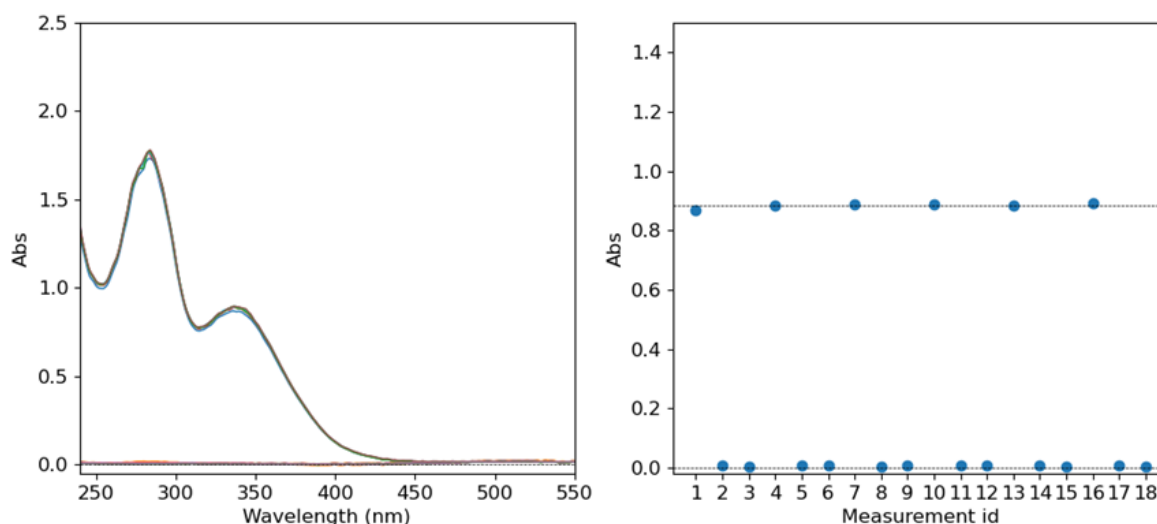


Figure S11. Validation of Nanodrop automatic cleaning. Eighteen samples were measured in the test where 6 vials (vial labels: 1, 4, 7, 10, 13, 16) were filled with a reagent with the same concentration, and the remaining vials were filled with pure ethanol. **a**, Overlapped spectra of the samples in 18 vials. **b**, Absorption of the 18 samples at wavelength of 340 nm.

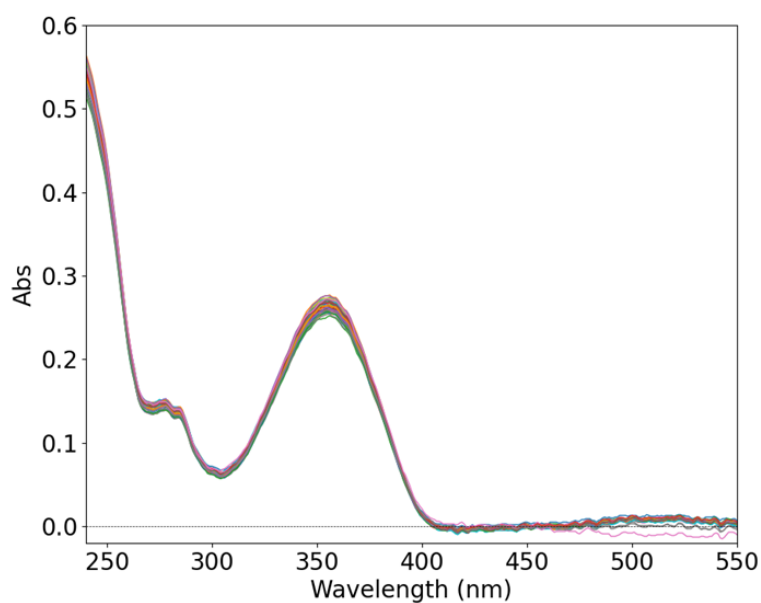


Figure S12 Reproducibility test. UV-Vis absorption spectrum of 0.1 mg/mL Hantzsch ester solution in ethanol is measured 54 times.

2.6. Containers and 3D-printed parts



Figure S13. Arrangement of items on the breadboard. 1, 2, 3, 4: slots for well plates. 4, 5: slots for stock solution in 20 mL bottles. 7, 8: slots for stock solution in 100 mL jars. 9, tip trash can. 10, 11, 12: tip racks for tips of 1000 μ L, 300 μ L and 50 μ L, respectively. All holders or racks are 3D printed, and the well plates are CNC machined. The CAD files for machined plates and 3D printed parts can be found the [Zenodo data archive](#) associated with this paper – specifically, in the directories “mechanical_designs\well_plates” and “mechanical_designs\holder_for_plates”.

2.7. Vial washer

In most of the experiments, the 2 mL reaction vials were reused. The reuse process involved a thorough washing step to remove any residual materials from previous reactions, followed by drying in an oven to ensure the vials were clean and free of moisture. The washing is done by a custom-designed vial washer (**Figure S14**). The washer washes the vial with two long needles, one for withdrawing liquid from the vial and another for dispensing clean solvents (mostly acetone). The liquid-withdrawing needle is connected to vacuum via a waste collecting Buchner flask, while the solvent-dispensing needle is supplied with solvent from a Buchner flask by an air pressure pump. The needles are mounted onto an XY drawing machine, replacing the position typically occupied by a drawing pen. The up-and-down movement of the needles along the Z-axis is controlled by a servo motor on the drawing machine. Thus, the washing process is achieved through the coordination of three operations: 1) XY movement of needles, 2) up-and-down movement of needles, and 3) solvent supply control. Experiments have shown that after

three washing cycles in each vial, the residuals from the previous reaction are reduced to undetectable levels.

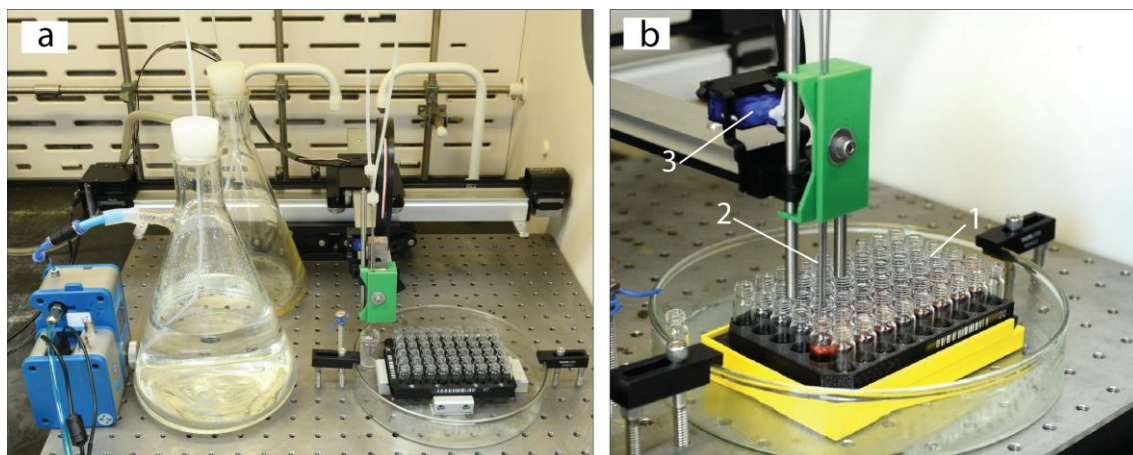


Figure S14. a, Vial washer. **b,** Zoom-in view of the washing section. 1, reaction vials. 2, needles: left needle for waste aspiration, right needle for clean solvent. 3, servo motor mounted on a drawing machine (AxiDraw). The green mount for needles visible in **b** can be 3D-printed using the CAD files (SolidWorks format, as well as .PLY file suitable for 3D printers) available from the Zenodo data archive associated with this paper – specifically, the directory “mechanical_designs\axidraw_z_caret”.

2.8. List of materials for the robotic system

Table S2 List of materials for the robotic system

Part#	Part Name	Supplier	Description	Qty
1	X-axis rail	FUYU	FPB50, 900 mm	1
2	Y-axis rail	FUYU	FPB50, 500 mm	1
3	Linear guide	HIWIN	HG15	1
4	Pipetting channel	Hamilton	Zeus	1
5	Spectrophotometer	Thermo Fisher	Nanodrop 2000c	1
6	USB-CAN cable	Kvaser	Leaf Light V2	1
7	Power supply (48V)	Delta Electronics	48V 480W	1
8	DC converter		48 to 24V	1
9	Arduino Uno	Arduino	/	2
10	Support post	Thorlabs	P50/M	4
11	End Switch	FC Sensor	F3N-18TN05-N	2
12	Analytical Balance	Mettler Toledo	XPE205	1
13	3D printed parts	Customized design	/	14
14	Solenoid valve	SMC	VZ110	3
15	Stepper motor	TowerPro	MG996R	1
16	XY plotter	AxiDraw	V3	1

2.9. Reaction mixing

The mixing of the reaction solution is achieved through jet flow during liquid dispensing and, more importantly, through orbital shaking. As shown in **Figure S15a**, the dispensing of 300 μL ethanol (colored with methyl orange) into 1000 μL ethanol demonstrates effective mixing, evidenced by the uniform dye distribution 60 seconds after injection. However, when injecting 80 μL ethanol into 1000 μL ethanol, a visibly uneven dye distribution remains at 60 seconds after injection (**Figure S15b**), indicating that injection flow alone is insufficient for thorough mixing in certain cases. Therefore, to ensure complete homogenization, all reactions undergo orbital shaking at 250 rpm for 5 minutes following pipetting. As detailed in **Section 4.4**, a comparison of reaction yields between this mixing method and stirring with a magnetic bar revealed a negligible difference in yield.

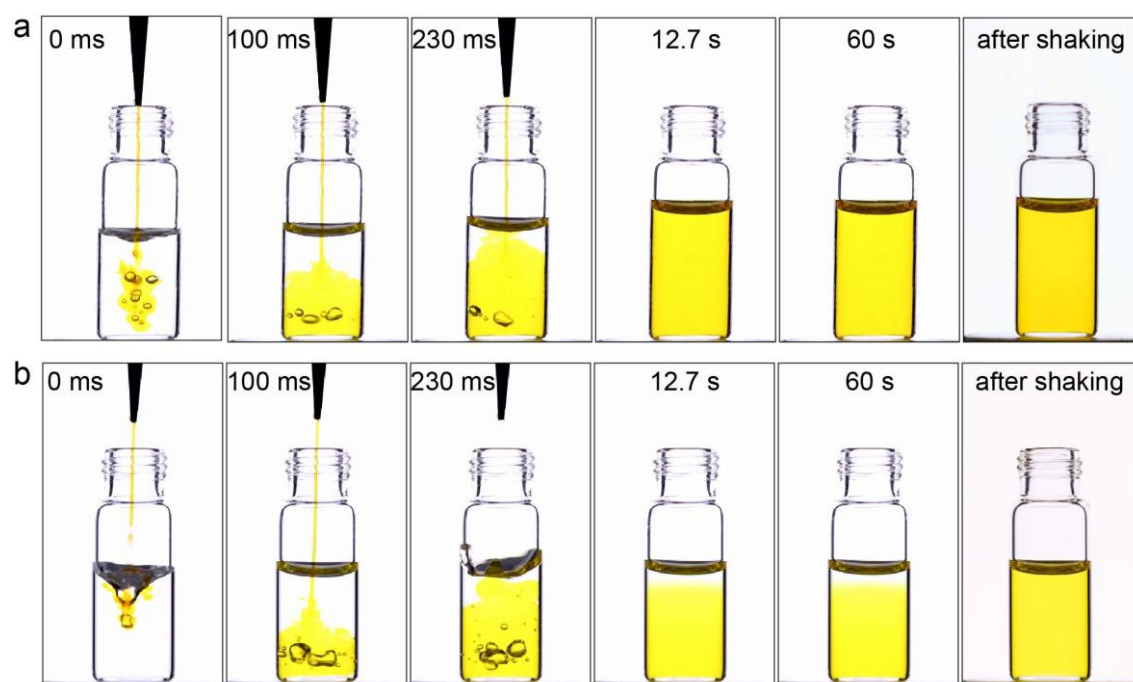


Figure S15. Mixing during solvent dispensing. **a**, Injecting 300 μL of ethanol dyed with methyl orange into 1000 μL ethanol, **b**, Injecting 80 μL of ethanol dyed with methyl orange into 1000 μL ethanol.

2.10. Influence of solvent evaporation

Solvent loss in 2-mL vials due to evaporation was measured over a 45-min duration, representing the typical time frame for pipetting one plate of reactions. The loss is measured by an analytical balance. The temperature is 23.5 °C and the relative humidity is 56.7%. Note that dimethylformamide has a negative loss due to its high hygroscopicity.

Table S3 Evaporative loss of solvents in 2-mL vials over 45 min.

Solvents	Volume (μL)	Loss (%)
Ethanol	500	1.59
Acetonitrile	500	2.68
Dioxane	500	1.11
Dimethylformamide	500	-0.11
1,2-dichloroethane	500	2.53

2.11. Uncertainty analysis

In this section, the Hantzsch reaction is used as an example to evaluate and analyze experimental uncertainties. The yield evaluation process begins with the preparation of stock solutions. The powdered substrates are weighed using an analytical balance (Mettler Toledo XPE205) and dissolved in a solvent measured with volumetric flasks. The resulting solutions are then sampled with a Hamilton syringe and further diluted. The stock concentration is calculated using the formula $c_m = \frac{m}{v_1} * \frac{v_2}{v_3}$, where m is the weighted mass, v_1 the solvent volume for dissolving, v_2 the solution volume for dilution and v_3 the final net volume. As an example, $m = 200.00 \pm 0.60\text{mg}$, $v_1 = 20.00 \pm 0.020\text{ mL}$, $v_2 = 1.00 \pm 0.010\text{ mL}$ and $v_3 = 100.00 \pm 0.10\text{ mL}$. Error propagation is performed using the linear approximation implemented in Python Uncertainties module, yielding $c_m = 0.10 \pm 0.0011(1.1\%) \text{ mg/mL}$. This relative uncertainty of 1.1% is applied to the stock concentrations of the three substrates: methoxybenzaldehyde, ethyl acetoacetate, and ammonium acetate. In this reaction, the molar concentrations for the three substrates are $c_{s1} = 0.333 \pm 0.006$, $c_{s2} = 0.333 \pm 0.006$ and $c_{s3} = 1.333 \pm 0.026$.

Subsequently, the substrates and the solvent ethanol are pipetted into reaction vials in specified volumes. The electronic pipette employs three types of tips—50 μL , 300 μL , and 1000 μL —with relative uncertainties of 2%, 1%, and 0.5%, respectively. For example, one reaction has the volumes $v_a = 162.8 \pm 1.6\text{ }\mu\text{L}$, $v_b = 18.1 \pm 0.4\text{ }\mu\text{L}$, $v_c = 162.8 \pm 1.6\text{ }\mu\text{L}$ and $v_d = 198.9 \pm 2.0\text{ }\mu\text{L}$ for methoxybenzaldehyde, ethyl acetoacetate, ammonium acetate, and ethanol, respectively. The total reaction volume is $v_t = 542.6 \pm 3.1\text{ }\mu\text{L}$. Therefore, in the reaction the substrate concentration for methoxybenzaldehyde is $c_1 = c_{s1} * v_a / v_t = 0.100 \pm 0.001\text{ M}$, $c_2 = 0.011 \pm 0.003\text{ M}$, and $c_3 = 0.400 \pm 0.006\text{ M}$. Before running HPLC analysis, the reaction crudes are diluted with ethanol using the same pipetting module. In this example, the diluting factor is $d = 19.1 \pm 0.2$. The limiting concentration in the reaction crude for Hantzsch ester (HE) is $c_{l1} = \min(c_1, c_2/2, c_3)/d = (291 \pm 7) \times 10^{-6}\text{ M}$ and for tetrahydropyridine (THP) $c_{l2} = \min(c_1/2, c_2, c_3/2)/d = (582 \pm 14) \times 10^{-6}\text{ M}$.

The known concentrations of the two products, HE and THP, are measured using HPLC to construct the calibration curves. Validation of the HPLC method confirms that this yield evaluation approach results in uncertainties of 5.1% for HE and 3.8% for THP, respectively. In this reaction example, the HPLC-measured concentrations of the HE and THP in the reaction crude are $c_{h1} = (8.4 \pm 0.4) \times 10^{-6}\text{ M}$, $c_{h2} = (3.4 \pm 0.1) \times 10^{-4}\text{ M}$. Therefore, the yields are $y_1 = c_{h1}/c_{l1} * 100 = 2.9 \pm 0.2\%$, $y_2 = c_{h2}/c_{l2} * 100 = 58.6 \pm 2.7\%$. Likewise, this method for error propagation is used for all the experimental data points.

This analysis was performed using Python script “robowski/misc_scripts/error_analysis.py” in the code repository associated with this article. This code can be run it to reproduce the calculations above.

3. Data processing

3.1. Spectral unmixing algorithm

Assuming the Beer-Lambert law is applicable, experimental light absorption spectrum $A(\lambda)$ of a crude mixture having N components can be modelled by a weighted linear sum of the reference spectra $A_{0,i}(\lambda)$ of individual components measured separately at reference concentrations $c_{0,i}$. In the simplest form, the model can be written as

$$A(\lambda) = \sum_i \frac{c_i}{c_{0,i}} A_{0,i}(\lambda) + \epsilon \quad (1)$$

where the model's free parameters set $\theta = \{c_i\}$ comprises the concentrations c_i of the components. The set of values c_i is then optimized so that the model follows the experimental spectrum as closely as possible. The residuals are expressed by ϵ , which is a normally distributed random variable reflecting the experimental uncertainties. This method of evaluating concentrations from absorption spectrum is known as “multivariate calibration”⁶⁴ or “spectral unmixing”^{22,65,66}, is widely used in hyperspectral imaging⁶⁷, and is superior to the more common method of evaluating concentrations from absorbance at a single wavelength (typically, the maximum of a certain absorption band)^{22,64,68–70}. For a direct comparison of spectral unmixing to single-wavelength method (sometimes called “univariate calibration”) for many substances, see the review by Brereton⁶⁴, who reports, for instance, that the error of pyrene concentration has reduced from 24.1% in the single-wavelength method to 7.88% with spectral unmixing. Similarly, Langergraber et al.⁶⁸ report a roughly two-fold reduction of uncertainty upon switching from a single-wavelength method to the spectral unmixing. Lastly, the ability of spectral unmixing to determine the abundances of spectral components precisely by capitalizing on data at multiple wavelengths instead of an integral signal from one spectral band – is the foundation of the entire field of hyperspectral/multispectral imaging⁶⁷, with wide applications from routine agricultural land surveys to military surveillance.

One advantage is that spectral unmixing works even when the spectra of individual components overlap. Second, from the information-theoretical standpoint, spectral unmixing utilizes all the available experimental information (i.e., absorbance values at all accessed wavelengths), which ultimately results in higher precision of the evaluated concentrations. Third, the unexpected appearance of components other than the expected ones in the absorption spectra is readily detected with the spectral unmixing method – it will be evident from the large discrepancy between the experimental spectrum and the best-fit model, because the model will be missing the unexpected absorption bands. In contrast, the appearance of unexpected components would have remained unnoticed by a more conventional single-wavelength method of evaluating concentrations.

To accurately model the biases and errors of a specific spectrophotometer, we expand the model (1) in four ways: (i) by modelling the baseline spectrum (of pure solvent) by two principal components $B_1(\lambda)$ and $B_2(\lambda)$, (ii) by allowing for a small (± 1.5 nm) spectral shift $\Delta\lambda$, (iii) by

using the separately measured dependencies of instrumental errors $\epsilon(\lambda, A(\lambda))$ on wavelength and absorbance, and (iv) by allowing a constant offset δ of spectra to compensate for geometric optics contributions to reproducibility:

$$A(\lambda) = \sum_i \frac{c_i}{c_{0,i}} A_{0,i}(\lambda + \Delta\lambda) + b_1 B_1(\lambda + \Delta\lambda) + b_2 B_2(\lambda + \Delta\lambda) + \delta + \epsilon(\lambda, A(\lambda)) \quad (2)$$

The set $\theta = \{c_i, b_1, b_2, \Delta\lambda, \delta\}$ of model's parameters now includes $\Delta\lambda$, δ , and the weights b_1 and b_2 of the baseline components. Baseline components are discussed in the next Section 3.2.

For a worked example of the spectral unmixing workflow using the real spectral data to reproduce the main-text Figures 1c,d, see the Jupyter Notebook located at “notebooks/spectral_unmixing_tutorial.ipynb” in the code repository associated with this article. In our high-throughput analysis code, spectral unmixing is implemented in the functions *spectrum_to_concentration()* and *multispectrum_to_concentration()* of the *process_wellplate_spectra.py* module located in “robowski/uv_vis_absorption_spectroscopy/” directory of the code repository associated with this article. The implemented definitions of spectral unmixing models can be found in the same module, in the functions *model_function()*, *preliminary_model_without_stoichiometric_inequalities()*, and *model_with_stoichiometric_inequalities()*.

For a summary of model options used for analysis of crude mixtures from specific reactions, see **Table S4**.

3.2. PCA-based model of the spectral baseline

Spectrum of a pure solvent measured by a given spectrophotometer slightly varies from measurement to measurement due to various imperfections of the optics and variations of the vial's geometry. This variation is not purely random noise, but instead has some internal structure that depends on the spectrophotometer used and, to a lesser extent, on the solvent for which the baseline is taken. We model these variations by taking 54 spectra of solvent and performing a principal component analysis (PCA) of the resulting spectra. PCA essentially synthesizes two new spectra (called “principal components”) such that all the 54 spectra can be approximated as weighted sums of the two principal components and a constant. We found that about 90% of the baseline variation can be described by using two principal components only. We further smooth the principal components by Savitsky-Golay filter before using them in the spectral unmixing model. Example of the baseline spectra and the extracted principal components is shown in **Figure S16** for the case of DMF as solvent and CRAIC microspectrometer.

Software implementation of this PCA analysis and use of the extracted components for subsequent spectral unmixing is demonstrated by a worked example in the Jupyter Notebook located at “notebooks/spectral_unmixing_tutorial.ipynb” in the associated code repository.

Scripts used for PCA analysis for different solvents can be found in the directory “robowski/uv_vis_absorption_spectroscopy/examples/background_pca” of the same repository. For a summary of which reactions studied in present work were analyzed with the help of this approach, see **Table S4**.

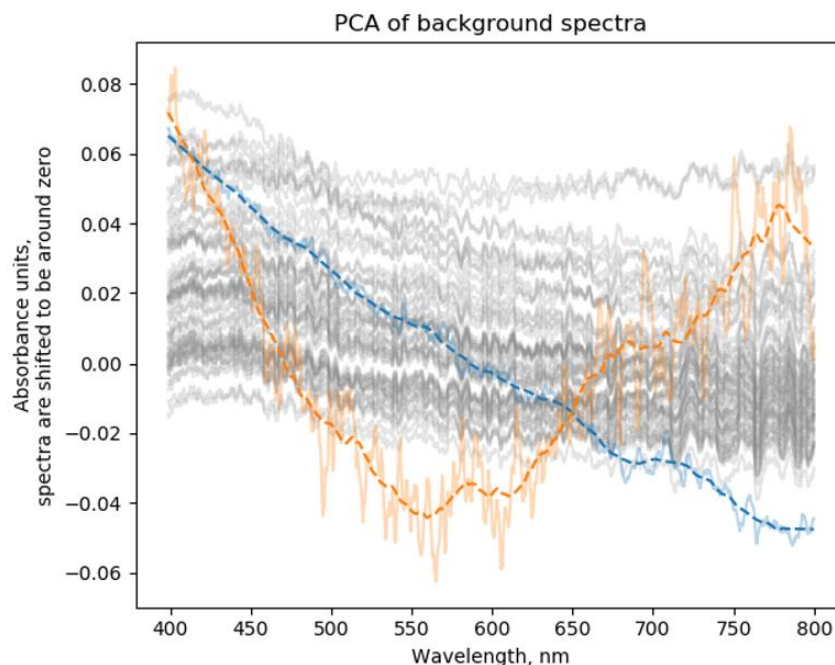


Figure S16. Multiple spectra of pure DMF (grey) taken with CRAIC spectrophotometer are subjected to principal component analysis to obtain the first principal component (*solid light blue*) and second principal component (*solid light orange*). Due to the basic nature of principal component analysis, extracted components are shown here in arbitrary vertical scale. They are then smoothed by linear (order = 1) Savitsky-Golay filter with window length 71 nm (smoothed curves are *dashed blue and dashed orange*).

3.3. Experimental map of spectrophotometer’s heteroscedastic uncertainties

The uncertainty of absorbance measured by the spectrophotometer depends on the wavelength and on the absorbance value itself: the errors are higher at the edges of the spectrophotometer's nominal spectral range, and at the upper edge of its dynamic range of absorbance. Higher absorbance means fewer photons reaching the detector, and therefore the signal-to-noise ratio is higher. To accurately reflect this heteroscedasticity, we mapped the NanoDrop spectrophotometer’s errors across wavelengths and absorbances.

Our strategy for mapping the errors was based on spectra of pure compounds measured at different concentrations. Spectra of the same compound at different concentrations are supposed to be linearly related via rescaling all the absorbances by the ratio of concentrations, and therefore the residuals after such rescaling (“autoprediction error”⁶⁴) can be viewed as the intrinsic noise of the spectrophotometer. We used this strategy for all the calibration

measurements performed for the components of the Hantzsch crude mixture to obtain the root-mean squared error map $\sigma(\lambda, A_{\text{exp}}(\lambda))$, which is shown in **Figure S17**. The region above absorbance 0.8 at 450 nm is not well-mapped because the Hantzsch reaction component's spectra lacked the absorbance bands in that region, but then the errors in this region are not relevant for Hantzsch reaction crudes. Furthermore, in our unmixing procedure we always ignore (mask) the data above absorbance 0.95, because of the deviation from the linearity of Beer-Lambert approximation. For reaction crude that may have strong absorption bands at 450 nm, we used homoscedastic model instead.

The map $\sigma(\lambda, A_{\text{exp}}(\lambda))$ is used in the heteroscedastic version of the loglikelihood function to be maximized:

$$L(\theta, A_{\text{exp}}(\lambda)) = - \sum_{\lambda} \frac{(A_{\text{exp}}(\lambda) - A_{\text{model}}(\theta, \lambda))^2}{\sigma^2(\lambda, A_{\text{exp}}(\lambda))} \quad (3)$$

where $A_{\text{exp}}(\lambda)$ is the experimentally measured spectrum of the mixture, $A_{\text{model}}(\theta, \lambda)$ is the theoretical spectrum of the unmixing model given the vector θ of model parameters.

In the code repository associated with this article, the generation of the autoprediction residuals is performed by the Python script located at “robowski/uv_vis_absorption_spectroscopy/examples/Hantzsch_reaction/calibration_for_hantzsch_for_residuals_mapping.py” relative to the code repository root. Analysis of these residuals and plotting of the error map is performed by the Python script “robowski/uv_vis_absorption_spectroscopy/absorbance_errorbar_model.py”.

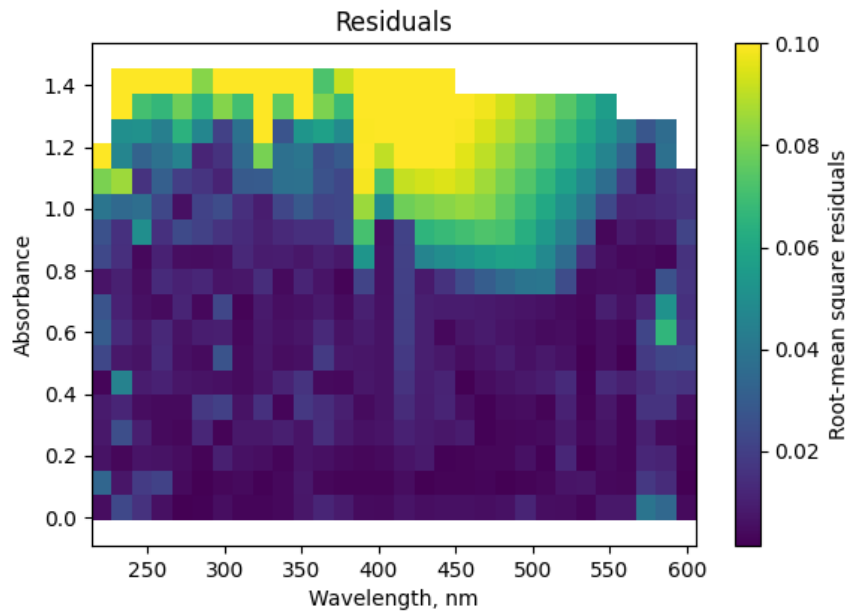


Figure S17. Map of the root-mean squared uncertainty of absorbance measured by NanoDrop spectrophotometer as function of wavelength and absorbance.

3.4. Stoichiometric inequalities

Given the possible overlaps between the spectra of individual compounds, we additionally constrained the unmixing model by enforcing the stoichiometric inequalities to preclude the sets of final concentrations that violate the mass conservation at given experimental condition.

As an illustration, consider a simple reaction $A + B \rightarrow C$. Stoichiometry of this reaction places constraints on possible concentrations of components in the final crude mixture. The difference between the starting concentration $[A]_{\text{starting}}$ and the final concentration $[A]_{\text{final}}$ of the substrate A cannot exceed the final concentration $[C]_{\text{final}}$ of the product: it would have been physically impossible. A similar constraint applies to substrate B in relation to $[C]_{\text{final}}$. In experimental practice, the values $[A]_{\text{starting}}$, $[A]_{\text{final}}$ and $[C]_{\text{final}}$ are known with uncertainties: $[A]_{\text{starting}}$ suffers from imprecision of liquid handling, $[A]_{\text{final}}$ and $[C]_{\text{final}}$ are subject to instrumental errors of the crude mixture analysis. Hence, this stoichiometric inequality is “soft”: the associated additional “cost function” is

$$H([A]_{\text{final}} + [C]_{\text{final}} - [A]_{\text{starting}}) \cdot \frac{([A]_{\text{final}} + [C]_{\text{final}} - [A]_{\text{starting}})^2}{\sigma_A^2} \\ + H([B]_{\text{final}} + [C]_{\text{final}} - [B]_{\text{starting}}) \cdot \frac{([B]_{\text{final}} + [C]_{\text{final}} - [B]_{\text{starting}})^2}{\sigma_B^2}$$

Where σ_A^2 and σ_B^2 are variances of concentrations $[A]_{\text{starting}}$ and $[B]_{\text{starting}}$, and $H(x)$ is the Heaviside function ($H(x) = 1$ if $x > 0$ and $H(x) = 0$ otherwise). The meaning of the Heaviside function is that the additional cost is zero if $[A]_{\text{final}} + [C]_{\text{final}} \leq [A]_{\text{starting}}$, and similarly for substrate B.

More generally, for a given prospective set $\{c_i\}$ of final concentrations considered by the unmixing algorithm, we compute the theoretical minimum concentrations $\widetilde{c}_{0,j}$ of each substrate j that would have to be present at the start in order to produce these substances in concentrations c_i by the end of the reaction. Namely, given the stoichiometric coefficients s_{ij} ,

$$\widetilde{c}_{0,j} = \sum_i \frac{c_i}{s_{ij}} \quad (4)$$

We then penalize the loglikelihood function (3) if the $\widetilde{c}_{0,j}$ exceeds the actual starting concentration $c_{0,j}$:

$$L(\theta, A_{\text{exp}}(\lambda)) = - \sum_{\lambda} \frac{(A_{\text{exp}}(\lambda) - A_{\text{model}}(\lambda))^2}{\sigma^2(\lambda, A_{\text{exp}}(\lambda))} - \\ - \sum_j H(\widetilde{c}_{0,j} - c_{0,j}) \cdot \frac{(\widetilde{c}_{0,j} - c_{0,j})^2}{\sigma_{c,j}^2} \quad (5)$$

where $H(x)$ is the Heaviside function, and $\sigma_{c,j}^2$ is the standard deviation associated with experimental uncertainty of preparing the solution with nominal initial concentration $c_{0,j}$.

In our codebase, the definitions of unmixing model with stoichiometric inequalities is located in the function *model_with_stoichiometric_inequalities()* in the *process_wellplate_spectra.py* module, which is in “robowski/uv_vis_absorption_spectroscopy/” directory of the code repository associated with this article.

While the spectral unmixing has been used for all the reactions in the present study, the stoichiometric inequalities were only enforced for Hantzsch reaction, as summarized in the **Table S4**. In contrast to Hantzsch reaction case, the crude mixture spectra for other reactions has only a few components, and therefore spectral unmixing for those reactions was not prone to producing results that violate the stoichiometric inequalities even when those inequalities were not explicitly enforced in the cost function.

3.5. Zero-dose extrapolation to compensate for photobleaching

Another factor one may need to consider is that a reaction crude could suffer from photobleaching during the acquisition of the spectrum. We actually observed this effect when the spectra of the Ugi-type reaction were taken on the CRAIC spectrophotometer. If we collect absorption spectra consecutively (once every 8 seconds) for the same crude mixture sample, absorbance decreases with time as shown in **Figure S18**.

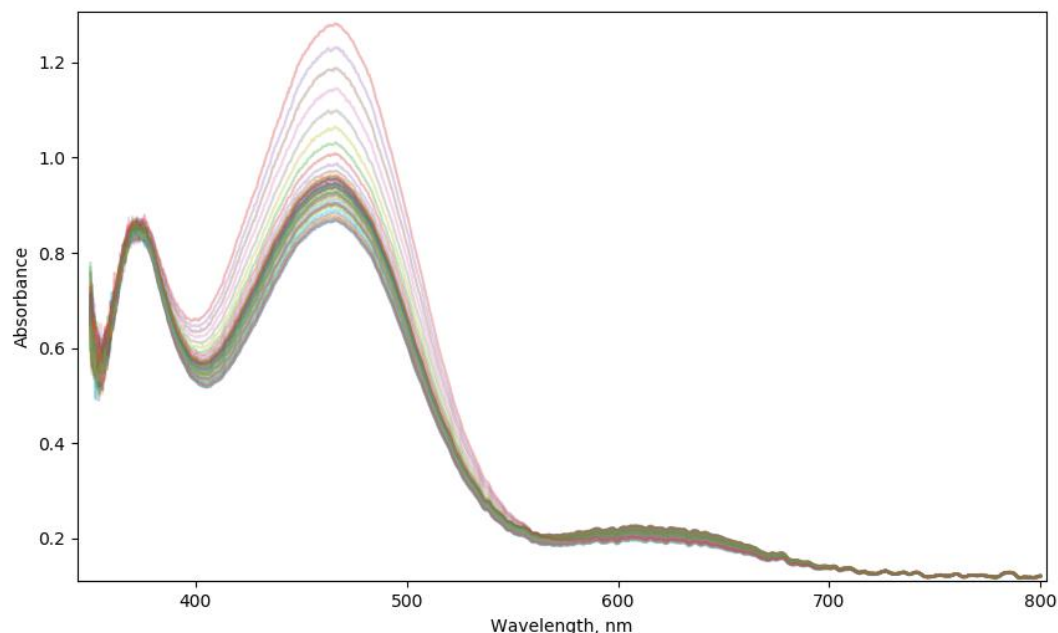


Figure S18. Photobleaching of the Ugi reaction crude. The plot shows absorbance spectra taken consecutively (every 8 seconds) for the same sample of crude mixture (Ugi reaction). The highest-lying spectrum is the one measured first.

The common way of compensating for the photobleaching is the zero-dose extrapolation: if one measures several spectra one after another, the net exposure to light (dose of absorbed photons) experienced by the mixture is known *a priori* for each spectrum, which allows one to extrapolate the spectrum to zero dose (exposure).

Let the exposure time be τ and the incident light intensity I_0 . Before the first measurement the dose is zero, linearly increases during the measurement and reaches τI_0 by the end of measurement. Average dose during the first measurement is then $0.5\tau I_0$. At the start of the second measurement, the dose is τI_0 and reaches $2\tau I_0$ by the end of it. The average dose during the second measurement is then $(\tau I_0 + 2\tau I_0)/2 = 1.5\tau I_0$. At the start of the photobleaching (small doses) the absorbance at each wavelength can be assumed to change linearly with dose due to first-order kinetics of photobleaching and linearity of absorbance with respect to concentrations. This linearity can also be seen from the first few spectra in **Figure S18**. This allows us to extrapolate the absorbance at each wavelength to zero dose as

$$A_{t=0}(\lambda) = A_{t=0.5\tau}(\lambda) - 0.5(A_{t=1.5\tau}(\lambda) - A_{t=0.5\tau}(\lambda)) \quad (6)$$

where $A_{t=0}(\lambda)$ is the absorbance spectrum obtained the first measurement, and $A_{t=1.5\tau}(\lambda)$ is the absorbance spectrum obtained in the second.

Accordingly, for the Ugi reaction, we were performing two consecutive measurements of the absorbance spectrum for each crude mixture, then applying zero-dose extrapolation and submitting the resulting spectrum to the spectral unmixing algorithm. Care was taken to avoid exposure of reaction mixtures to light during the reaction (in the temperature controlled chamber) or at any point before the spectral measurements.

In our codebase, the zero-dose extrapolation is implemented within the function `load_craic_spectrum_by_id()` defined in the `process_wellplate_spectra.py` module located in “robowski/uv_vis_absorption_spectroscopy/” directory of the code repository associated with this article.

For a summary of which reactions studied in present work were analyzed using zero-dose extrapolation technique, see **Table S4**.

3.6. Synthetic dynamic range from measurements at two dilutions

We do not use spectral data points with absorbances above 0.95 because the Beer-Lambert approximation begins to break down at high absorbance: the absorbance is no longer perfectly linear function of concentration. We should then choose the factor by which the crude mixture is diluted before the measurement so that the spectrum is not saturated (i.e., most of it is below absorbance 0.95) and not too weak (preferably above absorbance 0.1). However, choosing the optimal dilution factor may be problematic if the absorption bands due to different components

of the crude mixture (in different spectral regions) are of markedly different intensities. For example, in in case of the Hantzsch reaction, the absorption bands near the UV end of the spectral range (below 270 nm) are much stronger than the absorption bands above 270 nm. Our solution is to perform measurements of the same crude at two different dilutions: one optimized for measuring the spectrum below 270 nm, another for measurement above 270 nm. The spectral unmixing is then taking both measured spectra (at two different dilutions) and attempts to approximate both of them by a shared set of substance concentrations c_i . The baseline component weights b_1 and b_2 , wavelength offsets $\Delta\lambda$ and absorbance offsets δ are individual for each of the two measured spectra. The dilution factor for the first dilution is fixed and equal to its experimental value; the dilution factor of the second dilution is treated as a free parameter that is allowed to vary within the 5% of its nominal value to account for the imperfect precision of the automatic pipetting system.

This data processing approach is similar to “stitching” the two spectra to extend the effective dynamic range of usable absorbance (an approach used for constructing reference spectra, Section 3.8.1), but it is more robust than plain “stitching” because it does not imply identical baselines, wavelength offsets, absorbance offsets. It also treats the instrumental errors more consistently.

For a Python implementation of this method, see the function *multispectrum_to_concentration()* of the *process_wellplate_spectra.py* module located in “robowski/uv_vis_absorption_spectroscopy/” directory of the code repository associated with this article.

3.7. Quality control of the spectral data

We apply systematic quality control to exclude problematic wavelength regions that could compromise fitting accuracy. This masking strategy addresses multiple sources of instrumental artifacts and measurement limitations.

The quality control is implemented through the *create_spectrum_mask()* function of *process_wellplate_spectra.py* module. This function applies three sequential filtering criteria: (1) limiting the spectral range, (2) absorbance ceiling, and (3) removal of instrumental artefacts. These criteria are discussed in the sections below.

These masking criteria are combined (using logical “AND” operations on Boolean arrays/masks in Python) to create a comprehensive Boolean mask that identifies wavelength regions suitable for quantitative analysis. The resulting mask typically retains 60-80% of the original spectral data, focusing the analysis on high-quality measurements where the physical assumptions of the spectral model are valid. Yellow zones in the figures in Section 3.12 are showing the regions of spectral data excluded according to the combined criteria described here, as applied to examples of experimental spectra obtained in present work.

3.7.1. Limiting the spectral range

The parameters *cut_from* and *cut_to* define the usable spectral window. The *cut_from* parameter (typically 5-50 wavelength points, see **Table S4** below for specific values used for each reaction) excludes the blue end of the spectrum where the quality of the spectrum can be compromised by low intensity of light source at those wavelengths, by absorption of solvent itself, and by presence of impurities. The *cut_to* parameter removes red-end regions with negligible absorption signal that contribute only noise to the analysis. Note that this trimming of spectral range is applied in addition to the one performed on loading of NanoDrop files as described in Section 2.5 above to remove the sample-to-sample variability of the spectral range of NanoDrop.

3.7.2. Absorbance ceiling

The *upper_limit_of_absorbance* parameter (0.95 for NanoDrop spectrophotometer) prevents model fitting in regions where the linearity of the Beer-Lambert law breaks down and/or where the spectral noise is increased due to low amount of light reaching the detector. This threshold ensures that only data within the linear response regime is used for quantitative analysis. In the case of CRAIC spectrophotometer, whose performance gradually deteriorates with decreasing wavelength near the UV region due to low light intensity provided by the instrument in UV, we implemented the “absorbance ceiling” that decreases with wavelength in that region, as shown in **Figure S19**. Specifically, the threshold is a piecewise linear function passing through the following node points: [350 nm, 0.67 absorbance units], [375 nm, 0.75 absorbance units], [377 nm, 1.6 absorbance units]. It stays at 1.67 absorbance units for wavelengths above 377 nm. In our codebase, these node points are set by class variables of the “*SpectraProcessor*” class in “*process_wellplate_spectra.py*” module.

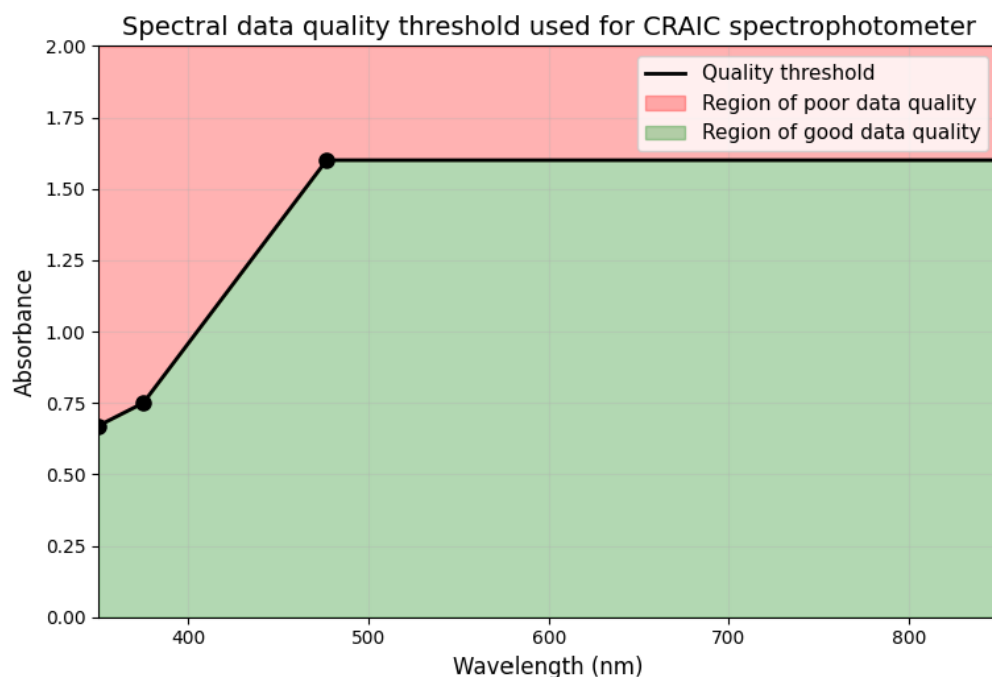


Figure S19. Quality control threshold applied to spectral data points for spectra obtained with CRAIC spectrophotometer. Quality threshold is a piecewise linear function passing through the following node points: [350 nm, 0.67 absorbance units], [375 nm, 0.75 absorbance units], [377 nm, 1.6 absorbance units]. It stays at 1.67 absorbance units for wavelengths above 377 nm.

3.7.3. Removal of instrumental artefacts

When the absorbance exceeds the value of 1.5 or 2 in NanoDrop spectrometer, the amount light reaching the detector becomes comparable to noise level. Unfortunately, the consequence of this low light level is that the apparent absorbance does not merely stay high, near the “absorbance ceiling”. Instead, the spectrometer produces spurious spectral features having apparent absorbance values anywhere between zero and 2 units. The logic of excluding regions is based on the following observation. Typically, the absorption at the red end of the spectrum (highest wavelength) is the weakest. With decreasing wavelength, the absorbance increases – suppose that for a given sample it eventually reaches a value of 1.5 at some wavelength λ_a . With further decrease of wavelength, the absorbance first goes higher, and then proceeds to generate spurious spectral features: where the true absorbance may be as high as 10, the spectrometer may report low absorbances around, say, 0.7. Hence, we exclude all wavelengths below λ_a from the analysis. Note that λ_a is determined individually for each spectrum. In our codebase, the *artefactogenic_upper_limit_of_absorbance* parameter (set to 1.5 for NanoDrop measurements, unused for CRAIC spectrometer) is controlling the threshold described above. In a given spectrum, the algorithm finds the highest wavelength λ_a where absorbance exceeds this threshold and excludes all shorter wavelengths from further downstream analysis.

3.7.4. Recommendations for optimizing the quality control parameters

We recommend that the values of the abovementioned parameters are chosen based on the inspection of the spectra of pure compounds during the calibration procedure. As demonstrated by a worked example in the Jupyter Notebook located at “notebooks/spectral_unmixing_tutorial.ipynb” in the associated code repository, the calibration algorithm attempts to scale the reference spectrum of a pure substance (e.g. spectrum collected at a specific “reference” concentration) to perfectly match a spectrum of the same substance at a different concentration. If one notices a substantial irremovable mismatch between the scaled reference spectrum and the measured spectra at different concentrations, we recommend to exclude the spectral range where this mismatch is present.

3.8. Construction of reference spectrum of a given compound

Calibration procedure establishes, for each pure component, the reference spectrum and the relationship between the concentration and the scaling of the reference spectrum. This process is implemented in the *perform_calibration()* function of the *calibrator.py* module located in “robowski/uv_vis_absorption_spectroscopy/” directory of the code repository associated with

this article. It is also demonstrated in the worked example in the form of Jupyter Notebook located at “notebooks/spectral_unmixing_tutorial.ipynb” in the associated code repository.

Multispectrum unmixing described in Section 3.6 above demands a high dynamic range and signal-to-noise ratio of the reference spectrum, because the same reference spectrum must be used to model the spectra of less dilute samples and the more dilute ones. This section describes considerations and techniques we have developed to improve the quality of the reference spectrum. **Table S5** and **Table S6** show whether each of these techniques has been used for calibration of a given substance studied in the present work.

3.8.1. Creating the reference spectrum by stitching spectra measured as several concentrations

A spectrum measured at a chosen reference concentration serves as the starting point. This concentration should not be too high or too low: it should be chosen to provide good signal-to-noise ratio (typically resulting in absorbance >0.1) while avoiding instrumental artifacts (typically keeping absorbance <1.0).

Reference spectrum can be constructed by “stitching” spectra measured at different concentrations. When “*do_reference_stitching*” parameter of “*perform_calibration()*” method is set to “*True*”, the algorithm iteratively improves the quality of sections of the reference spectrum by incorporating spectral data from higher-concentration samples where signal-to-noise ratio for weak absorption bands is superior, even though some spectral bands may be already above the absorbance limit of the spectrometer. The masking strategy described in Section 3.7 is used to identify the high-quality data in each of the measured spectra, and then these high-quality regions from different spectra are “stitched” together to yield the final reference spectrum. This requires scaling the magnitude of these pieces so that their overlapping parts match precisely before they are stitched. For achieving the best possible match, we use a linear fitting procedure instead of simply relying on the known concentrations at which the spectra were collected. This allows us to correct for small, but non-zero errors of the pipetting made during preparation of the solutions with required concentrations. Specifically, the “*fit_reference_model_to_target_spectrum()*” method of “*calibrator.py*” module operates on current-best reference spectrum $A_{\text{ref}}(\lambda)$ and a spectrum $A(\lambda, C_i)$ obtained at a higher concentration C_i . The method fits a linear model: $A(\lambda, C_i) = a \cdot A_{\text{ref}}(\lambda) + b$, where a and b are the free parameters: a is the scaling coefficient and b corrects for baseline differences. High-quality region of $A(\lambda, C_i)$ data is then identified as described in Section 3.7, and the data in the respective wavelength range of the reference spectrum $A_{\text{ref}}(\lambda)$ is overwritten with values $(A(\lambda, C_i) - b)/a$. Using progressively higher C_i , this process iteratively replaces noisy regions of $A_{\text{ref}}(\lambda)$ with higher-quality data from more concentrated samples.

3.8.2. Applying adaptive noise reduction to the reference spectrum

Next, we optionally apply adaptive noise reduction to the reference spectrum. This option is controlled by setting the parameter *do_smoothing_at_low_absorbance* of the *perform_calibration()* function. The purpose of this procedure is to smooth the spectrum at low absorbance regions in order to reduce instrumental noise, but to preserve precise spectral shape at high absorbance regions. Exponential weighting provides gradual transition of smoothing behavior between the low- and high-absorbance regions.

Let $A_{\text{ref}}(\lambda)$ be the original reference spectrum and $S(\lambda)$ be its Savitzky-Golay filtered version. Savitsky-Golay is performed with a 4th-order polynomial and window length set by *savgol_window* parameter.

The algorithm defines an exponential decay constant $\alpha = \tau \cdot A_{\text{max}}$, where τ is the value of *do_smoothing_at_low_absorbance* parameter, and $A_{\text{max}} = \max_{\lambda} A_{\text{ref}}(\lambda)$ is the maximum absorbance in the spectrum $A_{\text{ref}}(\lambda)$. The final adaptively smoothed reference spectrum $A_{\text{final}}(\lambda)$ is the weighted sum of original $A_{\text{ref}}(\lambda)$ and smoothed $S(\lambda)$, with weights depending on the absorbance level of the $A_{\text{ref}}(\lambda)$:

$$A_{\text{final}}(\lambda) = S(\lambda) e^{-A_{\text{ref}}(\lambda)/\alpha} + A_{\text{ref}}(\lambda)(1 - e^{-A_{\text{ref}}(\lambda)/\alpha})$$

The parameter τ serves as a relative absorbance threshold that controls the transition point between aggressive smoothing and feature preservation: in the spectral regions where absorbance is much smaller than α (i.e., typically at the red tail of the spectrum) the final spectrum $A_{\text{final}}(\lambda)$ is almost equal to the Savitsky-Golay filtered version $S(\lambda)$, whereas in the spectral regions where absorbance is much larger than α , $A_{\text{final}}(\lambda)$ is equal to the original (unsmoothed) spectrum $A_{\text{ref}}(\lambda)$.

This smoothing functionality is implemented in the “*_smooth_the_reference_spectrum()*” function of the “*calibrator.py*” module located in “*robowski/uv_vis_absorption_spectroscopy/*” directory of the code repository associated with this article.

3.9. Analysis of residuals with autocorrelation statistics

Residual spectrum is the difference between the observed spectrum of a complex mixture and the spectrum produced by the best-fit spectral unmixing model. In an ideal scenario, the residual spectrum should be statistically indistinguishable from white (Gaussian) noise of a perfect spectrophotometer. Because white noise is uncorrelated, statistical measures of autocorrelation^{71,72} of residual spectrum across wavelengths (and not across time) can indicate that the model misses underlying patterns, such as spectral bands of unexpected species present in the mixture. While traditionally associated with statistical analysis of time series, these autocorrelation tests apply to any regression residuals^{71,73,74}.

Classical definition of Durbin-Watson statistic⁷¹ considers correlation of preceding data point with the next. In our case, this means wavelength difference of 1 nm between the neighbouring points. However, when there are unexpected spectral bands in the spectrum, the characteristic widths of these bands is much larger than 1 nm. As a result, unexpected bands have less of an effect on the autocorrelation between nearest neighbours than on the correlation between points separated by typical width of the unexpected bands. With this in mind, to increase sensitivity of autocorrelation statistic to the appearance of new absorption bands, we computed the Durbin-Watson statistic at “lag” equal to 30 nm – a value chosen to be comparable to typical half-widths at half-maximum of the absorption bands in UV-Vis spectra.

An alternative statistic we have implemented is the Ljung-Box statistic⁷², which offers a more comprehensive approach by testing for autocorrelation at multiple lags instead. We evaluated at a lag equal to 20% of the spectral range, as recommended by the statsmodels package.

In practice, the noise of the spectrophotometer has some degree of autocorrelation as well. Therefore, the autocorrelation of residuals after fitting (unmixing) the spectrum of complex mixture should be compared not to an uncorrelated white noise, but to the autocorrelation of residuals obtained for “blank” sample (i.e., pure solvent) with the same spectrophotometer. In our experience, for the case of specific spectrophotometers we operated, the “baseline” autocorrelation is more stable and reproducible for Durbin-Watson metric than for Ljung-Box metric, so both metrics can be useful as indicators of anomalies, but Ljung-Box metric is more prone to “false positives”. Specifically, our rule of thumb was that an anomaly should be suspected when Durbin-Watson statistic is below 1 or above 3, or when Ljung-Box statistic is above 1000.

3.10. Tables of calibration and spectral unmixing options used in each experiment

Table S4. Table of experimental and spectral unmixing parameters used for a given reaction.

Reaction	E1	Simple SN1	SN1	Ugi	Hantzsch	PBA
Figure in the main text	2a	2b	3	4	5	6
Spectrometer	Nano Drop	Nano Drop	CRAIC	CRAIC	Nano Drop	Nano Drop
Photobleaching zero-dose extrapolation	No	No	Yes	Yes	No	No
Upper cut-off of spectral range (cut_to)*	Not set	400 nm	Not set	Not set	570 nm	350 nm
Lower cut-off of spectral range (cut_from)*	246 nm	262 nm	429 nm	Not set	221 nm	No
Background model with PCA components	No	No	Yes	Yes	Yes	No

Synthetic dynamic range from spectra at two dilutions	No	No	No	No	Yes	No
Stoichiometric inequalities enforced	No	No	No	No	Yes	No
Heteroscedastic instrumental errors	No	No	No	No	Yes	No
Wavelength shift $\Delta\lambda$ allowed (eq. (2))	No	No	No	No	Yes	No

* In the signal processing code, the “*cut_from*” and “*cut_to*” parameters are in wavelength indices, not in nanometers. For example, since the default range of NanoDrop spectrophotometer is from 220 nm to 600 nm, as set by variables “*nanodrop_lower_cutoff_of_wavelengths*” and “*nanodrop_upper_cutoff_of_wavelengths*” variables of the “*SpectraProcessor*” class in “*process_wellplate_spectra.py*” module (Section 2.5 above), setting *cut_from*=42 translates to lower cut-off wavelength equal to $220+42 = 262$ nm, and setting *cut_to*=180 translates to upper cut-off wavelength equal to $220+180 = 400$ nm.

Table S5. Table of experimental and software parameters used for calibration of Hantzsch reaction reactants and products

Species code in the main text	Name in the codebase	Reference stitching	Reference by Agilent Cary 5000	Reference smoothing at low absorbance	Upper limit of absorbance	Reference concentration (mol/L)
19a	ethyl_ace toacetate	No	Yes	Yes (0.02)	0.95	0.006
19c	methoxy benzaldehyde	Yes	Yes	No	0.95	0.006
19d	HRP01	No	Yes	No	0.7557	0.0003
19e	bb017	No	Yes	No	0.95	0.0003
19f	dm40_12	Yes	Yes	No	0.95	0.01
19g	dm40_10	Yes	Yes	No	0.95	0.01
19h	bb021	No	Yes	No	0.95	0.0002
19i	dm088_4	No	Yes	No	0.95	0.0003
19j	dm053	No	Yes	No	0.95	0.0003
19k	dm35_9	No	Yes	No	0.95	0.0003
19m	dm35_8	Yes	No	No	0.95	0.0003
19o	dm36	Yes	No	No	0.95	0.0005
19p	dm70	Yes	No	No	0.95	0.00128
19r	dm37	No	Yes	No	0.95	0.0005
EAB	EAB	No	Yes	No	0.95	0.0007
(E,E) isomer of 19h	bb021_f2	No	Yes	No	0.95	0.0005
Acetic acid	acetic_acid	No	No	Yes (0.02)	0.95	0.05978

Table S6. Table of experimental and software parameters used for calibration of reactants and products of reactions other than Hantzsch. Reaction names correspond to ones in **Table S4**, which has references to the figures in the main text.

Reaction	Species code in the main text	Name in the codebase	Reference stitching	Reference by Agilent Cary 5000	Reference smoothing at low absorbance	Upper limit of absorbance	Reference concentration (mol/L)
E1	13a	E1OH02	No	No	No	0.95	0.0006
E1	13b	E1DB02	No	No	No	0.95	0.0003
Simple SN1	14a	SN1OH03	No	No	No	0.95	0.00075
Simple SN1	14b	SN1Br03	No	No	No	0.95	0.00075
Simple SN1	HBr	HBr	No	No	No	0.95	0.00894
SN1	15a	SN1OH01	No	No	No	Curve	5.74E-05
SN1	15d	SN1Br01s1	No	No	No	Curve	0.002585
Ugi	16e	IIO029A	Yes	No	No	Curve	0.00011
Ugi	16a	ald001	No	No	No	Curve	0.01921

3.11. Implementation

While the simple model (1) is, mathematically, a vector decomposition problem (where vector components are absorbance at specific wavelengths) and therefore can be formulated in terms of standard linear regression algorithms, these algorithms can no longer be used for a more complex model (2). For solving the unmixing problem (2) (with or without the stoichiometric inequalities enforced), we use a more general Trust Region Reflective algorithm⁷⁵ with 2-point method of computing the Jacobian, as implemented in the `curve_fit` method of the SciPy library⁷⁶. The fitted parameter uncertainties are obtained as square roots of the diagonal elements of the covariance matrix returned by `curve_fit`; this corresponds to the standard linear approximation method for estimating the confidence intervals of model parameters^{77–79}. These uncertainties are propagated further into the analysis of the yield maps.

For Durbin-Watson and Ljung-Box statistic of autocorrelation in residuals, we used the methods from the StatsModels Python package⁸⁰.

3.12. Examples of unmixed spectra for specific reactions from Extended Data Fig. 1

Each figure in this section presents the diagnostic plot reporting the results of spectral unmixing performed for a spectrum of a specific reaction mixture.

These plots were produced by functions `“diagnostic_plot_of_spectrum_to_concentration_unmixing()”` and `“plot_result_of_multispectrum_unmixing()”` of the module `“robowski/uv_vis_absorption_spectroscopy/process_wellplate_spectra.py”` in the code repository associated with this article. Grey curve is the raw experimental data before subtracting the nominal “baseline”/“background” spectrum, and the black curve – the one after this subtraction was done. Red curve labelled “fit” in the legend is the theoretical spectrum produced by the best-fit model of spectral unmixing. Curves of other colors (blue, orange, green and so on) show contributions of the spectra of individual substances to the overall (red) spectrum of the unmixing model. If a linear drift term (i.e., absorbance growing linearly with wavelength) is included into the background model, the contribution of this term is labeled as “Line” in the legend. If PCA components of the background are included into the background model (described in Section 3.2 above), contribution of these PCA components is shown on the plot and labeled as “Bkg. PC1” (first component) and “Bkg. PC2” (second component). Yellow rectangles mark the wavelength ranges that are excluded from the spectral unmixing analysis – for description of exclusion criteria see Section 3.7 above.

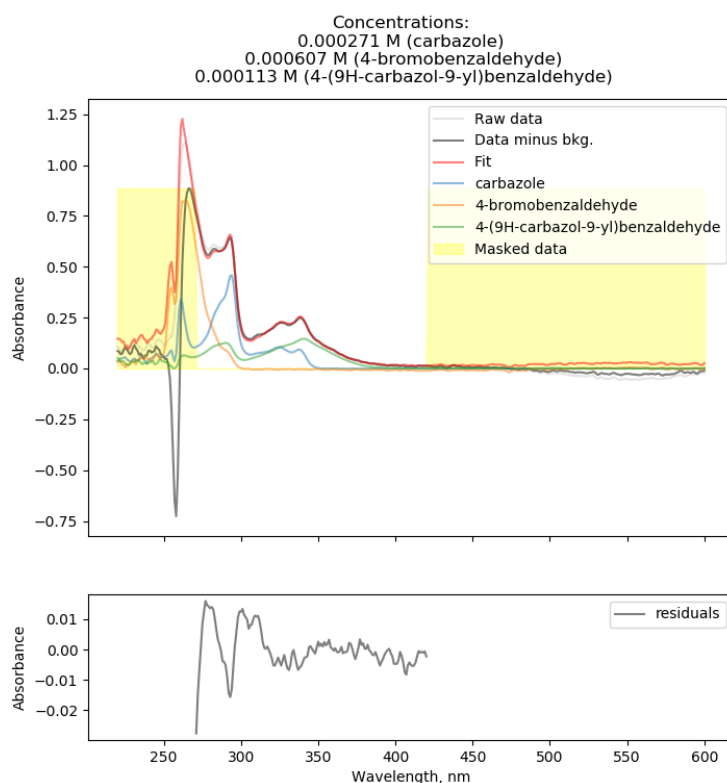


Figure S20. Spectral unmixing for Ullmann-type reaction 2

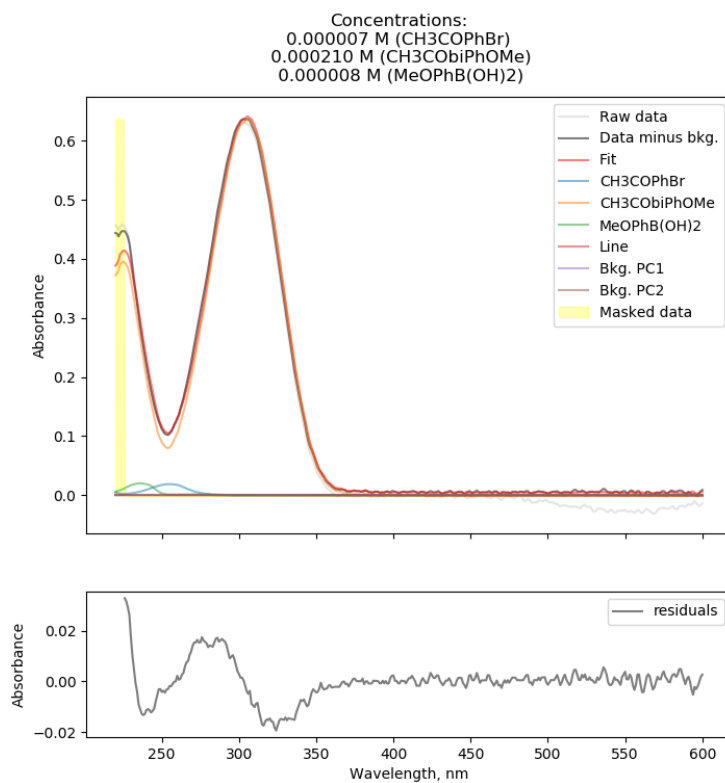


Figure S21. Spectral unmixing for Suzuki–Miyaura cross-coupling **3**

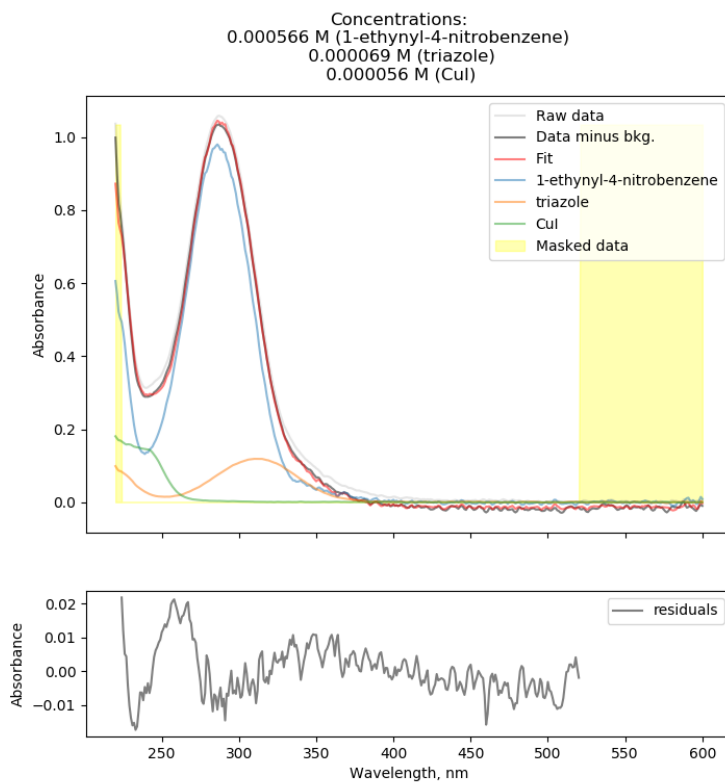


Figure S22. Spectral unmixing for Cu(I)-catalyzed alkyne–azide cycloaddition **4**

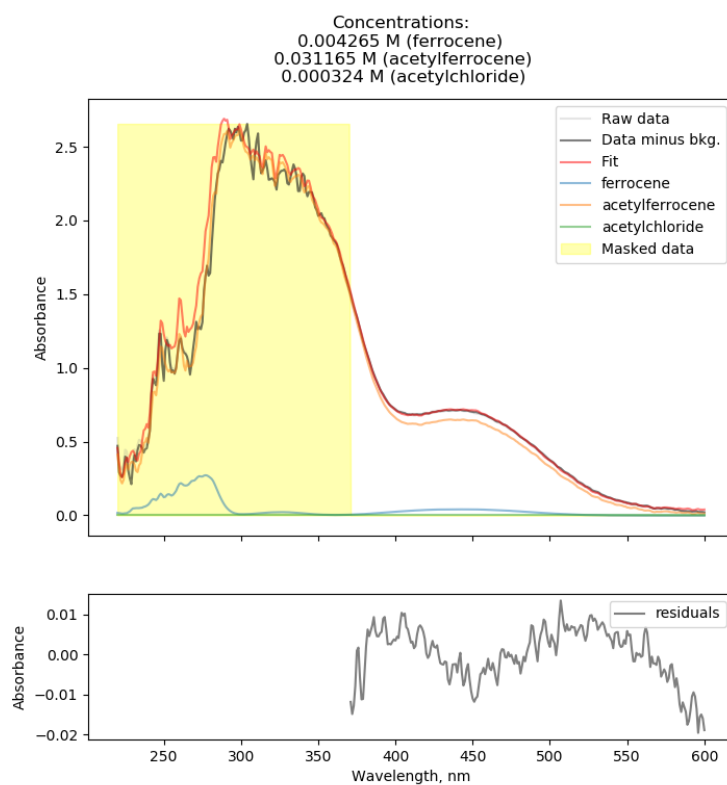


Figure S23. Spectral unmixing for Friedel–Crafts acylation **5**

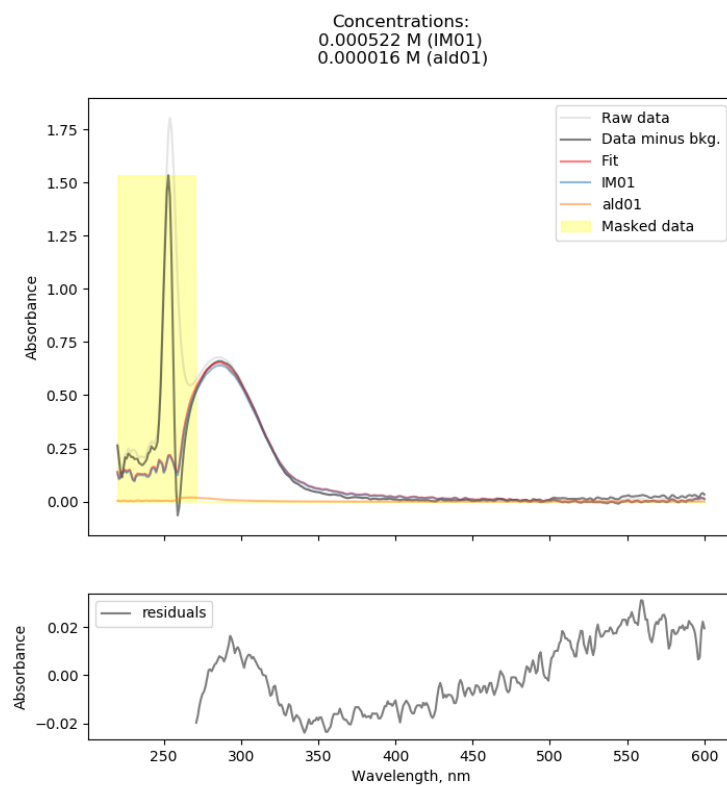


Figure S24. Spectral unmixing for imine condensation **6**

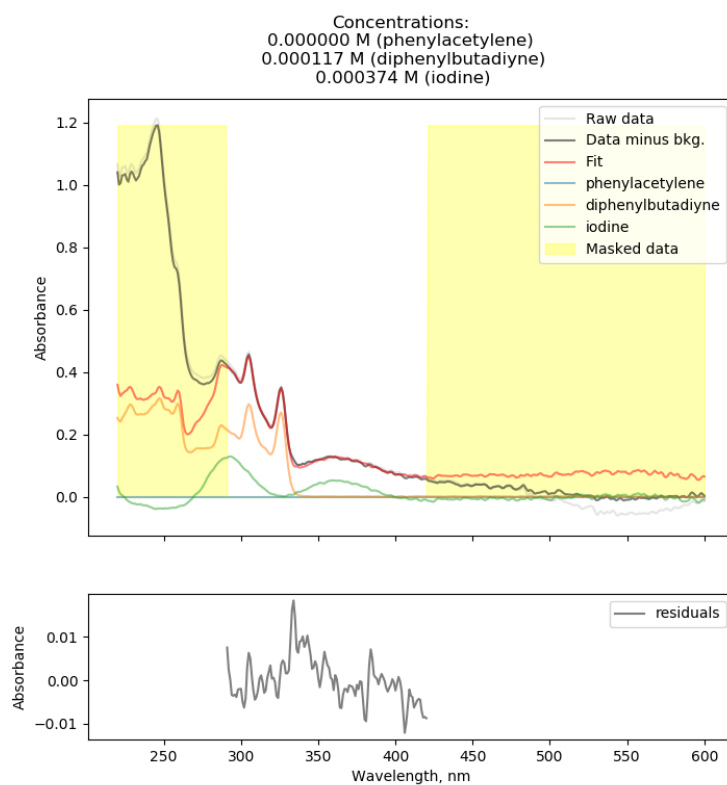


Figure S25. Spectral unmixing for Glaser coupling 7

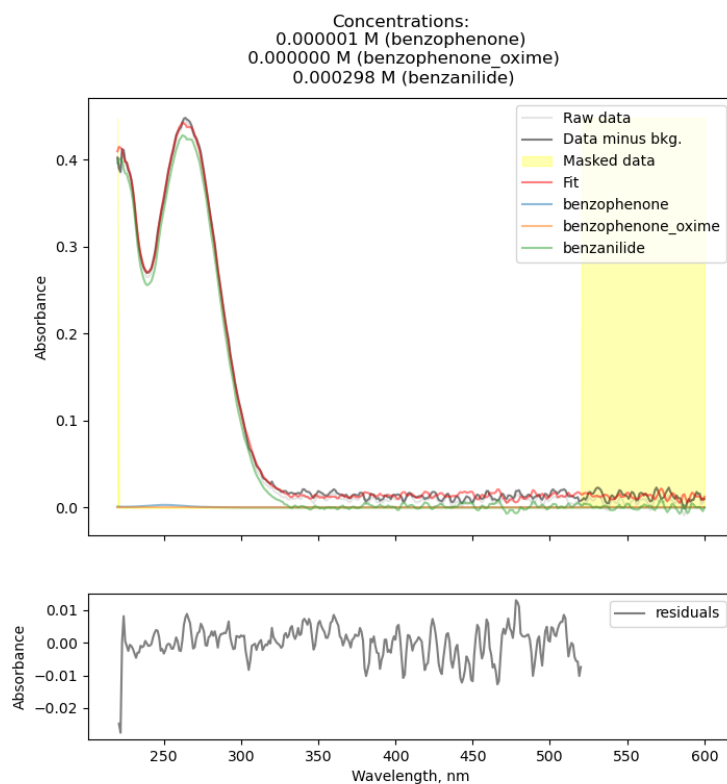


Figure S26. Spectral unmixing for Beckmann rearrangement 8

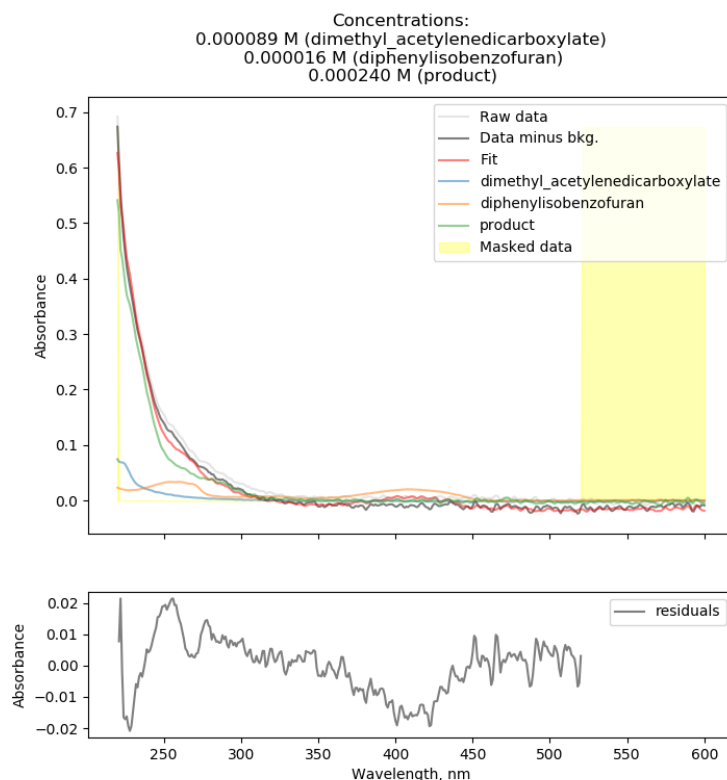


Figure S27. Spectral unmixing for Diels–Alder reaction **9**

3.13. Outlier filtering and smoothing of the Ugi reaction yield maps

Outliers in the Ugi reaction yield maps were filtered by smoothing the experimental yield profiles across the concentration of p-TSA, which had the highest density of sampling in our experiments. This filtering and smoothing was performed separately for each combination of starting concentration of amine, aldehyde and isocyanide, and the comparison of smoothed to raw data is presented in **Figure S126** – **Figure S138**. First, an outlier was defined as a point whose yield deviates in the same direction from the preceding and the next point by more than a threshold of 0.02. Each outlier's value was then replaced by the average of the preceding and the next point. Then, a weighted Savitsky-Golay filter of degree 2 and window length of 5 points was applied. Weighting was based on the 14.5% relative error. The result of Savitsky-Golay smoothing is shown by yellow curves in **Figure S126** – **Figure S138**. Next, weighted B-spline interpolation was applied with smoothing factor equal to 0.0005. That is, the sum of weighted squared residuals (differences between the B-spline and data points) must be less than or equal to 0.0005 (see SciPy `splrep` method for implementation details). Again, weighting was based on the 14.5% relative error.

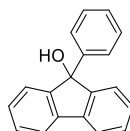
For a Python implementation of this procedure, see Python script located at “*robowski/visualize_results/smooth_results.py*” in the code repository associated with this article.

4. Experimental procedures and results

Mass spectra, NMR spectra and HPLC chromatograms mentioned in this section are shown in Section 11, 12 and 13, respectively.

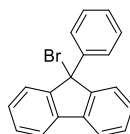
4.1. Compounds synthesized manually

4.1.1. E1 and S_N1 reactions



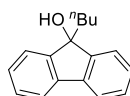
Compound 14a (9-phenyl-9H-fluoren-9-ol) was synthesized according to the literature procedure.¹

¹H NMR (400 MHz, chloroform-*d*): δ = 7.68 (d, J = 7.5 Hz, 2H), 7.52–7.57 (m, 4H), 7.37 (dt, J = 7.6, 1.1 Hz, 2H), 7.24–7.32 (m, 5H). **¹³C NMR** (100 MHz, chloroform-*d*): δ = 149.8, 141.3, 138.2, 129.1, 128.7, 128.5, 128.2, 127.6, 126.2, 120.5, 67.6. **HRMS** (ESI–TOF): m/z : [M + H – H₂O]⁺ Calcd for C₁₉H₁₃⁺: 241.1012; Found 241.1011.



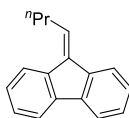
Compound 14b (9-bromo-9-phenyl-9H-fluorene) was synthesized according to the literature procedure.⁸¹

¹H NMR (400 MHz, chloroform-*d*): δ = 7.68 (d, J = 7.5 Hz, 2H), 7.52–7.57 (m, 4H), 7.37 (dt, J = 7.6, 1.1 Hz, 2H), 7.24–7.32 (m, 5H). **¹³C NMR** (100 MHz, chloroform-*d*): δ = 149.8, 141.3, 138.2, 129.1, 128.7, 128.5, 128.2, 127.6, 126.2, 120.5, 67.6. **HRMS** (ESI–TOF): m/z : [M + H – H₂O]⁺ Calcd for C₁₉H₁₃⁺: 241.1012; Found 241.1011.



Compound 13a (9-butyl-9H-fluoren-9-ol) was synthesized according to the literature procedure.⁸²

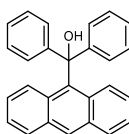
¹H NMR (400 MHz, chloroform-*d*): δ = 7.61 (d, J = 7.2 Hz, 2H), 7.51 (d, J = 7.2 Hz, 2H), 7.36 (td, J = 7.5, 1.3 Hz, 2H), 7.30 (td, J = 7.5, 1.3 Hz, 2H), 2.15 (m, 2H), 1.18 (sextet, J = 7.4 Hz, 2H), 0.87 (m, 2H), 0.75 (t, J = 7.4 Hz, 3H). **¹³C NMR** (100 MHz, chloroform-*d*): δ = 148.9, 139.7, 129.0, 128.1, 123.6, 120.0, 82.6, 39.3, 26.2, 23.0, 13.9. **HRMS** (ESI–TOF): m/z : [M + Na]⁺ Calcd for C₁₇H₁₈ONa⁺: 261.1250; Found 261.1245; m/z : [M + H – H₂O]⁺ Calcd for C₁₇H₁₇⁺: 221.1325; Found 221.1326.



Compound 13b (9-butylidene-9*H*-fluorene).

To the solution of 9-butyl-9*H*-fluoren-9-ol (1 g, 4.20 mmol) in 25 ml of acetonitrile, 1.5 ml of 48% water solution of HBr was added. The mixture was heated at 60 °C for 4 h. After cooling down to room temperature, the brine was added and the mixture was extracted three times with dichloromethane. Combined organic layers were dried over anhydrous sodium sulfate and evaporated under reduced pressure. The crude product was recrystallized from methanol (yield 50%).

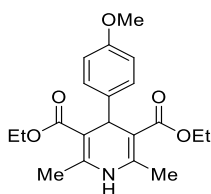
¹H NMR (400 MHz, chloroform-*d*): δ = 7.88 (d, J = 7.4 Hz, 1H), 7.77 (d, J = 7.0 Hz, 1H), 7.72 (d, J = 6.7 Hz, 1H), 7.68 (d, J = 6.7, 1H), 7.27–7.40 (m, 4H), 6.77 (t, J = 7.4 Hz, 1H), 2.84 (q, J = 7.5 Hz, 2H), 1.74 (sextet, J = 7.4, 2H), 1.10 (t, J = 7.4 Hz, 3H). **¹³C NMR** (100 MHz, chloroform-*d*): δ = 140.9, 139.5, 138.7, 137.7, 135.6, 131.2, 127.7, 127.4, 127.0, 126.9, 125.1, 119.9, 119.8, 119.6, 31.5, 23.0, 14.2. **HRMS** (ESI–TOF): m/z : [M + H]⁺ Calcd for C₁₇H₁₇⁺: 221.1325; Found 221.1327..



Compound 15a (anthracen-9-ylidiphenylmethanol) was synthesized according to the literature procedure.⁸³

¹H NMR (400 MHz, chloroform-*d*): δ = 8.46 (s, 1 H), 7.96 (d, J = 8.4 Hz, 2 H), 7.90 (dd, J = 9.2, 1.2 Hz, 2 H), 7.32 (m, 12 H), 7.03–7.07 (m, 2 H), 3.08 ppm (s, 1 H). **¹³C NMR** (100 MHz, chloroform-*d*): δ = 148.5, 139.2, 132.2, 130.9, 129.4, 128.8, 128.6, 128.2, 127.9, 127.6, 124.3, 124.1, 84.2. **HRMS** (ESI–TOF): m/z : [M + Na]⁺ Calcd for C₂₇H₂₀ONa⁺: 383.1407; Found 383.1399; m/z : [M + H – H₂O]⁺ Calcd for C₂₇H₁₉⁺: 343.1482; Found 343.1481.

4.1.2. Hantzsch reaction



Compound 19d (diethyl 4-(4-methoxyphenyl)-2,6-dimethyl-1,4-dihydropyridine-3,5-dicarboxylate) was synthesized according to the literature procedure.⁸⁴

In a pressure tube, 4-methoxybenzaldehyde (89.60 mg, 80 μ L, 0.65 mmol), ethyl acetoacetate (155.72 mg, 151 μ L, 1.17 mmol), and ammonium acetate (55.34 mg, 0.70 mmol) were dissolved in 10 mL of ethanol. The mixture was then heated at 80 $^{\circ}$ C for 48 h. After cooling to room temperature, the reaction mixture was diluted with water and extracted with ethyl acetate. The organic phase was dried over anhydrous sodium sulfate and then the solvent was removed under reduced pressure. The crude mixture was passed via silica column chromatography using 30% DCM in *n*-Hex to remove starting aldehyde and then subjected to RP-HPLC purification to obtain pure product (collection 1 in HPLC, 121 mg, 57%) as a colorless solid.

^1H NMR (400 MHz, chloroform-*d*): δ = 7.19 (d, J = 8.7 Hz, 2 H), 6.74 (d, J = 8.9 Hz, 2 H), 5.67 (s, 1 H), 4.92 (s, 1 H), 4.09 (m, 4 H), 3.75 (s, 3 H), 2.32 (s, 6 H), 1.22 (t, J = 7.2 Hz, 6 H).
 ^{13}C NMR (100 MHz, chloroform-*d*): δ = 167.89, 157.93, 143.96, 140.47, 128.99, 113.25, 104.22, 59.76, 55.19, 38.80, 19.46, 14.34.

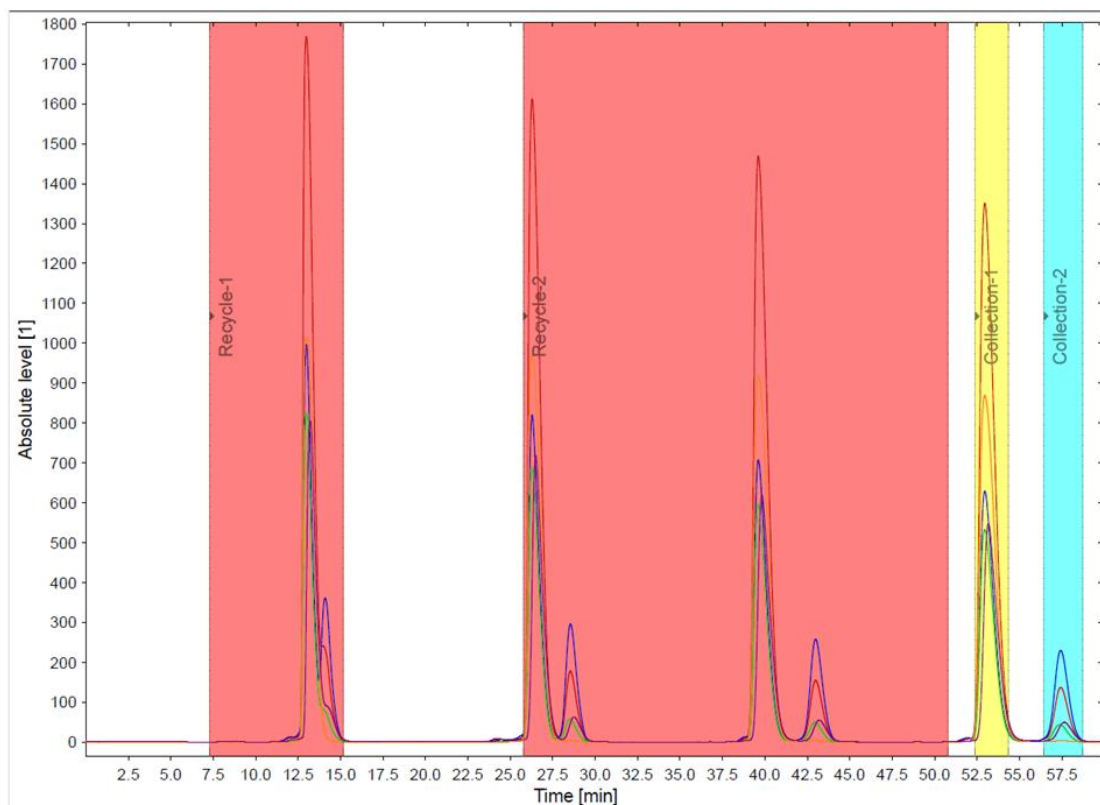
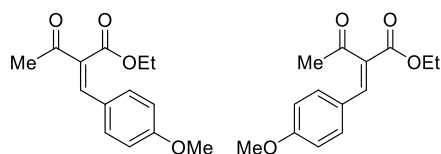
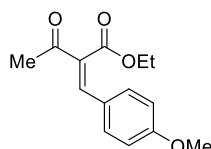


Figure S28. Chromatogram for purification of **19d** (RP-HPLC/MeOH).

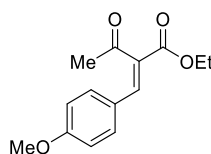


Compound 19f and **19g** (ethyl (*Z*)-2-(4-methoxybenzylidene)-3-oxobutanoate and ethyl (*E*)-2-(4-methoxybenzylidene)-3-oxobutanoate) were synthesized according to the literature procedure⁸⁴ and separated by RP–HPLC.



Compound 19f (ethyl (*Z*)-2-(4-methoxybenzylidene)-3-oxobutanoate).

¹H NMR (400 MHz, chloroform-*d*): δ 7.48 (s, 1H), 7.40 (d, J = 8.6 Hz, 2H), 6.87 (d, J = 8.9 Hz, 2H), 4.33 (q, J = 7.2 Hz, 2H), 3.81 (s, 3H), 2.37 (s, 3H), 1.29 (t, J = 7.2 Hz, 3H). **¹³C NMR** (100 MHz, chloroform-*d*): δ 194.75, 168.35, 161.85, 141.09, 132.32, 131.81, 125.32, 114.43, 61.67, 55.44, 26.43, 14.00.



Compound 19g (ethyl (*E*)-2-(4-methoxybenzylidene)-3-oxobutanoate).

¹H NMR (400 MHz, chloroform-*d*): δ 7.60 (s, 1H), 7.35 (d, J = 8.7 Hz, 2H), 6.87 (d, J = 8.8 Hz, 2H), 4.28 (q, J = 7.1 Hz, 2H), 3.82 (s, 3H), 2.37 (s, 3H), 1.31 (t, J = 7.1 Hz, 3H). **¹³C NMR** (100 MHz, chloroform-*d*): δ 204.12, 164.86, 161.59, 140.43, 131.88, 131.65, 125.56, 114.5, 61.47, 55.49, 31.36, 14.32. **HRMS** (ESI–TOF): m/z : $[M + H]^+$ Calcd for C₁₄H₁₇O₄⁺: 249.1122; Found 249.1118; m/z : $[M + Na]^+$: Calcd for C₁₄H₁₆O₄Na⁺: 271.0941; Found 271.0934.

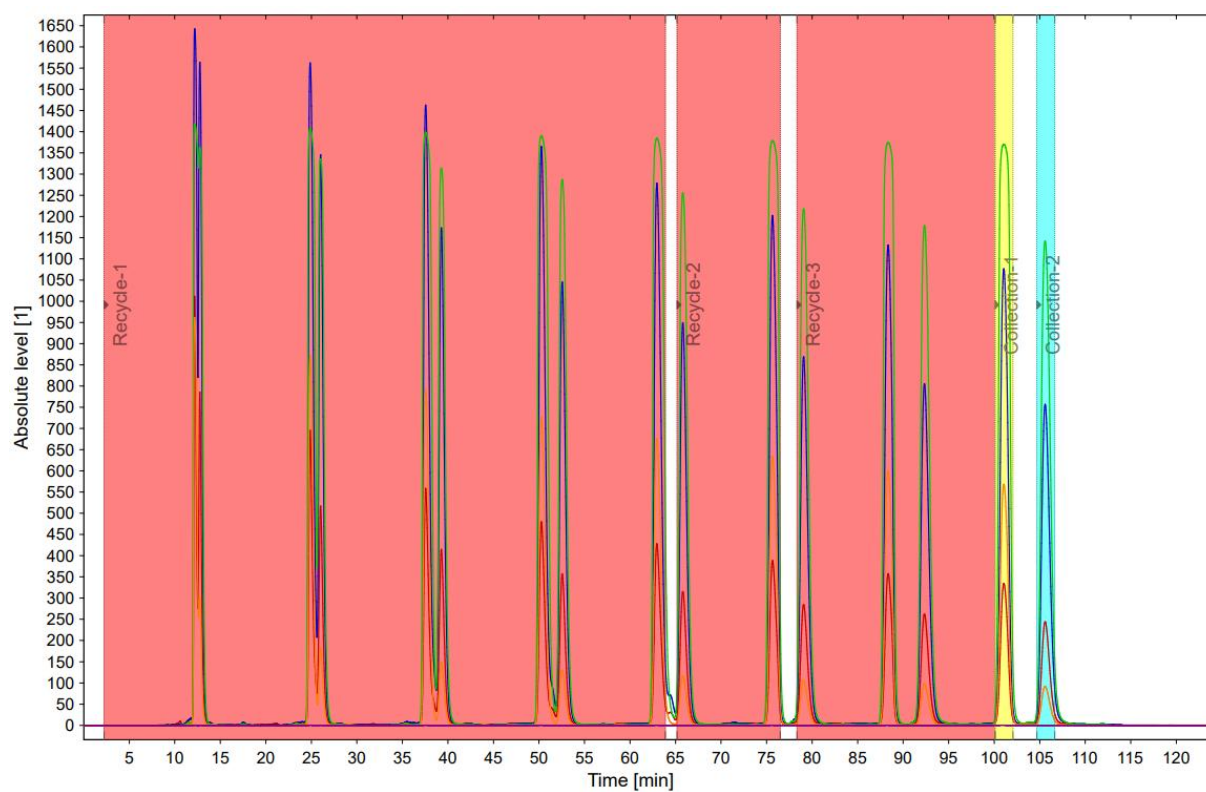
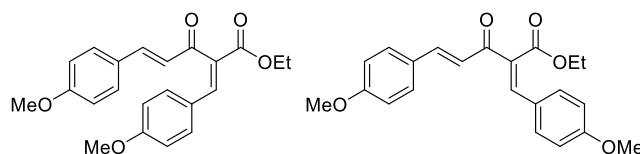


Figure S29. Chromatogram for purification of **19f** and **19g** (fraction 1 and 2, respectively, RP–HPLC/MeOH).



Compound 19h (mixture of ethyl (*E*)-2-((*Z*)-4-methoxybenzylidene)-5-(4-methoxyphenyl)-3-oxopent-4-enoate and ethyl (*E*)-2-((*E*)-4-methoxybenzylidene)-5-(4-methoxyphenyl)-3-oxopent-4-enoate).

To a solution of ethyl acetoacetate (1.15 mL, 9.01 mmol) and *p*-methoxybenzaldehyde (4.38 mL, 36.0 mmol) dissolved in benzene (20 mL), piperidine (0.36 mL, 3.6 mmol) and acetic acid (0.31 mL, 5.4 mmol) were added. The resulting solution was stirred and heated at 150 °C in the presence of the Dean–Stark apparatus. The reaction was monitored by ¹H NMR spectroscopy. At 4 hours, the reaction mixture contained mostly singly condensed products (**19f** and **19g**) and unreacted *p*-methoxybenzaldehyde. At 18 hours, the reaction mixture contained much smaller amounts of the singly condensed products. Then, the whole mixture was directly purified by column chromatography on silica (eluent: hexane:EtOAc = 5:1 gradient to 2:1). The early fractions contained benzene, *p*-methoxybenzaldehyde and singly condensed products (**19f** and **19g**). The desired products later eluted as yellow fractions. Upon solvent removal, a mixture of major (*E,Z*)-isomer (**19h**) and minor (*E,E*)-isomer was obtained as a brownish yellow paste (1.78 g, 56%). The mixture was subjected to HPLC purification to give the (*E,Z*)-isomer (pure) and the (*E,E*)-isomer (enriched). As shown by ¹H NMR spectroscopy, the (*E,Z*)-isomer (**19h**) is configurationally stable, while the (*E,E*)-isomer was found to slowly isomerize in solution.

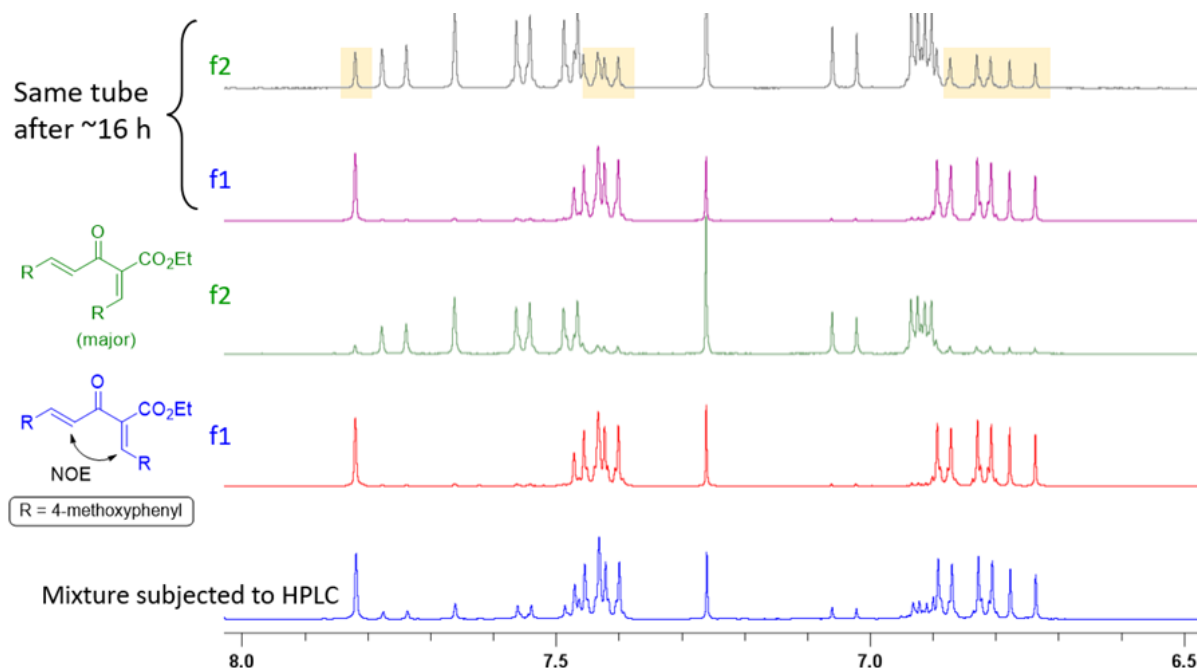
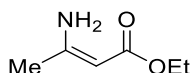


Figure S30. ¹H NMR spectra for comparing the purities from Hantzsch and independent synthesis.

¹H NMR (400 MHz, chloroform-*d*): δ = 7.76 (d, J = 15.5 Hz, 1H), 7.66 (s, 1H), 7.54–7.57 (m, 2H), 7.46–7.49 (m, 2H), 7.04 (d, J = 15.5 Hz, 1H), 6.89–6.94 (m, 4H), 4.40 (q, J = 7.1 Hz, 2H), 3.85 (s, 6H), 1.33 (t, J = 7.1 Hz, 3H). **¹³C NMR** (100 MHz, chloroform-*d*): δ = 186.24, 168.78, 161.89, 161.80, 144.58, 141.05, 132.73, 131.98, 130.46, 127.65, 125.98, 119.45, 114.56, 114.45, 61.79, 55.56, 14.18. **HRMS** (ESI–TOF): m/z : $[M + H]^+$ Calcd for C₂₂H₂₃O₅⁺: 367.1540; Found 367.1536.



Compound EAB (ethyl (*Z*)-3-aminobut-2-enoate) was synthesized according to the literature procedure.⁸⁴

In a round bottom flask, ethyl acetoacetate (200.00 mg, 1.51 mmol) and ammonium acetate (201.38 mg, 2.56 mmol) were dissolved in 1 mL of ethanol. The mixture was stirred at room temperature for 12 h. The reaction mixture was then diluted with water and extracted with dichloromethane. The organic phase was dried over anhydrous sodium sulfate and then the solvent was removed under reduced pressure. The crude mixture was passed via silica column chromatography to obtain pure product (113 mg, 58%) as a colorless oil.

¹H NMR (400 MHz, DMSO-*d*₆): δ = 6.85 (br, 1H), 6.09 (br, 1H), 3.43 (s, 1H), 3.10 (m, 2H), 0.95 (s, 3H), 0.28 (t, J = 7.2 Hz, 3H). **¹³C NMR** (100 MHz, DMSO-*d*₆): δ = 169.1, 161.2, 81.4, 57.4, 21.5, 14.6. **HRMS** (ESI–TOF): m/z : $[M + H]^+$ Calcd for C₆H₁₂NO₂⁺: 130.0863; Found 130.0854.

4.2. Compounds synthesized using the automated platform

4.2.1. S_N1 reaction

Reaction yield map for S_N1 (synthesis of 9-bromo-9-phenyl-9*H*-fluorene). The stock solutions of 9-phenyl-9*H*-fluoren-9-ol in 1,4-dioxane (0.395 M) and hydrobromic acid (48% in water) in acetic acid (1.500 M) were prepared, and then they were used by automatic liquid handler to prepare 6 sets of 155 reaction conditions from all combinations of 5 initial concentrations of 9-phenyl-9*H*-fluoren-9-ol (0.0300, 0.0975, 0.1650, 0.2325 and 0.3000 M) and 31 initial concentrations of hydrobromic acid (0.0000, 0.0300, 0.0393, 0.0486, 0.0579, 0.0672, 0.0765, 0.0859, 0.0952, 0.1044, 0.1138, 0.1231, 0.1324, 0.1417, 0.1510, 0.1603, 0.1697, 0.1790, 0.1883, 0.1976, 0.2069, 0.2162, 0.2255, 0.2348, 0.2441, 0.2534, 0.2628, 0.2721, 0.2814, 0.2907 and 0.3000 M). The order of compounds addition was as follows: 1,4-dioxane, solution of 9-phenyl-9*H*-fluoren-9-ol, solution of hydrobromic acid. The total volume of every reaction was 500 μ L. After sealing, every set of reactions was shaken for 5 min (250 rpm) and placed in a thermostatic oven (temperatures: 16, 26, 36, 46, 56, 66 $^{\circ}$ C) for 48 h. After that time, every reaction mixture was diluted 100 times and UV–Vis spectrum was recorded. Reaction mixtures from every set were combined, extracted with three portions of ethyl acetate, dried over anhydrous sodium sulfate and evaporated under reduced pressure.

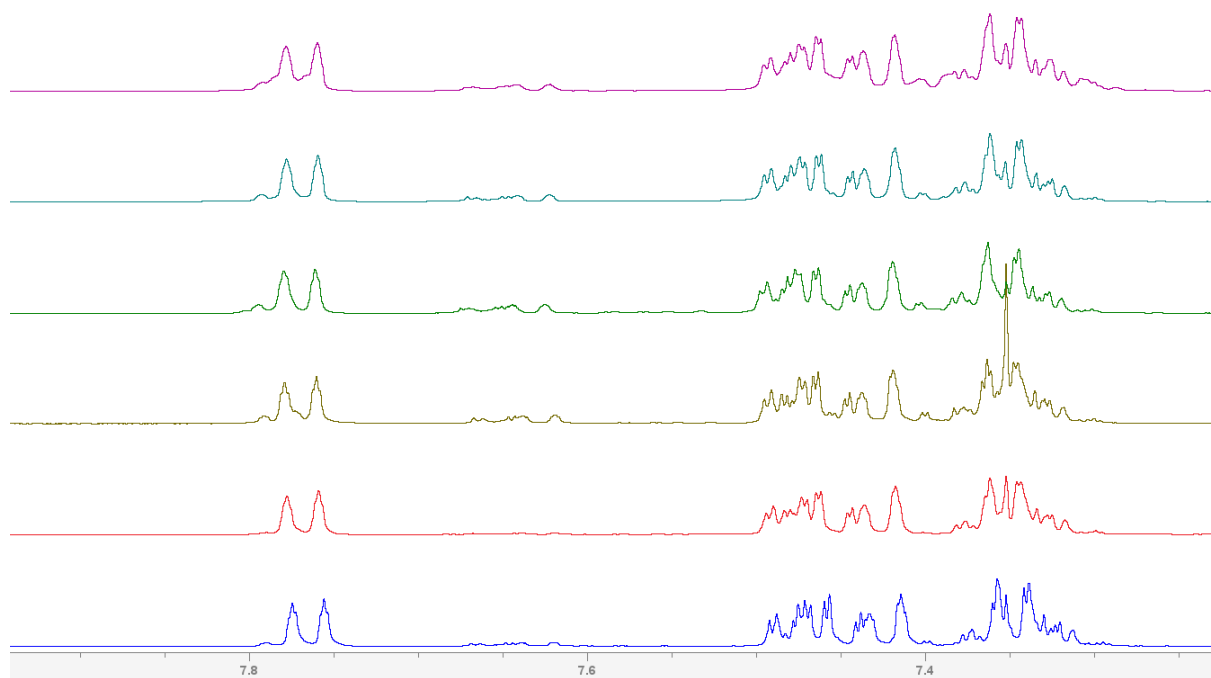


Figure S31. Stacked ¹H NMR spectra (aromatic region, chloroform-*d*) of combined reaction mixtures from every temperature (from bottom to top temperatures: 16, 26, 36, 46, 56, 66 $^{\circ}$ C).

4.2.2. E1 reaction

Reaction yield map for E1 (synthesis of 9-butylidene-9*H*-fluorene). The stock solutions of 9-butyl-9*H*-fluoren-9-ol in acetonitrile (19.7 mM) and a solution of hydrobromic acid (48% in water) in acetonitrile (50 mM, 5 mM, 0.5 mM) were prepared, and then they were used by automatic liquid handler to prepare 5 sets of 155 reaction conditions from all combinations of 5 initial concentrations of 9-butyl-9*H*-fluoren-9-ol (1.5, 4.9, 8.3, 11.6, 15.0 mM) and 31 initial concentrations of hydrobromic acid (0.0, 0.1, 0.4, 0.8, 1.1, 1.5, 1.8, 2.1, 2.5, 2.8, 3.2, 3.5, 3.9, 4.2, 4.5, 4.9, 5.2, 5.6, 5.9, 6.2, 6.6, 6.9, 7.3, 7.6, 8.0, 8.3, 8.6, 9.0, 9.3, 9.7 and 10.0 mM). The order of compounds addition was as follows: acetonitrile, solution of 9-butyl-9*H*-fluoren-9-ol, solution of hydrobromic acid. The total volume of every reaction was 500 μ L. After sealing, every set of reactions was shaken for 5 min (250 rpm) and placed in a thermostatic oven (temperatures: 16, 21, 26, 31, 36 $^{\circ}$ C) for 4 h. After that time, every reaction mixture was diluted 30 times and UV–Vis spectrum was recorded.

4.2.3. Rearrangement of anthracen-9-ylidiphenylmethanol

Reaction yield map for rearrangement of anthracen-9-ylidiphenylmethanol (synthesis of 10-(diphenylmethylene)-4*a*,9,9*a*,10-tetrahydroanthracen-9-ol). The stock solutions of anthracen-9-ylidiphenylmethanol in DMF (30 mM, 0.3 mM) and the stock solutions of hydrobromic acid (48% in water) in DMF ($3.4 \cdot 10^{-1}$ M, $3.4 \cdot 10^{-2}$ M, $3.4 \cdot 10^{-5}$ M) were prepared, and then they were used by automatic liquid handler to prepare 7 sets of 155 reaction conditions from all combinations of 5 initial concentrations of anthracen-9-ylidiphenylmethanol (0.02, 5.7, 11.4, 17.2, 22.9 mM) and 31 initial concentrations of hydrobromic acid (0.0000, 0.0229, 2.4, 4.7, 7.1, 9.5, 11.8, 14.2, 16.6, 18.9, 21.3, 23.6, 26.0, 28.4, 30.7, 33.1, 35.5, 37.8, 40.2, 42.6, 44.9, 47.3, 49.7, 52.0, 54.4, 56.8, 59.1, 61.5, 63.9, 66.2 and 68.6 mM). The order of compounds' addition was as follows: DMF, solution of anthracen-9-ylidiphenylmethanol, solution of hydrobromic acid. The total volume of every reaction was 500 μ L. After sealing, every set of reactions was shaken for 5 min (250 rpm) and placed in a thermostatic oven (temperatures: 6, 16, 26, 36, 46, 56, 66 $^{\circ}$ C) for 4 h. After that time, every reaction mixture was diluted 100 times and UV–Vis spectrum was recorded. Reaction mixtures were combined, extracted with three portions of dichloromethane, dried over anhydrous sodium sulfate and evaporated under reduced pressure. The pure product was obtained after flash chromatography (SiO₂, gradient from hexane:dichloromethane (1:1) to 7% MeOH in dichloromethane). Compounds **15b**, and **15d** were isolated. X-ray structure was determined for compound **15b** (Section 4.16).

Compound 15b (10-(diphenylmethylene)-4*a*,9,9*a*,10-tetrahydroanthracen-9-ol).

The molecular structure of **15b** was unambiguously confirmed by X-ray crystallography (Section 4.16).

¹H NMR (400 MHz, dichloromethane-*d*₂): δ = 7.60 (d, *J* = 7.6 Hz, 2H), 7.31 (d, *J* = 7.2 Hz, 4H), 7.25 (t, *J* = 7.6 Hz, 4H), 7.18 (m, 4H), 7.06 (d, *J* = 7.7 Hz, 2H), 6.91 (t, *J* = 7.6 Hz, 2H),

5.73 (d, $J = 6.7$ Hz, 1H), 2.51 (d, $J = 6.7$ Hz, 1H). ^{13}C NMR (100 MHz, dichloromethane- d_2): $\delta = 143.1, 141.1, 140.8, 133.9, 134.0, 129.8, 129.1, 128.6, 127.1, 126.8, 126.4, 124.3, 71.3$. HRMS (ESI-TOF): m/z : $[\text{M} + \text{Na}]^+$ Calcd for $\text{C}_{27}\text{H}_{20}\text{ONa}^+$: 383.1407; Found 383.1402; m/z : $[\text{M} + \text{H} - \text{H}_2\text{O}]^+$ Calcd for $\text{C}_{27}\text{H}_{19}^+$: 343.1482; Found 343.1485.

Compound 15d (anthraquinone)

^1H NMR (400 MHz, DMSO- d_6): $\delta = 8.21\text{--}8.24$ (m, 4H), $7.93\text{--}7.96$ (m, 4H).

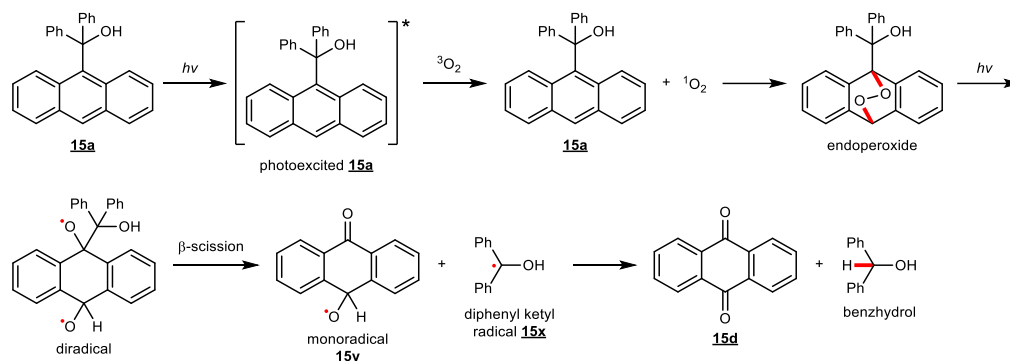


Figure S32. Mechanistic pathway from alcohol **15a** to the anthraquinone side product **15d** that was detected by UV-Vis spectroscopy, isolated and confirmed by ^1H NMR spectroscopy^{25,26}. In the accepted mechanism, the photoexcited state of compound **15a** can be quenched by triplet-state oxygen ($^3\text{O}_2$). The resultant ground-state **15a** and singlet oxygen ($^1\text{O}_2$) can undergo cycloaddition to yield an endoperoxide intermediate, that affords a diradical upon O–O bond cleavage. The diradical undergoes β -scission to furnish a monoradical **15y** and the diphenyl ketyl radical **15x**. The former monoradical **15y** loses a hydrogen atom to radical **15x** (or other non-radical quenchers) to yield the observed anthraquinone **15d**, with **15x** being converted into benzhydrol.

4.2.4. Reaction yield map for multicomponent reaction (synthesis of 3-butyl-4-(4-nitrophenyl)-1-(tosylmethyl)-1*H*-imidazol-3-ium-5-olate)

Following abbreviations will be used: TosMIC – *p*-toluenesulfonylmethyl isocyanide, *p*TSA – *p*-toluenesulfonic acid monohydrate.

The stock solutions of TosMIC, *n*-butylamine, *p*-nitrobenzaldehyde and *p*TSA in DMF were prepared, and then they were used by automatic liquid handler to prepare all combinations of reaction conditions from adaptive sampling: TosMIC (7 concentrations in a range 120–300 mM), *n*-butylamine (7 concentrations in a range 120–300 mM), *p*-nitrobenzaldehyde (7 concentrations in a range 120–300 mM), *p*TSA (21 concentrations in a range 0–300 mM). The order of compounds addition was as follows: DMF, solution of *p*-nitrobenzaldehyde, solution of *p*TSA, solution of *n*-butylamine and solution of TosMIC. The total volume of every reaction was 200 μL . After sealing, every set of reactions was shaken for 5 min (250 rpm) and placed in

a thermostatic oven (temperature: 26 °C) for 16 h. After that time, every reaction mixture was diluted 200 times and UV–Vis spectrum was recorded. Reaction mixtures were combined, extracted with three portions of dichloromethane, dried over anhydrous sodium sulfate and evaporated under reduced pressure. The pure product was obtained after purification on RP–HPLC.

Compound 16e (3-butyl-4-(4-nitrophenyl)-1-(tosylmethyl)-1*H*-imidazol-3-ium-5-olate).

¹H NMR (400 MHz, DMSO-*d*₆): δ = 8.41 (s, 1H), 8.10 (d, *J* = 9.3 Hz, 2H), 7.80 (d, *J* = 9.3 Hz, 2H), 7.66 (d, *J* = 8.3 Hz, 2H), 7.42 Hz (d, *J* = 8.3 Hz, 2H), 5.46 (s, 2H), 4.38 (t, *J* = 7.0 Hz, 2H), 2.39 (s, 3H), 1.56 (quintet, *J* = 7.4 Hz, 2H), 1.20 (sextet, *J* = 7.4 Hz, 2H), 0.85 (t, *J* = 7.5 Hz, 3H). **¹³C NMR** (100 MHz, DMSO-*d*₆): δ = 153.8, 145.6, 141.5, 138.4, 133.7, 130.1, 128.5, 127.8, 124.0, 122.4, 100.4, 59.9, 49.5, 30.5, 21.2, 18.6, 13.4. **HRMS** (ESI–TOF): *m/z*: [M + H]⁺ Calcd for C₂₁H₂₄N₃O₅S⁺: 430.1432; Found 430.1410.

4.2.5. Reaction yield map for Hantzsch reaction

The stock solutions of *p*-methoxybenzaldehyde (0.333 M), ethyl acetoacetate (0.333 M) and ammonium acetate (1.333 M, 0.133 M, 0.013 M) were prepared, and then they were used by automatic liquid handler to prepare 480 reaction conditions from all combinations of 4 initial concentrations of *p*-methoxybenzaldehyde (0.01, 0.04, 0.07 and 0.1 M), 4 initial concentrations of ethyl acetoacetate (0.01, 0.04, 0.07 and 0.1 M) and 30 initial concentrations of ammonium acetate (0.0010 M, 0.0148 M, 0.0285 M, 0.0423 M, 0.0560 M, 0.0698 M, 0.0836 M, 0.0973 M, 0.1111 M, 0.1248 M, 0.1386 M, 0.1523 M, 0.1661 M, 0.1799 M, 0.1936 M, 0.2074 M, 0.2211 M, 0.2349 M, 0.2487 M, 0.2624 M, 0.2762 M, 0.2899 M, 0.3037 M, 0.3174 M, 0.3312 M, 0.3450 M, 0.3587 M, 0.3724 M, 0.3862 M, 0.4000 M). The order of compounds addition was as follows: ethanol, solution of *p*-methoxybenzaldehyde, solution of ethyl acetoacetate and solution of ammonium acetate. The total volume of every reaction was 500 μ L. After sealing, every set of reactions was shaken for 5 min (250 rpm) and placed in a thermostatic oven (temperatures: 80 °C and 26 °C) for 48 h. After that time, every reaction mixture was diluted 20 and 200 times and UV–Vis spectrum was recorded. This procedure was triplicate for reaction in 80 °C and duplicate for reaction in 26 °C, in total 2440 datapoints from UV-Vis spectrophotometer. Crude reaction mixtures from every temperature were combined, evaporated under reduced pressure and separated by recycling HPLC (see **Figure S99** for an overview of the stepwise purification process). The structures of all isolated compounds are shown in **Table S7**, and the chromatograms are shown from **Figure S33** to **Figure S50**.

Table S7. Compounds isolated from crude Hantzsch reaction mixture.

Code	Systematic name	Structure
<u>19e</u>	Ethyl 4-amino-2,6-bis(4-methoxyphenyl)-1,2,5,6-tetrahydropyridine-3-carboxylate	
<u>19f</u>	Ethyl (Z)-2-(4-methoxybenzylidene)-3-oxobutanoate	
<u>19g</u>	Ethyl (E)-2-(4-methoxybenzylidene)-3-oxobutanoate	
<u>19h</u>	Ethyl (E)-2-((Z)-4-methoxybenzylidene)-5-(4-methoxyphenyl)-3-oxopent-4-enoate	
<u>19i</u>	Ethyl (Z)-6-(2-ethoxy-2-oxoethylidene)-4-(4-methoxyphenyl)-2-methyl-1,4,5,6-tetrahydropyridine-3-carboxylate	
<u>19k</u>	Diethyl (E)-4-(4-methoxyphenyl)-2-(4-methoxystyryl)-6-methyl-1,4-dihydropyridine-3,5-dicarboxylate	
<u>19l</u>	Diethyl (Z)-4-(4-methoxyphenyl)-2-(4-methoxystyryl)-6-methyl-1,4-dihydropyridine-3,5-dicarboxylate	

<u>19m</u>	Diethyl (<i>E</i>)-4-(4-methoxyphenyl)-2-(4-methoxystyryl)-6-methylpyridine-3,5-dicarboxylate	
<u>19n</u>	Diethyl (<i>Z</i>)-4-(4-methoxyphenyl)-2-(4-methoxystyryl)-6-methylpyridine-3,5-dicarboxylate	
<u>19o</u>	Ethyl 4-(4-methoxyphenyl)-2,6-dimethylnicotinate	
<u>19p</u>	Diethyl 2-(4-methoxyphenyl)-4,6-dimethyl-1,2-dihydropyridine-3,5-dicarboxylate	
<u>19r</u>	Diethyl 2-(4-methoxyphenyl)-4,6-dimethylpyridine-3,5-dicarboxylate	

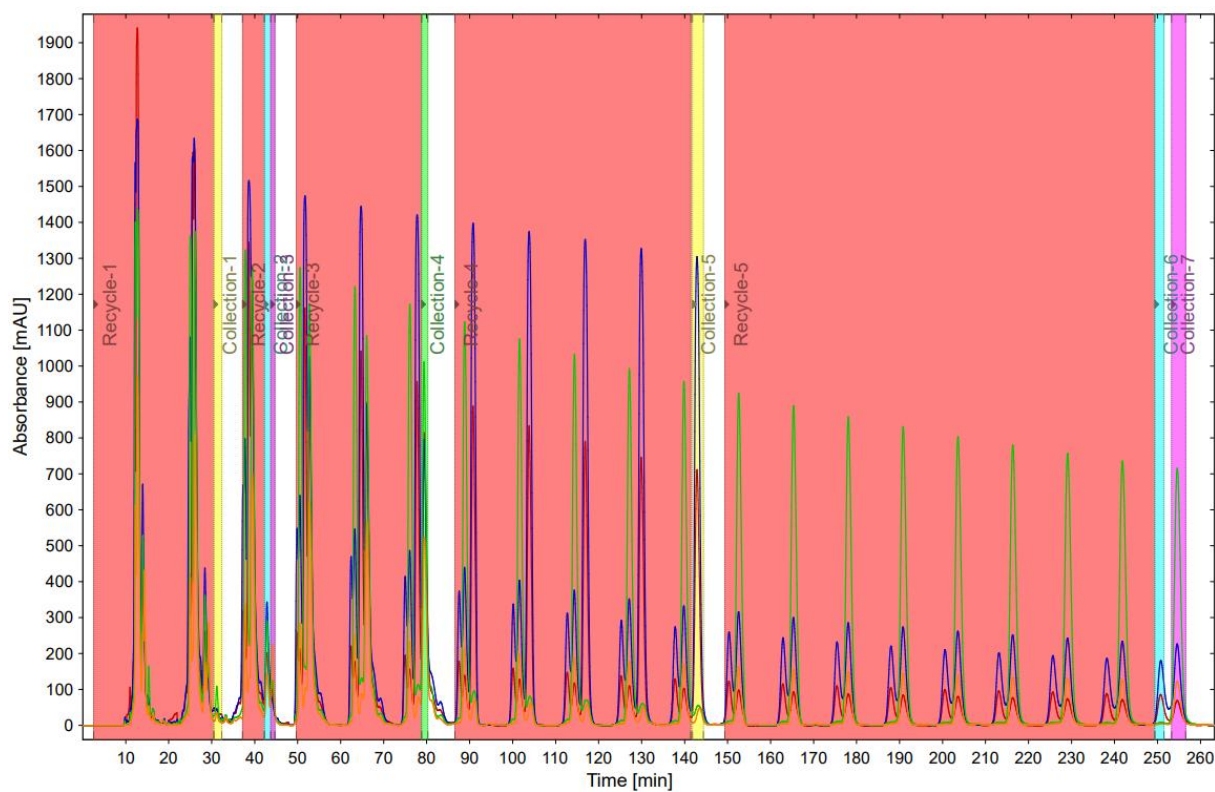


Figure S33. Chromatogram for initial separation of combined all crude reaction mixtures at 80 °C (RP-HPLC/MeOH).

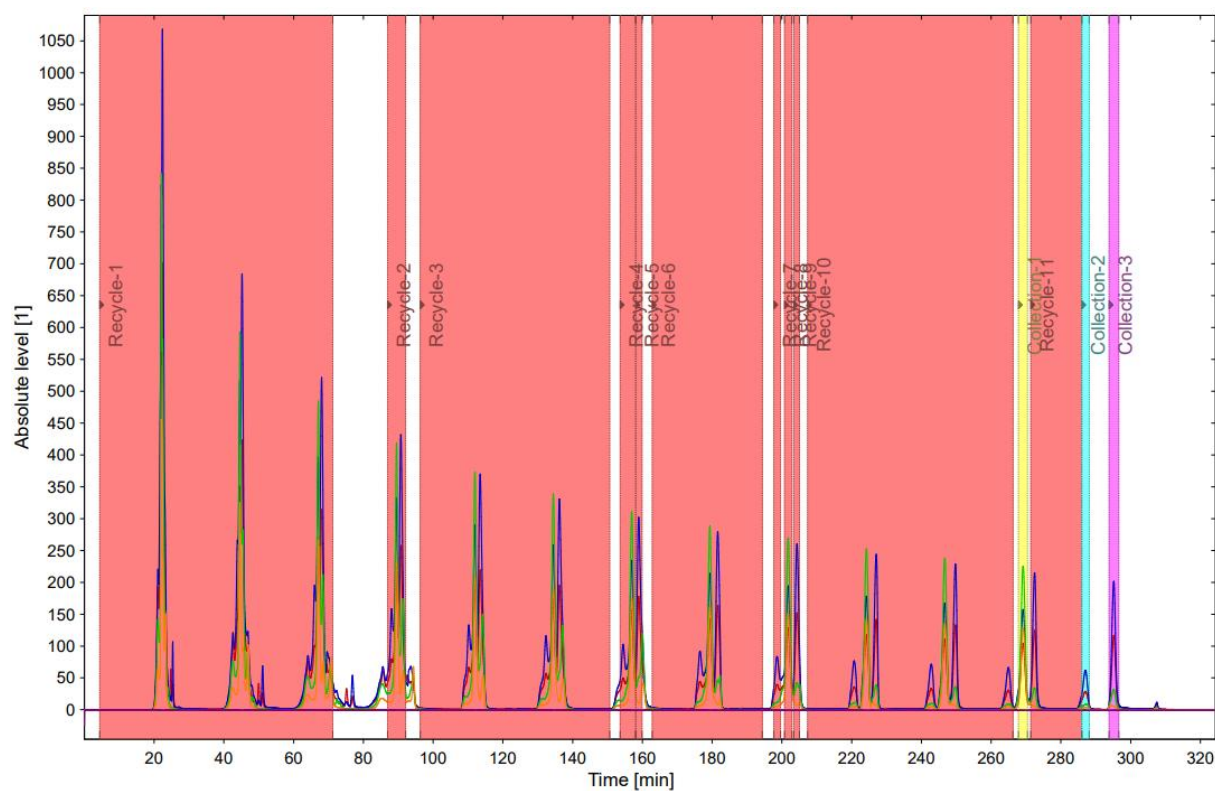


Figure S34. Chromatogram for separation of **fraction 2** from initial separation (*cf* **Figure S33**, GPC/chloroform). **Fractions 1** and **3** contain **19k** and **19p**, respectively.

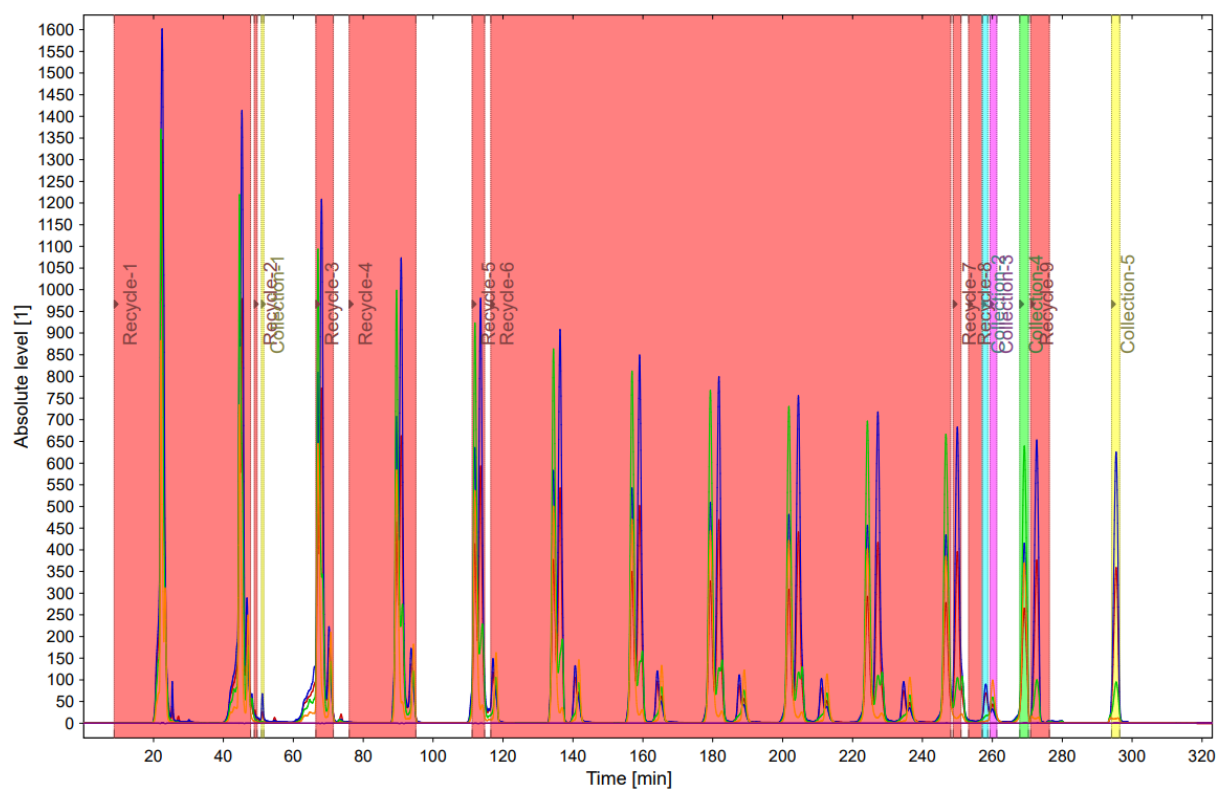


Figure S35. Chromatogram for separation of **fraction 3** from initial separation (*cf.* **Figure S33**, GPC/chloroform). **Fractions 2, 4 and 5** contain **19o**, **19k** and **19p**, respectively.

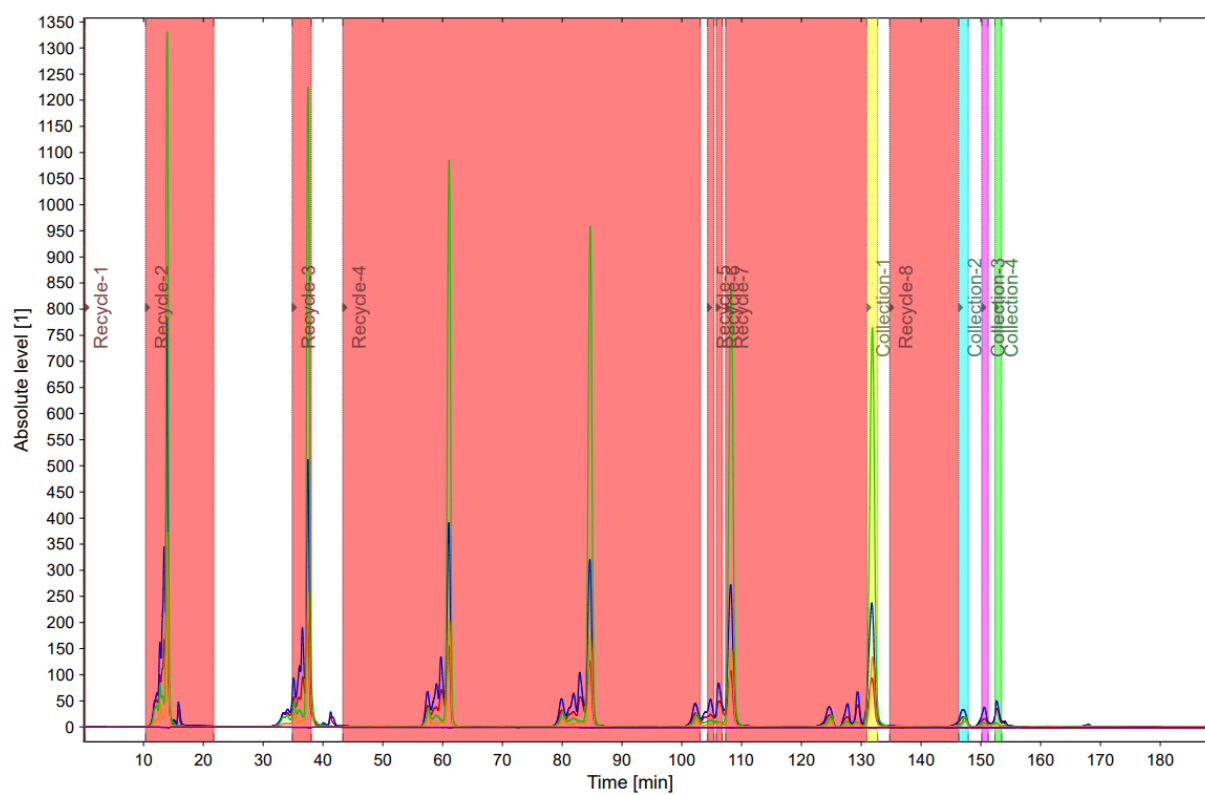


Figure S36. Chromatogram for separation of **fraction 6** from initial separation (*cf.* **Figure S33**, GPC/chloroform). **Fraction 1** contains mixture of **19f** and **19g**.

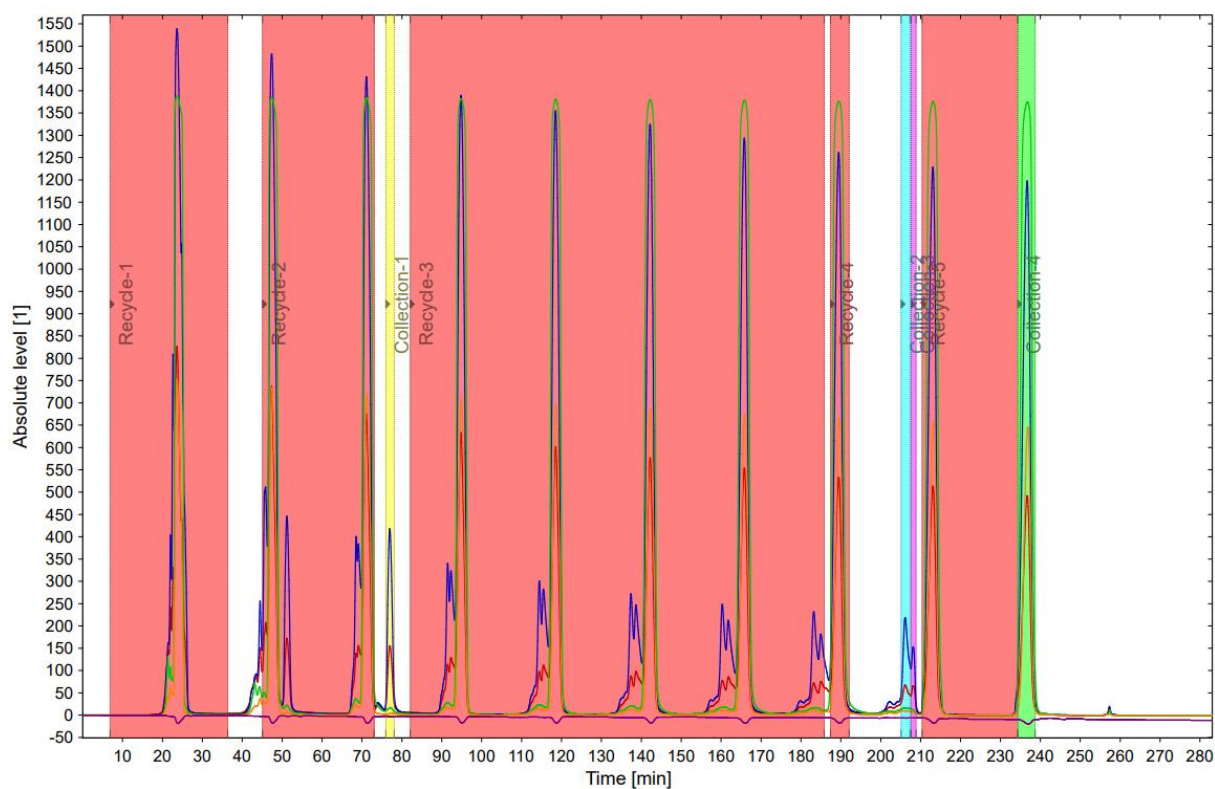


Figure S37. Chromatogram for separation of **fraction 7** from initial separation (*cf.* **Figure S33**, GPC/chloroform). **Fraction 4** contains mixture of **19f** and **19g**.

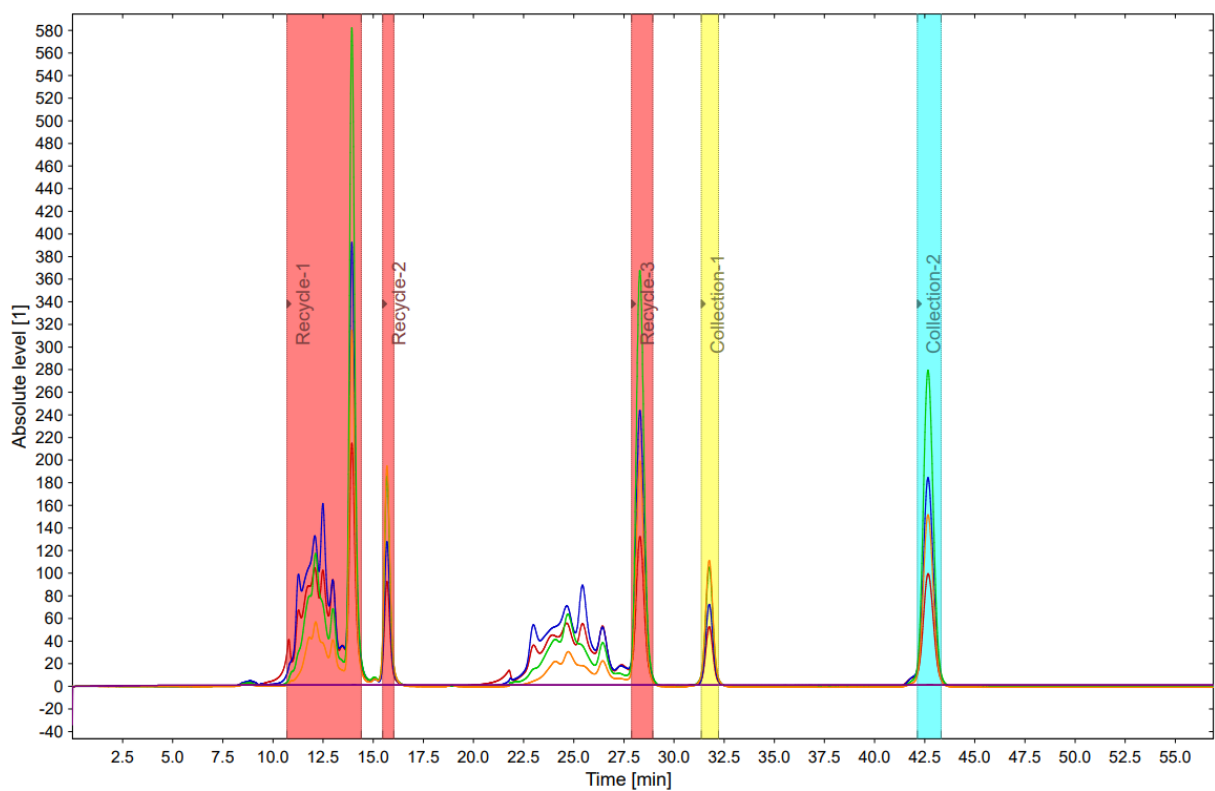
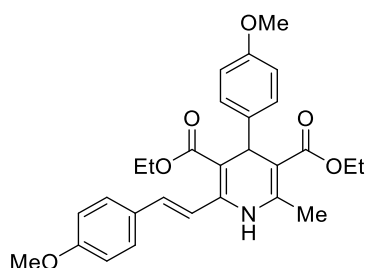


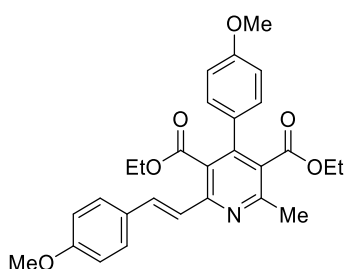
Figure S38. Chromatogram for separation of **fraction 4** from **Figure S35** to give **19m** (fraction 1) and **19k** (fraction 2) by RP-HPLC/MeOH.



Compound 19k (diethyl (*E*)-4-(4-methoxyphenyl)-2-(4-methoxystyryl)-6-methyl-1,4-dihydropyridine-3,5-dicarboxylate).

¹H NMR (400 MHz, chloroform-*d*): δ 7.91 (d, J = 16.8 Hz, 1H), 7.44 (d, J = 8.8 Hz, 2H), 7.23 (d, J = 8.7 Hz, 2H), 6.88 (d, J = 8.8 Hz, 2H), 6.71–6.83 (m, 3H), 6.12 (s, 1H), 5.03 (s, 1H), 4.05–4.20 (m, 4H), 3.82 (s, 3H), 3.75 (s, 3H), 2.42 (s, 3H), 1.25 (m, 6H). **¹³C NMR** (100 MHz, chloroform-*d*): δ 167.72, 167.54, 160.4, 158.13, 143.97, 141.69, 140.1, 130.56, 129.14, 128.78, 128.67, 120.66, 114.41, 113.42, 105.84, 104.31, 60.17, 59.88, 55.46, 55.27, 39.17, 19.99, 14.44.

HRMS (ESI–TOF): m/z : $[M + H]^+$ Calcd for C₂₈H₃₂NO₆⁺: 476.2223; Found 476.2222.



Compound 19m (diethyl (*E*)-4-(4-methoxyphenyl)-2-(4-methoxystyryl)-6-methylpyridine-3,5-dicarboxylate).

¹H NMR (400 MHz, chloroform-*d*): δ 7.91 (d, J = 15.5 Hz, 1H), 7.52 (d, J = 8.8 Hz, 2H), 7.21 (d, J = 8.7 Hz, 2H), 7.05 (d, J = 15.5 Hz, 1H), 6.90 (m, 4H), 4.01–4.12 (m, 4H), 3.84 (s, 3H), 3.83 (s, 3H), 2.65 (s, 3H), 0.99 (m, 6H). **¹³C NMR** (100 MHz, chloroform-*d*): δ 168.29, 168.16, 160.36, 159.93, 155.79, 151.93, 146.35, 136.24, 129.62, 129.45, 129.15, 129.07, 127.47, 126.21, 121.46, 114.28, 113.71, 61.61, 61.45, 55.48, 23.4, 13.94, 13.89. **HRMS** (ESI–TOF): m/z : $[M + H]^+$ Calcd for C₂₈H₃₀NO₆⁺: 476.2068; Found 476.2070.

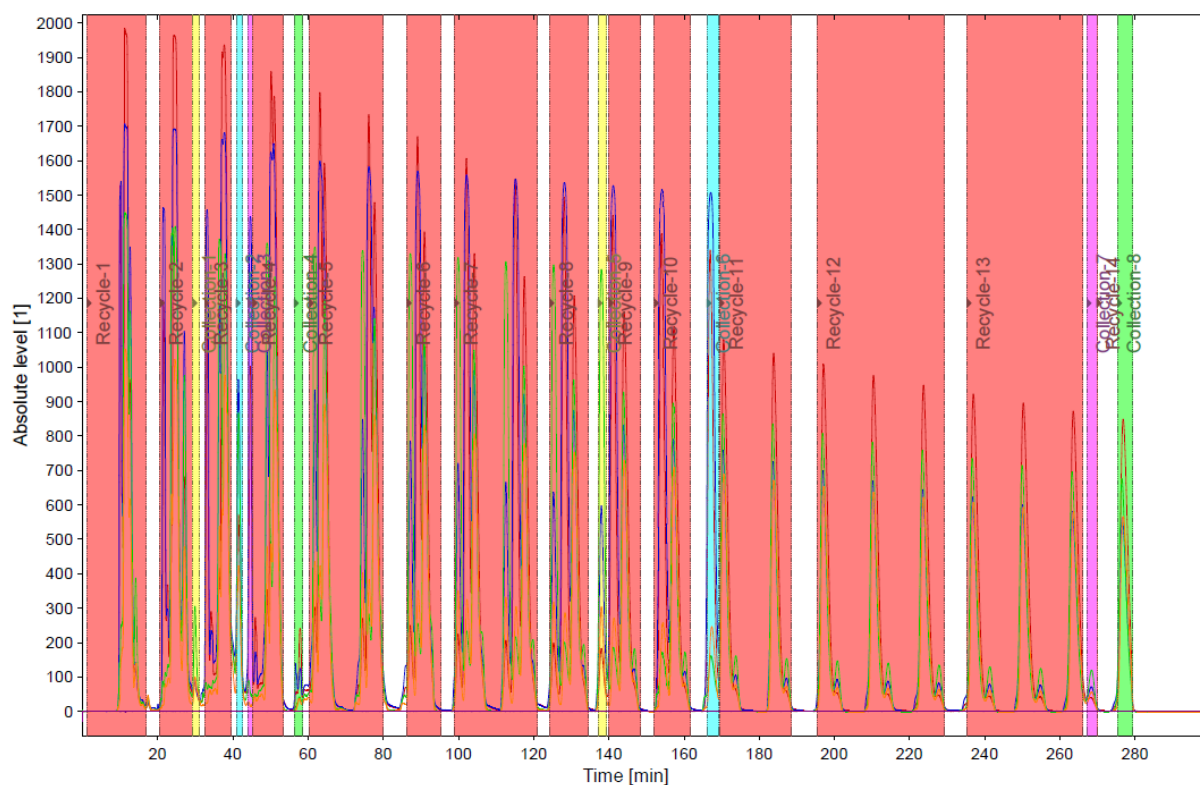


Figure S39. Chromatogram for initial separation of combined all crude reaction mixtures from a second run of the same conditions as in **Figure S35** at 80 °C (RP–HPLC/MeOH).

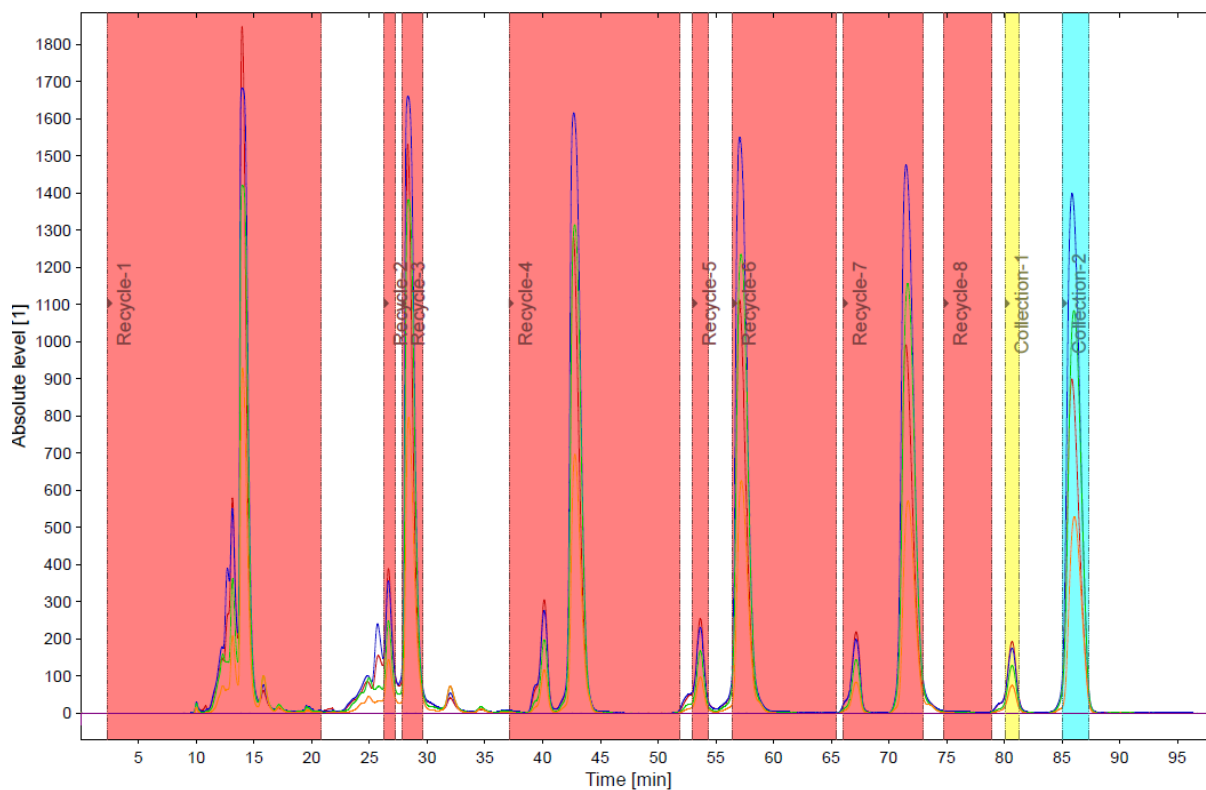


Figure S40. Chromatogram for separation of **fraction 2** from initial separation (*cf.* **Figure S39**, GPC/chloroform). **Fraction 1** contains **191**.

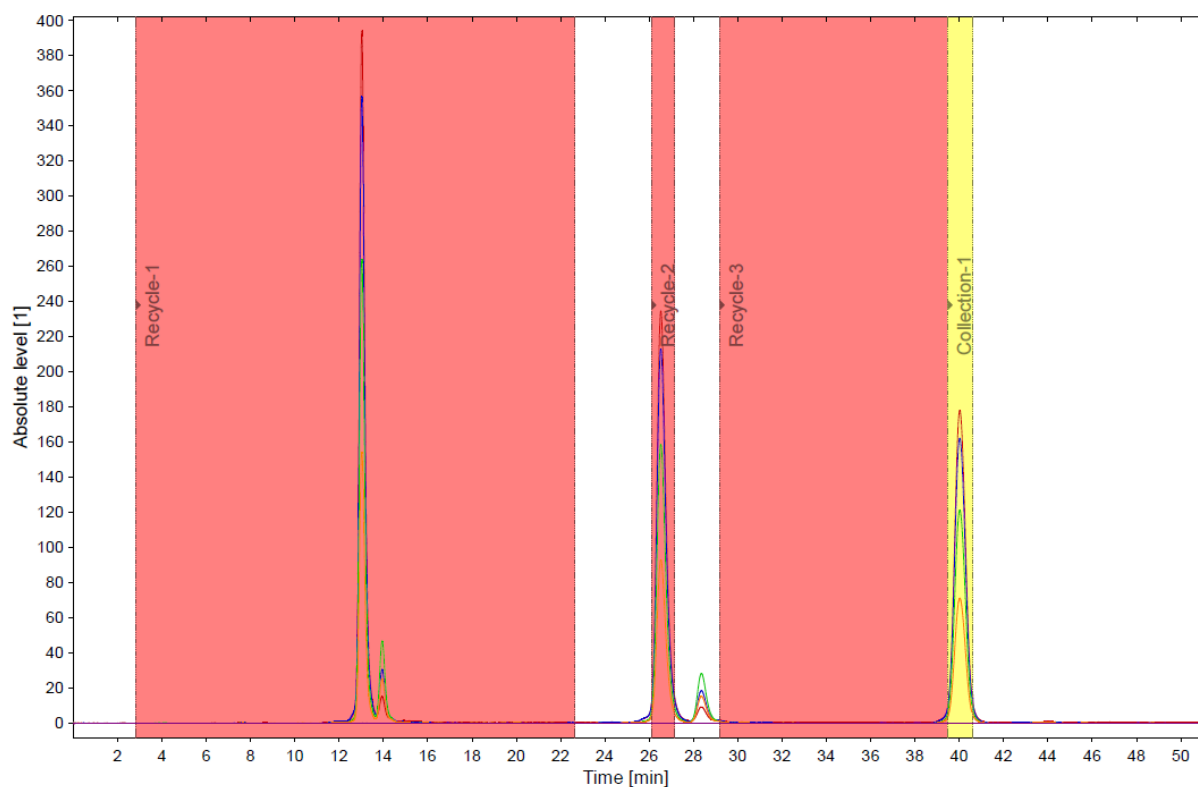
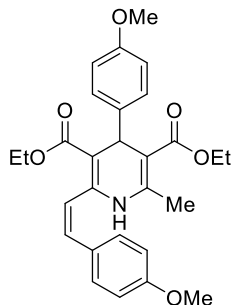


Figure S41. Final purification of **191** (fraction 1) (RP-HPLC/MeOH).



Compound 191 (diethyl (Z)-4-(4-methoxyphenyl)-2-(4-methoxystyryl)-6-methyl-1,4-dihydropyridine-3,5-dicarboxylate).

¹H NMR (400 MHz, chloroform-*d*): 7.22 (m, 4H), 6.85 (d, *J* = 8.8 Hz, 2H), 6.72–6.80 (m, 3H), 6.66 (d, *J* = 12.1 Hz, 1H), 5.54 (s, 1H), 5.04 (s, 1H), 4.04–4.21 (m, 4H), 3.82 (s, 3H), 3.77 (s, 3H), 1.99 (s, 3H), 1.20–1.30 (m, 6H). **¹³C NMR** (100 MHz, chloroform-*d*): δ 167.72, 167.21, 159.79, 158.18, 144.76, 142.42, 139.81, 132.13, 130.41, 129.13, 127.96, 124.19, 114.07, 113.45, 106.08, 103.76, 60.17, 59.88, 55.48, 55.31, 38.59, 19.22, 14.43. **HRMS** (ESI-TOF): *m/z*: [M – H][–] Calcd for C₂₈H₃₀NO₆[–]: 476.2078; Found 476.2067.

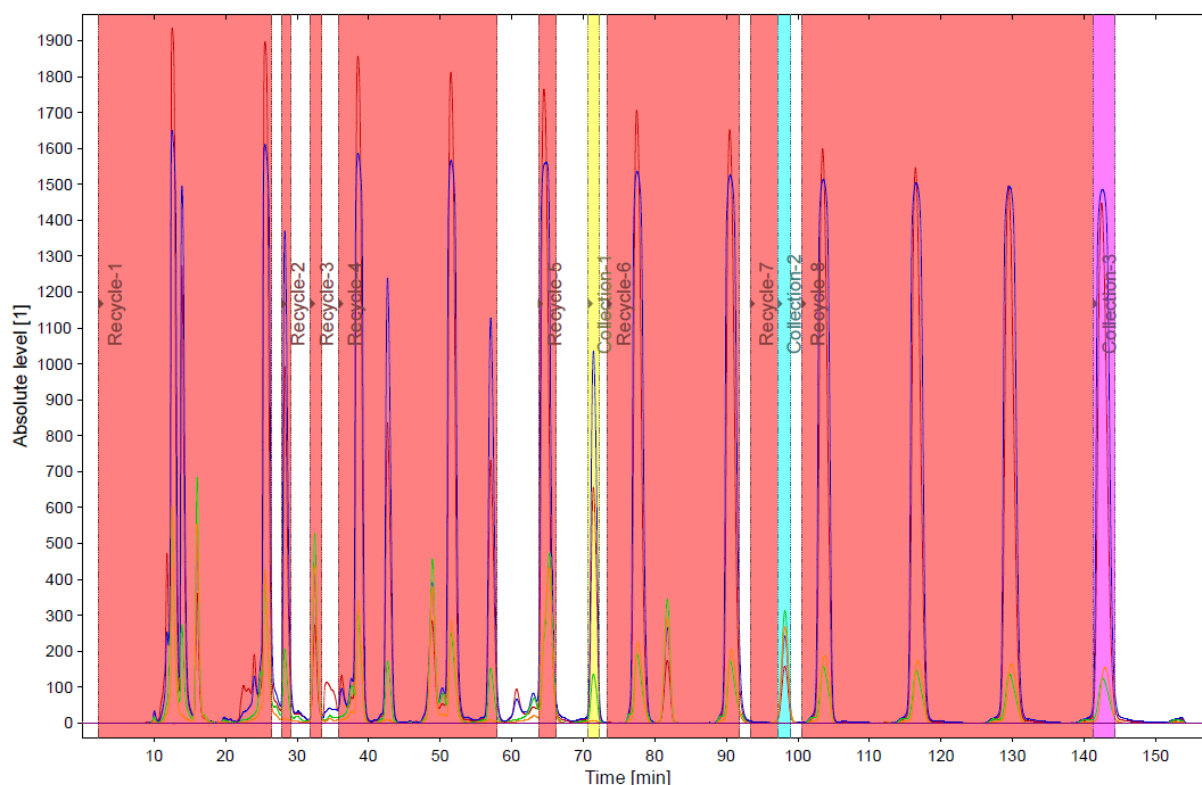
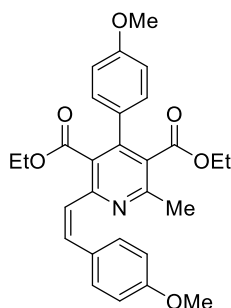


Figure S42. Chromatogram for separation of **fraction 6** from initial separation (*cf.* **Figure S39**, GPC/chloroform). **Fraction 1** contains **19n**.



Compound 19n (diethyl (Z)-4-(4-methoxyphenyl)-2-(4-methoxystyryl)-6-methylpyridine-3,5-dicarboxylate).

¹H NMR (400 MHz, chloroform-*d*): 7.93 (d, *J* = 15.5 Hz, 1H), 7.64 (d, *J* = 8.9 Hz, 2H), 7.50 (d, *J* = 8.7 Hz, 2H), 7.01 (d, *J* = 15.5 Hz, 1H), 6.97 (d, *J* = 8.9 Hz, 2H), 6.90 (d, *J* = 8.8 Hz, 2H), 4.50 (q, *J* = 7.1 Hz, 2H), 4.17 (q, *J* = 7.2 Hz, 2H), 3.86 (s, 3H), 3.83 (s, 3H), 2.36 (s, 3H), 1.45 (t, *J* = 7.1 Hz, 3H), 1.09 (t, *J* = 7.1 Hz, 3H). **¹³C NMR** (100 MHz, chloroform-*d*): δ 168.93, 168.56, 160.54, 160.35, 156.10, 151.80, 143.27, 136.21, 132.60, 130.04, 129.41, 129.06, 127.04, 126.79, 121.51, 114.30, 113.92, 61.95, 61.69, 55.52, 55.47, 17.13, 14.50, 13.95. **HRMS** (ESI-TOF): *m/z*: [M + H]⁺ Calcd for C₂₈H₃₀NO₆⁺: 476.2068; Found 476.2067.

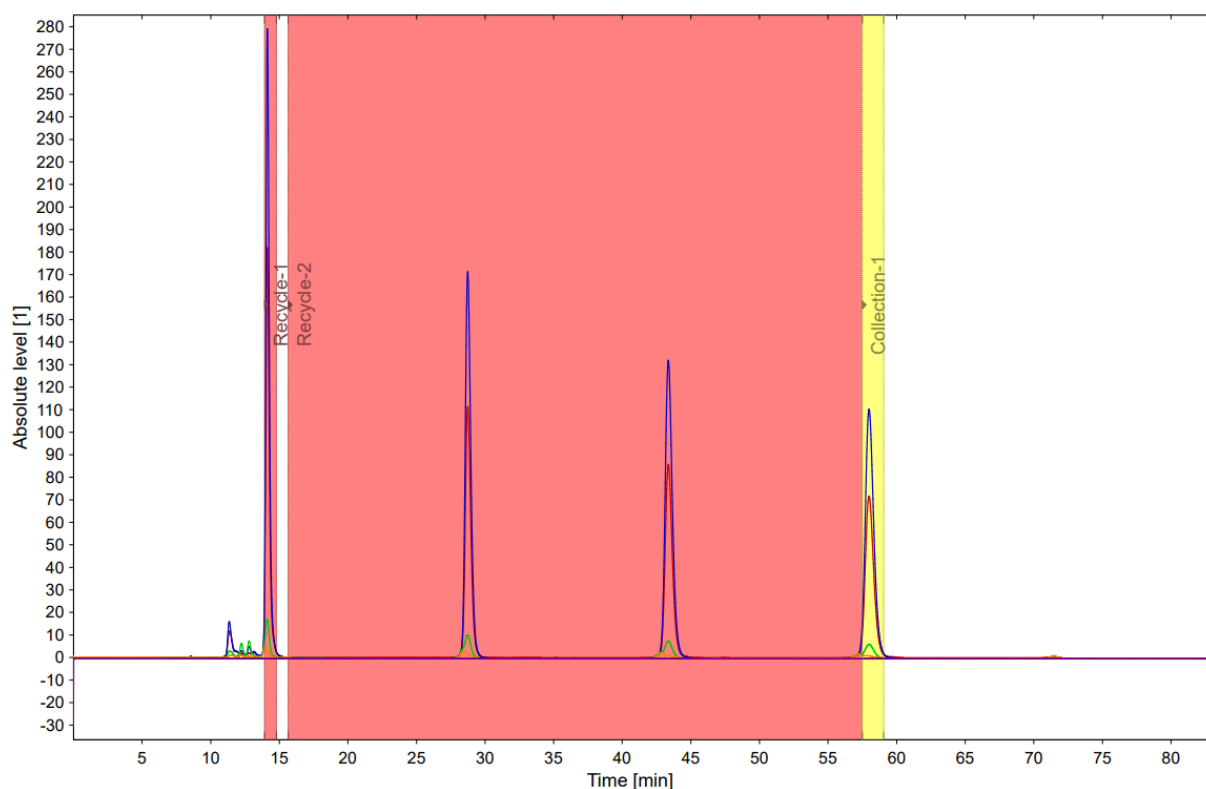
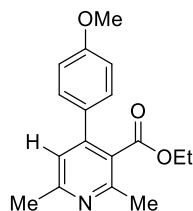


Figure S43. Final purification of **19o** (RP-HPCL/MeOH)



Compound 19o (ethyl 4-(4-methoxyphenyl)-2,6-dimethylnicotinate).

^1H NMR (400 MHz, chloroform-*d*): δ 7.30 (d, J = 8.8 Hz, 2H), 6.99 (s, 1H), 6.93 (d, J = 8.9 Hz, 2H), 4.14 (q, J = 7.1 Hz, 2H), 3.84 (s, 3H), 2.59 (s, 3H), 2.56 (s, 3H), 1.07 (t, J = 7.1 Hz, 3H). **^{13}C NMR** (100 MHz, chloroform-*d*): δ 169.51, 160.05, 158.68, 155.07, 148.07, 131.18, 129.31, 125.90, 121.19, 114.13, 61.41, 55.49, 24.62, 22.96, 13.96. **HRMS** (ESI-TOF): m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{17}\text{H}_{20}\text{NO}_3^+$: 286.1438; Found 286.1438.

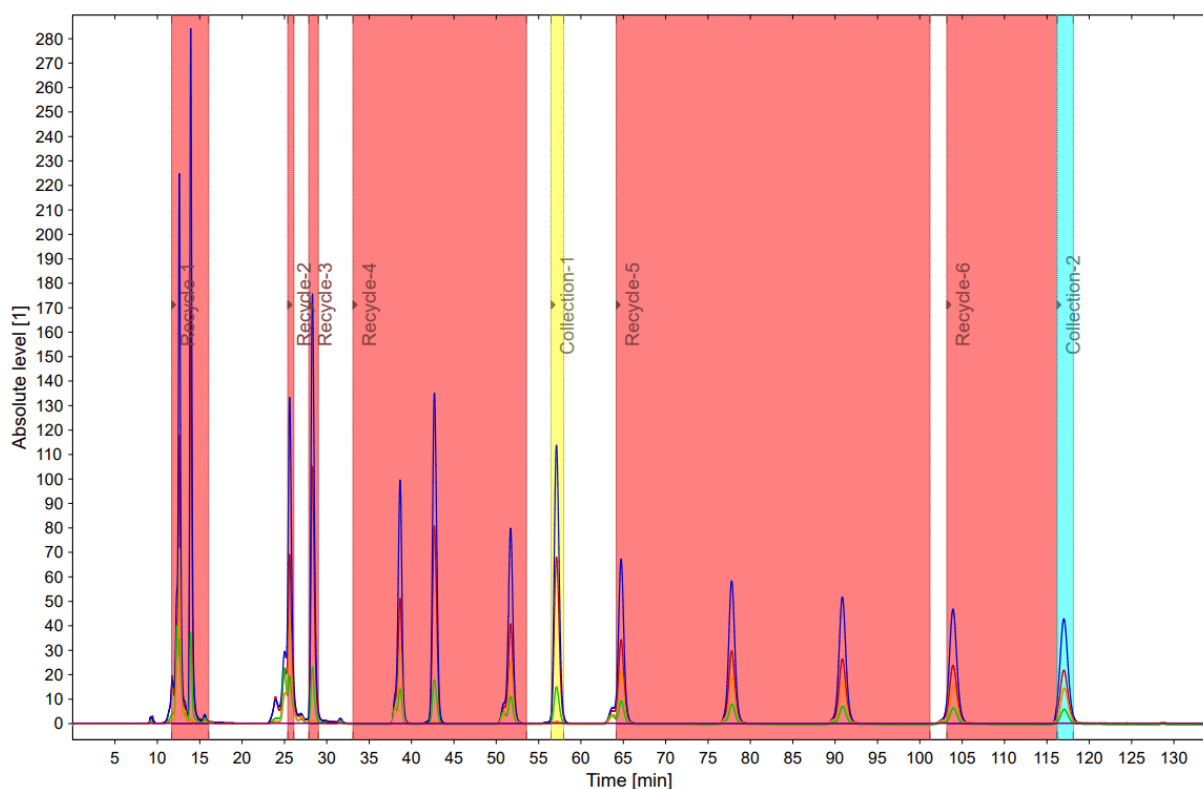
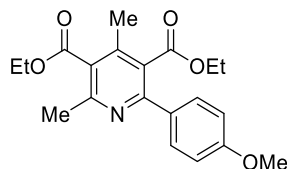
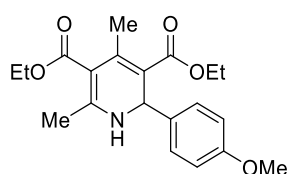


Figure S44. Final purification (RP–HPLC/MeOH) of **19r** (fraction 1) and **19p** (fraction 2).



Compound 19r (diethyl 2-(4-methoxyphenyl)-4,6-dimethylpyridine-3,5-dicarboxylate).

¹H NMR (400 MHz, chloroform-*d*): δ 7.53 (d, J = 8.9 Hz, 2H), 6.93 (d, J = 8.9 Hz, 2H), 4.43 (q, J = 5.4 Hz, 2H), 4.15 (q, J = 7.1 Hz, 2H), 3.82 (s, 3H), 2.59 (s, 3H), 2.33 (s, 3H), 1.40 (t, J = 7.2 Hz, 3H), 1.07 (t, J = 7.1 Hz, 3H). **¹³C NMR** (100 MHz, chloroform-*d*): δ 168.80, 168.53, 160.41, 155.97, 155.27, 142.78, 132.28, 129.82, 128.03, 127.02, 113.94, 61.76, 61.63, 55.46, 23.25, 16.98, 14.31, 13.89. **HRMS** (ESI–TOF): m/z : $[M + H]^+$ Calcd for C₂₀H₂₄NO₅⁺: 358.1649; Found 358.1647.



Compound 19p (diethyl 2-(4-methoxyphenyl)-4,6-dimethylpyridine-3,5-dicarboxylate).

¹H NMR (400 MHz, chloroform-*d*): δ 7.24 (d, J = 8.7 Hz, 2H), 6.82 (d, J = 8.7 Hz, 2H), 5.53 (d, J = 3.9 Hz, 1H), 5.20 (d, J = 2.8 Hz, 1H), 4.07–4.28 (m, 4H), 3.77 (s, 3H), 2.38 (s, 3H), 2.21 (s, 3H), 1.31 (t, J = 7.1, 3H), 1.21 (t, J = 7.1, 3H). **¹³C NMR** (100 MHz, chloroform-*d*): δ 168.02, 167.00, 159.37, 153.93, 145.82, 135.90, 127.86, 114.07, 111.24, 103.58, 60.01, 59.77, 55.39, 54.80, 21.54, 19.66, 14.52, 14.44. **HRMS** (ESI–TOF): m/z : $[M - H_2 + H]^+$ Calcd for C₂₀H₂₄NO₅⁺: 358.1649; Found 358.1650.

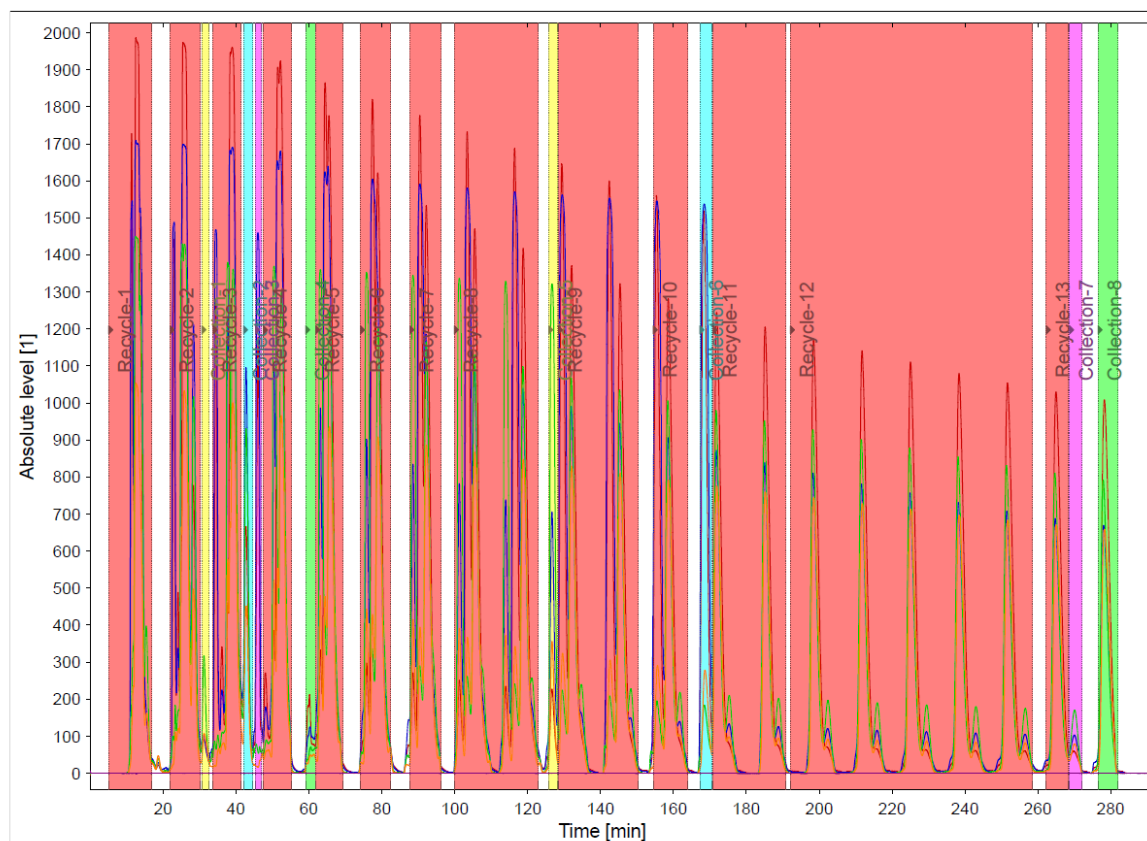


Figure S45. Chromatogram for initial separation of combined all crude reaction mixtures of a third run of the same conditions as in **Figure S35** at 80 °C (RP–HPLC/MeOH).

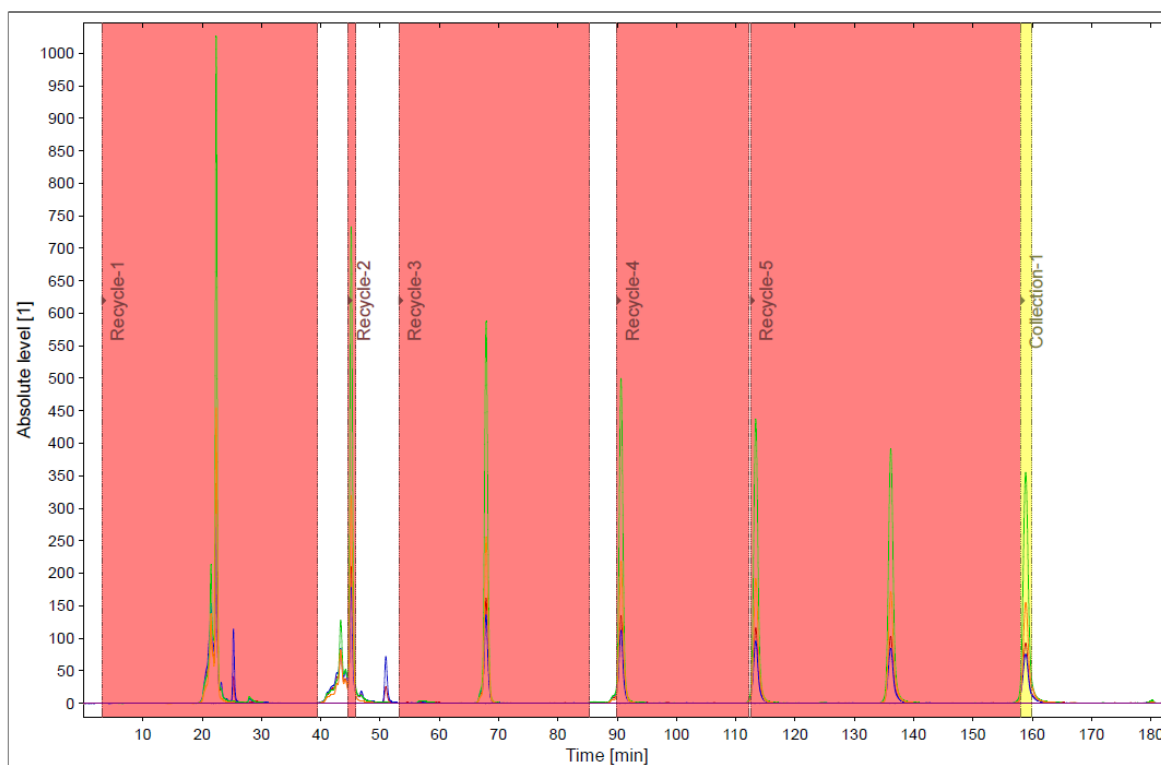


Figure S46. Chromatogram for separation of **fraction 1** from initial separation (*cf.* **Figure S45**, GPC/chloroform). **Fraction 1** contains **19j**.

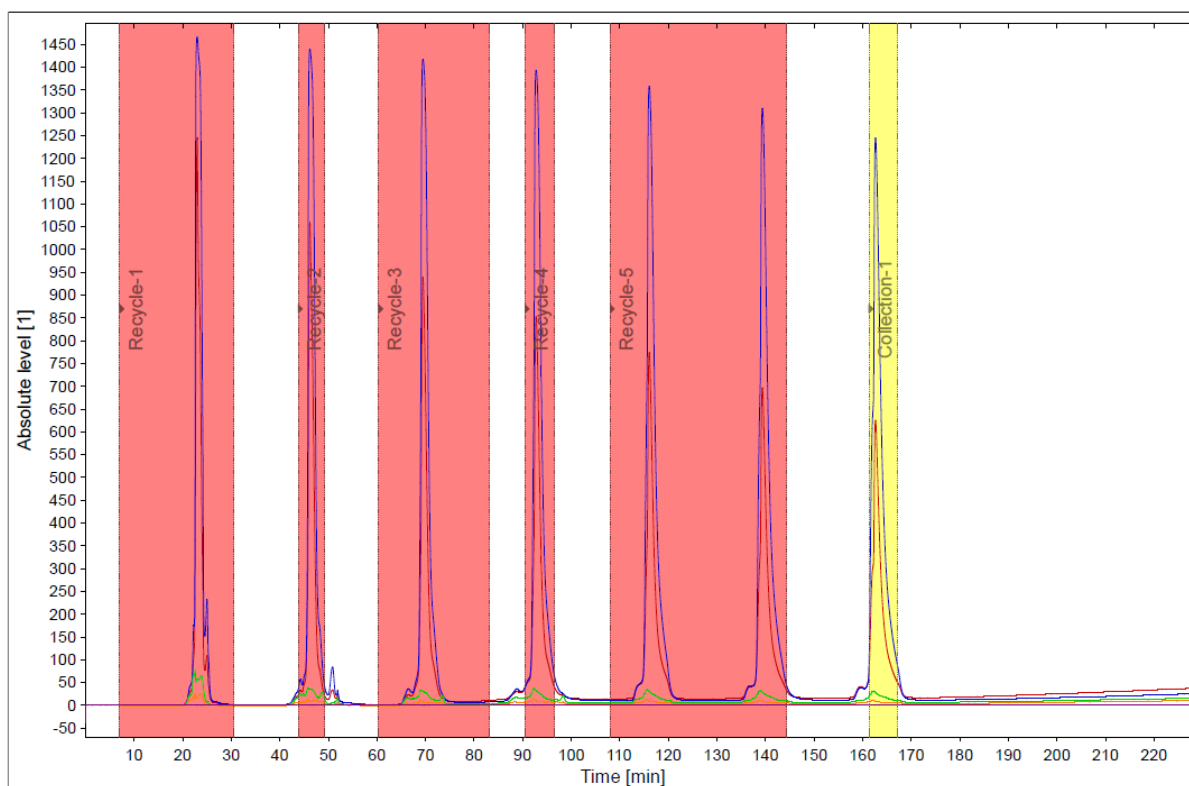


Figure S47. Chromatogram for separation of **fraction 3** from initial separation (*cf.* **Figure S45**, GPC/chloroform). **Fraction 1** contains **19j**.

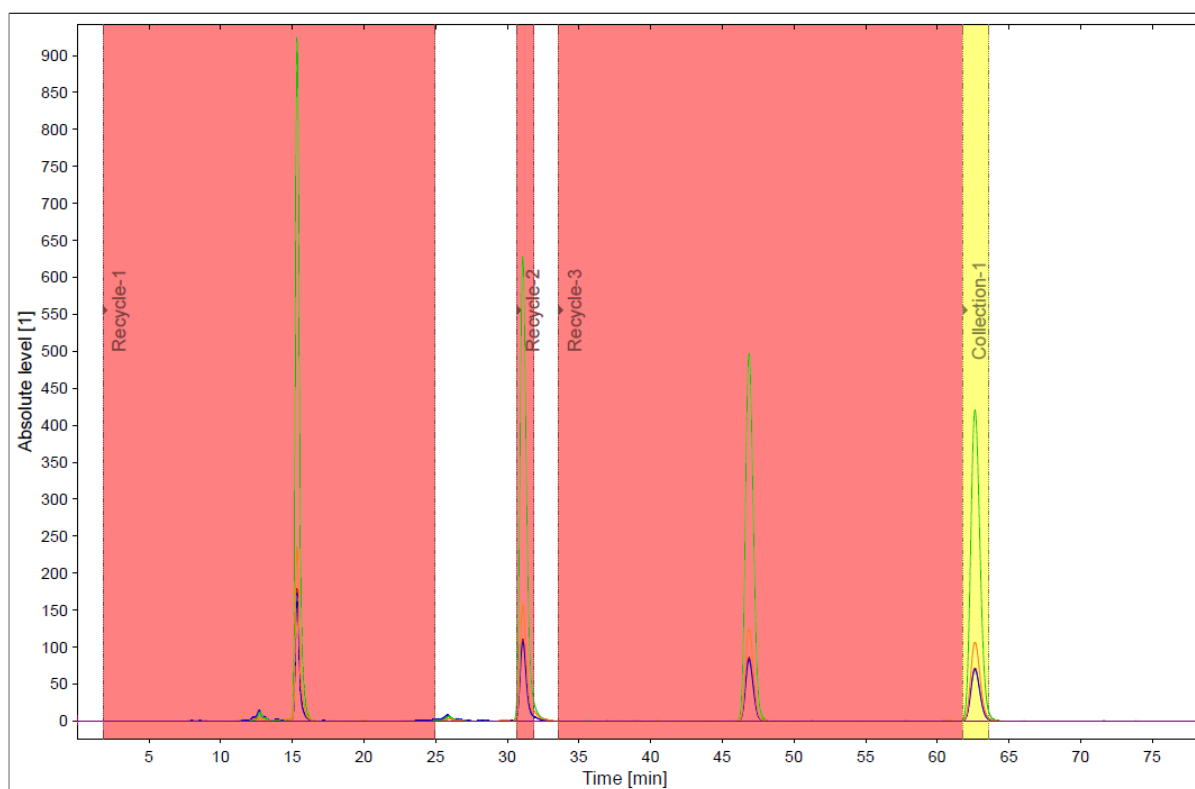
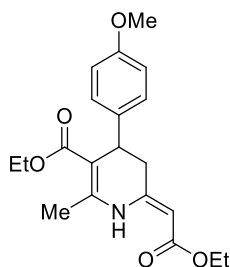


Figure S48. Final purification (RP-HPLC/MeOH) of **19j** (fraction 1).



Compound 19j (ethyl (Z)-6-(2-ethoxy-2-oxoethylidene)-4-(4-methoxyphenyl)-2-methyl-1,4,5,6-tetrahydropyridine-3-carboxylate)

¹H NMR (400 MHz, chloroform-*d*): δ 9.94 (s, 1H), 7.04 (d, J = 8.7 Hz, 2H), 6.78 (d, J = 8.7 Hz, 2H), 4.66 (d, J = 1.6 Hz, 1H), 4.00–4.19 (m, 5H), 3.76 (s, 3H), 2.91 (ddd, J = 1.6, 7.1, 15.5 Hz, 1H), 2.45 (m, 4H), 1.25 (t, J = 7.1 Hz, 3H), 1.18 (t, J = 7.1 Hz, 3H). **¹³C NMR** (100 MHz, chloroform-*d*, 300 K): δ 169.79, 167.48, 158.31, 152.62, 145.54, 135.84, 128.19, 113.94, 105.39, 90.81, 59.86, 59.58, 55.30, 35.98, 35.90, 20.29, 14.50, 14.41. **HRMS** (ESI-TOF): m/z : [M + H]⁺ Calcd for C₂₀H₂₆NO₅⁺: 360.1806; Found 360.1803.

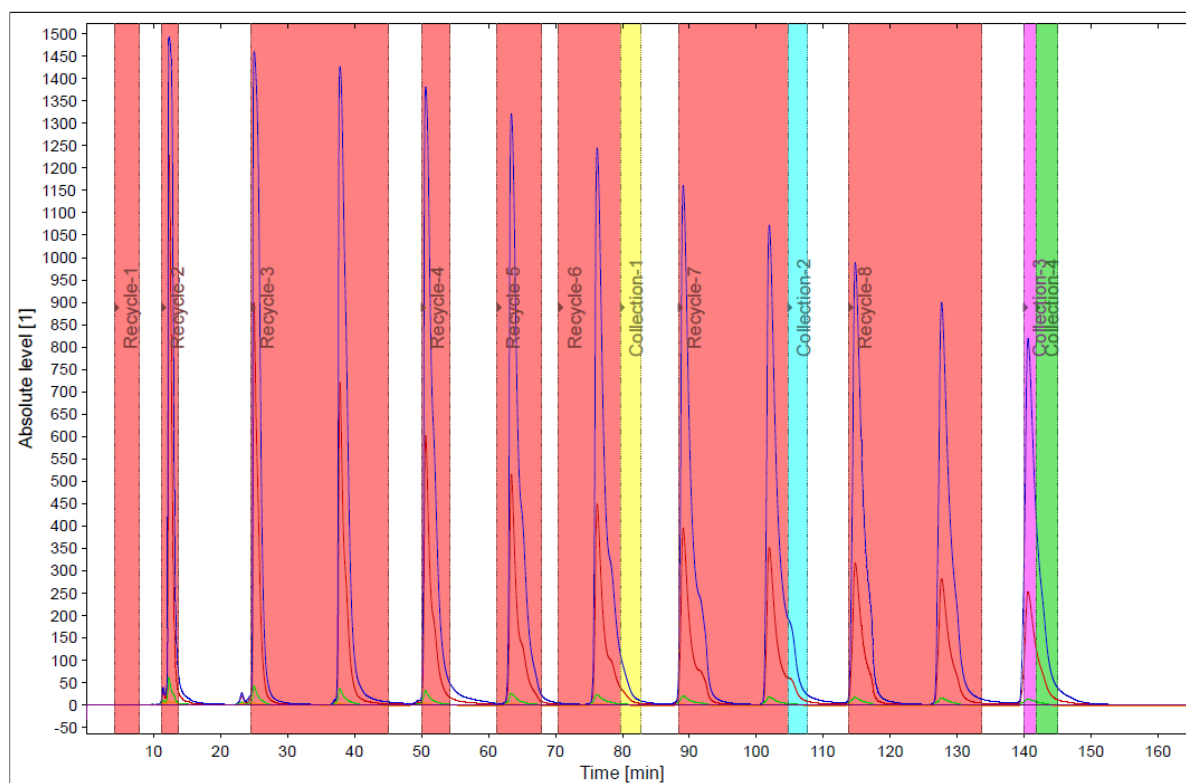
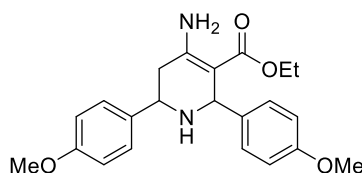


Figure S49. Final purification (RP-HPLC/MeOH) of **19e** (fraction 3).



Compound 19e (ethyl 4-amino-2,6-bis(4-methoxyphenyl)-1,2,5,6-tetrahydropyridine-3-carboxylate).

¹H NMR (400 MHz, chloroform-*d*): δ 7.20 (m, 4H), 6.82 (m, 4H), 5.08 (s, 1H), 3.92–4.02 (m, 2H), 3.86 (m, 1H), 3.80 (s, 3H), 3.76 (s, 3H), 2.56 (m, 1H), 2.34 (m, 1H), 1.94 (br, 1H), 0.97 (t, $J = 7.1$ Hz, 3H). **¹³C NMR** (100 MHz, chloroform-*d*, 300 K): δ 168.94, 159.10, 158.40, 156.29, 137.85, 135.60, 128.66, 127.91, 114.01, 113.37, 93.17, 58.92, 55.93, 55.39, 55.33, 49.84, 38.57, 14.35. **HRMS** (ESI-TOF): m/z : $[M - H]^+$ Calcd for $C_{22}H_{25}N_2O_4$: 381.1809; Found 381.1909, m/z : $[M + 2H]^+$ Calcd for $C_{22}H_{27}N_2O_4^+$: 383.1966; Found 383.1902, m/z : $[M - H + H_2O]^+$ Calcd for $C_{22}H_{27}N_2O_5^+$: 399.1915; Found 399.2016.

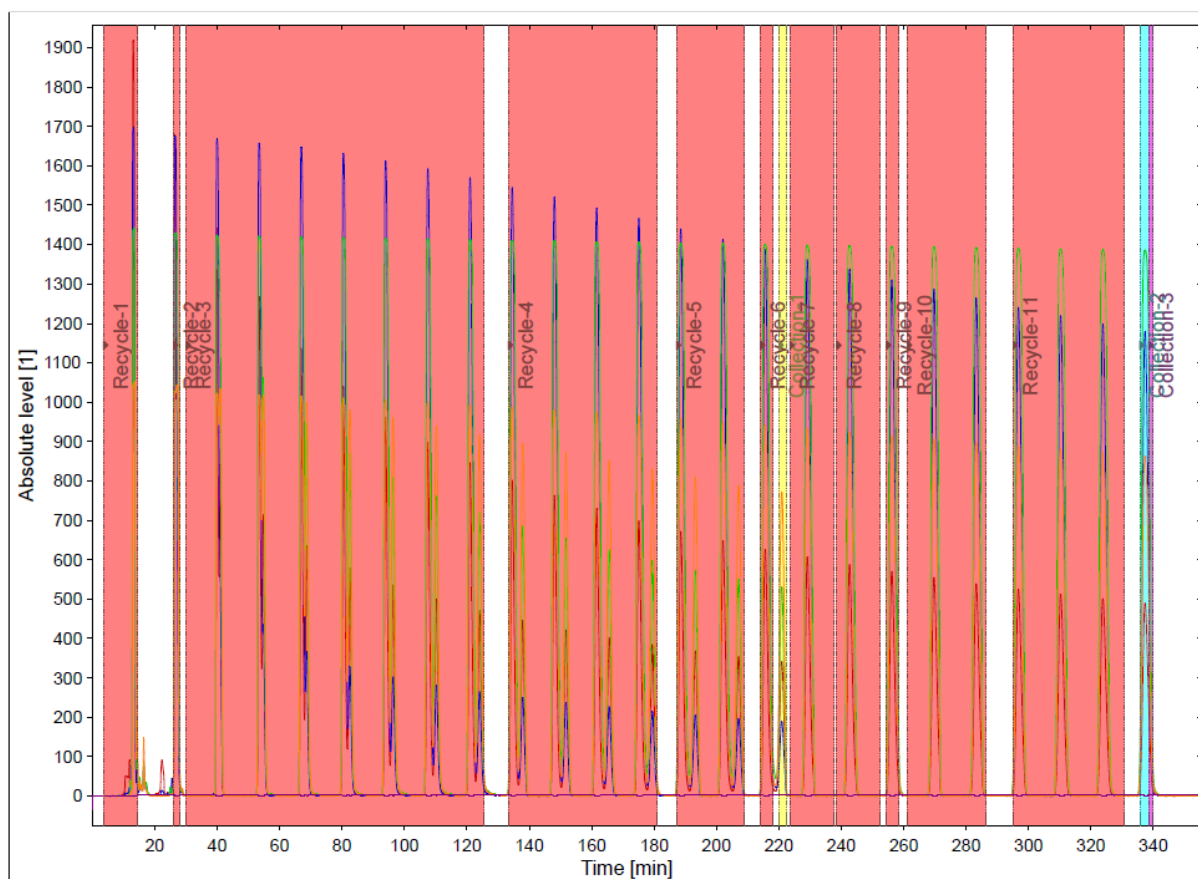
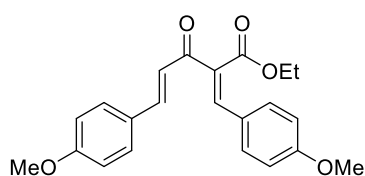


Figure S50. Chromatogram for separation of **fraction 7** from initial separation (cf. **Figure S45**, RP-HPLC/MeOH). **Fraction 2** contains **19h**.

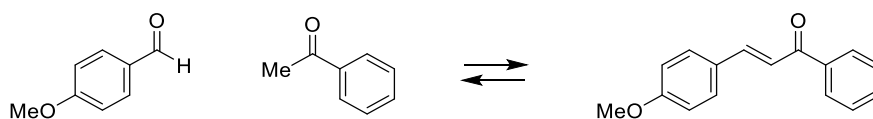


Compound 19h (ethyl (*E*)-2-((*Z*)-4-methoxybenzylidene)-5-(4-methoxyphenyl)-3-oxopent-4-enoate).

¹H NMR (400 MHz, chloroform-*d*): δ = 7.82 (s, 1H), 7.43, (m, 5H), 6.88 (d, *J* = 8.7 Hz, 2H), 6.82 (d, *J* = 8.9 Hz, 2H) 6.76 (d, *J* = 16.3 Hz, 1H), 4.27 (m, 2H) 3.83 (s, 3H), 3.78 (s, 3H), 1.27 (t, *J* = 7.2 Hz, 3H). **¹³C NMR** (100 MHz, chloroform-*d*): δ = 196.37, 165.61, 162.11, 161.48, 146.26, 141.97, 132.40, 130.54, 129.24, 127.08, 125.75, 125.19, 114.57, 114.41, 61.51, 55.57, 55.45, 14.35. **HRMS** (ESI-TOF): *m/z*: [M + Na]⁺ Calcd for C₂₂H₂₂O₅Na⁺: 389.1360; Found 389.1366.

4.3. Synthetic procedures for versatility test (Extended Data Fig. 1)

4.3.1. Claisen–Schmidt condensation **1**: synthesis of 3-(4-methoxyphenyl)-1-phenylprop-2-en-1-one⁸⁵



p-Methoxybenzaldehyde (136 mg, 1.00 mmol), acetophenone (120 mg, 1.00 mmol) and potassium *tert*-butoxide (224 mg, 2.00 mmol) were dissolved in ethanol to make up a volume of 10 mL. The resulting mixture was stirred at room temperature with the flask stoppered. After 3 h, an aliquot of reaction mixture was withdrawn for analysis on UV–Vis NanoDrop. The remaining reaction mixture was concentrated in vacuo and then diluted with dichloromethane. The organic solution was washed successively with 1.2 M hydrochloric acid, water and brine. After solvent removal, the residue was analyzed by ¹H NMR spectroscopy. To this end, chromatographic isolation of the three components was not pursued.

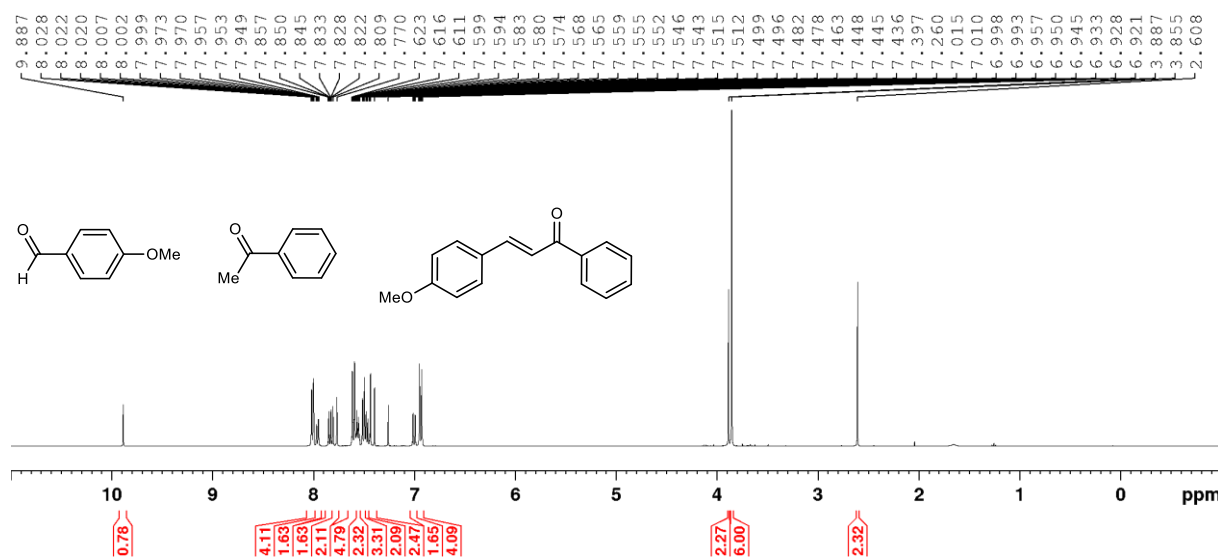
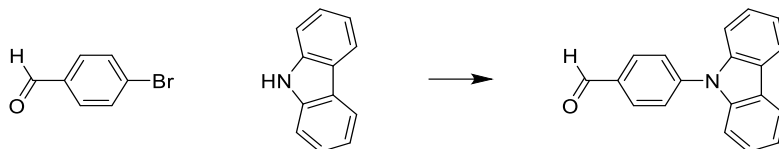


Figure S51. ¹H NMR spectrum of a crude mixture after Claisen–Schmidt condensation (400 MHz, chloroform-*d*).

4.3.2. Ullmann-type coupling 2: synthesis of 4-(9*H*-carbazol-9-yl)benzaldehyde



9*H*-carbazole (0.50 g, 2.87 mmol), 4-bromobenzaldehyde (1.07 g, 5.74 mmol), copper(I) iodide (94 mg, 0.49 mmol) and potassium carbonate (0.60 g, 4.31 mmol) were mixed with 4 mL of DMF in a pressure tube. The mixture was heated at 140 °C for 15 h. After the reaction was completed, the solution was diluted and analyzed using UV–Vis NanoDrop. The crude mixture was diluted with water. The mixture was extracted with dichloromethane three times and the organic layer was dried over anhydrous sodium sulfate. The solvent was evaporated, and the crude mixture was purified via silica column chromatography to give light yellow solid (213 mg, 27%).

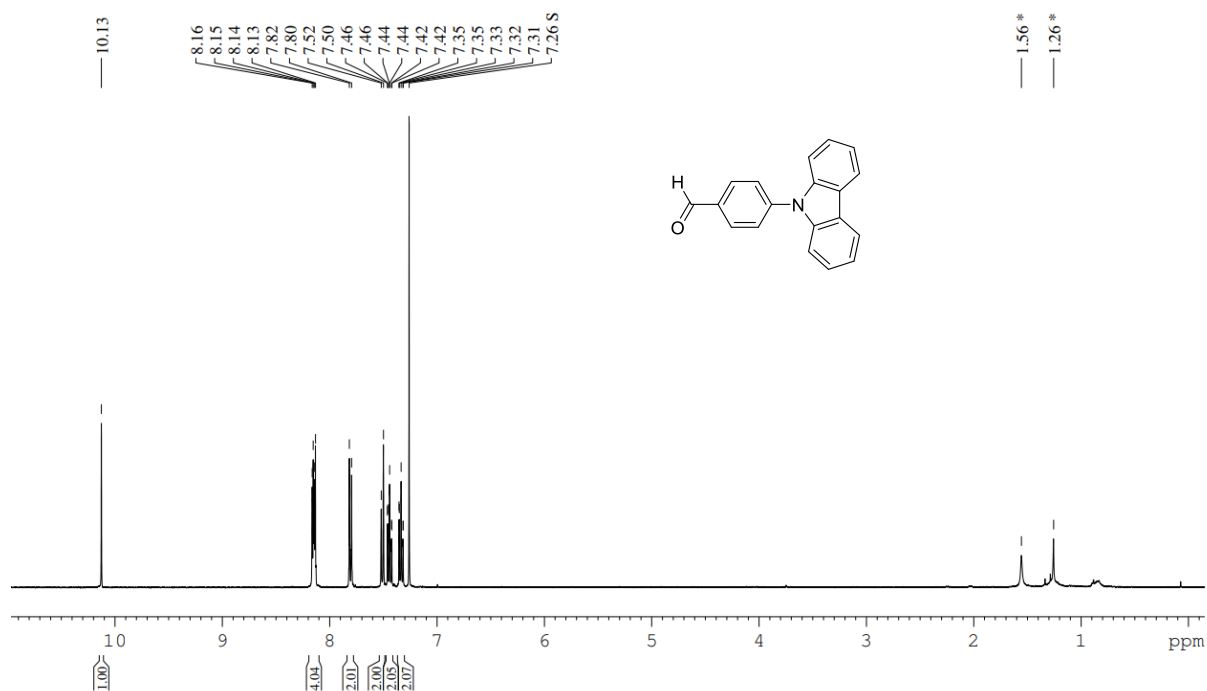
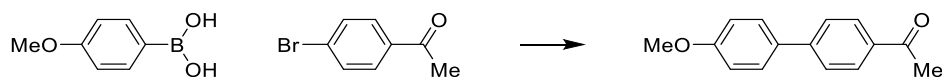


Figure S52. ¹H NMR spectrum of 4-(9*H*-carbazol-9-yl)benzaldehyde (400 MHz, chloroform-*d*).

4.3.3. Suzuki-Miyaura cross-coupling **3**: synthesis of 1-(4'-methoxy-[1,1'-biphenyl]-4-yl)ethan-1-one.



Stock solutions:

S1: 4-methoxyphenylboronic acid (109.1 mg, 0.72 mmol) was dissolved in 5 mL of °Cassed ethanol (0.144 M).

S2: 4-bromoacetophenone (143.3 mg, 0.72 mmol) was dissolved in 5 mL of °Cassed ethanol (0.144 M).

S3: Potassium carbonate (398 mg, 2.88 mmol) was dissolved in 20 mL of °Cassed water (0.144 M).

S4: Tetrakis(triphenylphosphine)palladium(0) (8.32 mg, 7.2 μ mol) was dissolved in 1 mL of °Cassed chloroform (7.2 mM).

Reaction: 0.25 mL of **S1**, **S2**, 0.5 mL of **S3**, 0.05 mL of **S4**, and 0.45 mL of ethanol were dispensed into a vial using a pipetting robot. The vial was shaken for 10 minutes and placed in the oven for 12 hours at 53 °C. After the reaction was completed, the solution was diluted (Dilution Factor: 78.65) and analyzed using UV–Vis NanoDrop.

For ^1H NMR analysis, the crude product was dissolved in chloroform (2 mL) and washed with water (1 mL). The organic phase was collected and dried over anhydrous sodium sulfate and the solvent was removed under reduced pressure. ^1H NMR (400 MHz, chloroform-*d*) δ 8.01 (d, J = 8.4 Hz, 2H), 7.65 (d, J = 8.4 Hz, 2H), 7.58 (d, J = 8.8 Hz, 2H), 7.00 (d, J = 8.8 Hz, 2H), 3.87 (s, 3H), 2.63 (s, 3H).

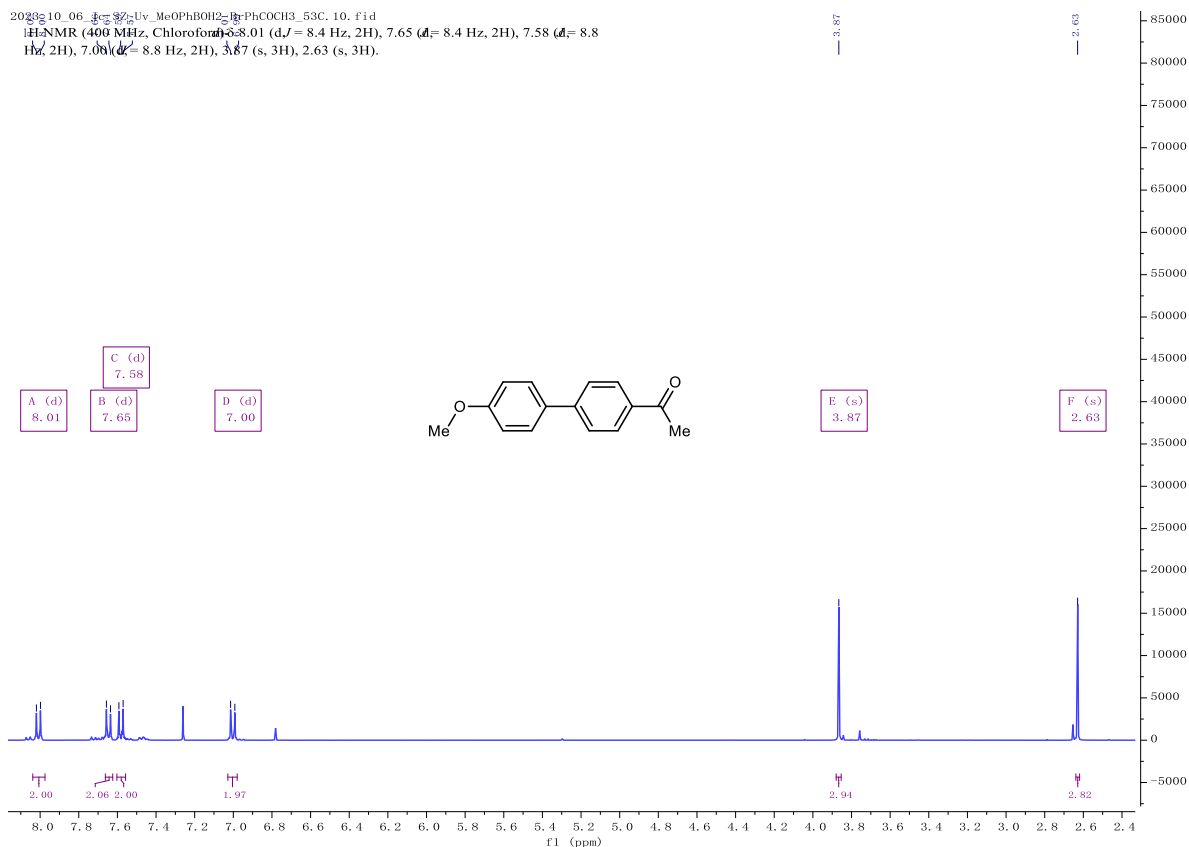
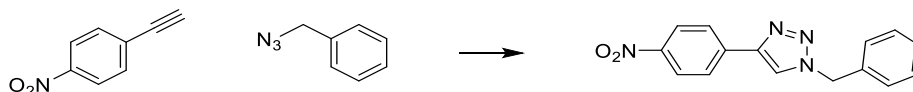


Figure S53. ¹H NMR spectrum of 1-(4'-methoxy-[1,1'-biphenyl]-4-yl)ethan-1-one (400 MHz, chloroform-*d*).

4.3.4. Copper(I)-catalyzed alkyne-azide cycloaddition **4**: synthesis of 1-benzyl-4-(4-nitrophenyl)-1*H*-1,2,3-triazole



Solution of benzyl azide (0.5 M in dichloromethane, 730 μ L, 0.346 mmol) was placed in a vial equipped with magnetic stirring bar, and purged with nitrogen for 15 min. Then, 1-ethynyl-4-nitrobenzene (52 mg, 0.346 mmol), copper(I) iodide (6.7 mg, 34.6 μ mol) and 1 mL of acetonitrile were added, and the reaction mixture was stirred at 40 $^{\circ}$ C for 14 h. Then, the solution was diluted and analyzed using UV-Vis NanoDrop. After that time solvent was removed under reduced pressure and pure product was isolated by column chromatography (silica gel, dichloromethane) as a colorless solid (13 mg, 13%).

¹H NMR (400 MHz, chloroform-*d*): δ = 8.26 (d, *J* = 8.8 Hz, 2H), 7.97 (d, *J* = 8.9 Hz, 2H), 7.81 (s, 1H), 7.37–7.46 (m, 3H), 7.31–7.37 (m, 2H), 5.61 (s, 2H).

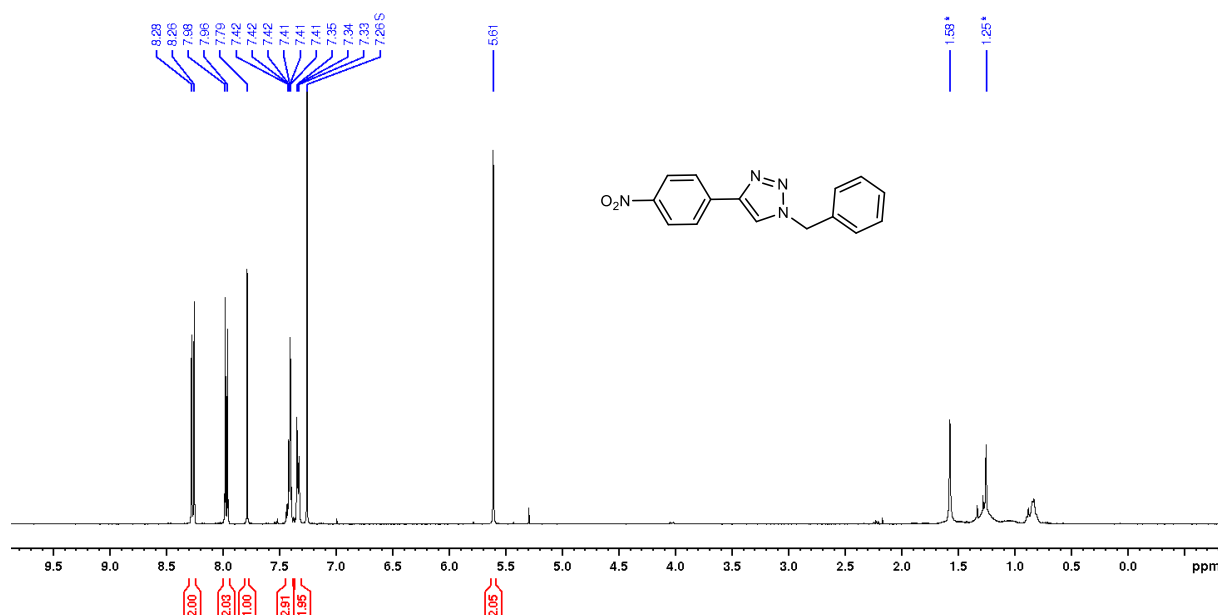
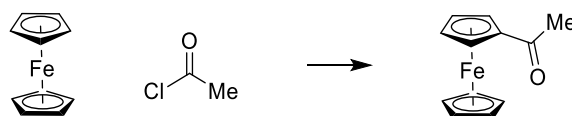


Figure S54. ¹H NMR spectrum of 1-benzyl-4-(4-nitrophenyl)-1*H*-1,2,3-triazole (400 MHz, chloroform-*d*).

4.3.5. Friedel–Crafts acylation **5**: synthesis of acetylferrocene



Acetyl chloride (85 μ L, 1.2 mmol) was added to a solution of ferrocene (186 mg, 1.0 mmol) in 20 mL 1,4-dioxane in two-necked round-bottom flask. Then the reaction mixture was stirred for 1 h at room temperature under argon atmosphere or simply protected from air moisture with CaCl₂ tube. After that time, aluminum chloride (200 mg, 1.5 mmol) was added to the mixture, which immediately becomes dark violet. Resulting mixture was stirred for 1 h. An aliquot of reaction mixture (passed through the syringe filter) was withdrawn for analysis on UV–Vis NanoDrop. The crude mixture was then diluted with 20 mL of diethyl ether and poured into a cold 1M solution of KOH (40 mL) to remove aluminum, and unreacted acetyl chloride. The organic layer was separated and the aqueous one was extracted with an additional 20 mL of ether. The organic layers were collected and washed with deionized water, brine, and dried over anhydrous sodium sulfate. After evaporation of the solvents, the obtained raw reaction mixture was purified by column chromatography (SiO₂/toluene then toluene/ethyl acetate 9:1). The first fraction contained ferrocene, the second – monoacetyl derivative ($y = 85\%$) and the third one – diacetylferrocene.

¹H NMR (400 MHz, chloroform-*d*): $\delta = 4.77$ (s, 2H), 4.50 (s, 2H), 4.20 (s, 5H), 2.40 (s, 3H).

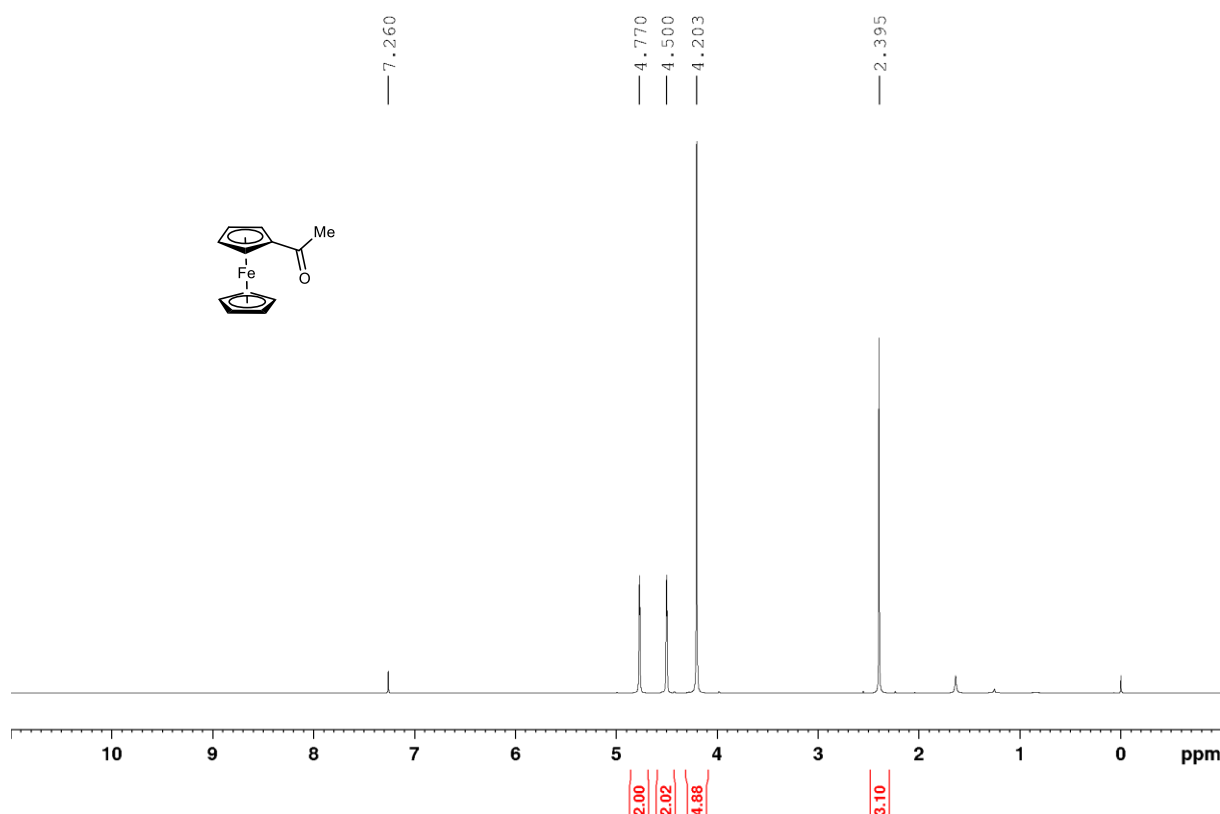
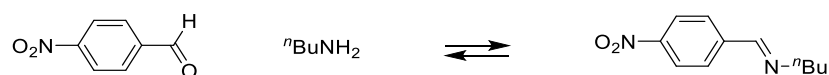


Figure S55. ^1H NMR spectrum of acetylferrocene (400 MHz, chloroform- d).

4.3.6. Imine condensation **6**



The reaction was performed according to the modified literature procedure.⁸⁶ One drop of acetic acid was added to the mixture of *n*-butylamine (261.6 μL , 2.65 mmol) and *p*-nitrobenzaldehyde (400 mg, 2.65 mmol) in 25 mL of DMF. The solution was stirred at room temperature for 24 h. An aliquot of the stirred reaction mixture was placed onto the sample holder of the NanoDrop spectrometer for the measurement of absorbance spectrum. Then, the mixture was evaporated under reduced pressure.

^1H NMR (400 MHz, toluene- d_8): δ = 7.77 (d, J = 8.8 Hz, 2H), 7.69 (s, 1H), 7.33 (d, J = 8.8 Hz, 2H), 3.38 (dt, J = 6.9, 1.4 Hz, 2H), 1.59 (qt, J = 7.2 Hz, 2H), 1.32 (sextet, J = 7.6 Hz, 2H), 0.90 (t, J = 7.4 Hz, 3H).

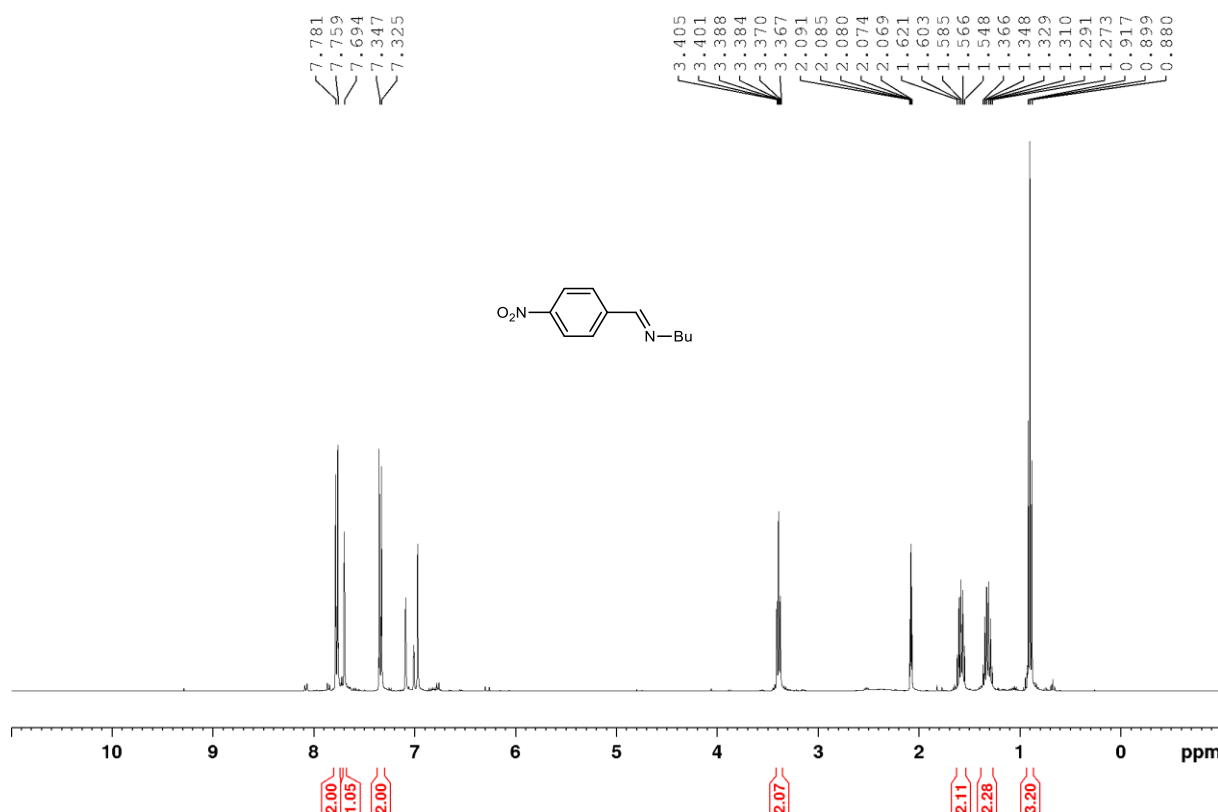
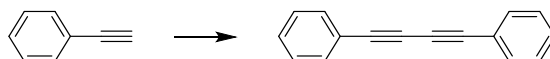


Figure S56. ¹H NMR spectrum of (*E*)-*N*-butyl-1-(4-nitrophenyl)methanimine (400 MHz, toluene-*d*₈).

4.3.7. Glaser coupling of phenylacetylene **7**: synthesis of 1,4-diphenylbutadiyne



The reported procedure⁸⁷ was modified by changing the reaction solvent from DMF to acetonitrile because the absorption maximum of DMF (270 nm) could interfere with unmixing procedure. To a stirred solution of phenylacetylene (110 μ L, 1.00 mmol) dissolved in acetonitrile (1 mL) copper(I) iodide (190 mg, 1.00 mmol), sodium carbonate (212 mg, 2.00 mmol) and iodine (254 mg, 1.00 mmol) were successively added. The flask was equipped with reflux condenser and the resulting mixture was stirred and heated at 80 °C under air. After 3 h, stirring was ceased such that the solid material settled on the bottom of the flask. The reaction mixture was allowed to cool to room temperature. An aliquot of the reaction mixture was withdrawn for analysis on UV–Vis NanoDrop. The remaining reaction mixture was filtered through Celite. The filtrate was diluted with ethyl acetate. The organic solution was washed with aqueous sodium thiosulfate solution, water and brine, and then concentrated under reduced pressure. The residue was purified by column chromatography on silica (hexane) to afford 1,4-diphenylbutadiyne (91 mg, 90%) as a colorless solid.

¹H NMR (400 MHz, chloroform-*d*): δ = 7.52–7.55 (m, 4H), 7.32–7.40 (m, 6H).

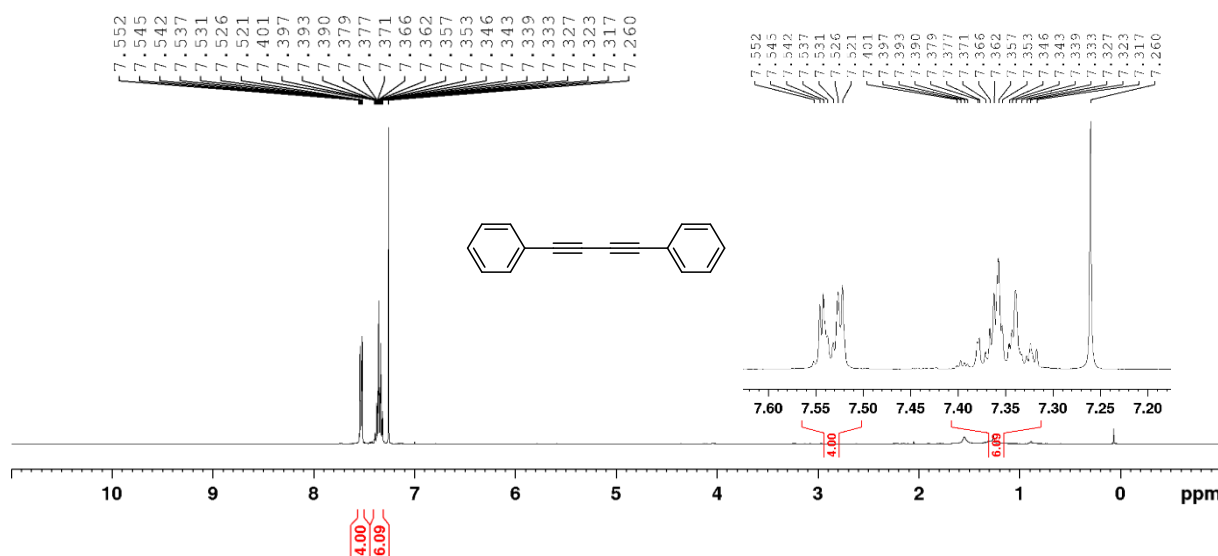
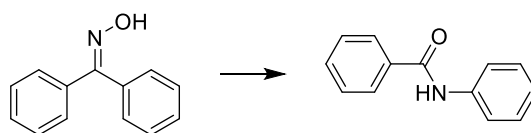


Figure S57. ^1H NMR spectrum of 1,4-diphenylbutadiyne (400 MHz, chloroform- d).

4.3.8. Beckmann rearrangement of benzophenone oxime **8**: synthesis of benzanilide



Benzophenone oxime was prepared as reported,⁸⁸ and the reported Lewis acid-mediated Beckmann rearrangement conditions were employed.⁸⁹ Aluminum chloride (41 mg, 0.307 mmol) was added to the stirred solution of benzophenone oxime (600 mg, 3.04 mmol) in acetonitrile (6 mL). The resulting mixture was stirred under nitrogen atmosphere and heated to gentle reflux (90 °C). After 18 h, an aliquot of the stirred reaction mixture was withdrawn for analysis on UV-Vis NanoDrop. The remaining reaction mixture was poured into 1.2 M hydrochloric acid. The resulting biphasic mixture was extracted with ethyl acetate. The organic solution was washed with water and brine, and then concentrated in vacuo. The residue was purified by column chromatography on silica (dichloromethane) to afford benzanilide (546 mg, 91%) as a colorless solid.

^1H NMR (400 MHz, chloroform- d): δ = 7.86–7.88 (m, 3H), 7.64–7.66 (m, 2H), 7.53–7.57 (m, 1H), 7.47–7.50 (m, 2H), 7.35–7.39 (m, 2H), 7.14–7.18 (m, 1H).

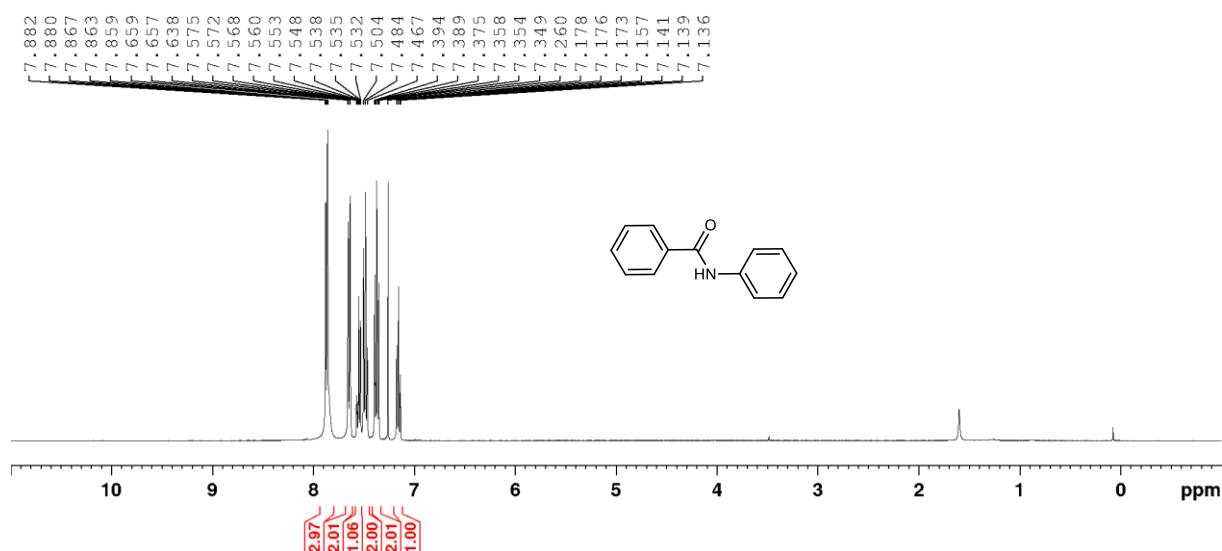
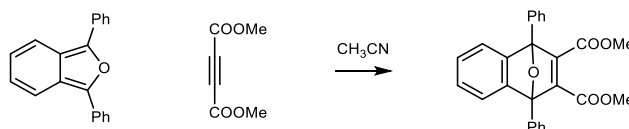


Figure S58. ^1H NMR spectrum of benzanilide (400 MHz, chloroform-*d*).

4.3.9. Diels–Alder reaction **2** between 1,3-diphenylisobenzofuran (DPBF) and dimethyl acetylenedicarboxylate (DMAD)



The reported procedure⁹⁰ was modified by changing from refluxing in toluene to stirring in acetonitrile at room temperature. Dimethyl acetylenedicarboxylate (61.4 μL , 0.499 mmol) was added to the stirred suspension of 1,3-diphenylisobenzofuran (135 mg, 0.499 mmol) in acetonitrile (1 mL). The flask was protected from light using aluminum foil. The resulting suspension was stirred at room temperature under nitrogen atmosphere. After 3 h, an aliquot of the stirred reaction mixture was withdrawn for analysis on UV–Vis NanoDrop. The remaining reaction mixture was concentrated under reduced pressure. The residue was purified by column chromatography on silica (hexane:dichloromethane = 2:1 gradient to 1:2) to afford the DPBF–DMAD cycloadduct (82 mg, 40%) as a colorless solid.

^1H NMR (400 MHz, chloroform-*d*): δ = 7.72–7.75 (m, 4H), 7.51–7.55 (m, 2H), 7.41–7.50 (m, 6H), 7.14–7.18 (m, 2H), 3.69 (s, 6H).

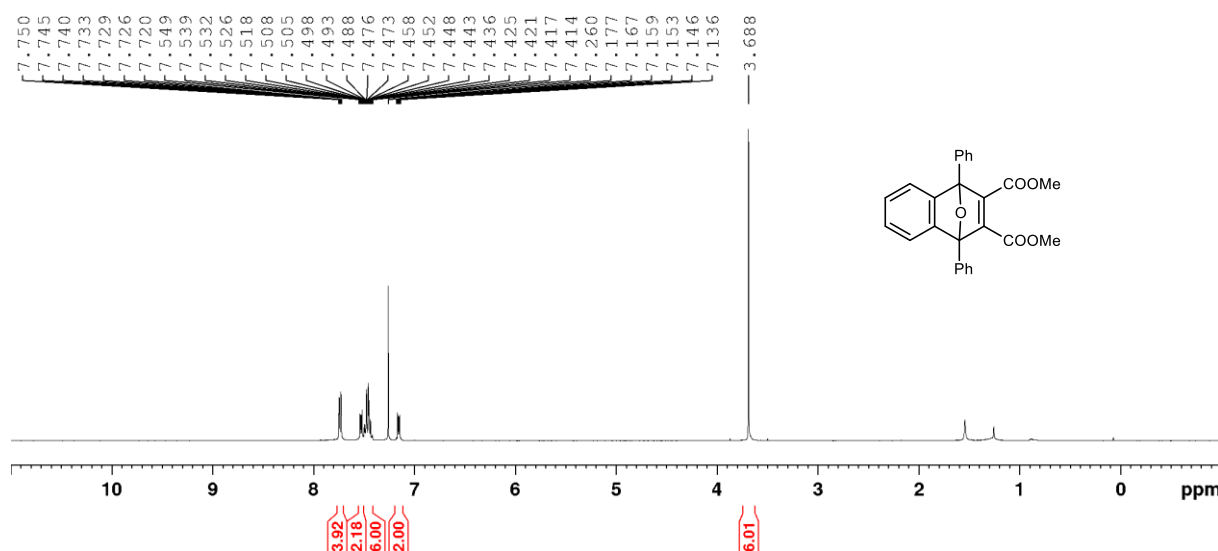
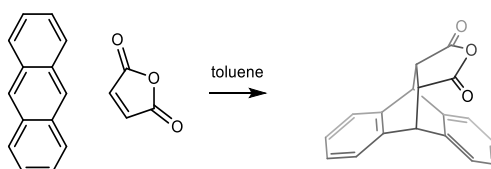


Figure S59. ¹H NMR spectrum of DPBF-DMAD cycloadduct (400 MHz, chloroform-*d*).

4.3.10. Diels–Alder reaction **10** between anthracene and maleic anhydride



The reaction was performed according to the modified literature procedure.⁹¹ Maleic anhydride (0.550 g, 5.61 mmol) was added to the solution of anthracene (1.00 g, 5.61 mmol) in toluene (50 mL) and then refluxed for 6 h until all the diene disappeared on TLC. An aliquot of the stirred reaction mixture was withdrawn for analysis on UV–Vis NanoDrop. After the removal of the solvent at vacuum, the occurred colorless solid was dried. The crude product was crystallized from a mixture of dichloromethane:petroleum ether (3:2), resulting the pure product in 97% yield (1.50 g).

¹H NMR (400 MHz, DMSO-*d*₆) δ = 7.47 (dd, J = 3.2, 2.4 Hz, 2H), 7.33 (dd, J = 3.2, 2.4 Hz, 2H), 7.18 (ddd, J = 3.2, 2.4, 1.6 Hz, 4H), 4.86 (s, 2H), 3.65 (dd, J = 2.0, 1.4 Hz, 2H).

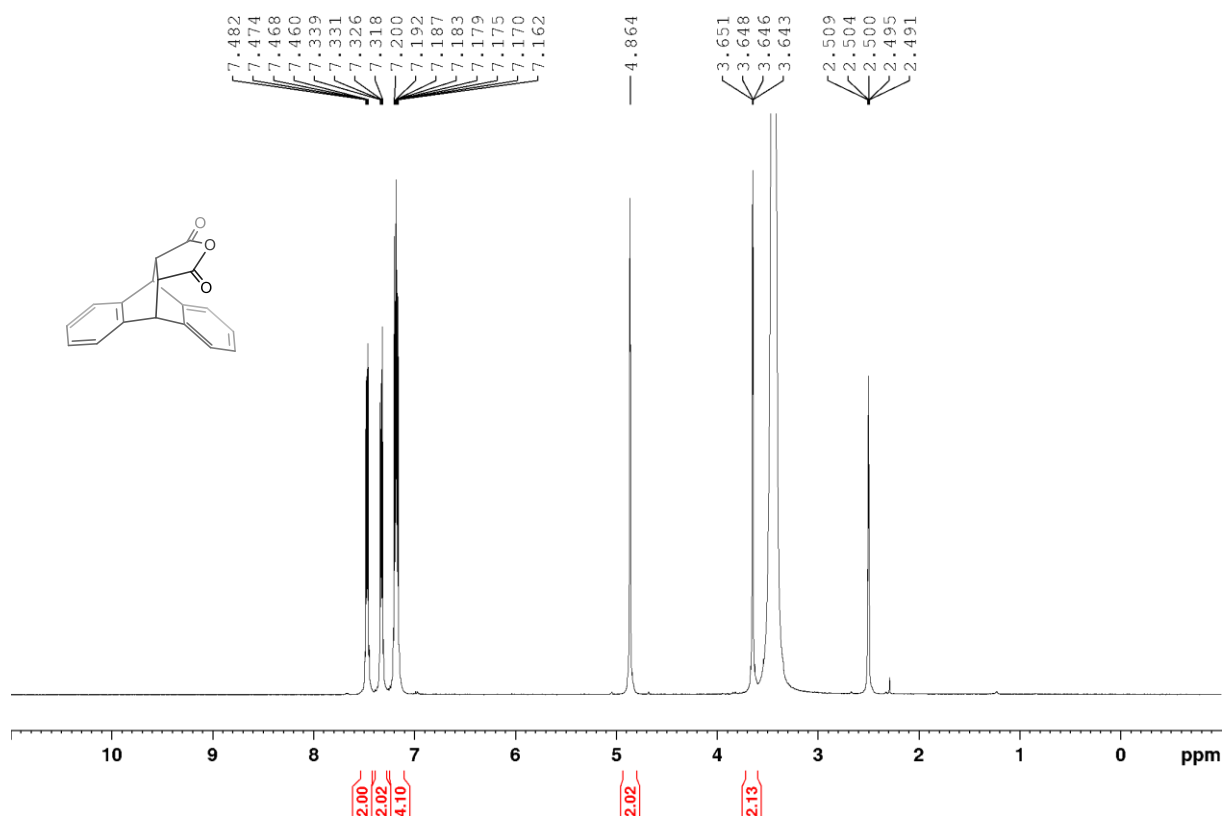


Figure S60. ¹H NMR spectrum of anthracene-maleic anhydride cycloadduct (400 MHz, DMSO-*d*₆).

Table S8. Comparison of yields from UV–Vis and NMR/isolated yields for reactions above.

	Reaction name	Optical yield	^a NMR yield ^b Isolated yield
<u>1</u>	Claisen–Schmidt condensation	0.486	0.567 ^a
<u>2</u>	Ullmann-type coupling	0.640	0.270 ^b
<u>3</u>	Suzuki–Miyaura cross-coupling	1.022	1.000 ^a
<u>4</u>	Copper(I)-catalyzed alkyne–azide cycloaddition	0.105	0.130 ^b
<u>5</u>	Friedel–Crafts acylation	0.880	0.850 ^b
<u>6</u>	Imine condensation	0.994	0.985 ^a
<u>7</u>	Glaser coupling	0.935	0.900 ^b
<u>8</u>	Beckmann rearrangement	1.000	0.910 ^b

4.4. Comparison of yields

Reactions were performed on the robot according to the procedure in Reaction yield map for E1 (synthesis of 9-butyldiene-9*H*-fluorene) (at 26 °C, **Section 4.2**) and Reaction yield map for multicomponent reaction (synthesis of 3-butyl-4-(4-nitrophenyl)-1-(tosylmethyl)-1*H*-imidazol-3-ium-5-olate, **Section 4.2**). Then after collection of UV–Vis spectrum the concentration was measured by analytical HPLC.

Table S9. Initial concentrations of 9-butyl-9*H*-fluoren-9-ol (**13a**) and HBr for E1 reaction and yields from UV-Vis spectroscopy and HPLC.

Entry	[13a] ₀ (mM)	[HBr] ₀ (mM)	UV-Vis yield (%)	HPLC yield (%)
0	11.63	8.290	74.4	67.1
1	11.63	8.290	74.7	66.8
2	11.63	8.290	75.4	68.8
3	4.875	1.124	42.3	44.9
4	4.875	1.124	42.7	44.8
5	4.875	1.124	43.3	46.2
6	11.63	5.903	76.1	71.2
7	11.63	5.903	75.8	71.9
8	11.63	5.903	76.0	72.2
9	1.500	2.490	65.9	80.1
10	1.500	2.490	67.0	78.7
11	1.500	2.490	65.8	78.7
12	11.63	3.514	61.6	60.0
13	11.63	3.514	61.5	60.2
14	11.63	3.514	61.1	58.6

Table S10. Initial concentrations (mol/L) of substrates for Ugi-type MCR reaction and yields from UV-Vis unmixing vs. HPLC.

	aldehyde	<i>p</i> TSA	amine	isocyanide	UV-Vis yield	HPLC yield
15	0.2550	0.1808	0.1200	0.3000	0.156	0.167
16	0.2550	0.1808	0.1200	0.3000	0.156	0.171
17	0.2550	0.1808	0.1200	0.3000	0.156	0.166
18	0.2550	0.1808	0.1200	0.3000	0.156	0.135
19	0.2550	0.1808	0.1200	0.3000	0.156	0.130
20	0.2550	0.1808	0.1200	0.3000	0.156	0.128

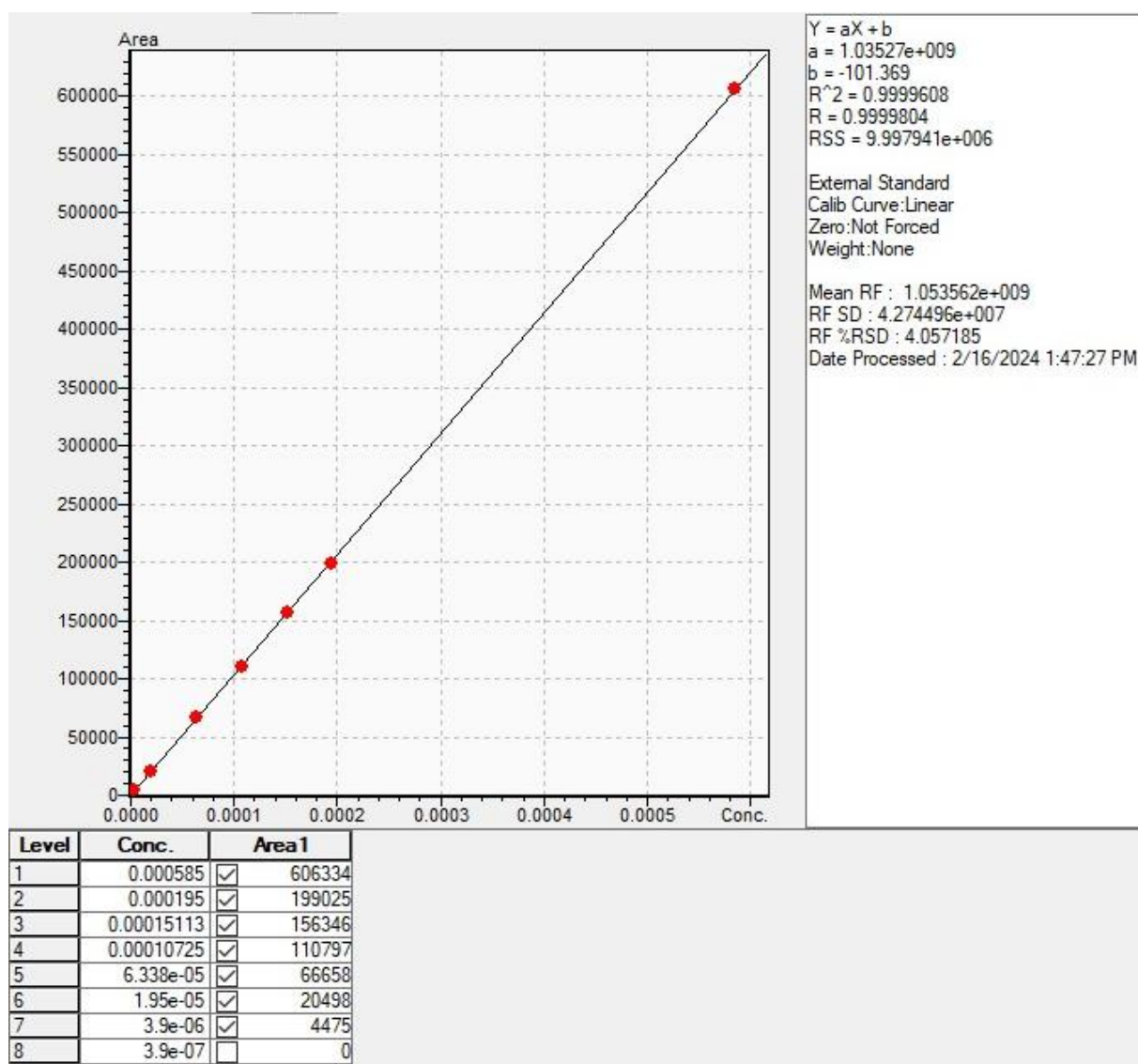


Figure S61. HPLC calibration curve for 9-butyl-9*H*-fluoren-9-ol for data in **Table S9** (Retention Time = 3.976 min).

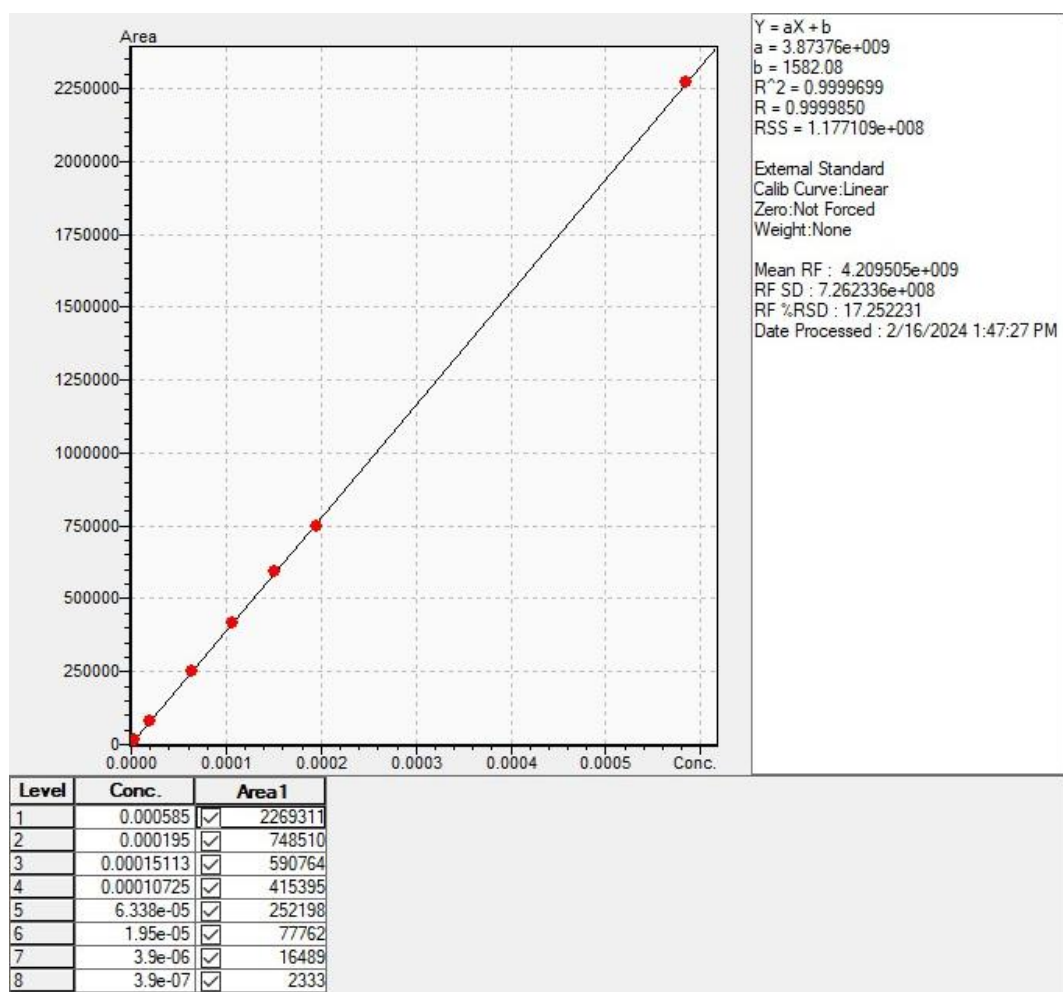


Figure S62. HPLC calibration curve for 9-butyldiene-9*H*-fluorene for data in **Table S9** (Retention Time = 6.407 min).

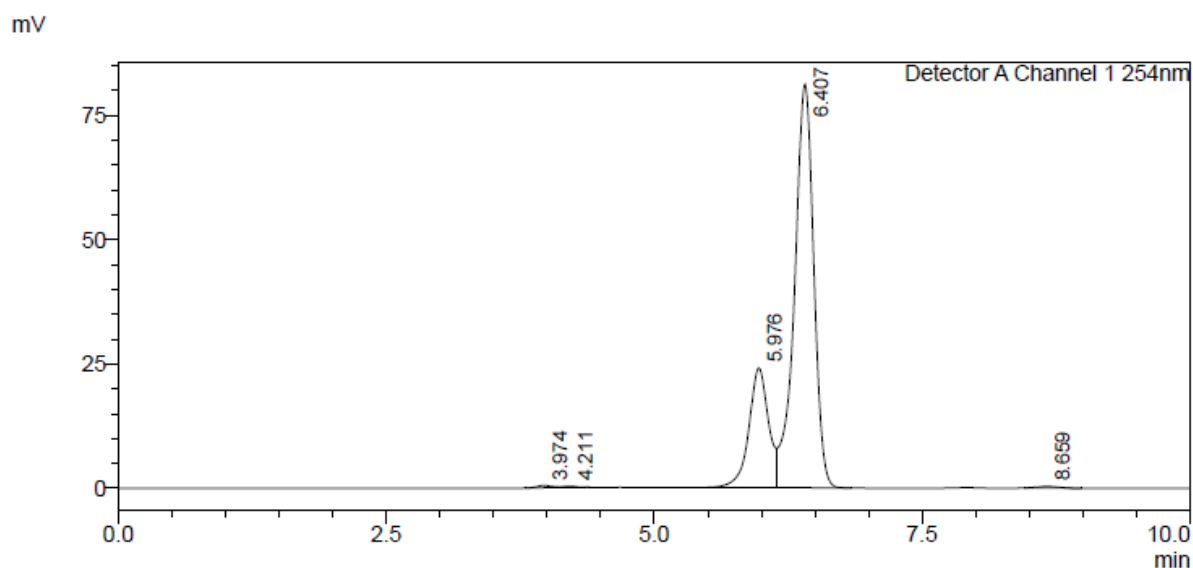


Figure S63 HPLC chromatogram of sample 0 from **Table S9**. For all the chromatograms of sample 1 to 14 in **Table S9**, see Section 13.

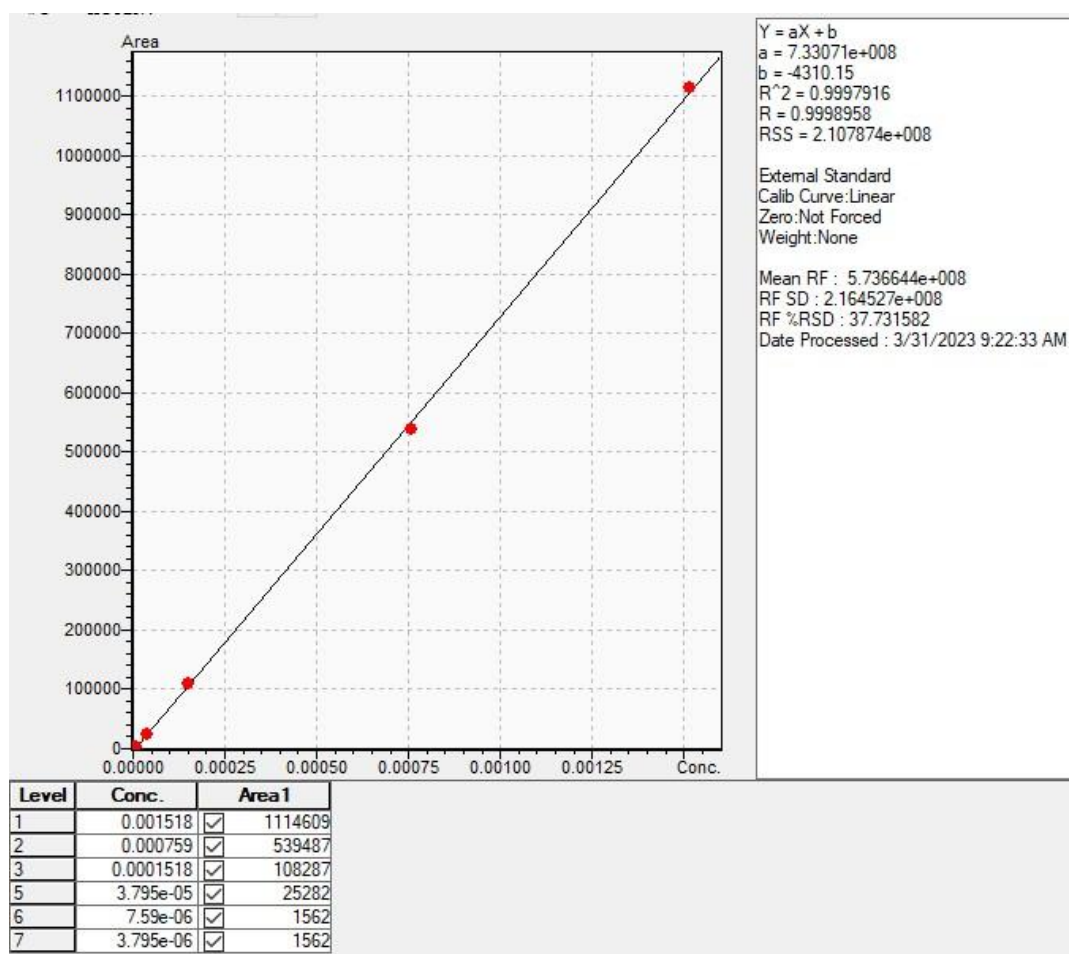


Figure S64. HPLC calibration curve for 3-butyl-4-(4-nitrophenyl)-1-(tosylmethyl)-1*H*-imidazol-3-ium-5-olate for data in

Table S10 (Retention Time = 26.980 min).

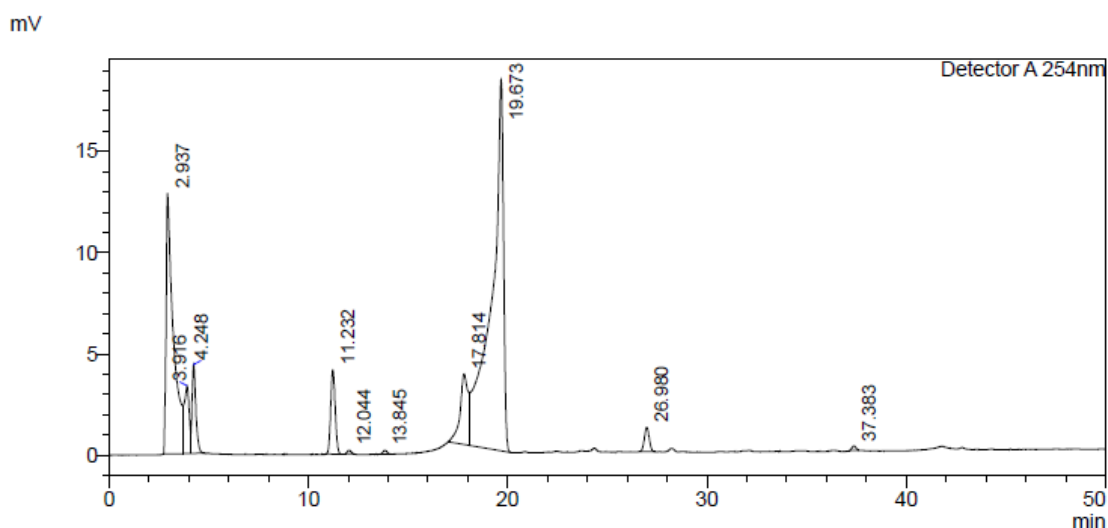


Figure S65 HPLC analysis of sample 15 from

Table S10. For all the chromatograms of sample 16 to 20 in, see Section 13.

4.5. Stirred and unstirred reactions yield comparison

Reactions were performed on the robot according to the procedure in **Reaction yield map for E1 (synthesis of 9-butylidene-9*H*-fluorene)** at 26 °C with and without stirring. Stirring was performed by placing magnetic stirring bar in every vial and magnetic stirrer inside thermostatic oven for a reaction time.

Table S11. Initial concentrations of 9-butyl-9*H*-fluoren-9-ol (**13a**) and HBr for E1 reaction and yields from stirred and unstirred experiments.

Entry	[13a] ₀ (mM)	[HBr] ₀ (mM)	Stirred yield (%)	Unstirred yield (%)
0	11.63	8.293	73.7	75.3
1	11.63	8.293	76.1	75.7
2	11.63	8.293	75.9	75.7
3	4.875	1.124	44.4	45.2
4	4.875	1.124	44.2	42.6
5	4.875	1.124	44.6	40.7
6	11.63	5.903	76.1	73.9
7	11.63	5.903	75.9	74.6
8	11.63	5.903	75.8	74.3
9	1.500	2.490	61.8	61.1
10	1.500	2.490	61.6	62.5
11	1.500	2.490	61.9	61.6
12	4.875	3.514	72.0	71.0
13	4.875	3.514	71.5	70.9
14	4.875	3.514	71.2	71.1
15	8.250	8.980	69.7	73.7
16	8.250	8.980	70.3	74.3
17	8.250	8.980	74.5	74.2
18	15.00	4.879	64.6	59.7
19	15.00	4.879	64.8	59.9
20	15.00	4.879	64.8	60.4

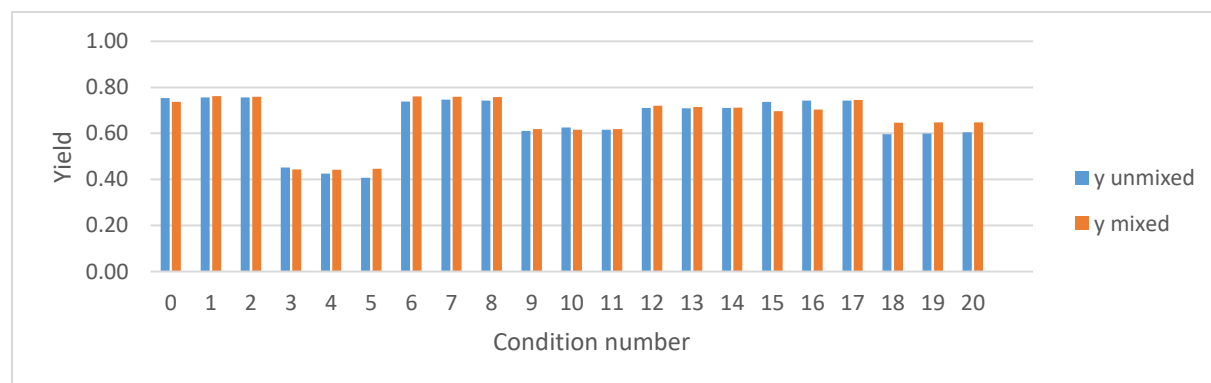


Figure S66. Graphical representation of data from **Table S11**.

4.6. ^1H NMR benchmarks for Hantzsch reaction

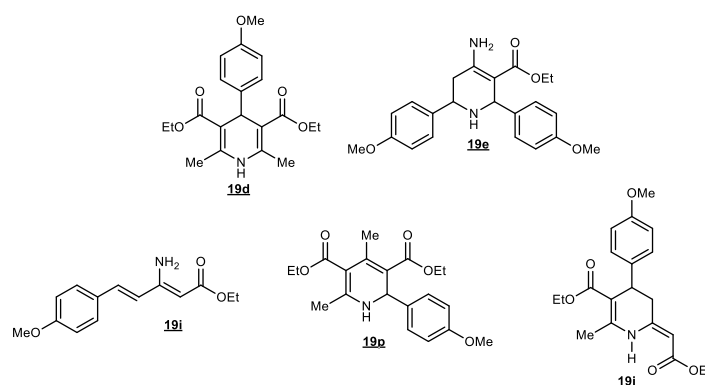


Figure S67. Substances for which the concentration was calculated from the crude ^1H NMR spectra.

Manual reactions:

Stock solutions of *p*-methoxybenzaldehyde (0.3333 M), ethyl acetoacetate (0.3333 M) and ammonium acetate (1.0000 M) in ethanol were prepared. Solvent and the stock solutions were then transferred to pressure tubes in following order and amounts:

- for conditions corresponding to maximum yield of **19d**: ethanol (270 μL), *p*-methoxybenzaldehyde (30 μL), ethyl acetoacetate (300 μL) and after 15 min ammonium acetate (400 μL)
- for conditions corresponding to maximum yield of **19e**: ethanol (405 μL), *p*-methoxybenzaldehyde (165 μL), ethyl acetoacetate (30 μL) and after 15 min ammonium acetate (400 μL)

Then, pressure tubes were sealed and the reactions were stirred at 80 $^{\circ}\text{C}$ for 48 h. After this time, reactions were cooled down to room temperature and UV–Vis spectra on NanoDrop were collected. Then, aliquot from every crude reaction mixture was taken and dried under vacuum. Subsequently, samples were quantitatively transferred into 1 mL volumetric flasks by dissolving in CDCl_3 with internal standard (0.2 vol% of dibenzyl ether) and ^1H NMR spectra were measured.

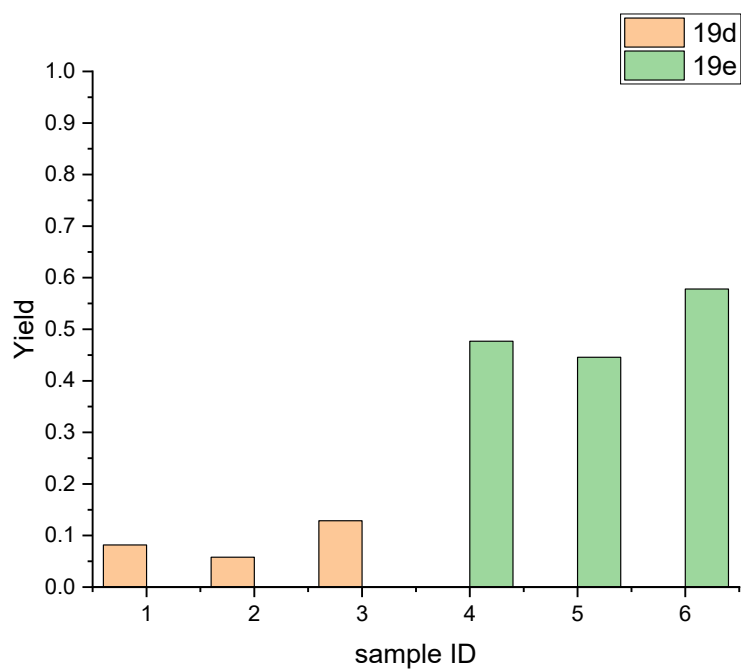


Figure S68. Yield of reactions performed manually.

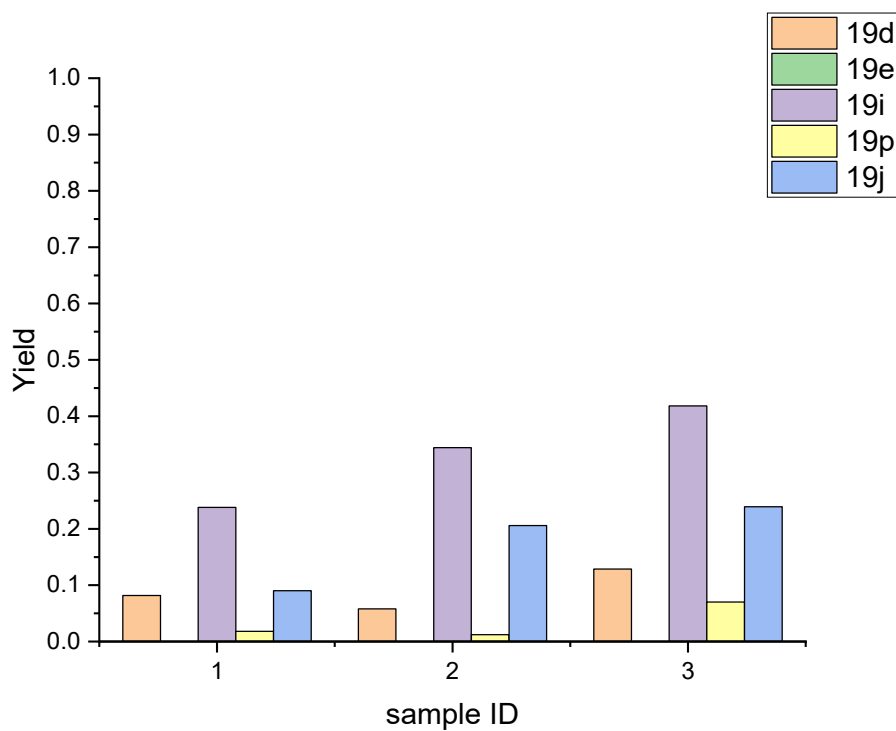


Figure S69. Yield of reactions performed manually.

4.7. HPLC data for Hantzsch reaction maps

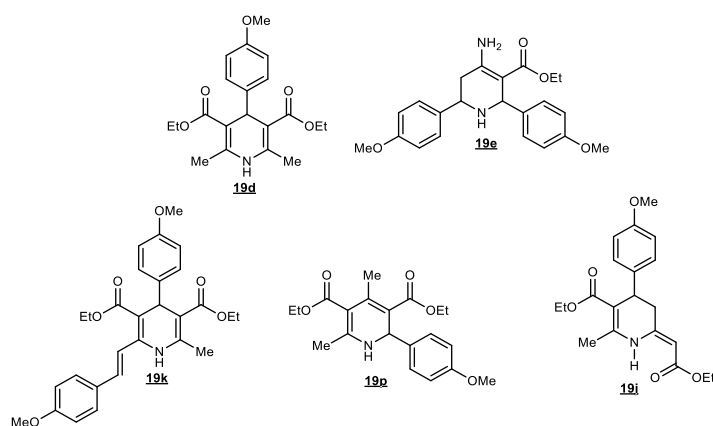


Figure S70. Substances used for calibrations in HPLC.

Robotic reactions and HPLC analysis:

Reactions were performed on the robot according to the procedure described in **Reaction yield map for Hantzsch reaction**. Each crude reaction mixture was diluted 20 times using robotic platform. Each diluted sample was filtered through 0.2 μm syringe filter and subjected to HPLC analysis. The injection volume was 50 μL , and the eluent was 10 mM aqueous ammonium acetate solution to acetonitrile (gradient from 50% to 10%), on JAIGEL-ODS-BP-A, SP-120-10 (Reversed Phase Column, 6.0 $\times$ 250 mm) at 30 $^{\circ}\text{C}$, at overall flow rate of 1.8 mL/min.

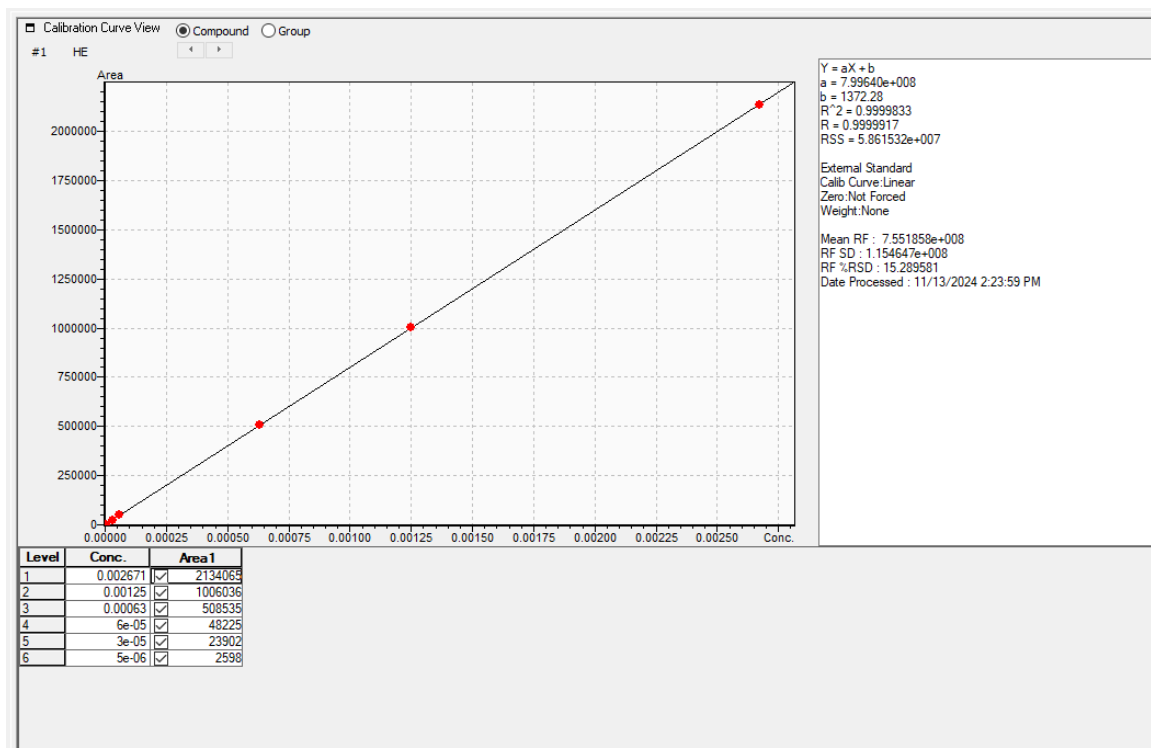


Figure S71. HPLC calibration curve for compound **19d** (Retention Time = 16.409 min).

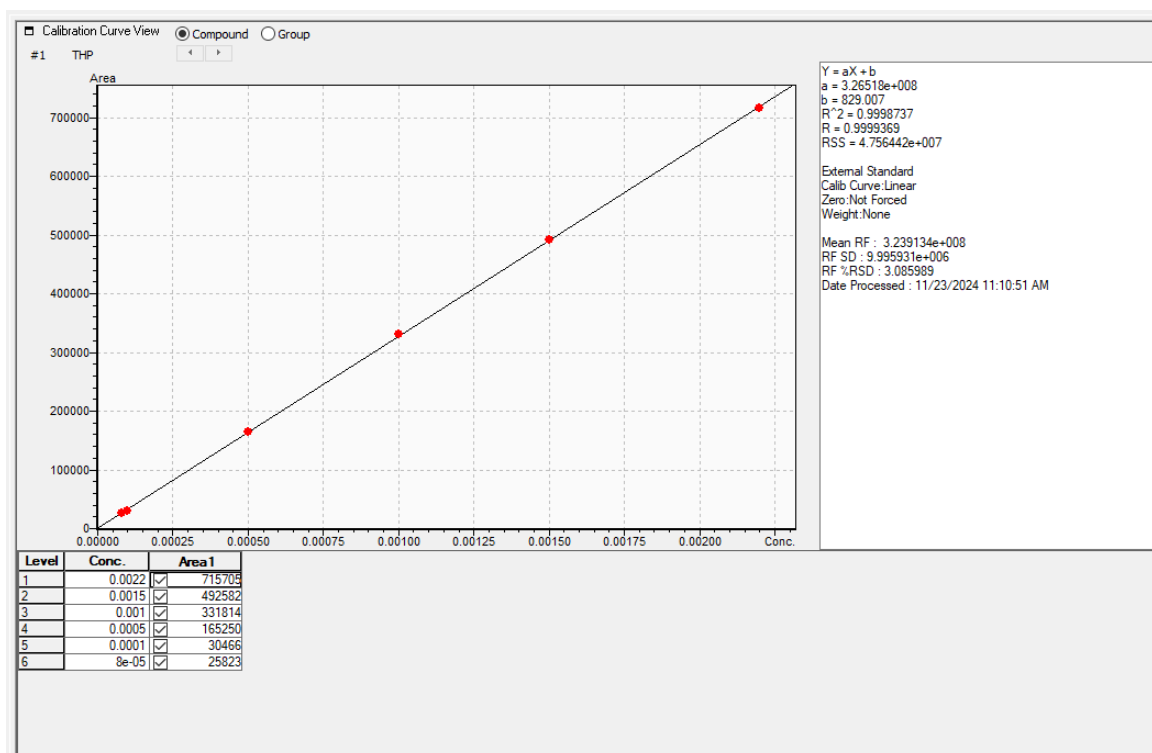


Figure S72. HPLC calibration curve for compound **19e** (Retention Time = 12.191 min).

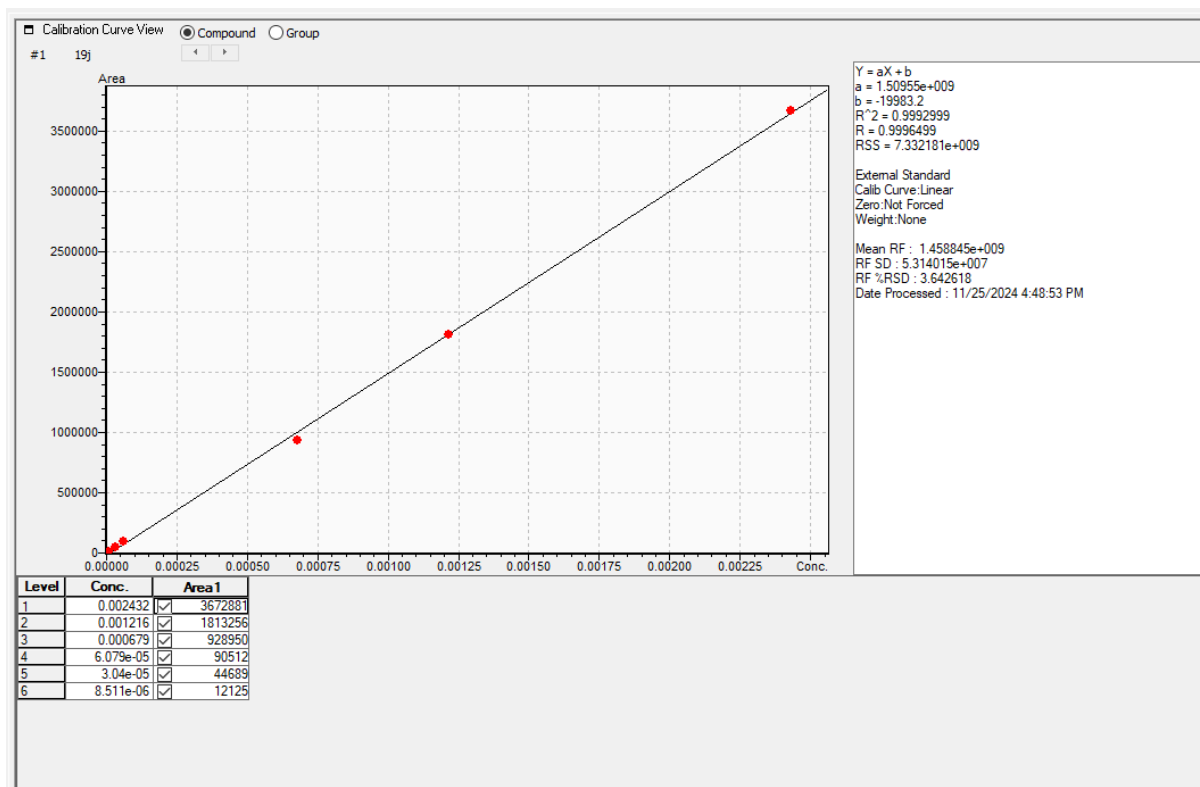


Figure S73. HPLC calibration curve for compound **19i** (Retention Time = 22.389 min).

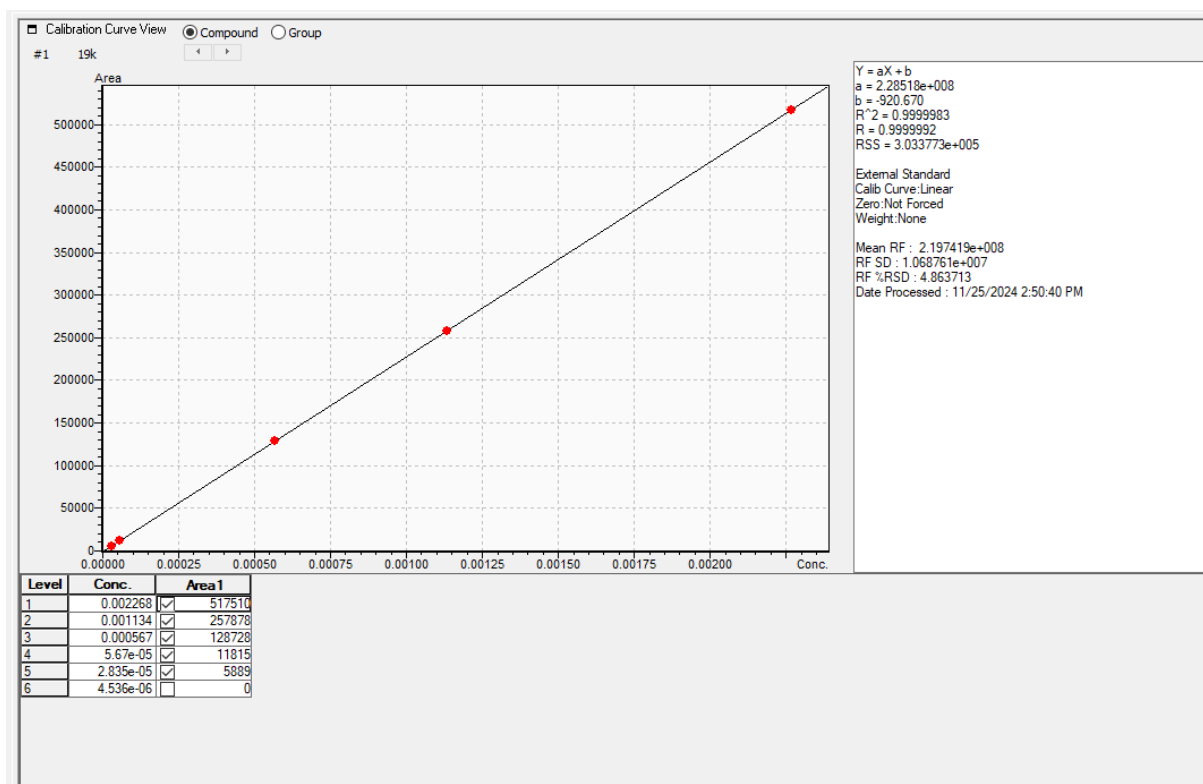


Figure S74. HPLC calibration curve for compound **19k** (Retention Time = 23.695 min).

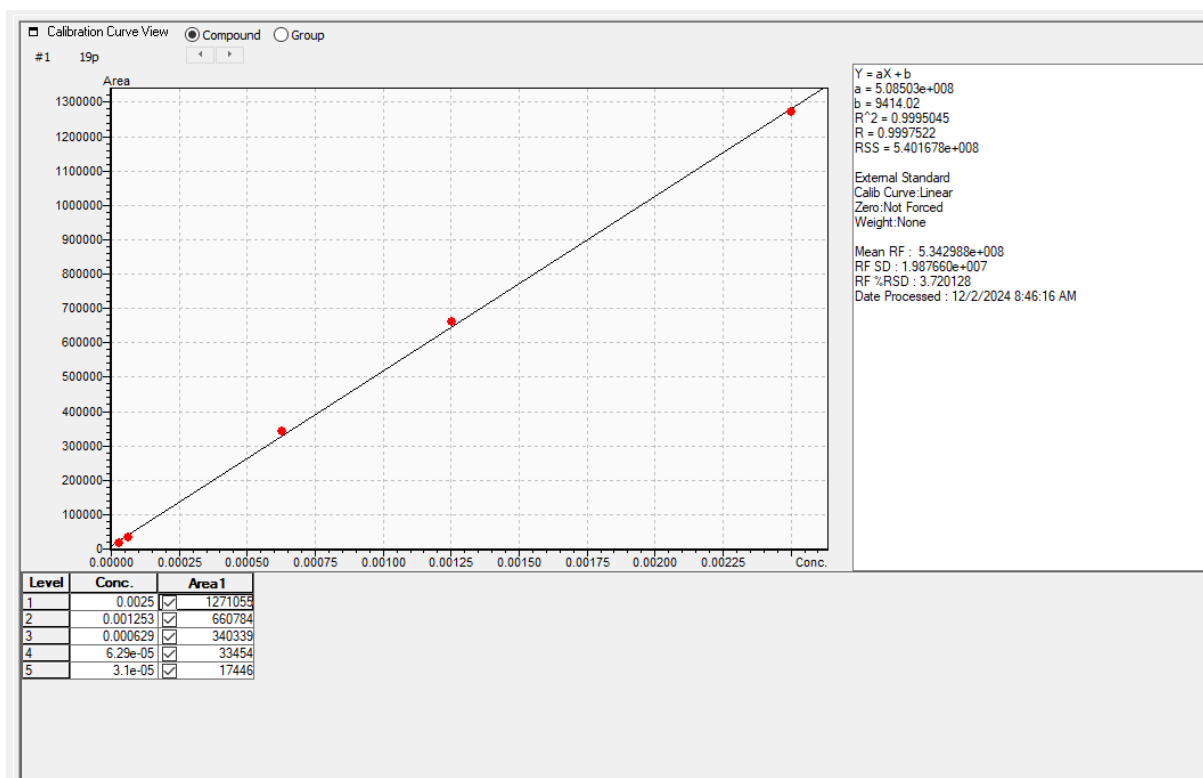


Figure S75. HPLC calibration curve for compound **19p** (Retention Time = 16.384 min).

Because of coelution of compounds **19d** and **19p** the ratio between them was established from NMR crude spectrum using internal standard (Bn_2O). As **19p** is a minor component, with yield below 3% for reaction at 80 °C, we used HPLC calibration for **19d** to calculate its concentration.

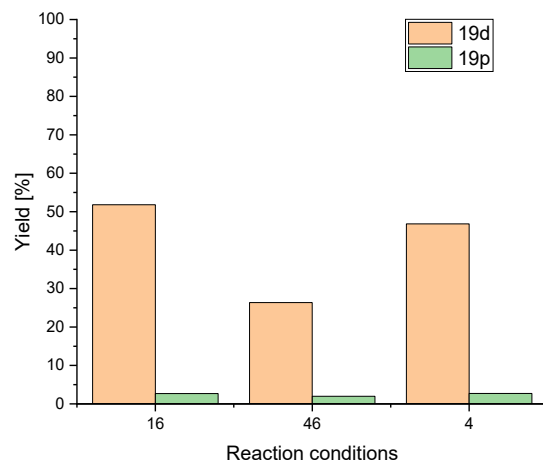


Figure S76. Yields of products **19d** and **19p** from NMR for reaction conditions: 16, 46, 4 at 80 °C.

Table S12. Initial concentrations (mol/L) of substrates **19a**, **19b** and **19c** used for Hantzsch reaction at 80 °C, and average yields of products **19d**, **19e**, **19k** and **19j** calculated from concentrations established by HPLC.

index	19c	19a	19b	19d	19e	19k	19j
1	0.011	0.011	0.106	0.0	25.6	0.0	0.0
4	0.011	0.101	0.105	63.5	0.0	13.0	3.0
7	0.027	0.054	0.105	6.8	28.3	18.0	1.8
10	0.069	0.075	0.105	25.1	16.5	46.9	2.3
13	0.101	0.011	0.105	0.9	63.3	0.0	0.0
16	0.100	0.100	0.104	41.0	10.3	61.8	2.2
19	0.011	0.011	0.403	0.0	35.1	0.0	2.5
22	0.011	0.100	0.400	10.4	9.5	11.6	9.3
25	0.027	0.054	0.399	2.3	37.3	9.2	2.7
28	0.069	0.074	0.397	3.1	44.3	14.8	4.0
31	0.100	0.011	0.400	3.1	65.0	1.0	2.1
34	0.100	0.100	0.399	6.2	37.0	30.6	5.6
37	0.011	0.057	0.105	11.1	18.0	9.3	3.1
40	0.057	0.011	0.105	0.0	46.3	2.8	6.7
43	0.057	0.057	0.105	9.9	29.8	28.0	2.6
46	0.057	0.100	0.104	36.4	7.3	50.4	1.6
49	0.100	0.057	0.104	11.1	27.1	57.2	3.5
52	0.011	0.011	0.258	0.0	23.7	0.0	0.0
55	0.011	0.057	0.256	4.4	26.3	6.3	4.2
58	0.011	0.100	0.256	15.7	8.0	13.8	7.1
61	0.057	0.011	0.256	0.3	61.4	0.0	4.7
64	0.056	0.056	0.255	3.9	41.4	13.5	2.9
67	0.056	0.099	0.254	6.7	31.9	29.6	4.1
70	0.100	0.011	0.256	2.7	66.8	2.3	0.0
73	0.099	0.056	0.254	2.6	42.1	18.6	3.0
76	0.099	0.099	0.254	8.1	33.3	38.3	4.9
79	0.011	0.057	0.401	3.9	20.0	6.6	5.2
82	0.057	0.011	0.401	1.4	65.1	0.0	13.2
85	0.056	0.056	0.397	2.2	42.3	9.5	3.3
88	0.056	0.099	0.397	4.5	32.9	22.1	4.7
91	0.099	0.056	0.397	2.0	46.8	14.2	2.8

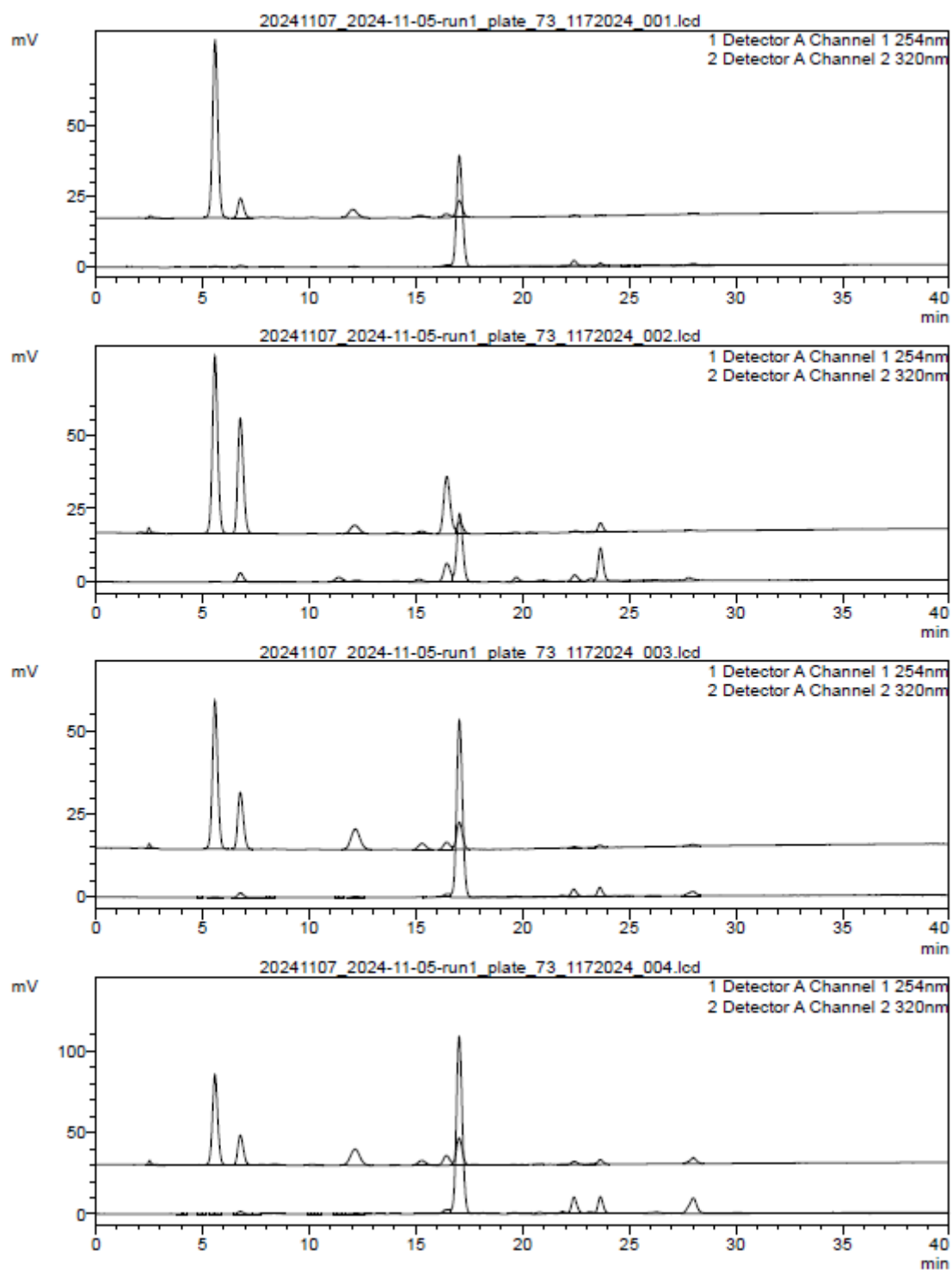


Figure S77 Chromatograms for the conditions 25, 10, 64, 34 from **Table S12**, top to bottom. For all the chromatograms of the remaining conditions in **Table S12**, see **Section 13**.

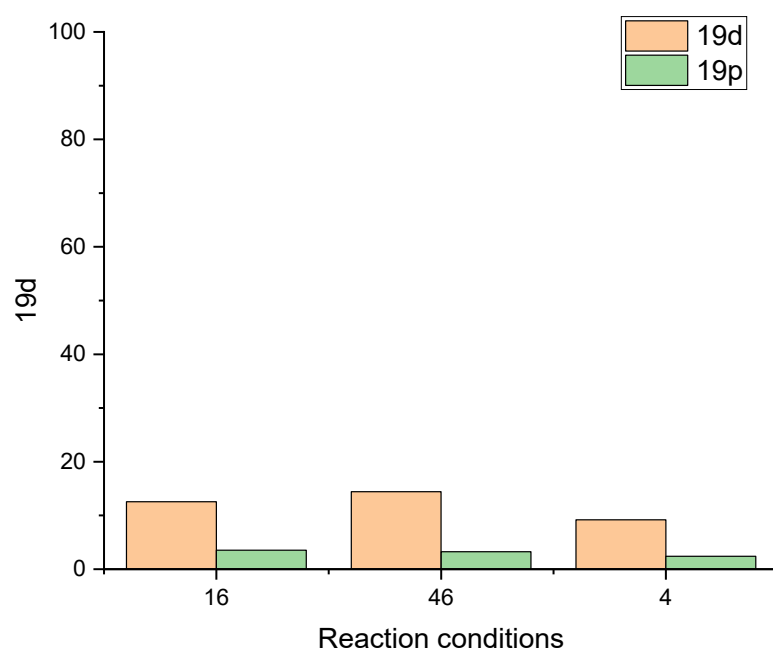


Figure S78. Yields of products **19d** and **19p** from NMR for reaction conditions: 16, 46, 4 at 26 °C.

Table S13. Initial concentrations (mol/L) of substrates **19a**, **19b** and **19c** used for Hantzsch reaction at 26 °C, and average yields of products **19d**, **19e**, **19k** and **19j** calculated from concentrations established by HPLC.

index	19c	19a	19b	19d	19e	19k	19j
1	0.011	0.011	0.106	0.0	5.8	0.0	0.0
4	0.011	0.101	0.105	17.0	4.1	0.0	0.0
7	0.027	0.054	0.105	3.0	16.5	0.7	0.0
10	0.069	0.075	0.105	7.4	16.1	0.0	0.0
13	0.101	0.011	0.105	0.0	24.5	0.0	0.0
16	0.100	0.100	0.104	12.1	12.9	0.0	0.0
19	0.011	0.011	0.403	0.0	12.3	0.0	0.0
22	0.011	0.100	0.400	5.8	36.7	14.1	0.0
25	0.027	0.054	0.399	0.6	30.1	4.0	0.0
28	0.069	0.074	0.397	1.6	34.4	1.5	0.0
31	0.100	0.011	0.400	0.0	32.0	0.0	0.0
34	0.100	0.100	0.399	2.9	41.4	1.0	0.0
37	0.011	0.057	0.105	5.2	11.7	0.0	0.0
40	0.057	0.011	0.105	0.0	14.4	0.0	0.0
43	0.057	0.057	0.105	4.2	17.0	0.0	0.0
46	0.057	0.100	0.104	9.6	12.6	0.0	0.0
49	0.100	0.057	0.104	4.5	16.5	0.0	0.0
52	0.011	0.011	0.258	0.0	9.8	0.0	0.0
55	0.011	0.057	0.256	1.6	22.8	11.3	0.0
58	0.011	0.100	0.256	8.7	25.0	2.9	0.0
61	0.057	0.011	0.256	0.0	21.4	0.0	0.0
64	0.056	0.056	0.255	1.2	28.3	1.2	0.0
67	0.056	0.099	0.254	4.0	35.5	1.5	0.0
70	0.100	0.011	0.256	0.0	29.7	0.0	0.0
73	0.099	0.056	0.254	1.1	27.7	0.0	0.0
76	0.099	0.099	0.254	4.9	35.2	0.5	0.0
79	0.011	0.057	0.401	1.0	28.0	13.1	0.0
82	0.057	0.011	0.401	0.0	25.3	0.0	0.0
85	0.056	0.056	0.397	0.7	30.7	0.0	0.0
88	0.056	0.099	0.397	2.6	40.6	2.3	0.0
91	0.099	0.056	0.397	0.4	31.7	1.0	0.0

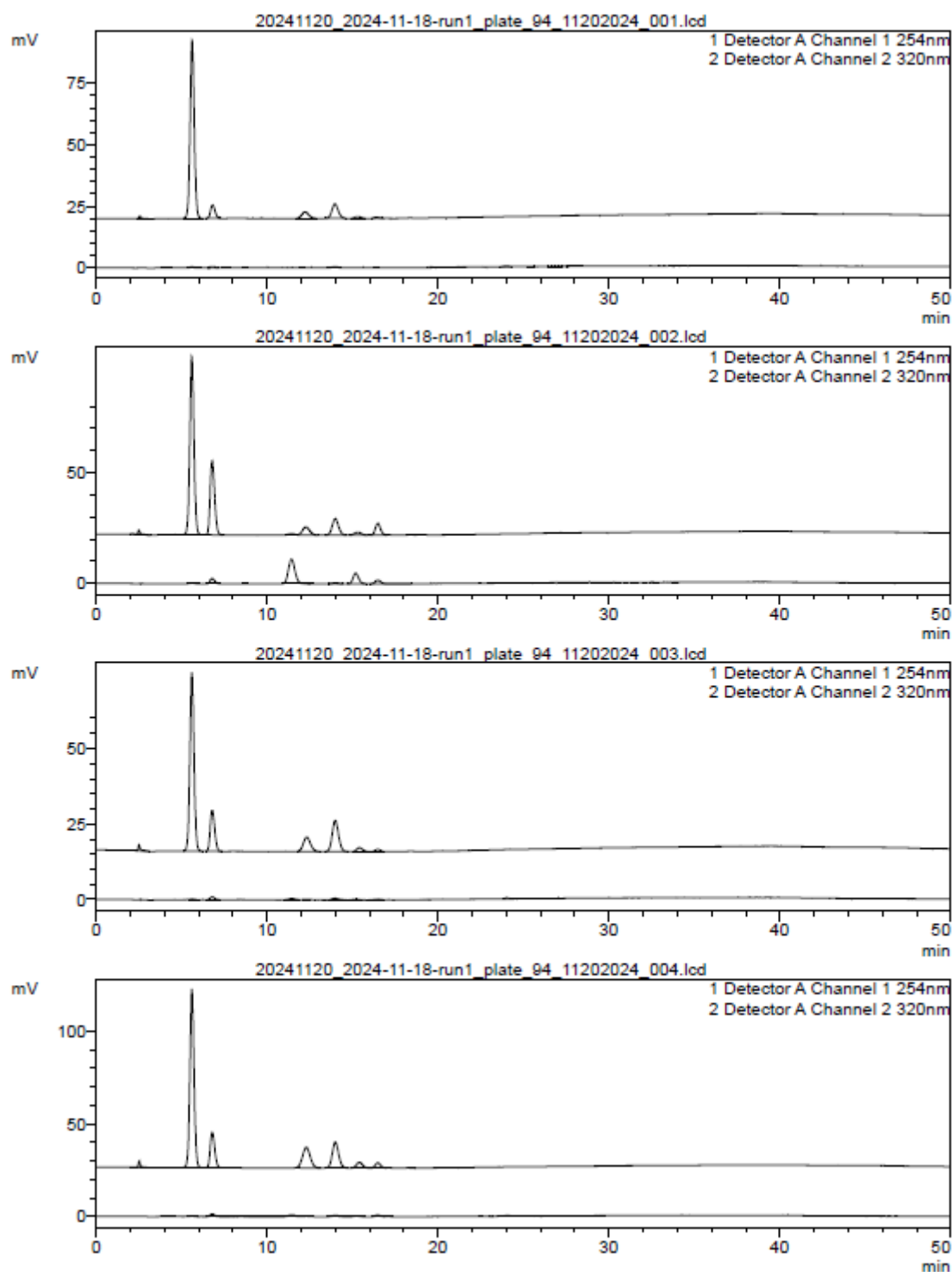


Figure S79. Chromatograms for the conditions 25, 10, 64, 34, from **Table S13**, top to bottom. For all the chromatograms of the remaining conditions in **Table S13**, see **Section 13**.

4.8. Mechanistic study of the Ugi-type Multicomponent Reaction

4.8.1. Global yield maximum

The stock solutions of TosMIC (1.2 M), *n*-butylamine (2.0 M), *p*-nitrobenzaldehyde (1.2 M), *p*TSA (1.5 M) in DMF were prepared, and then, they were used by automatic liquid handler to prepare reaction with the following initial substrates concentration: *p*-nitrobenzaldehyde 0.27 M, *p*TSA 0.120 M, *n*-butylamine 0.12 M and TosMIC 0.30 M. The order of compounds addition was as follows: DMF, solution of *p*-nitrobenzaldehyde, solution of *p*TSA, solution of *n*-butylamine and solution of TosMIC. The total volume of reaction was 200 μ L. After sealing, the reactions were shaken for 5 min (250 rpm) and placed in a thermostatic oven at 26 $^{\circ}$ C for 16 h. After 1 h and 16 h an aliquot was withdrawn for analysis on APCI-MS.

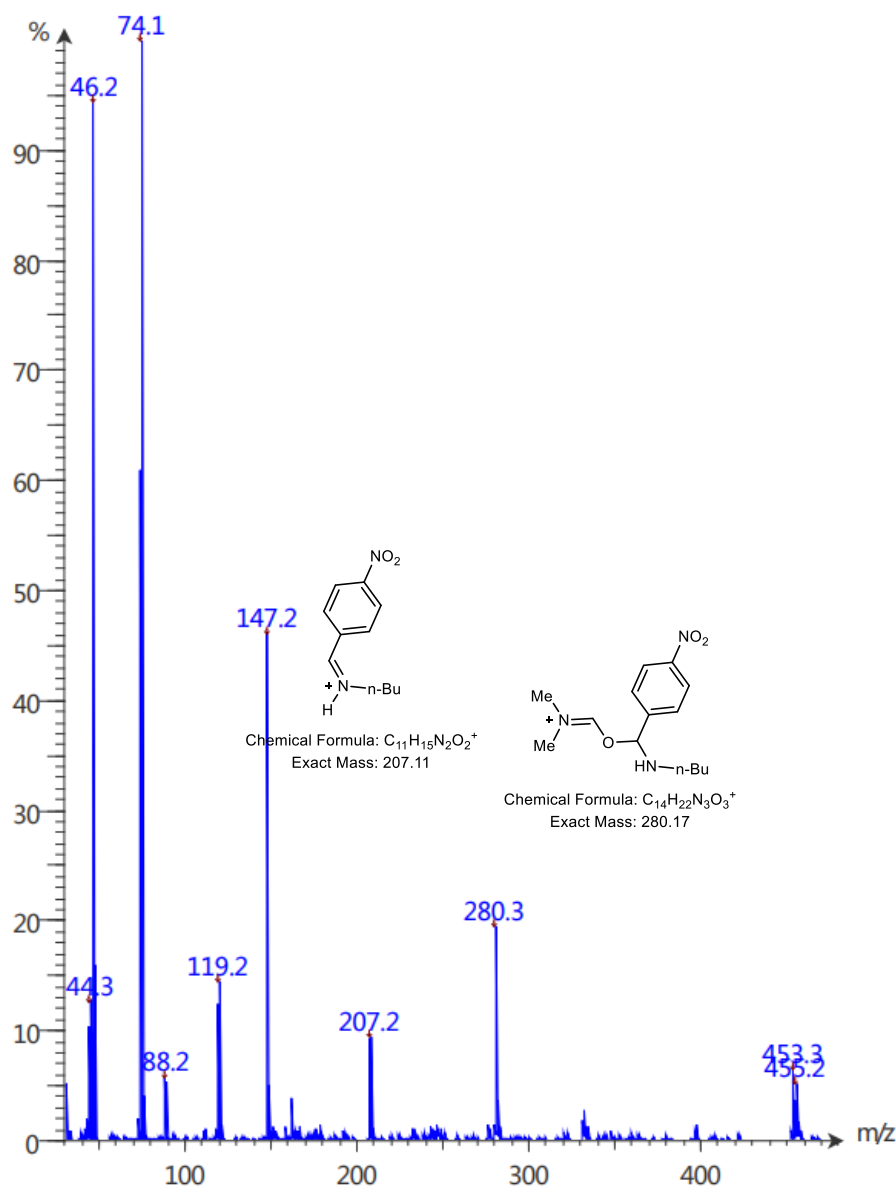


Figure S80. APCI-MS spectrum of multicomponent reaction in global yield maximum after 1 h.

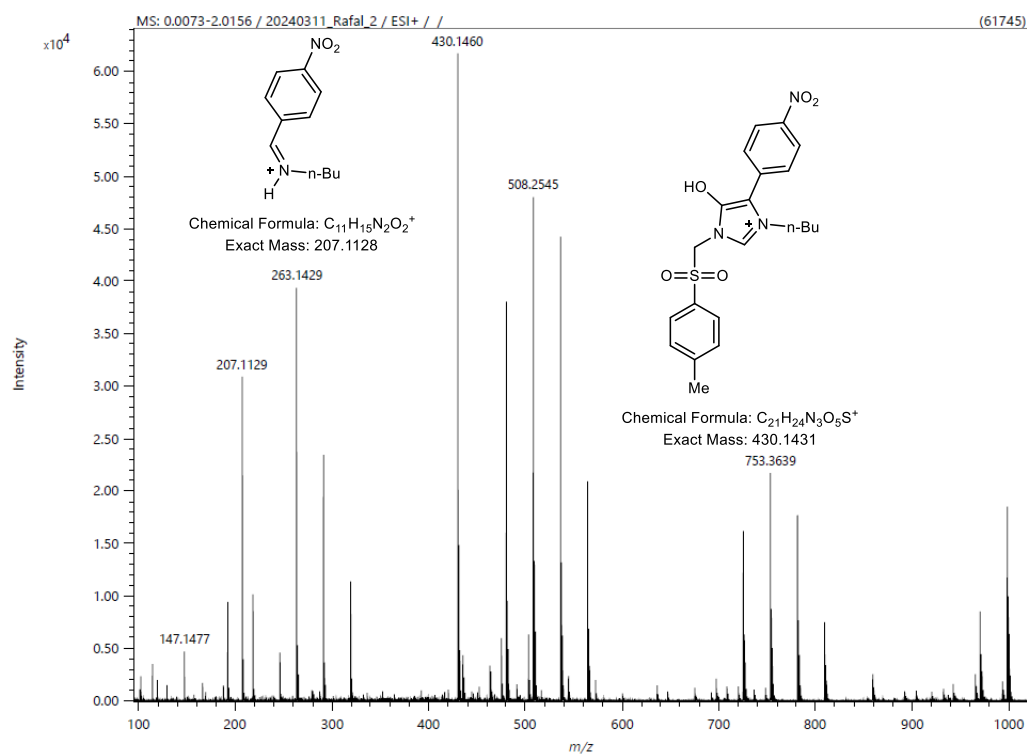


Figure S81. ESI-MS spectrum of multicomponent reaction in global yield maximum after 1 h.

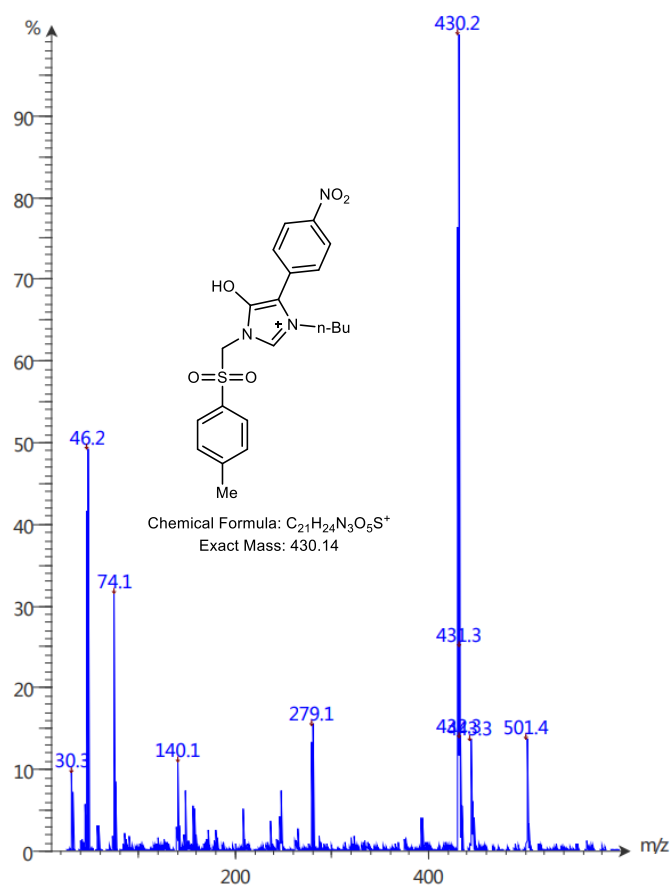


Figure S82. APCI-MS spectrum of multicomponent reaction in global yield maximum after 16 h.

4.8.2. Local yield maximum

The stock solutions of TosMIC (1.2 M), *n*-butylamine (2.0 M), *p*-nitrobenzaldehyde (1.2 M), *p*TSA (1.5 M) in DMF, were prepared, and then they were used by automatic liquid handler to prepare reaction with the following initial substrates concentration: *p*-nitrobenzaldehyde 0.15 M, *p*TSA 0.286 M, *n*-butylamine 0.30 M and TosMIC 0.30 M. The order of compounds addition was as follows: DMF, solution of *p*-nitrobenzaldehyde, solution of *p*TSA, solution of *n*-butylamine and solution of TosMIC. The total volume of reaction was 200 μ L. After sealing, the reactions were shaken for 5 min (250 rpm) and placed in a thermostatic oven at 26 $^{\circ}$ C for 16 h. After 1 h and 16 h an aliquot was withdrawn for analysis on APCI-MS.

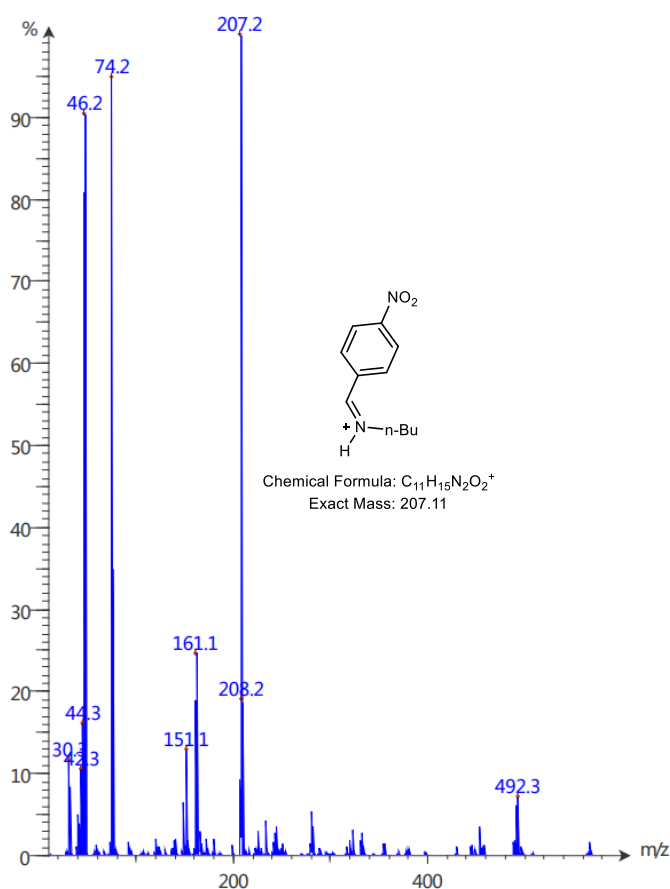


Figure S83. APCI-MS spectrum of multicomponent reaction in local yield maximum after 1 h.

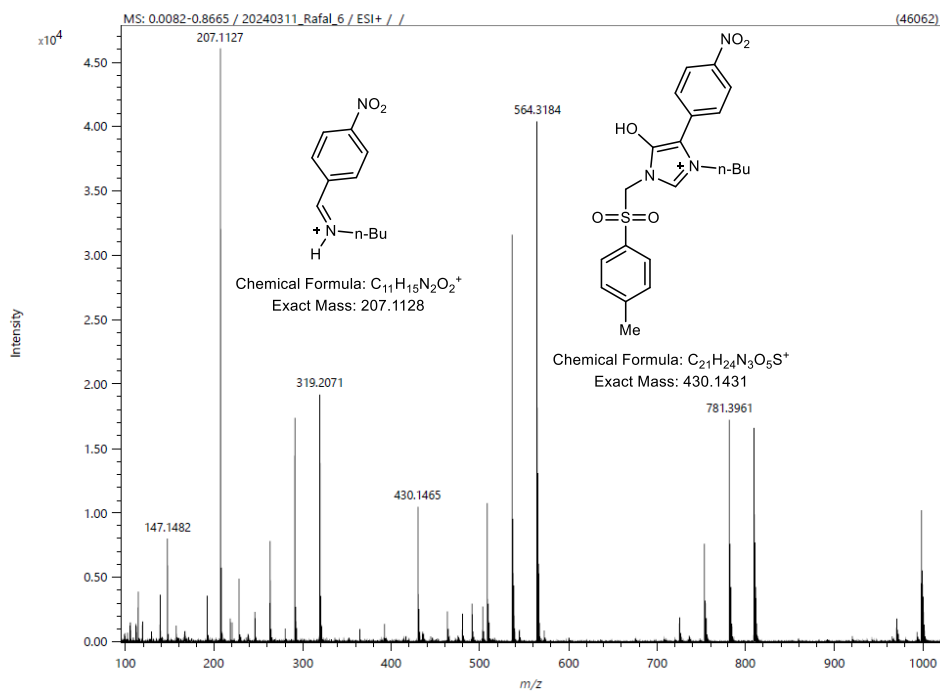


Figure S84. ESI–MS spectrum of multicomponent reaction in local yield maximum after 1 h.

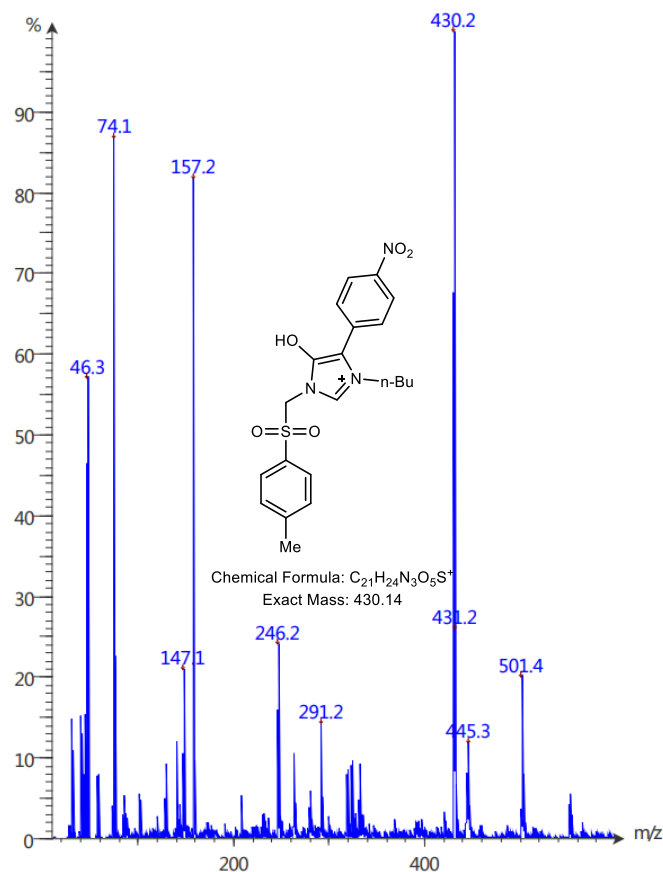


Figure S85. APCI–MS spectrum of multicomponent reaction in local yield maximum after 16 h.

4.8.3. Charge tag strategy

In an attempt to observe the reactive intermediates for the four-component Ugi-type reaction, the charge tag strategy for mass spectrometry^{84,92,93} was employed. Specifically, the 4-nitrobenzaldehyde (**16a**) component was replaced by the salt 4-formyl-*N,N,N*-trimethylbenzenaminium iodide. In principle, the cationic NMe_3^+ group introduces a positive charge to any intermediates and products formed. Hence, an ordinarily unobservable intermediate of minute concentration can potentially give a mass spectral signal upon charge tagging.

Global yield maximum:

The stock solutions of TosMIC (1.2 M), *n*-butylamine (2.0 M), 4-formyl-*N,N,N*-trimethylbenzenaminium iodide (0.6 M), *p*TSA (1.5 M) in DMF were prepared, and then they were used by automatic liquid handler to prepare reaction with the following initial substrates concentration: 4-formyl-*N,N,N*-trimethylbenzenaminium iodide 0.27 M, *p*TSA 0.120 M, *n*-butylamine 0.12 M and TosMIC 0.30 M. The order of compounds addition was as follows: DMF, solution of 4-formyl-*N,N,N*-trimethylbenzenaminium iodide, solution of *p*TSA, solution of *n*-butylamine and solution of TosMIC. The total volume of reaction was 200 μL . After sealing, the reactions were shaken for 5 min (250 rpm) and placed in a thermostatic oven at 26 °C for 16 h. After 1 h an aliquot was withdrawn for analysis on ESI-MS.

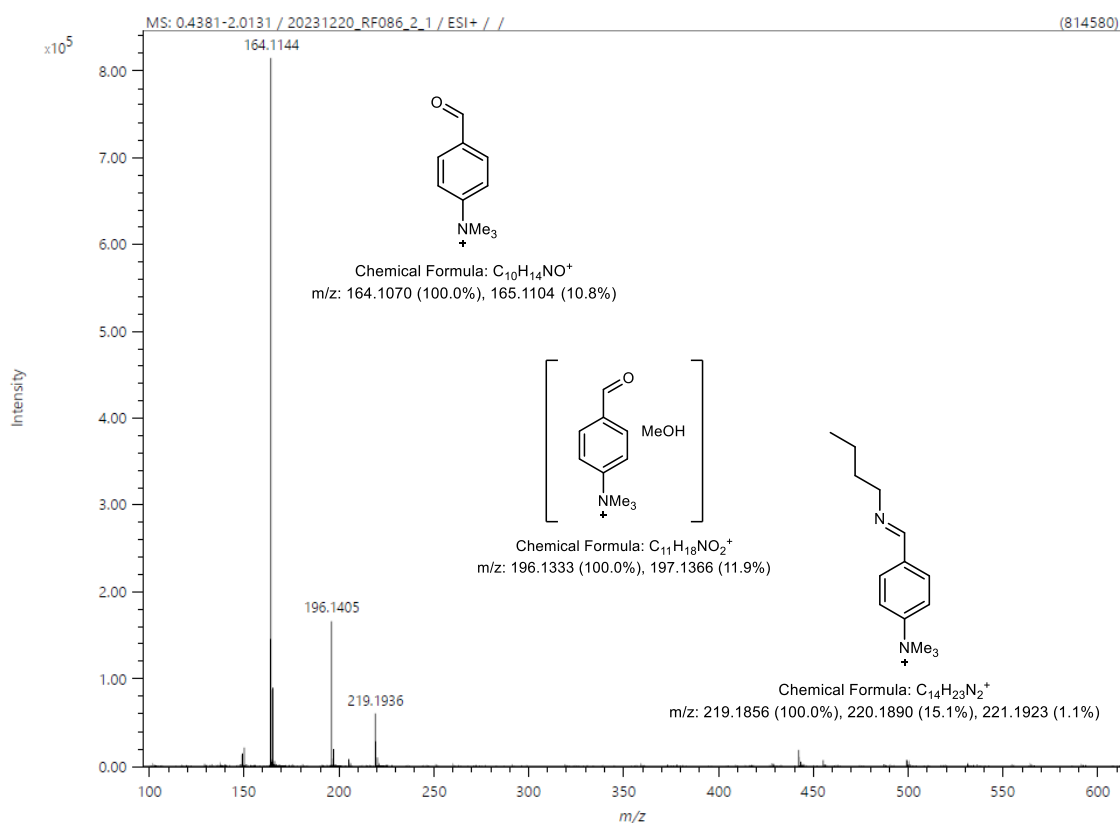


Figure S86. ESI-MS spectrum of multicomponent reaction in global yield maximum after 1 h.

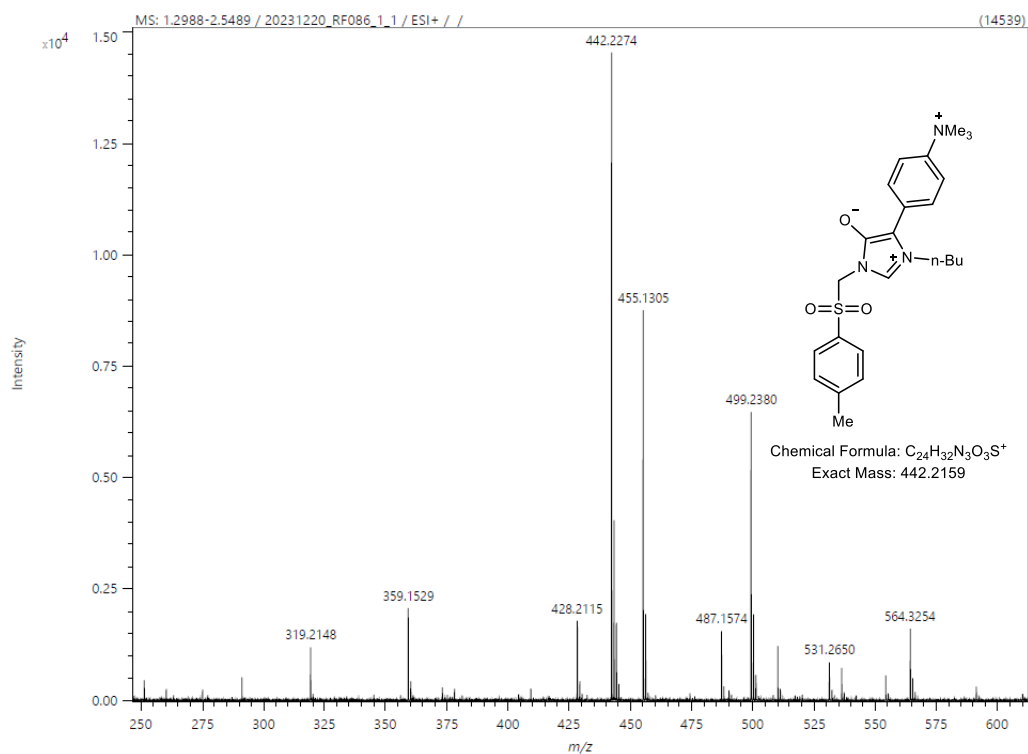


Figure S87. ESI-MS spectrum of multicomponent reaction in global yield maximum after 1 h (zoom in).

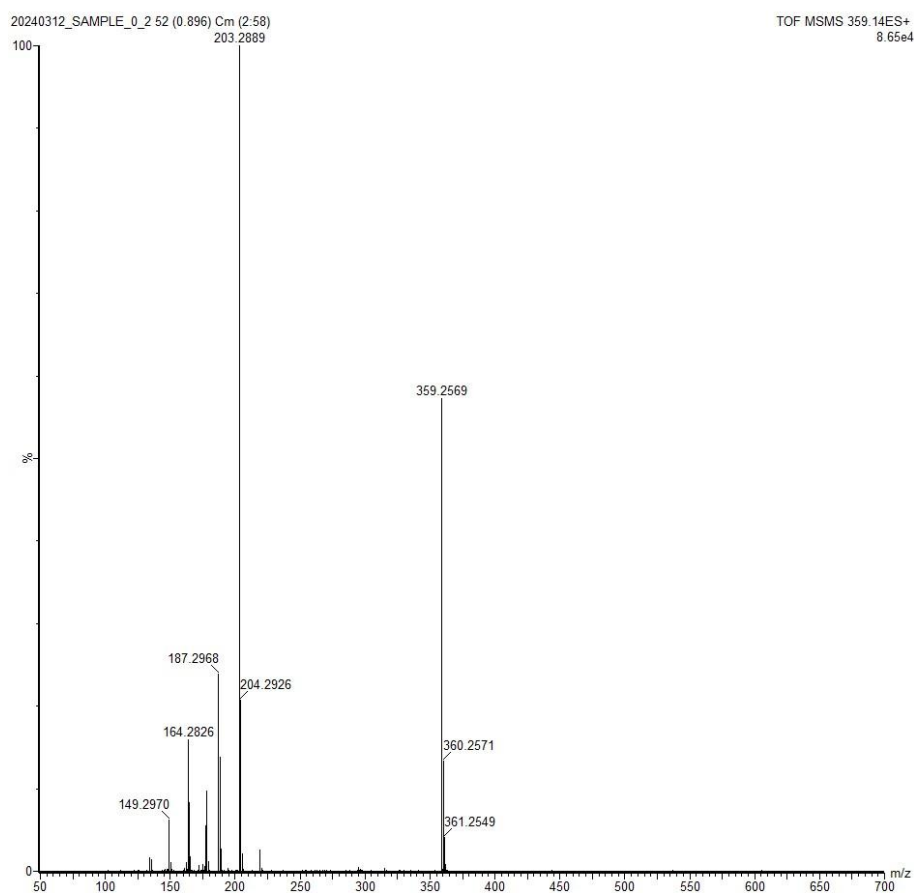


Figure S88. MS/MS experiment for ion 359.14 in global yield maximum after 1 h.

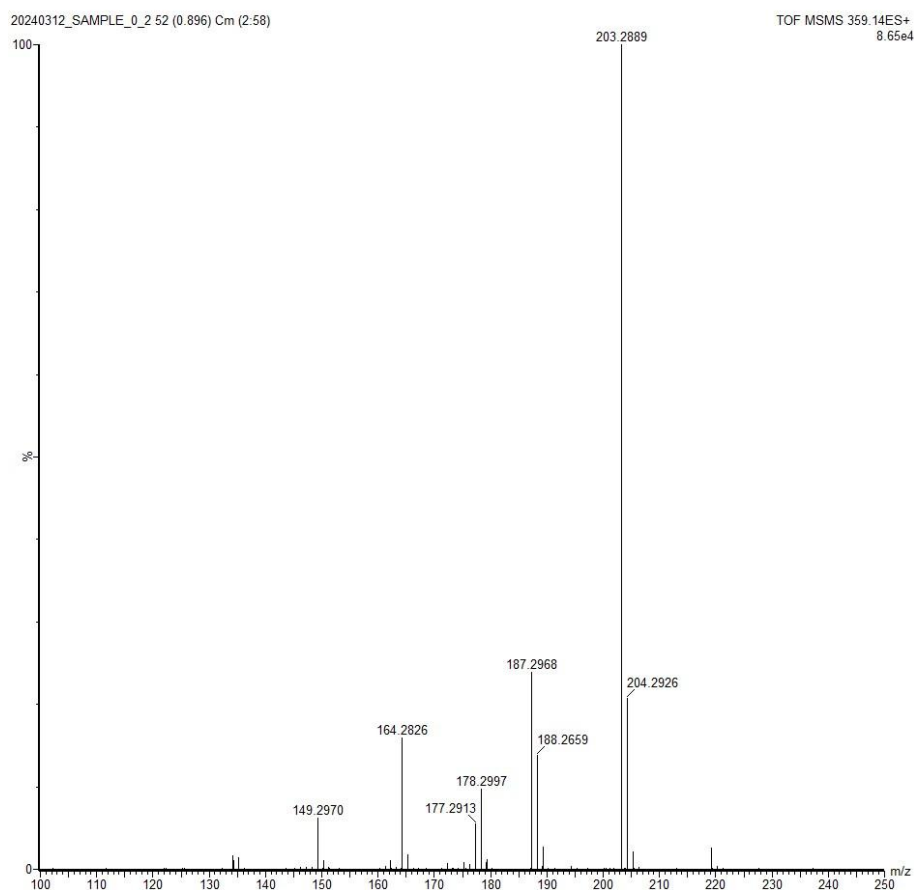


Figure S89. MS/MS experiment for ion 359.14 in global yield maximum after 1 h (zoom in).

Local yield maximum:

The stock solutions of TosMIC (1.2 M), *n*-butylamine (2.0 M), 4-formyl-*N,N,N*-trimethylbenzenaminium iodide (0.6 M), *p*TSA (1.5 M) in DMF, were prepared and then they were used by automatic liquid handler to prepare reaction with the following initial substrates concentration: 4-formyl-*N,N,N*-trimethylbenzenaminium iodide 0.15 M, *p*TSA 0.286 M, *n*-butylamine 0.30 M and TosMIC 0.30 M. The order of compounds addition was as follows: DMF, solution of 4-formyl-*N,N,N*-trimethylbenzenaminium iodide, solution of *p*TSA, solution of *n*-butylamine and solution of TosMIC. The total volume of reaction was 200 μ L. After sealing, the reactions were shaken for 5 min (250 rpm) and placed in a thermostatic oven at 26 $^{\circ}$ C for 16 h. After 1 h an aliquot was withdrawn for analysis on ESI-MS.

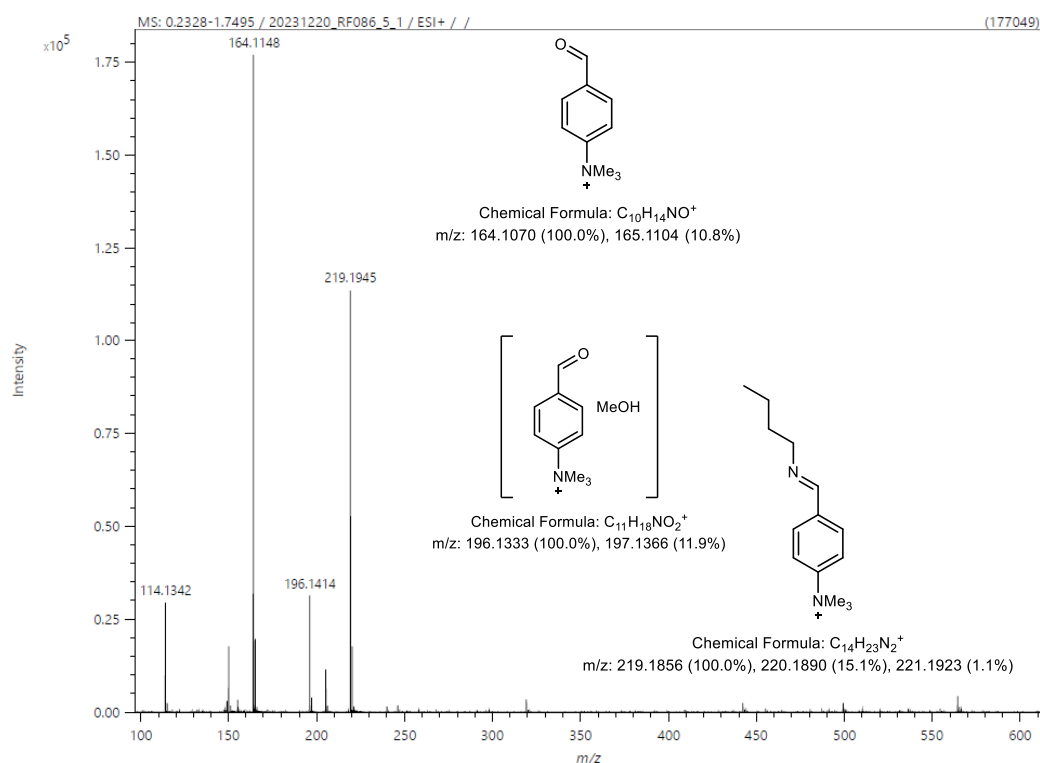


Figure S91. ESI-MS spectrum of multicomponent reaction in a local yield maximum after 1 h.

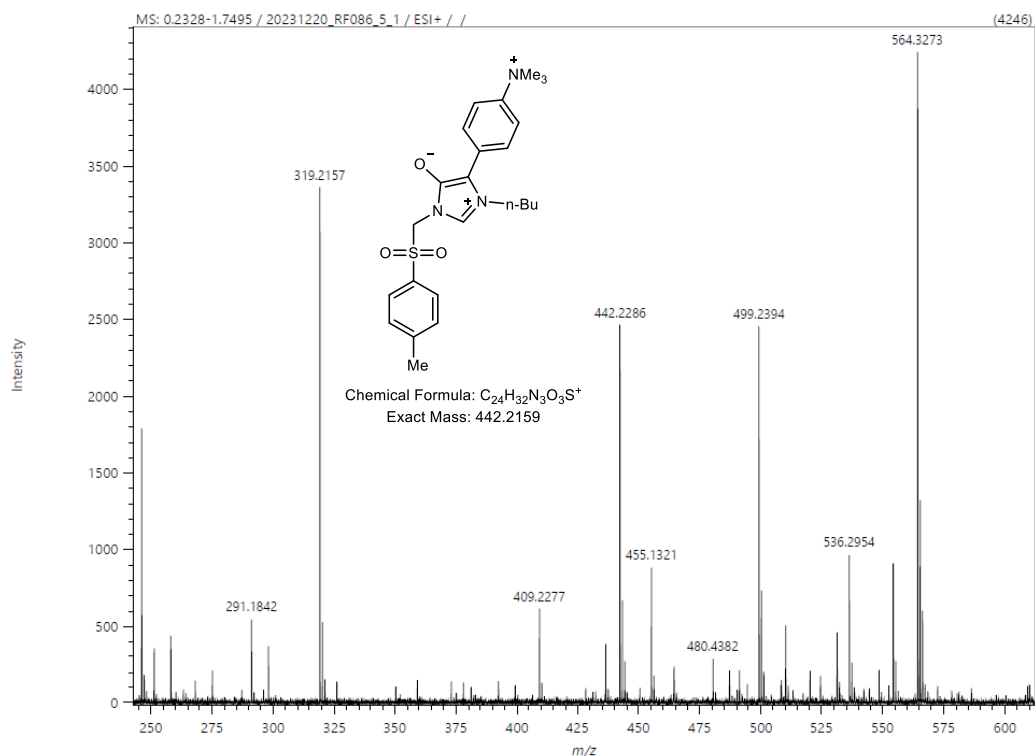
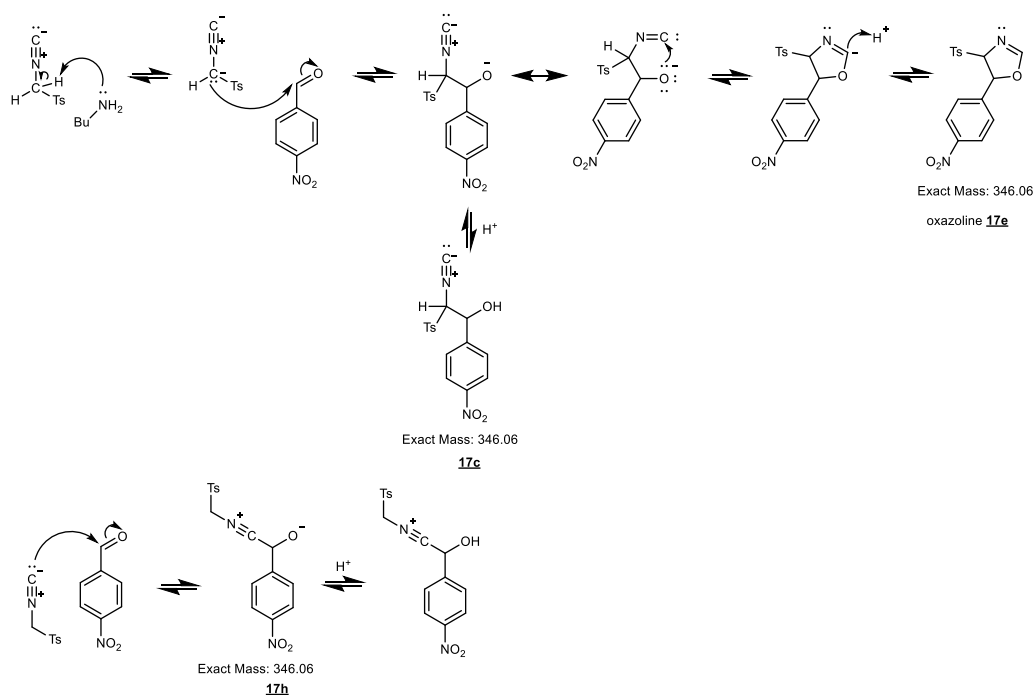


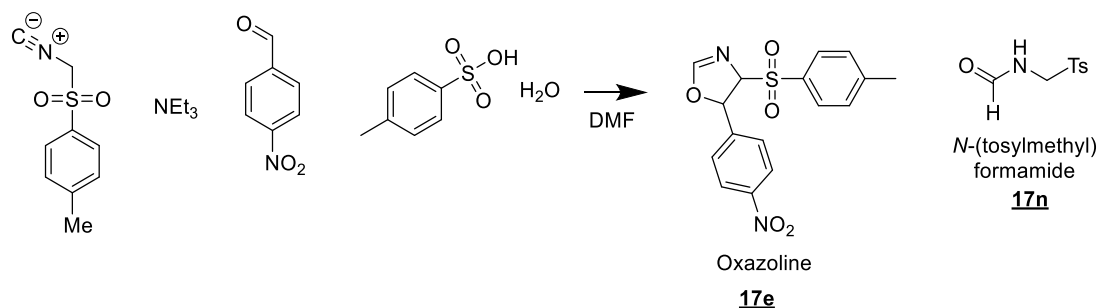
Figure S92. High-resolution mass spectrum (ESI-TOF) of multicomponent reaction in a local yield maximum after 1 h (zoom in).

4.8.4. NMR study of the global yield maximum



In the MS charge-tag experiments, the presence of a nitrilium ion hydride **17h** or **17c** was proved (either directly or indirectly by the detection of oxazoline **17e**, which requires the formation of

a nitrilium ion hydrin **17c**). In the following experiment, we recreated the reaction conditions corresponding to the global yield maximum, but we substituted *n*-butylamine by triethylamine to rule out the mechanisms involving an iminium ion **18c**. The formation of oxazoline **17e** is following the appearance of the global yield maximum, and its dependence on a *p*TSA concentration. However, the conversion of oxazoline to compound **16e** could not be proven.



The stock solutions of TosMIC (1.2 M), triethylamine (2.0 M), *p*-nitrobenzaldehyde (1.2 M), *p*TSA (1.5 M) in DMF were prepared, and then they were used by automatic liquid handler to prepare reaction with the following initial substrates concentration: *p*-nitrobenzaldehyde 0.27 M, *p*TSA (0.10523, 0.10823, 0.11123, 0.11423, 0.11723, 0.12023, 0.12323 and 0.12623 M), triethylamine 0.12 M and TosMIC 0.30 M. The order of compounds addition was as follows: DMF, solution of *p*-nitrobenzaldehyde, solution of *p*TSA, solution of triethylamine and solution of TosMIC. The total volume of reaction was 500 μ L. After sealing, the reactions were shaken for 5 min (250 rpm) and placed in a thermostatic oven at 26 $^{\circ}$ C for 16 h. Subsequently, every reaction mixture was lyophilized and dissolved in DMSO-*d*₆.

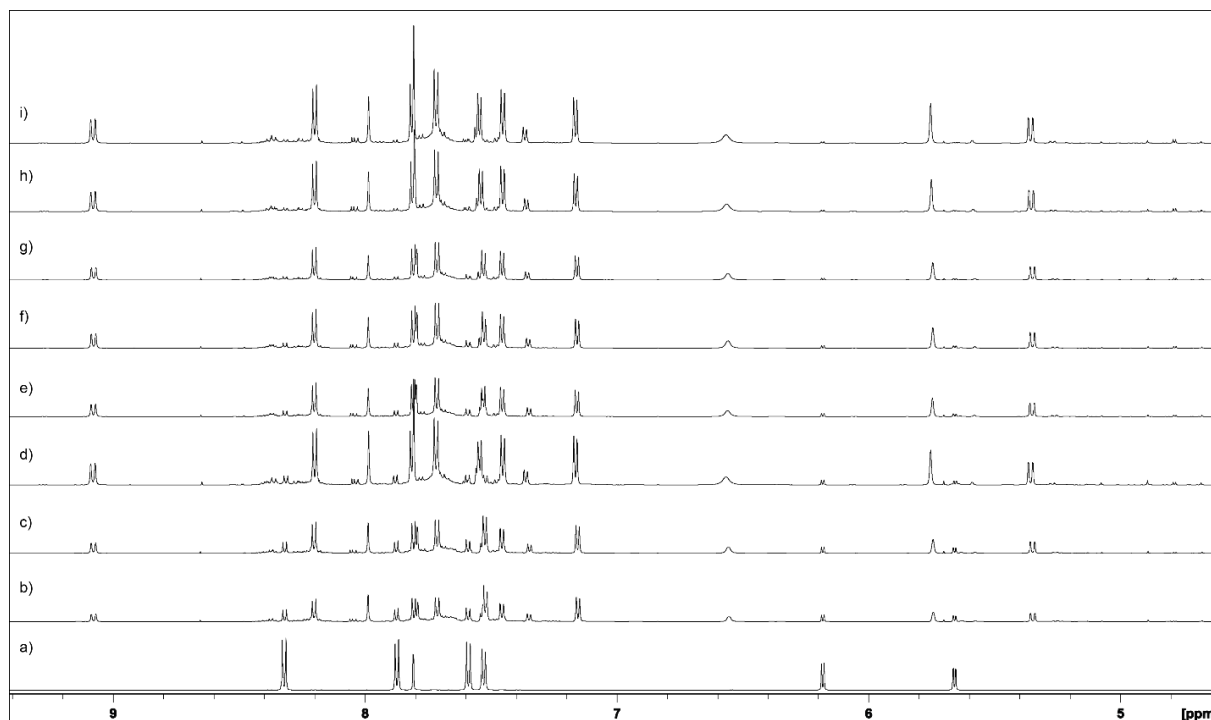
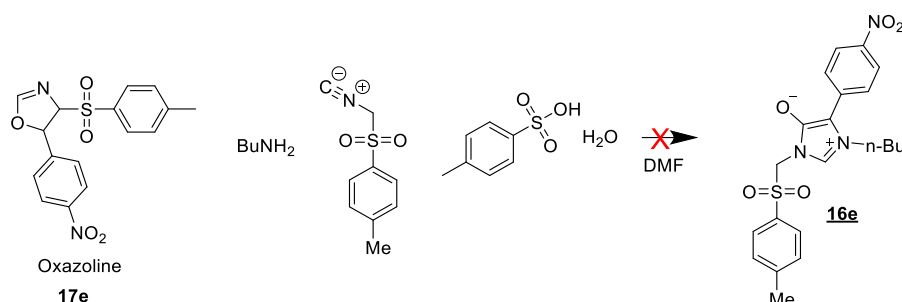


Figure S93. Stacked ^1H NMR spectra (400 MHz, $\text{DMSO}-d_6$): a) pure oxazoline **17e**; crude reaction mixtures for initial acid concentration: b) 0.10523 M, c) 0.10823 M, d) 0.11123 M, e) 0.11423 M, f) 0.11723 M, g) 0.12023 M, h) 0.12323 M, i) 0.12623 M.



The stock solutions of oxazoline **17e** (0.6 M), *n*-butylamine (2.0 M), TosMIC (1.2 M), *p*TSA (1.5 M) in DMF were prepared, and then they were used by automatic liquid handler to prepare reaction with the following initial substrates concentration: oxazoline **17e** (0.27 M), PTSA (0, 0.03023, 0.04523, 0.06023, 0.07523, 0.09023, 0.10523 and 0.12023 M), *n*-butylamine (0.12 M) and TosMIC (0.30 M). The order of compounds addition was as follows: DMF, solution of oxazoline **17e**, solution of *p*TSA, solution of *n*-butylamine and solution of TosMIC. The total volume of reaction was 500 μL . After sealing, the reactions were shaken for 5 min (250 rpm) and placed in a thermostatic oven at 26 $^{\circ}\text{C}$ for 16 h. Subsequently, every reaction mixture was lyophilized and dissolved in $\text{DMSO}-d_6$.

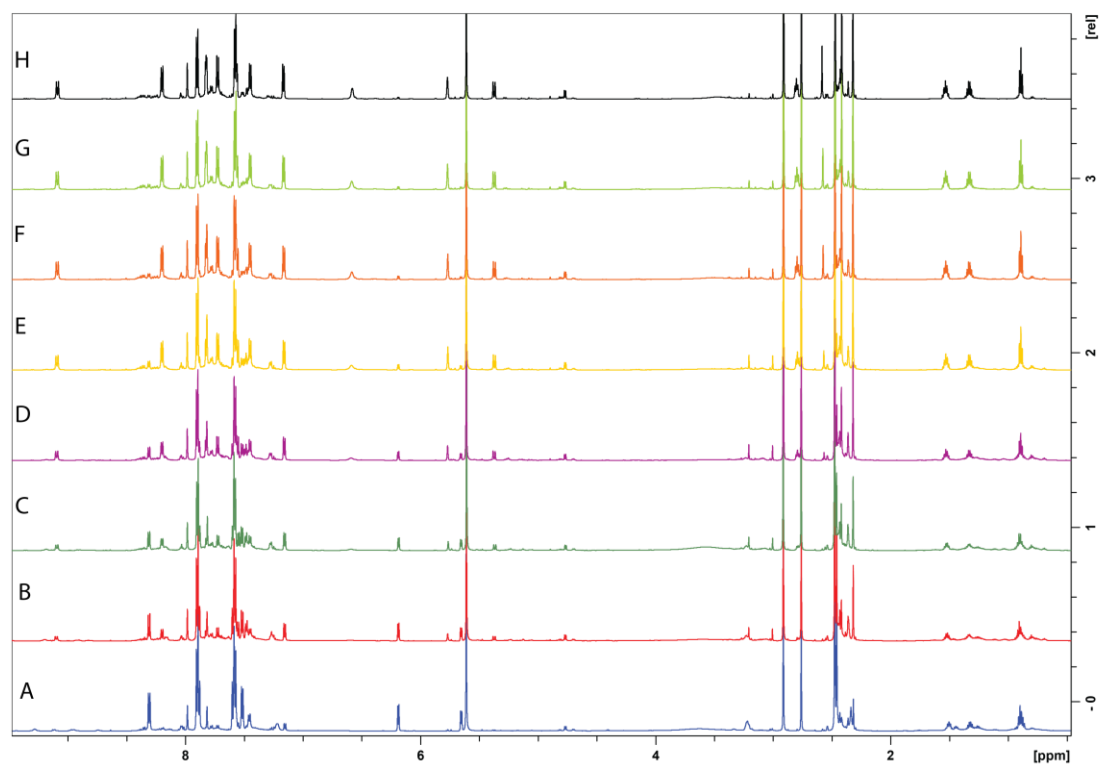
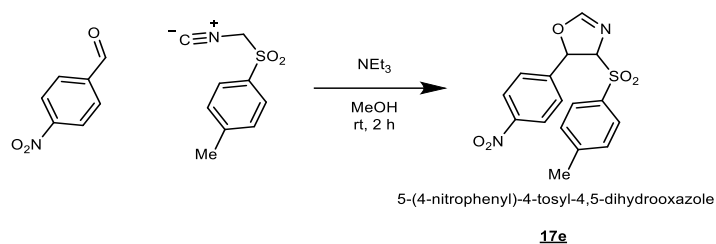


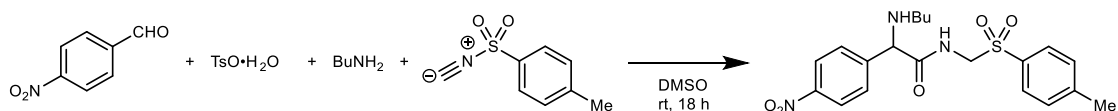
Figure S94. Stacked ¹H NMR spectra (400 MHz, DMSO-*d*₆) of crude reaction mixtures for initial acid concentration: A) 0 M, B) 0.03023 M, C) 0.04523 M, D) 0.06023 M, E) 0.07523 M, F) 0.09023 M, G) 0.10523 M, H) 0.12023 M.

4.8.5. Additional reactions



Compound **17e**, 5-(4-nitrophenyl)-4-tosyl-4,5-dihydrooxazole (oxazoline) was synthesized according to the literature procedure.⁹⁴

¹H NMR (400 MHz, chloroform-*d*): δ = 8.28 (d, *J* = 8.8 Hz, 2H), 7.85 (d, *J* = 8.3 Hz, 2H), 7.58 (d, *J* = 8.6 Hz, 2H), 7.40 (d, *J* = 7.9 Hz, 2H), 7.25 (d, *J* = 1.5 Hz, 1H), 6.16 (d, *J* = 6.7 Hz, 1H), 4.98 (dd, *J* = 6.3, 1.7 Hz, 1H), 2.47 (s, 3H). ¹³C NMR (100 MHz, chloroform-*d*): δ = 159.12, 148.13, 146.11, 144.61, 132.66, 130.01, 129.56, 126.07, 124.32, 92.49, 78.16, 21.76.



To a solution of 4-nitrobenzaldehyde (151 mg, 1.00 mmol) and *p*TSA (190 mg, 1.00 mmol) dissolved in DMSO (3 mL) was added *n*-butylamine (98 μ L, 1.0 mmol). The resulting solution was stirred at room temperature for 15 min. Afterward, TosMIC (195 mg, 1.00 mmol) was added to the reaction mixture. The resulting mixture was stirred at room temperature for 18 h. The mixture was diluted with dichloromethane (10 mL) and washed with saturated sodium bicarbonate (2 $\times$ 10 mL) and water (10 mL). The organic solution was dried over anhydrous sodium sulfate. After removal of solvent, the residue was purified by HPLC to give the pure product as a colorless solid (47 mg, 75%).

2-(Butylamino)-2-(4-nitrophenyl)-*N*-(tosylmethyl)acetamide

¹H NMR (400 MHz, DMSO-*d*₆): δ = 9.17 (t, *J* = 6.1 Hz, 1H), 8.16–8.18 (m, 2H), 7.54–7.55 (m, 2H), 7.52–7.53 (m, 2H), 7.23–7.25 (m, 2H), 4.78 (dd, *J* = 14.1, 6.4 Hz, 1H), 4.68 (dd, *J* = 14.1, 5.7 Hz, 1H), 4.32 (s, 1H), 2.33 (s, 3H), 2.26–2.29 (m, 2H), 1.49–1.52 (m, 1H), 1.30–1.38 (m, 2H), 1.21–1.30 (m, 2H), 0.84 (t, *J* = 7.2 Hz, 3H). ¹³C NMR (100 MHz, DMSO-*d*₆): δ = 171.39, 147.10, 146.82, 144.40, 134.30, 129.53, 128.54, 128.30, 123.20, 69.72, 64.78, 59.77, 46.75, 40.15, 39.94, 39.73, 39.52, 39.31, 39.10, 31.49, 26.09, 21.03, 19.84, 13.90. HRMS (ESI-TOF) *m/z*: [M + H]⁺ Calcd for C₂₀H₂₆N₃O₅S⁺: 420.1588; Found 420.1584.

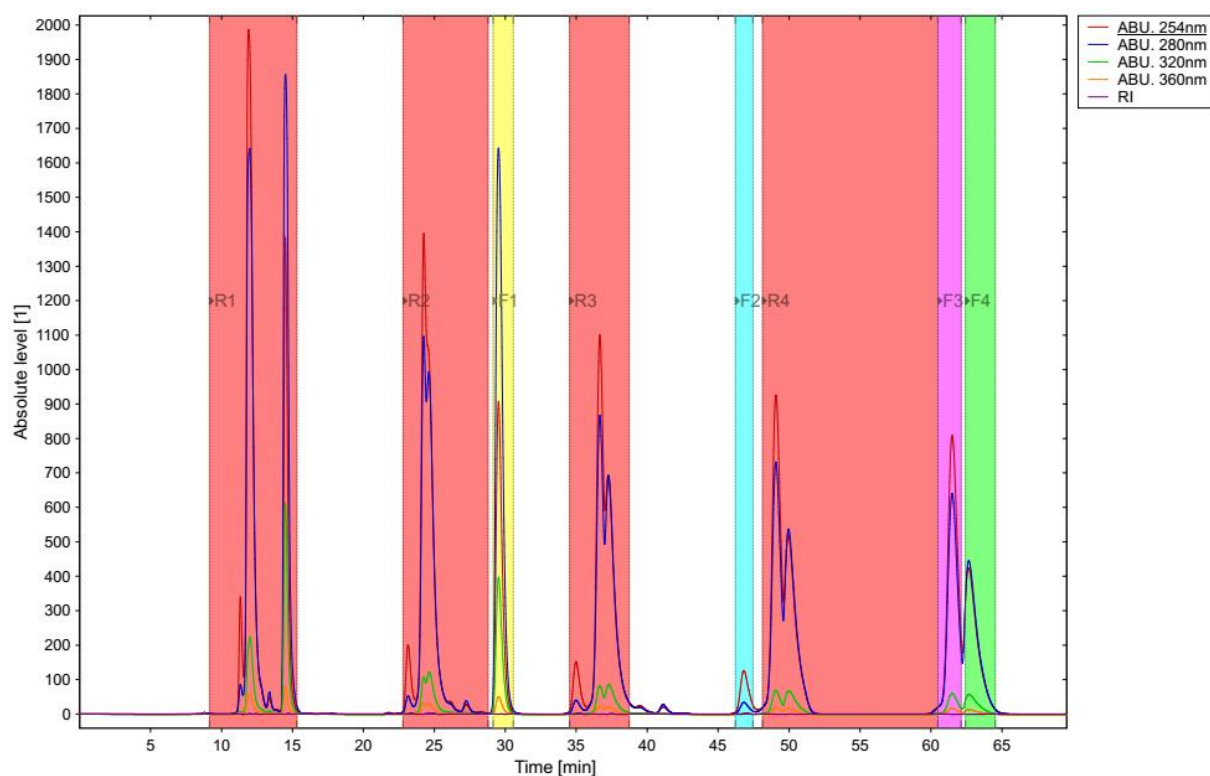


Figure S95. Chromatogram, (RP-HPLC/MeOH). **Fraction 4** contains impure 2-(butylamino)-2-(4-nitrophenyl)-*N*-(tosylmethyl)acetamide, which was then subjected to GPC purification.

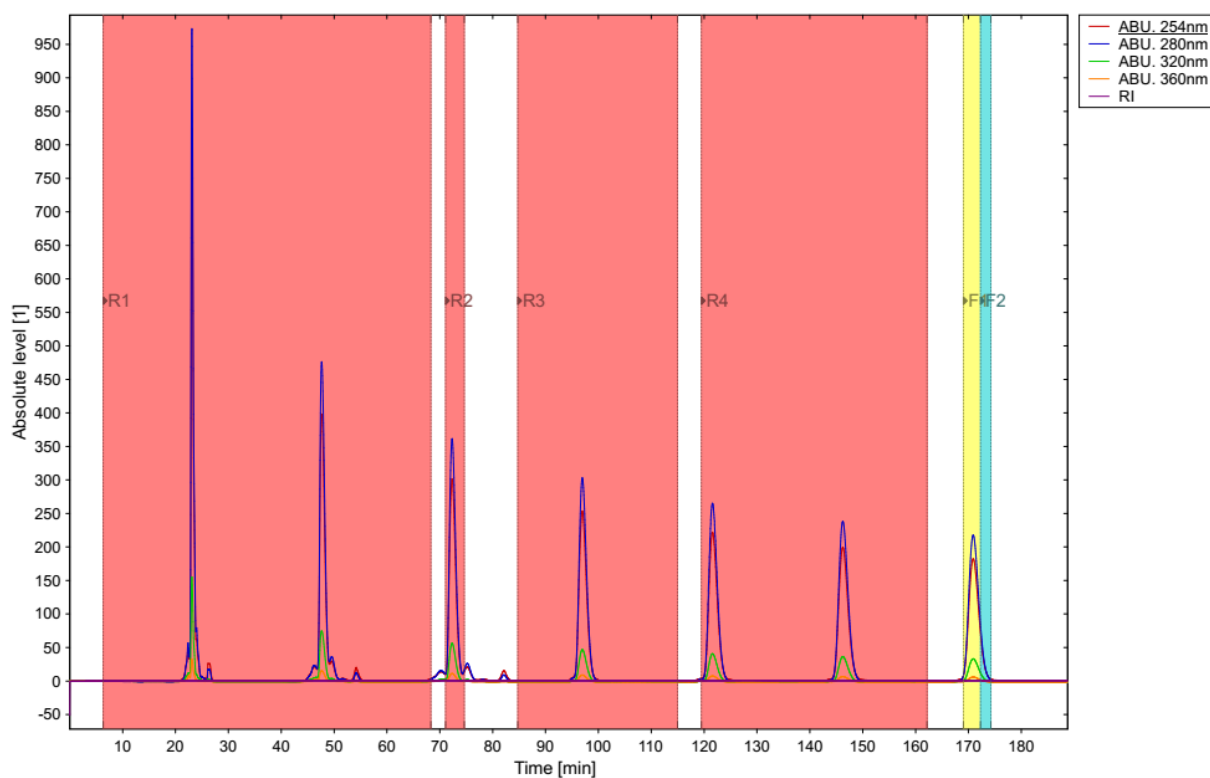
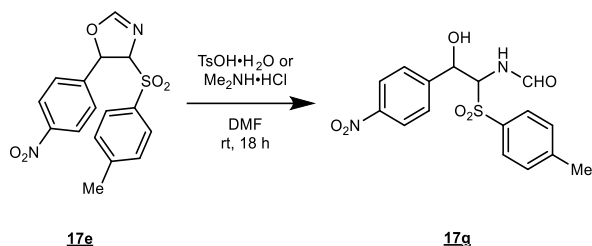


Figure S96. Final purification of 2-(butylamino)-2-(4-nitrophenyl)-*N*-(tosylmethyl)acetamide (GPC/chloroform).

In the presence of a proton source, oxazoline undergoes hydrolytic ring-opening where water attacks C5 and the C5–O1 bond is cleaved. This ring-opened product is formed also in the presence of TosMIC and *n*-butylamine.



To a solution of 5-(4-nitrophenyl)-4-tosyl-4,5-dihydrooxazole (35 mg, 0.10 mmol) dissolved in DMF (1 mL) either *p*TSA (19 mg, 0.10 mmol) or dimethylamine hydrochloride (8 mg, 0.1 mmol) was added. The resulting solution was stirred at room temperature for 18 h. Afterward, TosMIC (195 mg, 1.00 mmol) was added to the reaction mixture. The resulting mixture was stirred at room temperature for 18 h. The mixture was diluted with ethyl acetate (10 mL) and washed with saturated sodium bicarbonate (2×10 mL) and water (10 mL). The organic solution was dried over anhydrous sodium sulfate. After removal of solvent, the residue was purified by HPLC to give the pure product as a colorless solid (26 mg, 71%). This compound decomposes to give additional peaks in the ^1H NMR spectrum in roughly half a day.

Compound 17g, *N*-(2-Hydroxy-2-(4-nitrophenyl)-1-tosylethyl)formamide

^1H NMR (600 MHz, DMSO- d_6): 9.03 (d, J = 10.9 Hz, 1H), 8.16–8.17 (m, 2H), 7.75–7.77 (m, 2H), 7.74 (d, J = 1.0 Hz, 1H), 7.66–7.68 (m, 2H), 7.41–7.42 (m, 2H), 6.52 (d, J = 5.0 Hz, 1H), 5.70 (d, J = 4.6 Hz, 1H), 5.30 (dd, J = 1.1, 10.5 Hz, 1H), 2.39 (s, 3H). ^{13}C NMR (150 MHz, DMSO- d_6): δ = 160.87, 148.98, 146.95, 144.66, 134.88, 129.64, 129.13, 127.60, 123.17, 71.45, 67.73, 21.16. HRMS (ESI–TOF) m/z : $[\text{M} + \text{Na}]^+$ Calcd for $\text{C}_{16}\text{H}_{16}\text{N}_2\text{NaO}_6\text{S}^+$: 387.0622; Found 387.0621.

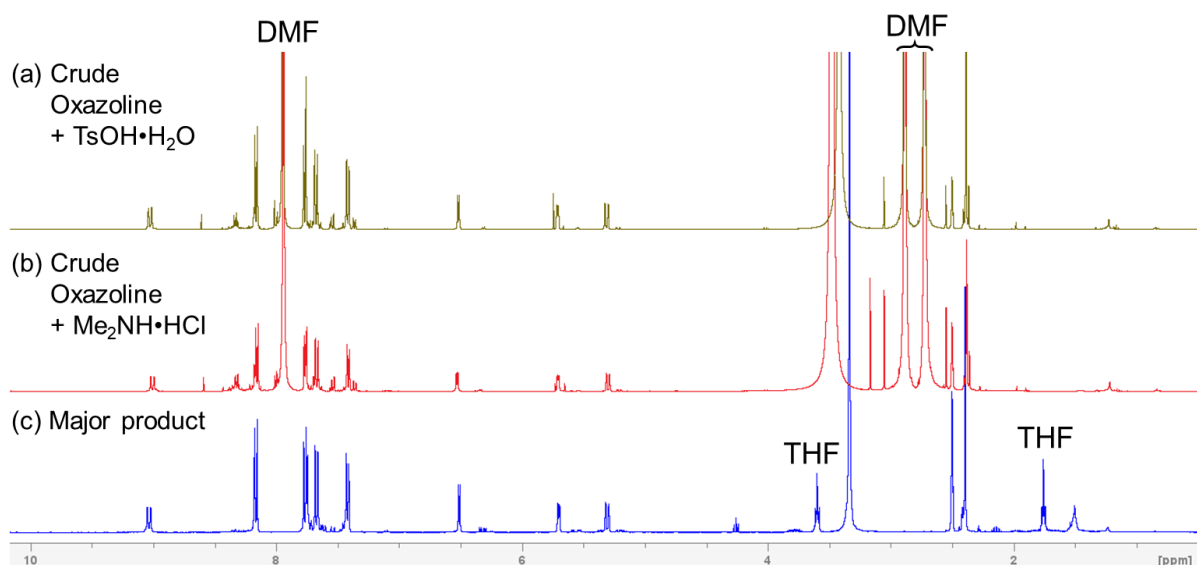


Figure S97. Stacked ^1H NMR spectra (400 MHz, $\text{DMSO-}d_6$) of (a) crude product mixture from treatment of oxazoline **17e** with *p*TSA in DMF for 18 h, (b) crude mixture from treatment of oxazoline **17e** with $\text{Me}_2\text{NH}\cdot\text{HCl}$ in DMF for 18 h, (c) **17g** after GPC purification.

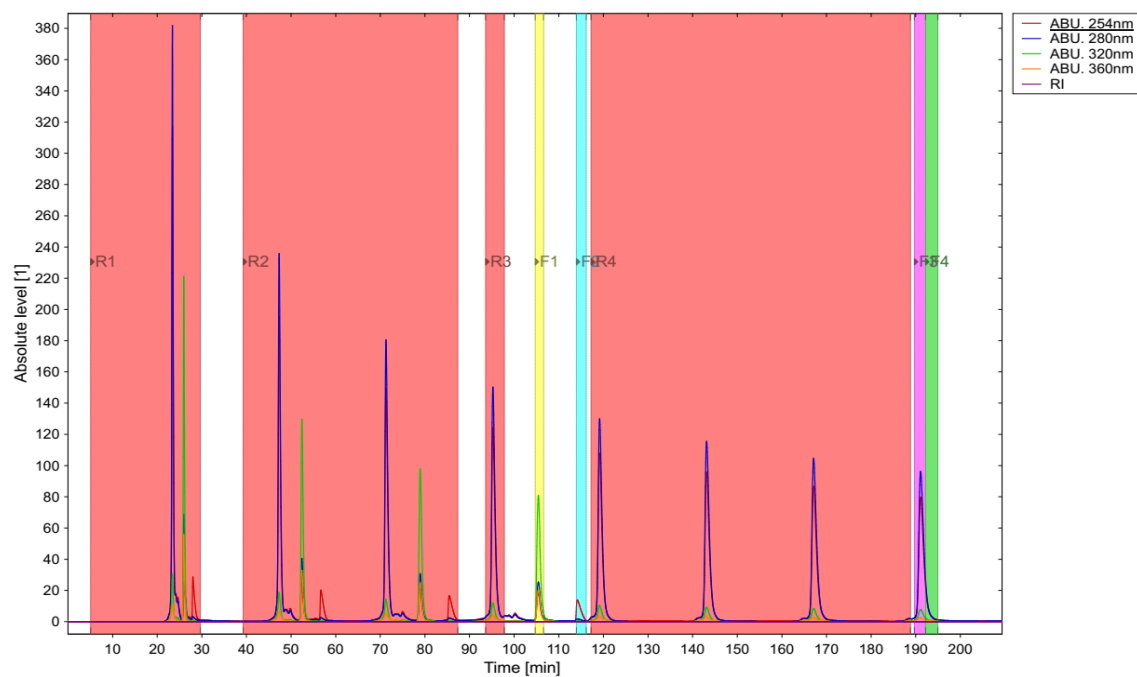


Figure S98. Chromatogram (GPC/THF). **Fractions 3 and 4** contain **17g**.

4.9. Iterative discovery in Hantzsch reaction network

The iterative discovery (**Figure S99**) started with measuring UV–Vis spectra for every condition within the hypercube, and then fitting the UV–Vis spectra with just three compounds **19a**, **19b** and **19d**, as labelled by Component Discovered ‘3’ in **Fig. 5a**. Obviously, the fitting residual was large since more components from the crude were needed for a proper fit. Therefore, another run of all hypercube reactions was performed, and the reactions were again combined to give one reaction crude. This reaction crude was used for isolating more fitting components by three rounds of recycling preparative HPLC.

Upon three rounds of purification, the total number of fractions for analysis would exceed 50 (**Figure S99b**), and each fraction contains one compound or a mixture of compounds resulting from the reaction. It is possible for a specific compound to appear in multiple fractions. Analyzing all of the fractions is highly time-consuming. As illustrated in **Figure S99a**, we use NMR and UV-Vis fitting to guide the process of fraction analysis. We first analyze a fraction using NMR to determine the presence of any new compounds. If new compounds are detected, the fraction is further purified, and the isolated compounds are subsequently fed to the UV-Vis fitting model. If significant discrepancy between the measured UV-Vis spectra and the fitting curves remains, the search for new compounds continues with additional fractions. This iterative process is repeated until the fitting residual reaches the noise level. In summary, NMR analysis guides the selection of fractions for further purification, while UV-Vis fitting determines the endpoint of the analytical process.

After identifying all the key components in the 80 °C reaction crude, we repeated all hypercube reactions four additional times, in total three times at 80 °C and twice at 26 °C (**Figure S100**). The reactions performed at 80 °C not only confirmed the experimental reproducibility but evidenced that our understanding of the reaction’s hyperspace is complete to the level of experimental noise – that is, for all replicates, the 16 compounds we had initially identified were sufficient to fit the hypercubes’ UV-Vis spectra to within experimental uncertainty (*green* bands in **Figure S98a,b**). The replicates of the hypercube at 26 °C were also reproducible. At the same time, we did not attempt a separate HPLC purification campaign for these 26 °C hypercubes but, instead, fitted the hypercubes’ spectra using the component spectra of the 16 compounds discovered at 80 °C. Interestingly, only 12 out of these 16 compounds were relevant in the sense that the residuals converged at these 12 species and addition of compounds 13-16 did not improve the fits. We observe that although the residuals were very small, they were slightly above the level of experimental noise indicating the present of some yet-unidentified albeit minor components in the 26 °C product mixture.

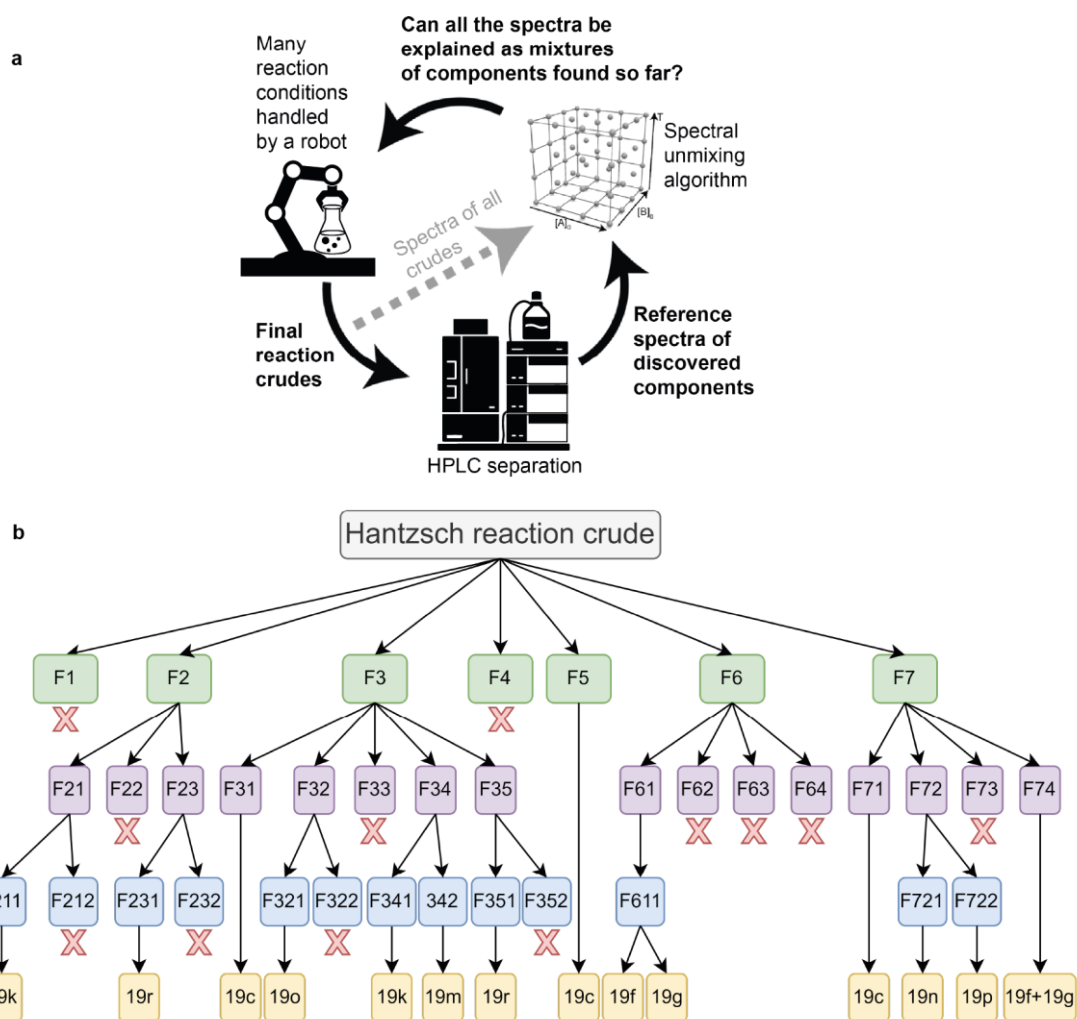


Figure S99. a, Scheme illustrating iterative elucidation of the Hantzsch reaction network. **b**, HPLC isolation process for fitting the components. Discontinued fractions are marked by *red* cross. Fractions were further pursued if they featured new signals, and the process was continued until the UV-Vis spectra over the hypercube corresponded to the fitted spectra to the noise level (see **Figure S98**). In the first-round HPLC purification of the reaction crude, seven separate fractions were obtained from 20 cycles of HPLC (**Figure S33** and **Figure S99**), with each fraction a mixture of compounds. These seven fractions were further purified by HPLC one by one. Typically, each of the seven fractions needed two more rounds of HPLC cycles before pure compounds were obtained. For example, the first-round fraction #6 was split into 4 new fractions (**Figure S36**). From these 4 fractions, NMR measurements identified two new compounds **19f** and **19g**, but still as mixtures. These second-round fractions were further purified by HPLC for a third round, and two compounds **19f** and **19g** were obtained from one of the third-round fractions (**Figure S38**). Generally, in the second round, each fraction was further divided into 3 to 5 sub-fractions, followed by an additional division each into 2 to 3 sub-fractions.

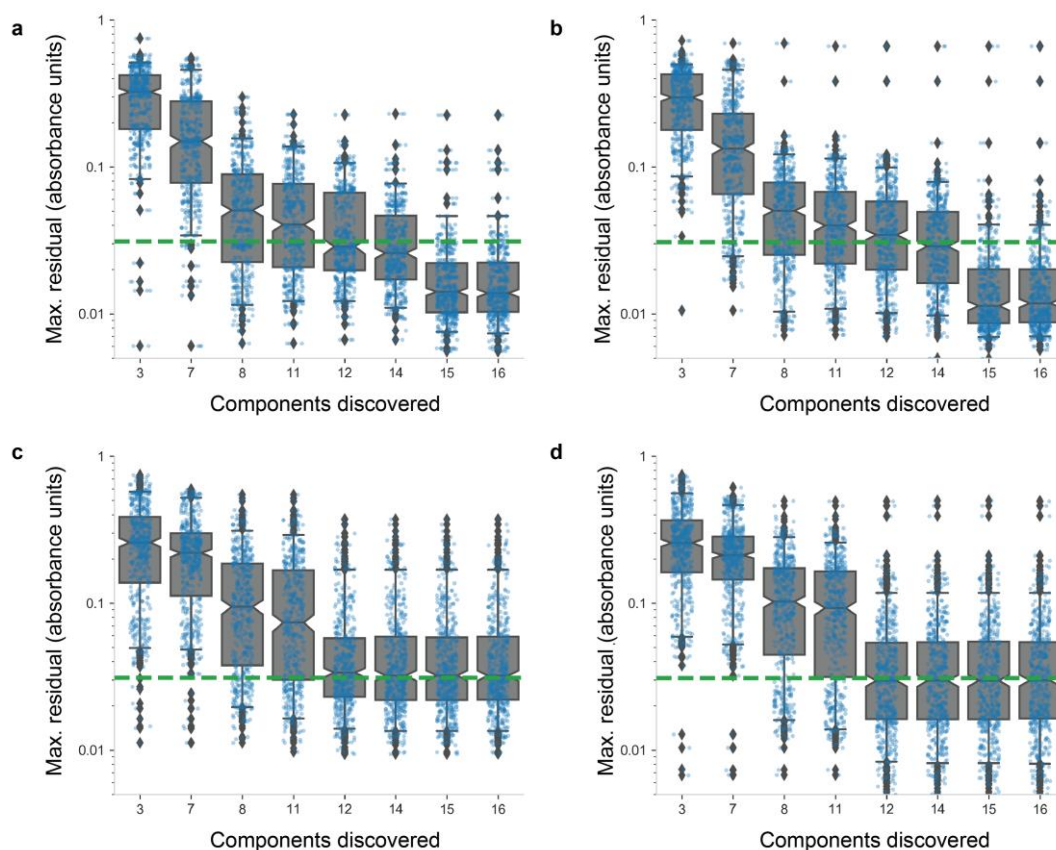
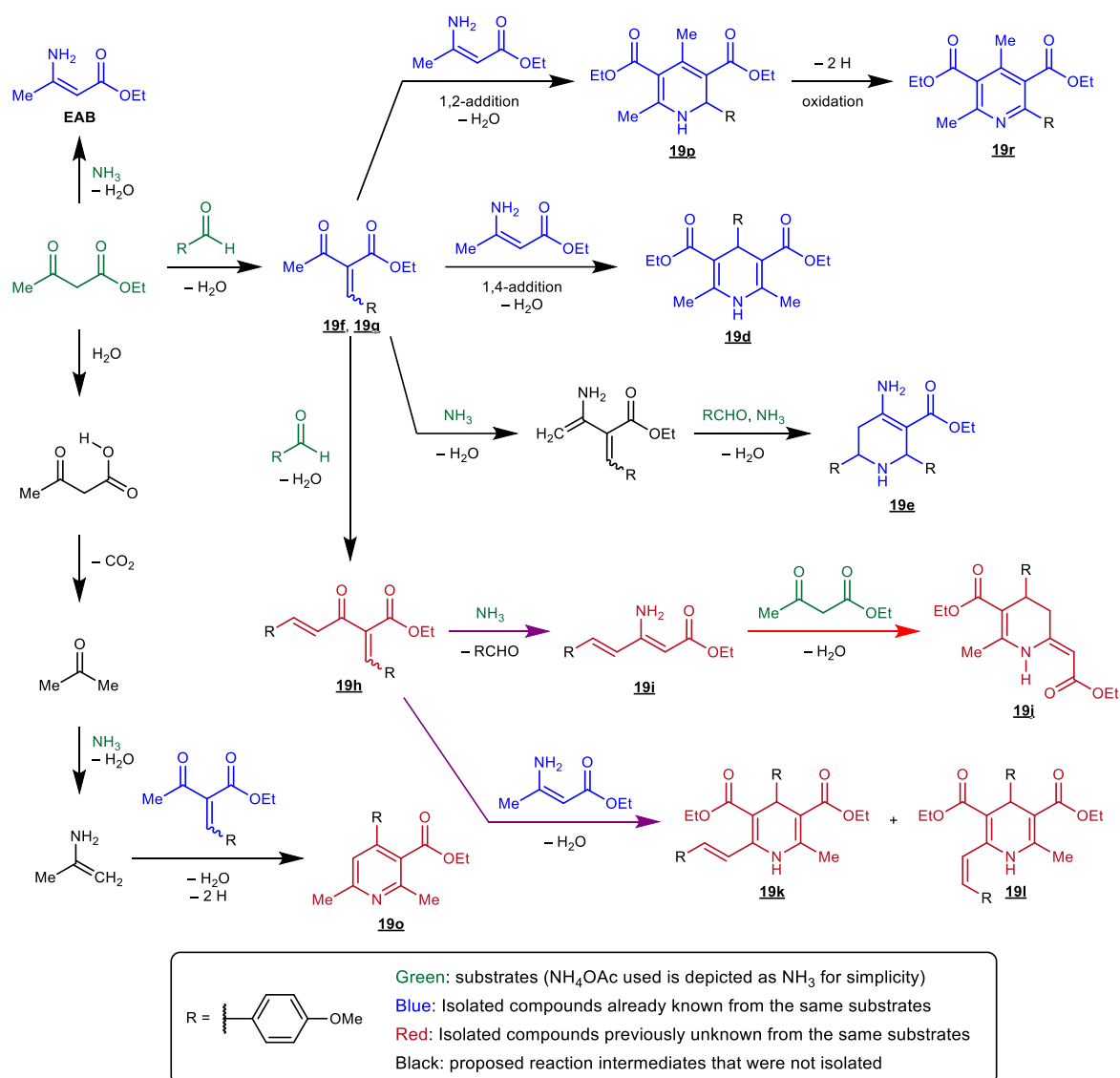
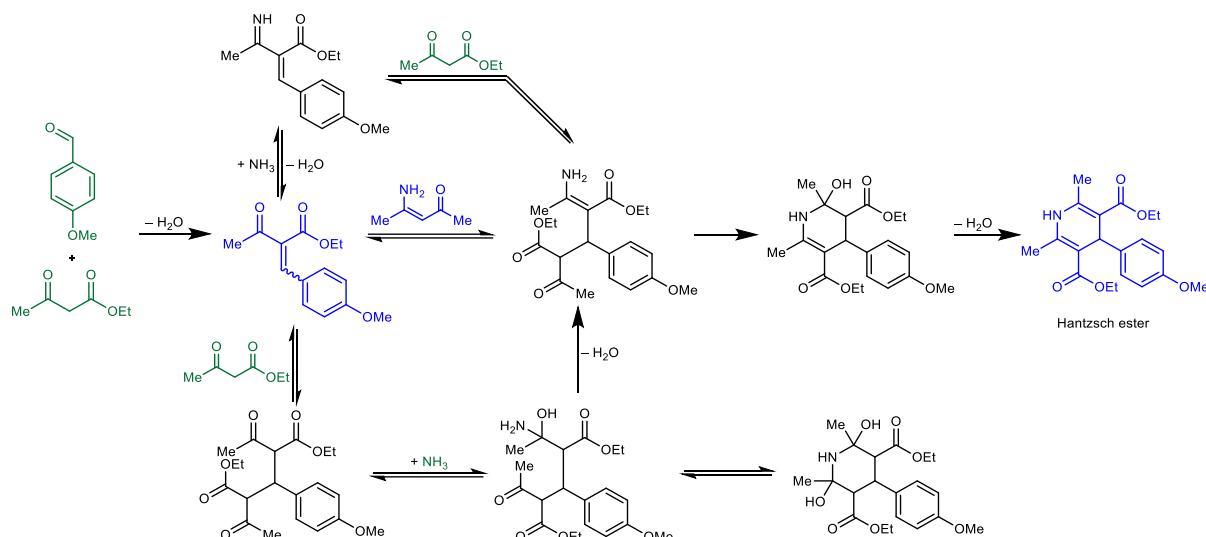


Figure S100. Additional Hantzsch hypercube replicates and evolution of residuals as a function of components included in the fitting. **a,b**, Two 80 °C hypercube repeats to supplement **Fig. 5a**. **c,d**, Two hypercube repeats at 26 °C with the residuals calculated for different numbers of products discovered at 80 °C. Green stripes correspond to experimental noise. Vertical axes quantify the distribution of the highest mismatch (highest difference across all spectral range) between experimental spectra and the unmixing model: one datapoint corresponds to the crude spectrum for a single condition probed. Boxplot elements are the first quartile, the median, and the third quartile. Whiskers are the 5% and 95% percentiles, with points outside this range (outliers) shown as black diamonds. Blue scatterplot shows all individual datapoints.

4.10. Hantzsch reaction network studies

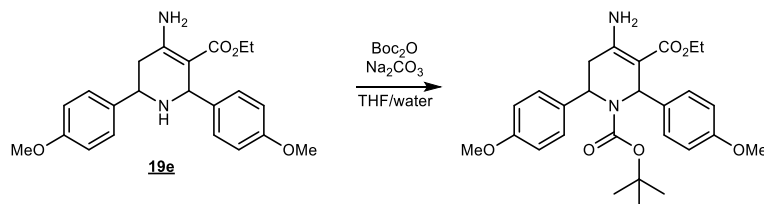


Scheme S1. Expanded Hantzsch reaction network from main text **Figure 5** showing also the mechanistic steps forming the Hantzsch ester **19d** and tetrahydropyridine **19e**.



Scheme S2. Three productive pathways of a Hantzsch ester formation known from previous studies ⁸⁴.

Because the amino protons of tetrahydropyridine **19e** are too broad to be well integrated in the ¹H NMR spectrum, we prepared the Boc-protected derivative **19e-BOC** for additional proof of the molecular structure of **19e**. The amino protons of **19e-BOC** are well integrated and are found to vanish upon addition of D₂O.



19e-BOC (1-(*tert*-Butyl) 3-ethyl 4-amino-2,6-bis(4-methoxyphenyl)-5,6-dihydropyridine-1,3(2*H*)-dicarboxylate):

To a solution of **19e** (50 mg, 0.131 mmol) and sodium carbonate (17 mg, 0.160 mmol) dissolved in THF (4 mL) and water (4 mL) di-*tert*-butyl dicarbonate (ca. 30% solution in THF, 0.20 mL, ca. 0.27 mmol) was added. The resulting mixture was stirred at room temperature and reaction was monitored by TLC. After 1 h the mixture was diluted with ethyl acetate (30 mL) and washed with water (2 × 20 mL), brine (10 mL) and dried over sodium sulfate. After removal of solvent, the organic residue was purified by HPLC to give the titled compound as a colorless solid (47 mg, 75%). ¹H NMR (400 MHz, DMSO-*d*₆): δ = 7.84 (br, 1H), 7.24–7.25 (m, 3H), 6.93 (d, *J* = 8.5 Hz, 2H), 6.81 (d, *J* = 8.6 Hz, 2H), 6.72 (d, *J* = 8.4 Hz, 2H), 6.33 (s, 1H), 4.57 (d, *J* = 8.6 Hz, 1H), 4.05 (m, 2H), 3.75 (s, 3H), 3.67 (s, 3H), 2.39 (dd, *J* = 15.1, 4.5 Hz, 1H), 2.17 (t, *J* = 13.8 Hz, 1H), 1.11–1.13 (m, 12H). ¹³C NMR (150 MHz, DMSO-*d*₆): δ = 166.77, 159.23, 158.06, 157.85, 155.59, 136.91, 135.67, 127.97, 127.05, 113.74, 113.30, 90.11, 79.04,

58.34, 57.43, 55.02, 55.00, 52.66, 37.87, 27.81, 14.61. **HRMS** (ESI-TOF) m/z : $[M + Na]^+$
 Calcd for $C_{27}H_{34}N_2O_6Na$: 505.2310; Found 505.2308.

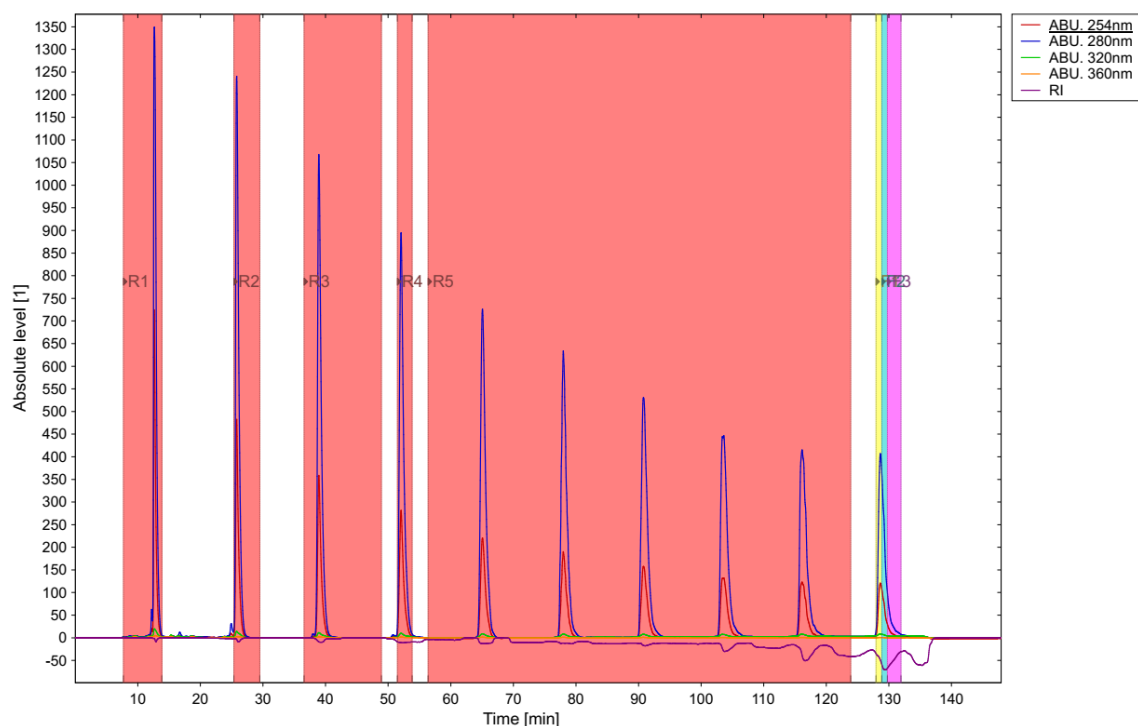


Figure S101. Chromatogram for purification of **19e-BOC** (RP-HPLC/MeOH). **Fractions 1–3** contain pure product.

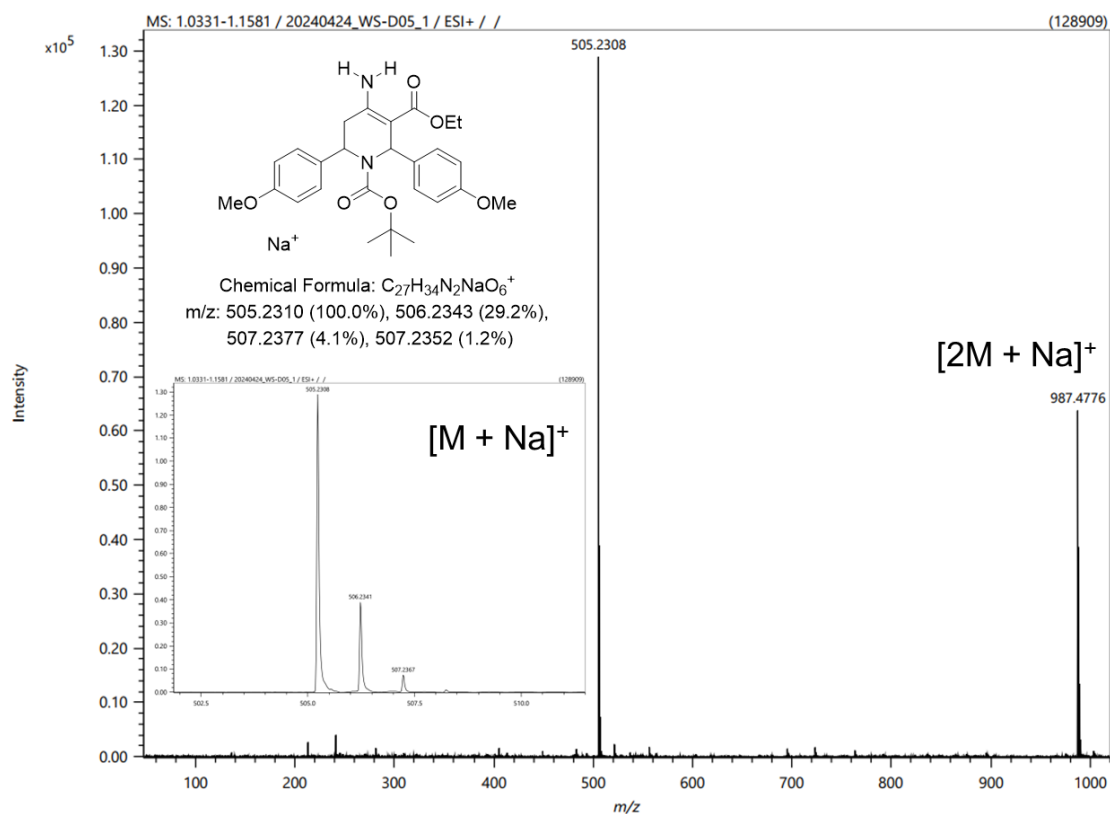
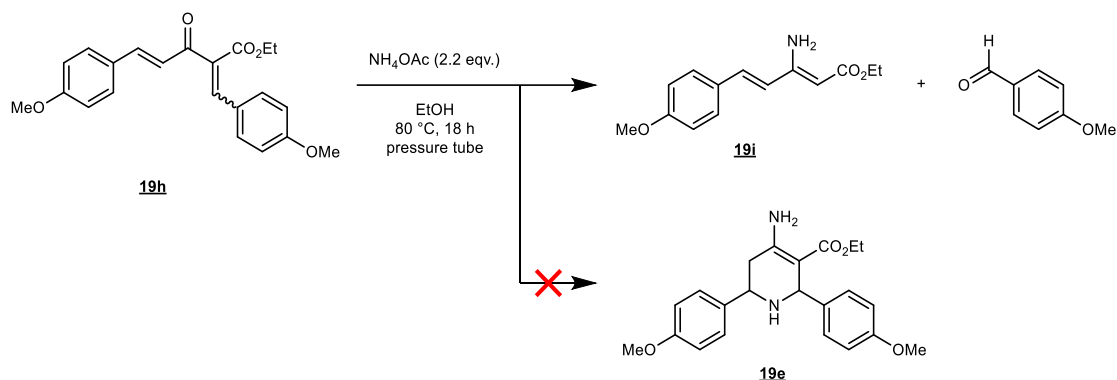


Figure S102. High resolution (ESI-TOF) mass spectrum of **19e-BOC**.

Next, we disproved formation of a product **19e** from initial Michael addition with ammonia (to the intermediate **19h**). Instead, we observed the benzylidene cleavage to aldehyde and intermediate **19i** (species that is also present in a crude reaction mixture from robot).



A solution of the mixture of dienone **19h** and its isomer (217 mg, 592 μmol) and ammonium acetate (100 mg, 1.30 mmol) in glacial acetic acid (8 mL) was placed in a pressure tube and stirred at 80 °C for 18 h. Then, the reaction mixture was cooled to room temperature and the solvent was removed under reduced pressure. The crude residue was then purified by HPLC to afford two major products: ethyl (2*Z*,4*E*)-3-amino-5-(4-methoxyphenyl)penta-2,4-dienoate (**19i**) (52 mg, 36%) as a pale yellow solid and *p*-methoxybenzaldehyde.

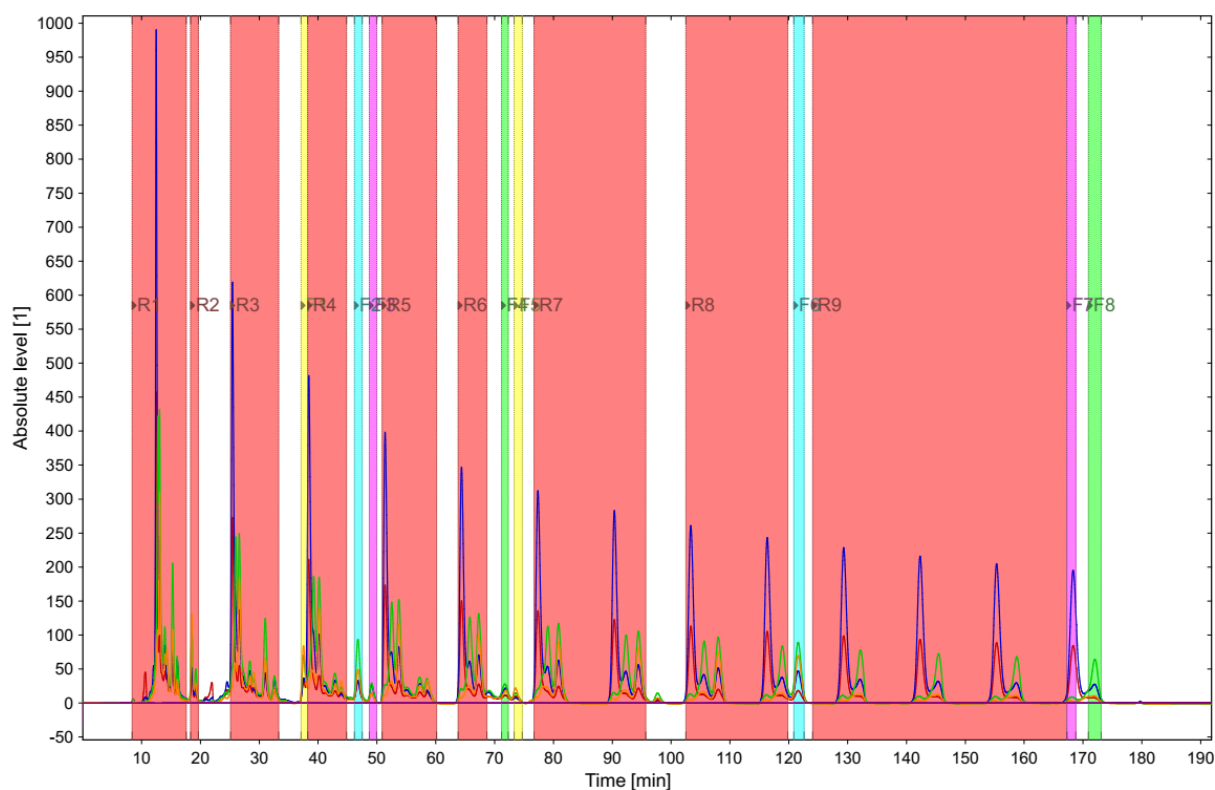


Figure S103. Chromatogram for initial purification of the crude mixture (RP-HPLC/MeOH). Fraction 6 contains impure **19i**.

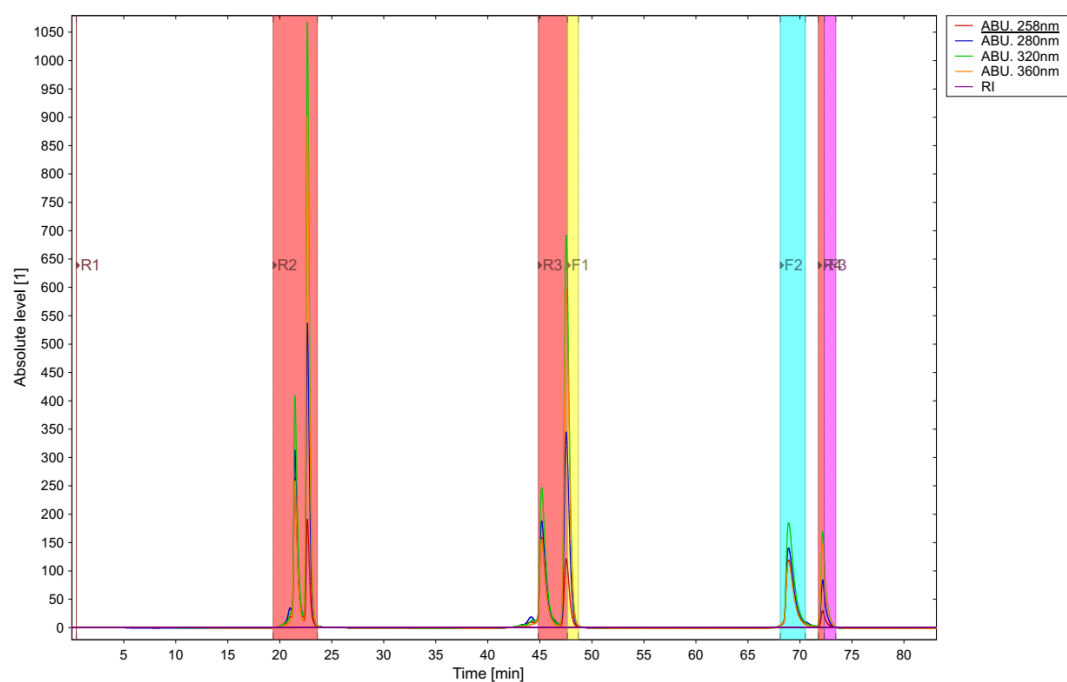


Figure S104. Chromatogram for final purification of **19i** (GPC/THF). **Fractions 1 and 2** contain pure product.

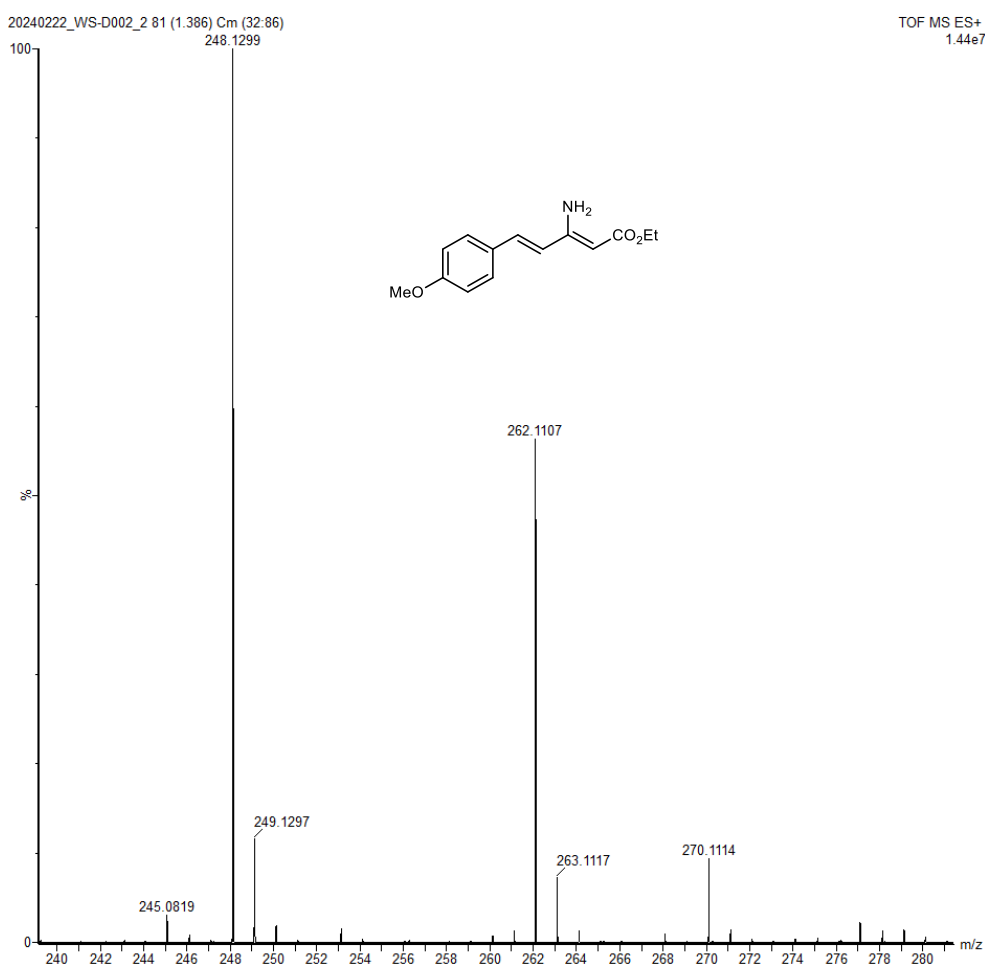
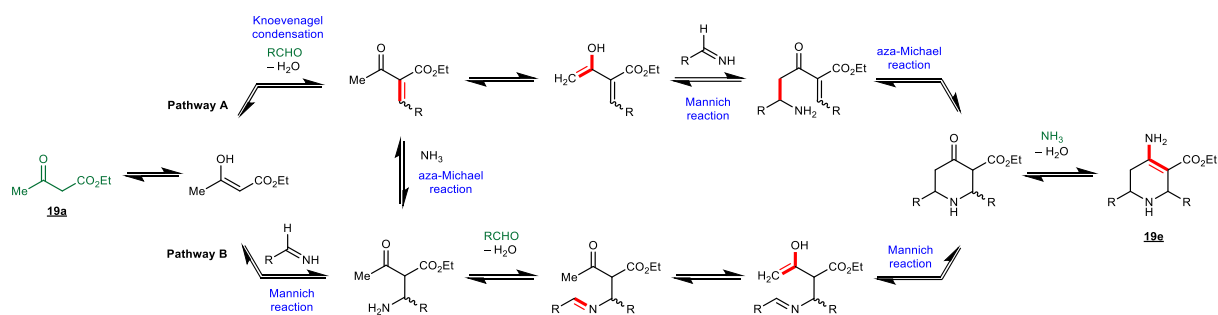
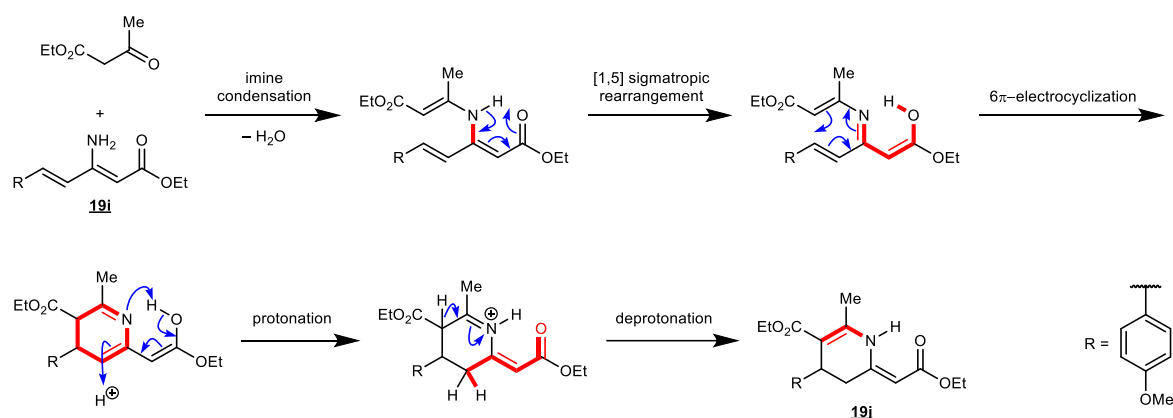


Figure S105. High resolution (ESI-TOF) mass spectrum of **19i**.

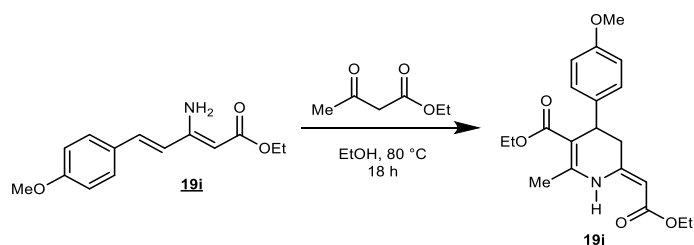


Scheme S3. The proposed mechanisms for the formation of product **19e**. In pathway **A**, the Knoevenagel condensation product, in its enol form, undergoes a Mannich–aza-Michael sequence to give the cyclized piperidone species, which further condenses with ammonia to give **19e**⁹⁵. In an alternative pathway **B**, the enol tautomer of ethyl acetoacetate (**19a**) annulates in a double-Mannich fashion to give the piperidone species that can lead to **19e**. The two pathways are connected via the Knoevenagel condensation product and the Mannich base. The rate of enolization plays a critical role. This rate is significantly influenced by the solvent and the amount of water present⁹⁶.



Scheme S4. The proposed mechanisms for the formation of product **19j**.

Experimental confirmation of the proposed mechanisms for ethoxyoxoethylidene-extended hydropyridine formation (**19j**).



A solution of **19i** (30 mg, 0.121 mmol) and ethyl acetoacetate (32 μ L, 0.25 mmol) dissolved in ethanol (2 mL) in a pressure tube was heated at 80 °C for 18 h. The reaction mixture was allowed to cool to room temperature and then the solvent was removed in vacuo. The residue was purified by HPLC to furnish the annulation product **19j** (13 mg, 30%) as a pale yellow solid.

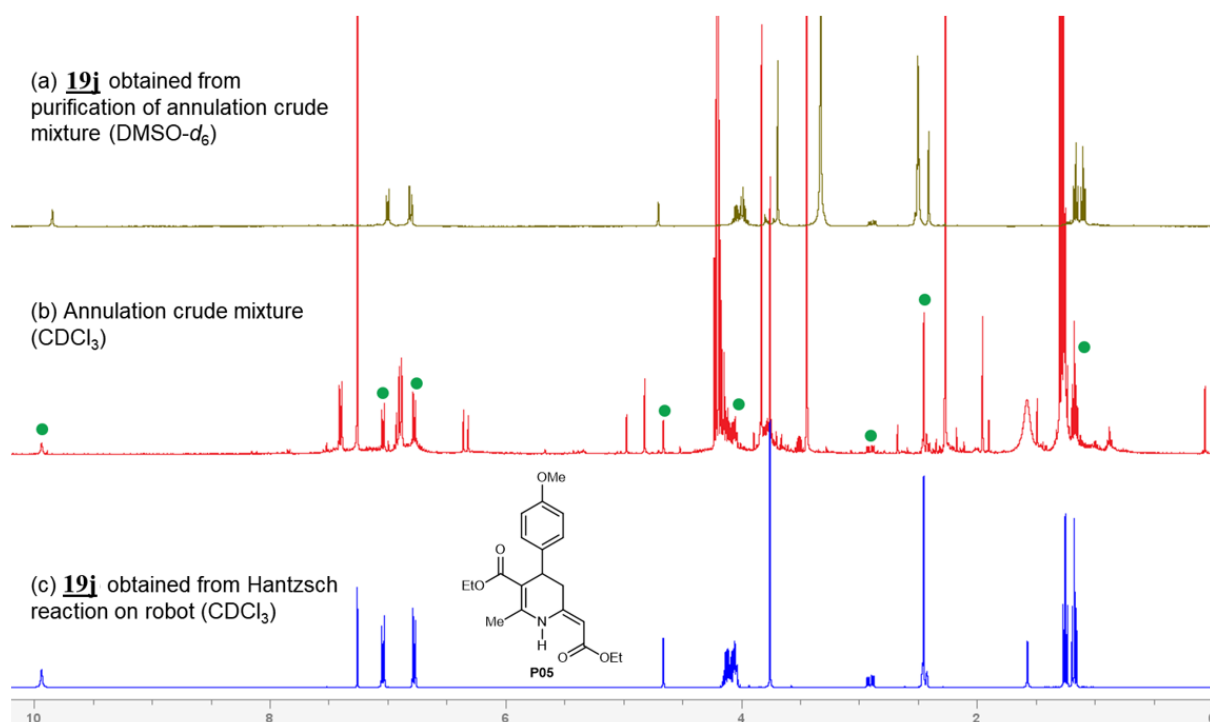


Figure S106. Stacked ^1H NMR spectra of (a) **19i** isolated from the annulation crude mixture by HPLC (400 MHz, DMSO- d_6), (b) annulation crude mixture (400 MHz, chloroform- d), and (c) pure **19i** obtained from Hantzsch reaction (400 MHz, chloroform- d). The green circles indicate the peaks originating from **19i** present in the crude mixture.

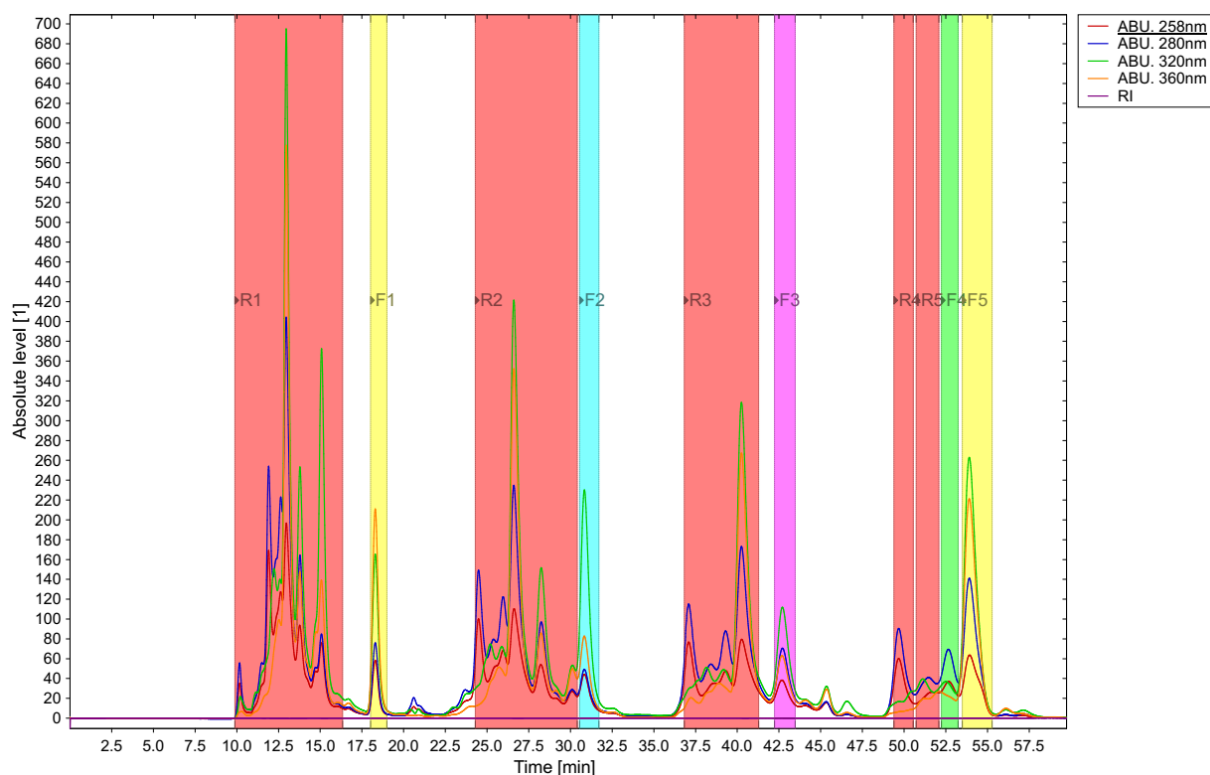
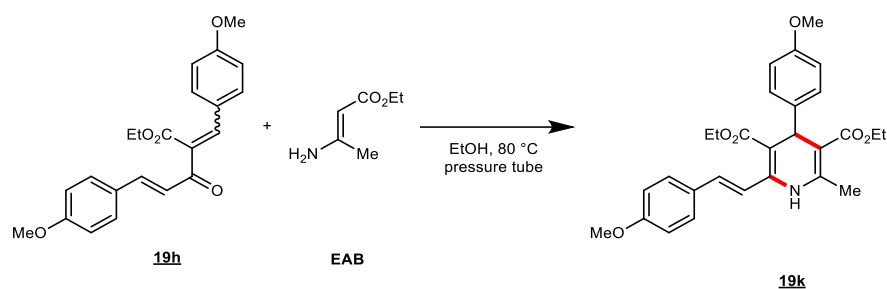


Figure S107. Chromatogram, (RP-HPLC/MeOH). **Fraction 2** contains **19i**, while **fraction 5** contains unreacted dienamine **19i**.

Experimental confirmation of proposed mechanisms for benzylidene-extended 1,4-dihydropyridine formation (**19k**) from initial Michael addition with enamino ester.



Ethyl (*Z*)-3-aminobut-2-enoate (175 μL , 1.36 mmol) was added to the solution of (*E*)-2-((*Z*)-4-methoxybenzylidene)-5-(4-methoxyphenyl)-3-oxopent-4-enoate (497 mg, 1.36 mmol) in 2.5 mL anhydrous ethanol. The mixture was purged with nitrogen for 10 min and stirred for 24 h at 80 $^\circ\text{C}$. Then, solvent was removed under reduced pressure and crude product was purified on RP-HPLC resulting in yellow glassy solid (41%).

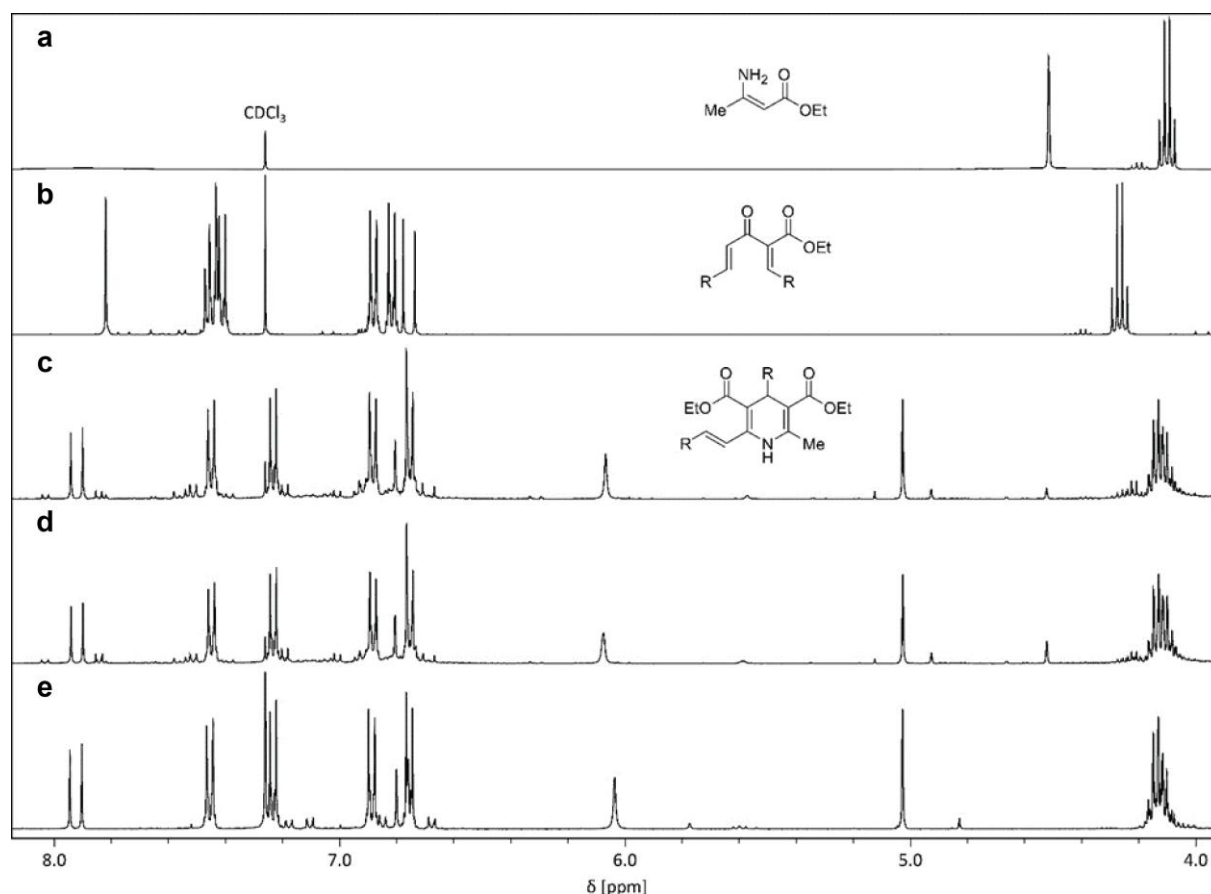


Figure S108. Partial ^1H NMR spectra (400 MHz, chloroform-*d*) of: **a**, **EAB**, **b**, **19h**, **c**, **19k** isolated from the robot, **d**, crude reaction mixture after 24 h, **e**, crude reaction mixture after 48 h.

4.11. Computer-driven analysis of Hantzsch mechanistic networks

The reaction networks such as those for the Hantzsch reaction (cf. main-text **Figure 5**) can be reconstructed based on expert insight or, as illustrated in this Section, based on computer-driven analysis of mechanistic networks. In brief, we have recently described the “MECH” algorithm that has been taught >9,000 mechanistic steps (roughly at the level of “arrow pushing”) and, for a given set of substrates, propagates “forward” networks of mechanistic steps. In ref. ⁹⁷, we showed how such networks can be pruned to trace sequences of mutually compatible steps, and how some of these sequences can give rise to mechanistically unprecedented multicomponent reactions. More recently, in ref. ⁴⁹ we showed how similar analysis can explain formation of unexpected and complex products, including MCRs involving as many as five components. These and other MECH analyses can be performed via the academically available WebApp (for access instructions and tutorials, please see source publications^{49,97})

Here, we illustrate the use of MECH to reconstruct the network of the Hantzsch reaction. The analysis begins by specifying the substrates (ammonia, ethyl acetoacetate and *p*-anisaldehyde) as well as admissible reaction conditions (here, ranging from basic to mildly acidic). In **Figure S109**, these substrates are in the bottom row, referred to as synthetic generation G_0 . The algorithm then applies matching mechanistic transforms to generate synthetic generation G_1 (17 species in the second-from-the-bottom row). Then, G_0 and G_1 species are combined and transforms are applied again. This process is iteratively repeated until some user-specified generation is reached. Here, the network is propagated up to G_7 and is comprised of >2,000 species, the majority of which are unstable or charged intermediates (nodes colored *pink* in **Figure S109**).

Once the network is generated, it is queried for the desired species – here, the products obtained from the robotized scan of the Hantzsch reaction hyperspace. When a given product is found within the network, the corresponding node is marked and mechanistic pathway(s) of mutually compatible steps (i.e., compatible with respect to conditions) are traced back from this node to the substrates. This generates an image such as that in **Figure S110** whereby the miniatures of the products are showed next to the nodes (and are color-coded as in the main text: molecules on *red* background = previously unreported products under Hantzsch reaction conditions; molecules on *blue* background = known products under Hantzsch reaction conditions).

Subsequently, this network is pruned to retain only the isolated products and their mechanistic pathways (in default setting, the shortest, but if alternative or longer pathways exist, the software also tracks them, see **Figure S111** and **Figure S112**; for this and other display options, see user manuals in source publications^{49,97}). Limiting to only selected products and pathways leaves a subgraph such as the one illustrated in **Figure S113**. For clarity, the network of the Hantzsch reaction is redrawn manually in **Figure S114** and provides a mechanistically more complete view than the network shown in the main-text **Figure 5**.

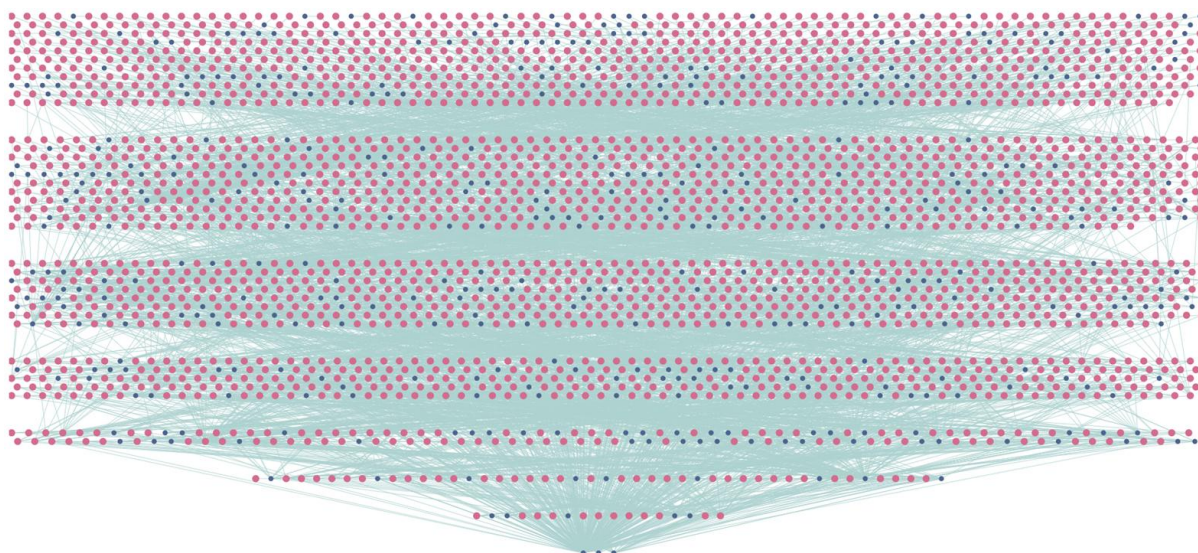


Figure S109. MECH screenshot of the network of mechanistic steps propagated from the ammonia, ethyl acetoacetate and *p*-anisaldehyde substrates (three nodes in the bottom row) of the Hantzsch reaction. The network was expanded to seven synthetic generations with limit for molecular weight of generated molecules set to 500 g/mol. Edges of the graph represent reactions (i.e., mechanistic steps), while nodes correspond to all possible products and intermediates. Here, nodes representing highly reactive or charged intermediates are enlarged and colored *pink* (85% of the molecules). These intermediates include, for instance, enols, hydrated ketones, imines, or non-stabilized enamines. While enols typically exist in equilibrium with their keto tautomers and some charged molecules can be stable and successfully isolated as salts, this display modality is standardized to classify only keto tautomers and uncharged moieties as “stable molecules” (represented by *blue* nodes).

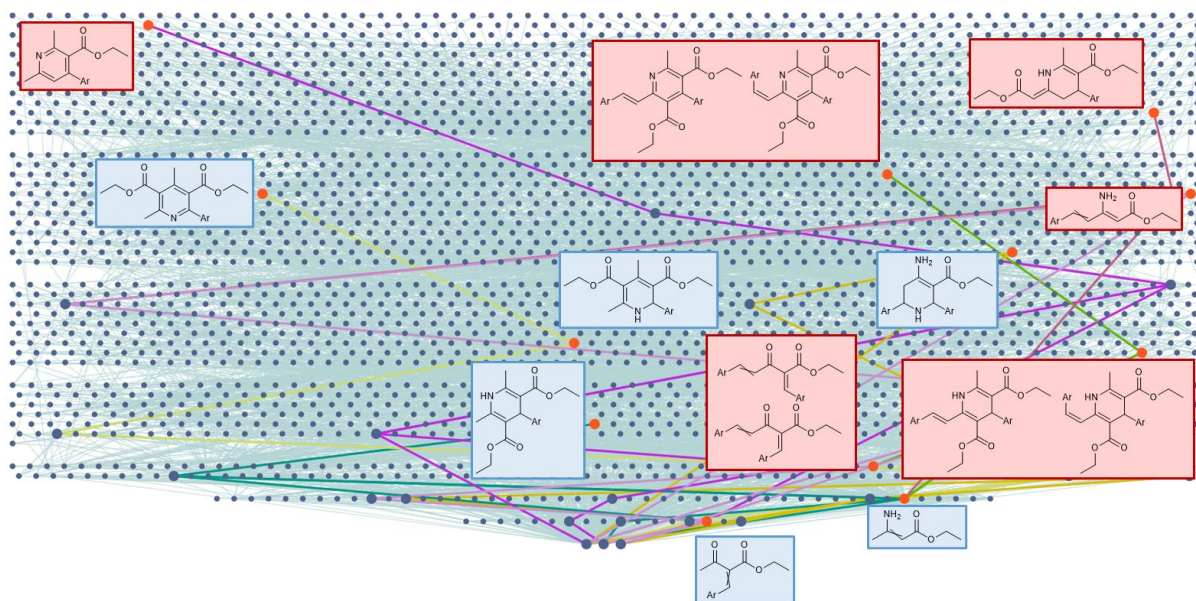
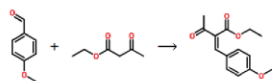
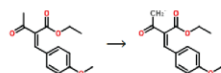


Figure S110. The same mechanistic network as in **Figure S109** but with the products of the Hantzsch reaction (isolated in the robotic hyperspace scan) identified in and overlaid over the network (*orange* nodes and miniatures next to them; *red* background = previously unreported products under Hantzsch reaction conditions; *blue* background = known products under Hantzsch reaction conditions). The shortest mechanistic pathways leading to these products are traced by colored bolded edges. Molecules placed together in one frame are represented by one node corresponding to a mixture of *E/Z* stereoisomers. Ar = *p*-methoxyphenyl.

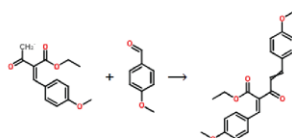
Reaction name: Stereoselective Knoevenagel condensation
 Reaction conditions: piperidine
 allow temp:H, allow AB: WB
 Solvent: EtOH
 Literature reference: 10.1016/j.tet.2017.03.041 and 10.3762/bjoc.9.215
[show similar reaction\(s\)](#)
[show competing steps](#)



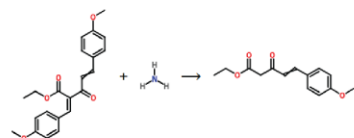
Reaction name: Deprotonation of unactivated ketone
 Reaction conditions: nBuLi or NaH
 allow temp:VL, allow AB: SB.B
 Solvent: THF
 Alternative Solvent: t-Butyl ethyl ether
 Literature reference: 10.1002/anie.200351413 and 10.1021/ja052164r (SI p.13) and 10.1016/j.tet.2011.03.116
[show competing steps](#)



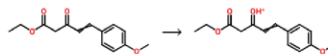
Reaction name: Addition of aldehydes to enolates
 Reaction conditions: THF
 allow temp:L,rt, allow AB: SB.B
 Solvent: THF
 Alternative Solvent: t-Butyl ethyl ether
 Literature reference: 10.1021/op020029i and 10.1016/j.bmc.2013.11.007 and 10.1016/j.bmcl.2005.02.066 and 10.1021/ja990400q and 10.1016/j.tet.2013.09.006
[show competing steps](#)



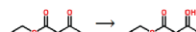
Reaction name: Retro-knoevenagel
 Reaction conditions: solvent
 allow temp:rt,H, allow AB: LA.WB.LA+WB.WA.B.N
 Solvent: DCM
 Literature reference: 10.1021/jo0602118 and 10.1021/acs.joc.2c00087 and 10.1021/ol401042b
[show competing steps](#)



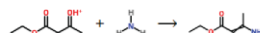
Reaction name: Ketone protonation
 Reaction conditions: H2SO4
 allow temp:L,rt,H.VH, allow AB: A.WA
 Solvent: EtOH
 Literature reference: Organic Chemistry, 7th Edition L. G. Wade, Jr p. 842 and 10.1021/oi500421k (SI S5)
[show competing steps](#)



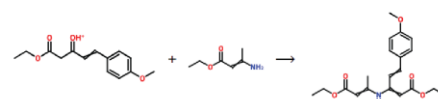
Reaction name: Ketone protonation
 Reaction conditions: H2SO4
 allow temp:L,rt,H.VH, allow AB: A.WA
 Solvent: EtOH
 Literature reference: Organic Chemistry, 7th Edition L. G. Wade, Jr p. 842 and 10.1021/oi500421k (SI S5)
[show competing steps](#)



Reaction name: Ammonia addition to protonated ketone
 Reaction conditions: AcOH
 allow temp:rt,H, allow AB: WA
 Solvent: EtOH
 Literature reference: 10.1007/s00044-017-2104-6 and 10.1021/oi500421k (SI S5)
[show competing steps](#)



Reaction name: Enamine addition to protonated ketone
 Reaction conditions: AcOH
 allow temp:rt,H, allow AB: WA
 Solvent: EtOH
 Literature reference: 10.1016/j.ejmech.2020.112538 and 10.1016/j.dyepig.2020.108370 10.3390/molecules23102530
[show competing steps](#)



Reaction name: Condensation of enamine with Michael acceptor
 Reaction conditions: EtOH, heating
 allow temp:rt,H, allow AB: WB.LA+WB.N.WA
 Solvent: EtOH
 Literature reference: 10.1007/BF00546737
[show competing steps](#)

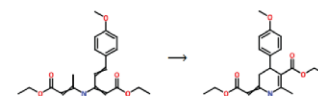
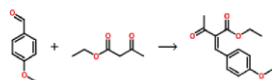
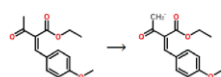


Figure S111 Upon clicking on any node in the network, MECH displays, in a separate window, details of the mechanistic route(s) leading to this particular species. The pathway shown in this screenshot is for product **19j**. Each step provides information about reagents and conditions, as well as links to illustrative literature references. If competing steps are found in the network, they can also be displayed by clicking on “show competing steps” hyperlink. For this and other display options, please see user manuals in the source publications^{49,97}.

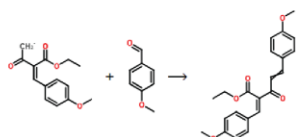
Reaction name: Stereoselective Knoevenagel condensation
 Reaction conditions: piperidine
 allow temp:H, allow AB: WB
 Solvent: EtOH
 Literature reference: 10.1016/j.tet.2017.03.041 and 10.3762/bjoc.9.215
 show similar reaction(s)
 show competing steps



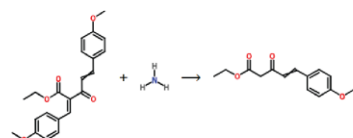
Reaction name: Deprotonation of unactivated ketone
 Reaction conditions: nBuLi or NaH
 allow temp:V.L.L, allow AB: SB.B
 Solvent: THF
 Alternative Solvent: t-Butyl ethyl ether
 Literature reference: 10.1002/anie.200351413 and 10.1021/ja052164r (SI p.13) and 10.1016/j.tet.2011.03.116
 show competing steps



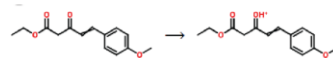
Reaction name: Addition of aldehydes to enolates
 Reaction conditions: THF
 allow temp:L.r.t, allow AB: SB.B
 Solvent: THF
 Alternative Solvent: t-Butyl ethyl ether
 Literature reference: 10.1021/op020029i and 10.1016/j.bmc.2013.11.007 and 10.1016/j.bmcl.2005.02.066 and 10.1021/ja990400q and 10.1016/j.tet.2013.09.006
 show competing steps



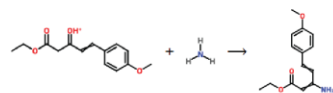
Reaction name: Retro-knoevenagel
 Reaction conditions: solvent
 allow temp:rt.H, allow AB: LA.WB.LA+WB.WA.B.N
 Solvent: DCM
 Literature reference: 10.1021/jo0602118 and 10.1021/acs.joc.2c00087 and 10.1021/oi401042b
 show competing steps



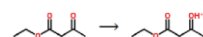
Reaction name: Ketone protonation
 Reaction conditions: H2SO4
 allow temp:L.r.t.H.VH, allow AB: A.WA
 Solvent: EtOH
 Literature reference: Organic Chemistry, 7th Edition L. G. Wade, Jr p. 842 and 10.1021/oi500421k (SI S5)
 show competing steps



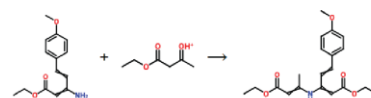
Reaction name: Ammonia addition to protonated ketone
 Reaction conditions: AcOH
 allow temp:rt.H, allow AB: WA
 Solvent: EtOH
 Literature reference: 10.1007/s00044-017-2104-6 and 10.1021/oi500421k (SI S5)
 show competing steps



Reaction name: Ketone protonation
 Reaction conditions: H2SO4
 allow temp:L.r.t.H.VH, allow AB: A.WA
 Solvent: EtOH
 Literature reference: Organic Chemistry, 7th Edition L. G. Wade, Jr p. 842 and 10.1021/oi500421k (SI S5)
 show competing steps



Reaction name: Enamine addition to protonated ketone
 Reaction conditions: AcOH
 allow temp:rt.H, allow AB: WA
 Solvent: EtOH
 Literature reference: 10.1016/j.ejmech.2020.112538 and 10.1016/j.dyepig.2020.108370 10.3390/molecules23102530
 show competing steps



Reaction name: Condensation of enamine with Michael acceptor
 Reaction conditions: EtOH, heating
 allow temp:rt.H, allow AB: WB.LA+WB.N.WA
 Solvent: EtOH
 Literature reference: 10.1007/BF00546737
 show competing steps

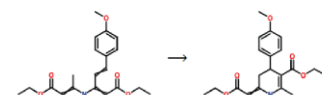


Figure S112 For some products more than one pathway can exist. Here, MECH displays such an alternative mechanistic route leading to the same product **19i** as in the preceding figure.

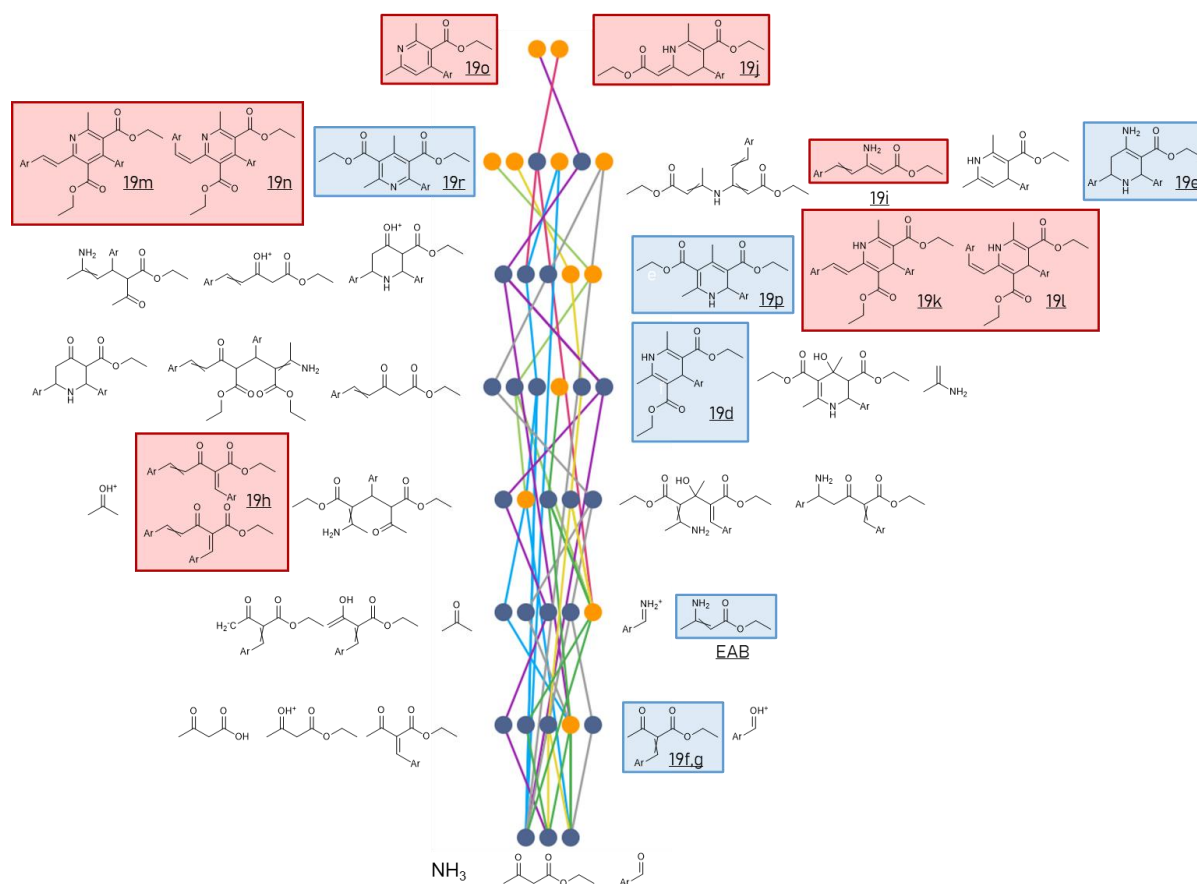


Figure S113. MECH screenshot of a subgraph of the network from **Figure S110**. This subgraph is limited to the products isolated in the robotic scan of the Hantzsch reaction hyperspace and to their shortest mechanistic pathways. Structures of all products (*orange* nodes) and intermediates (*blue* nodes) of each synthetic generation are shown next to the graph on the same level as their corresponding nodes. Isolated molecules are denoted by frames (as before, *red* background = previously unreported products under Hantzsch reaction conditions; *blue* background = known products under Hantzsch reaction conditions). Numbers next to the structures of isolated molecules correspond to the numbering in the main text **Figure 5**. Molecules placed together in one frame are represented by one node and correspond to a mixture of *E/Z* stereoisomers. Ar = *p*-methoxyphenyl.

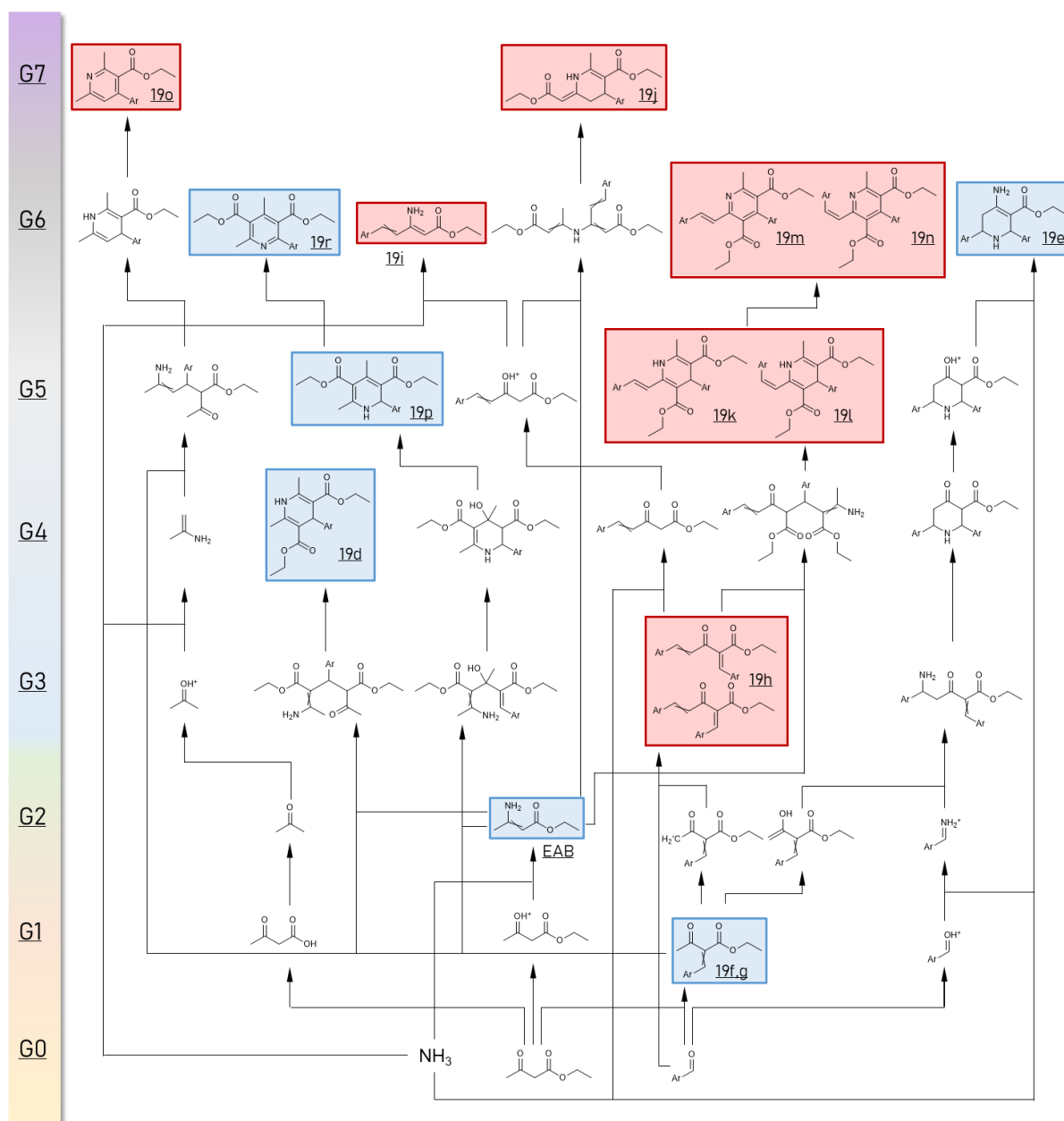


Figure S114. Full mechanistic network redrawn for clarity. Structures of substrates, products and intermediates are placed according to the synthetic generations from the “parent” network (Figure S109 and Figure S110) and indicated by the sidebar on the left. Ar = *p*-methoxyphenyl.

4.12. Condition hyperspace boundaries between distinct outcomes of Hantzsch reaction

To further investigate the relationship between the yield reaction space of the three main products of the Hantzsch reaction (**19k**, **19d** and **19e**) and the underlying mechanistic reaction network, we also analyzed the boundaries (in the condition space) between the regions yielding these products. Because the three yields in a 3D space of conditions form a 3D vector field, the geometry of these boundaries (if such clean boundaries even exist) may potentially be quite complex and hard to visualize. Here, we pursued two approaches: one based on differential yields, and another grounded in information theory and Shannon entropy.

4.12.1. Maps of differential yield as a tool for analyzing the reaction hyperspaces

In analyzing various modalities of visualizing yield distributions over reaction hyperspaces, we have found useful an approach focusing on differential yields of a particular species with respect to the sum of other products. This metric allows for the visualization of boundaries between zones of selective product formation and for linking these observations to the underlying reaction network. For example, **Figure S115d** shows a map of differential yields of the Hantzsch product **19k** with respect to products **19d** and **19e** – here, the yield of **19k** is subtracted from the sum of yields of products **19d** and **19e**. The positive (red) values on the map show the zones where **19d** or/and **19e** dominate and negative values (blue) clearly show the zone selective towards formation of **19k**. White color depicts transition area where none of the products dominates. Interestingly, the red zone in **Figure S115d** is further subdivided into two parts. Based on the raw yield maps in **Figure S115b,c** we can assign those separate areas to the products **19d** and **19e** respectively. Specifically, the underlying reaction network (**Figure S116**) indicates that **19d** is formed in a reaction of chalcone **19f,g** (which reacts as a Michael acceptor) with enamine while **19e** requires enolate form of **19f,g** (which reacts as a nucleophile) in both pathways (with aldehyde to give **19h** or with iminium cation). With this in mind, the map in **Figure S115d** visualizes two zones with different regimes of **19f,g** reactivity.

Next, let us analyze the formation of **19k**. It can be formed from keto or enol forms of **19f,g** (from product **19d**^{98–102} or intermediate **19h**), which implies that **19k** can be formed by both of the above mechanisms. Subtracting of yield of product **19k** from the combined yields of **19d** and **19e** and plotting the map of this difference across the condition space (**Figure S115d**) reveals two distinct maxima (red zones), separated by a zone of lower values (marked by green circle in **Figure S115d**). This is expected if the **19f,g** reactivity mode branches as illustrated in the reaction network in **Figure S116** and in the Venn diagram in **Figure S115g**. In real condition space, the zones corresponding to yields of **19d** and **19e** in the Venn diagram have blurred boundaries and are overlapping, but subtracting the presumed intersection (purple **Figure S115g**) from the sum map (**Figure S115h**) should theoretically produce a dip (green dot in **Figure S115i**) between the two resolved maxima. As noted above, the respective dip in the experimental data is shown by the green circle in **Figure S115d**. Presence of this dip is evidence for the conjectured branching of **19f,g** reactivity.

In contrast, maps produced by subtracting either **19d** or **19e** from their respective sums (**19e** + **19k** and **19d** + **19k**) do not feature multiple disjoint local maxima (red zones of **Figure S115e,f**).

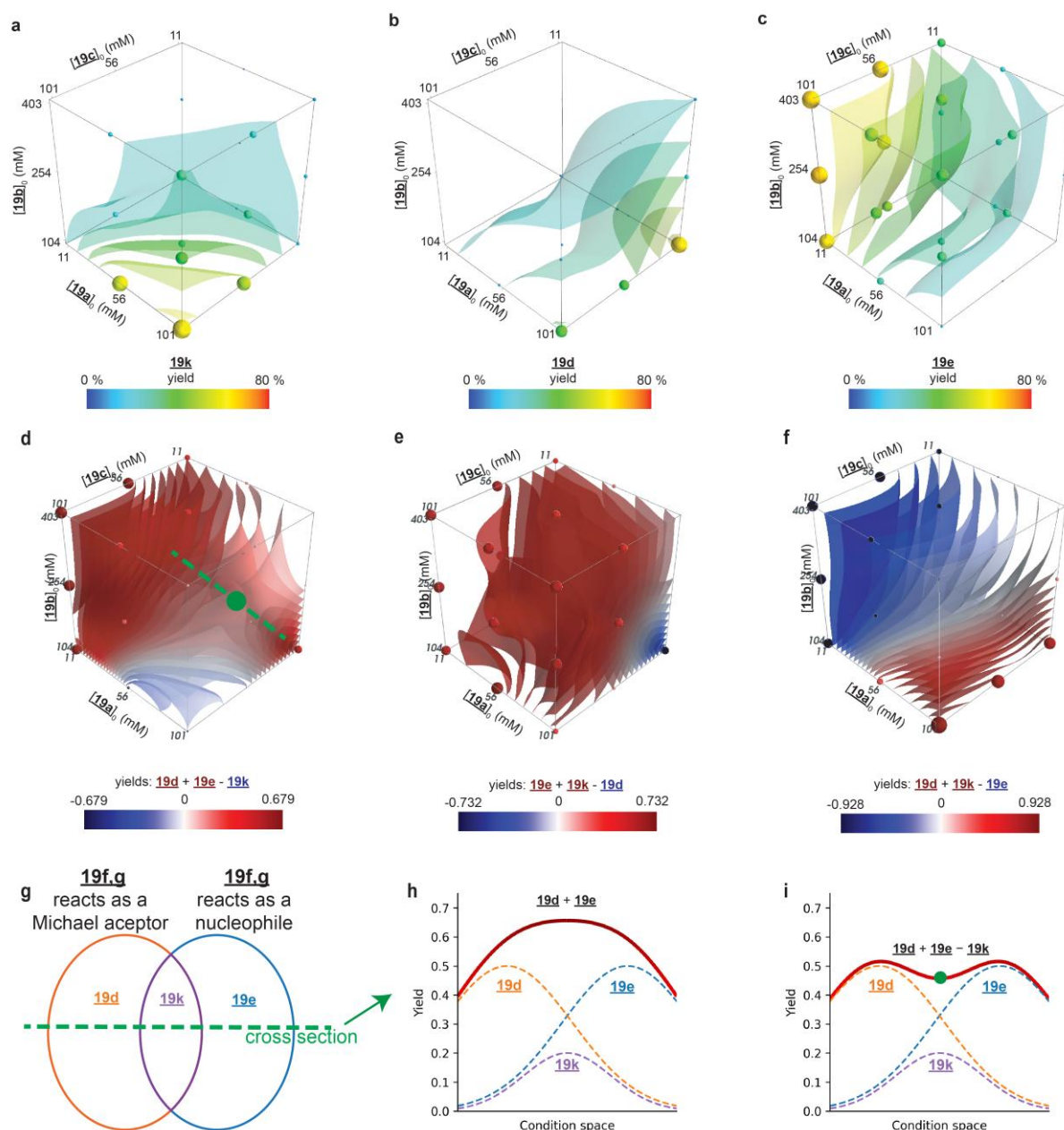


Figure S115. **a-c**, Raw yield maps of the Hantzsch reaction products: **19k**, **19d** and **19e**. **d**, A differential yield map, in which the yield of **19k** is subtracted from the sum of yields of **19d** and **19e**. The positive values (colored red) show areas where **19d** or/and **19e** dominate and negative values (blue) show the area selective towards formation of **19k**. **e**, Another differential yield map, in which the yield of **19d** is subtracted from the sum of yields of **19e** and **19k**. The positive values (colored red) show areas where **19e** or/and **19k** dominate and negative values (blue) show the area selective towards formation of **19d**. **f**, A differential yield map, in which

4.12.2. Information boundaries in hyperspace between distinct reaction outcomes

Next, we use Shannon entropy¹⁰³ as a measure of the informational complexity of the reaction outcome. Let Y_{19d} , Y_{19e} , and Y_{19k} be the experimental yields of products **19d**, **19e** and **19k** at a given condition, and let $Y_{\Sigma} = Y_{19d} + Y_{19e} + Y_{19k}$ be the sum of the yields. To obtain Shannon entropy of the reaction outcome at a given condition, we follow the standard definition by taking the “event probabilities” p_i , where $i = 1, 2, 3$, to be equal to normalized yields:

$$S = - \sum_{i\text{-th product}} p_i \ln(p_i) = - \left[\frac{Y_{19d}}{Y_{\Sigma}} \ln \left(\frac{Y_{19d}}{Y_{\Sigma}} \right) + \frac{Y_{19e}}{Y_{\Sigma}} \ln \left(\frac{Y_{19e}}{Y_{\Sigma}} \right) + \frac{Y_{19k}}{Y_{\Sigma}} \ln \left(\frac{Y_{19k}}{Y_{\Sigma}} \right) \right]$$

This definition of p_i is justified by the formation of all these products from the intermediate **19f,g** in the mechanism (**Figure S116**). The map of thus computed Shannon entropy in the condition space is shown in **Figure S117**.

There are two extreme values of entropy seen in the map: low entropy (≈ 0): when one of the products dominates (reaction is selective) and high entropy (up to $\ln(3) \approx 1.098$) when reaction is non-selective and many reaction pathways are competing. **Figure S117** shows two areas of low entropy (blue), where the formation of **19d** and **19e** is selective. They are clearly separated by an informationally dense zone (red color and gap between isosurfaces) that comprises competing formation of **19k** and transition between two regimes of **19f,g** reactivity. We find it somewhat striking that such a simple and robust tool as the information-theoretical Shannon entropy manifests so clearly the existence of a boundary between distinct reaction outcomes in a geometric object as mathematically complex as a 3D vector field. We expect that the utility of this metric in hyperspace analysis may be even greater for reactions producing more (say, ~ 10) different products. Anticipating this usefulness, we incorporate Shannon entropy as one of the visualization modalities in our HyperspaceViewer software (see caption to **Figure S117** and **Supplementary Video 4**).

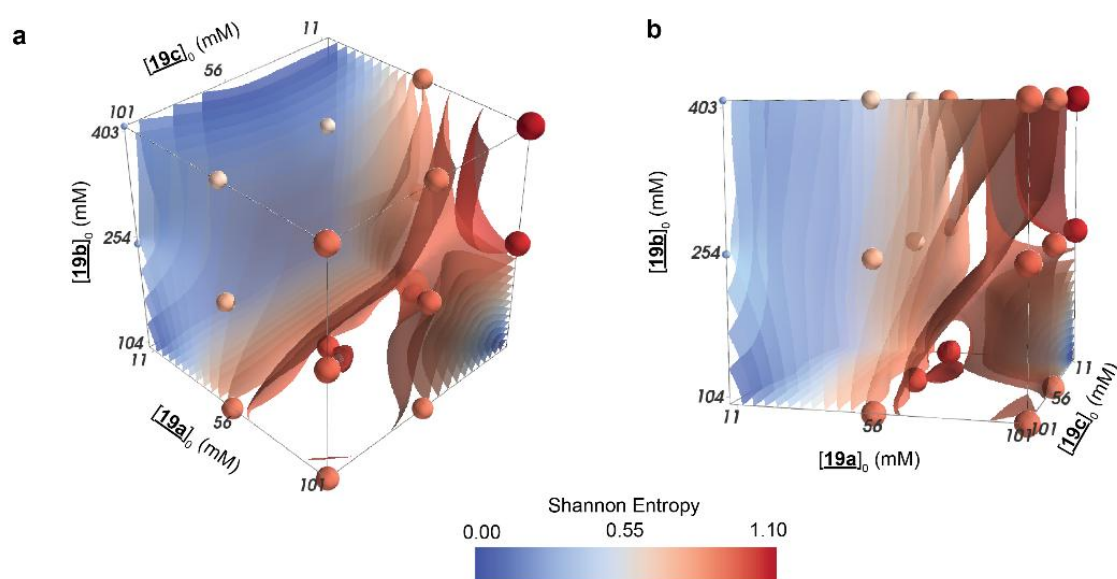


Figure S117. Map of Shannon entropy of the reaction outcomes in the condition space of the Hantzsch reaction. See text for the definition of Shannon entropy. Panels **a** and **b** show the same map viewed from two different perspectives. Surfaces of equal Shannon entropy were computed by the marching cubes algorithm⁵⁸ operating on a 50×50×50 regular grid that was obtained with the Radial Basis Functions (RBF) interpolator applied to the computed entropy values, which are shown by the spheres. For interactive versions of these plots, one way is to use the HyperspaceViewer released with this article: load the “5_Hantzsch_yields.csv” data file from CSV folder supplied with the HyperspaceViewer and then select “Entropy” as the function to be plotted. Another method is to run the script “robowski/visualize_results/examples/plot_3d_map_for_hantzsch_shannon_entropy_by_HPLC.py” in the code repository associated with this article.

4.13. Anomaly identification of S_N1 reaction by mass spectrometry

Mass spectra of crude mixture manifesting the unexpected spectral feature (cf. main text Figure 3e) were obtained with Advion CMS mass spectrometer operating in atmospheric pressure chemical ionization mode (APCI). Background spectrum was collected over 15 min integration time. Then, a 10 μ L droplet of reaction crude was deposited on a clean capillary of the “solid sample probe” and introduced into the APCI chamber. The resulting spectrum is shown in **Figure S118**. The peaks with m/z ratios of 685.227 and 686.251 can be attributed to the carbocation **15f** and the radical cation **15g**, respectively. They are both mechanistically viable products (in solution) or fragments (in gas phase) from the carbocation **15e** (most likely to be the pink species according to TD-DFT calculation). For details of the mechanistic pathways and calculated absorption spectra, see **Section 6.2**.

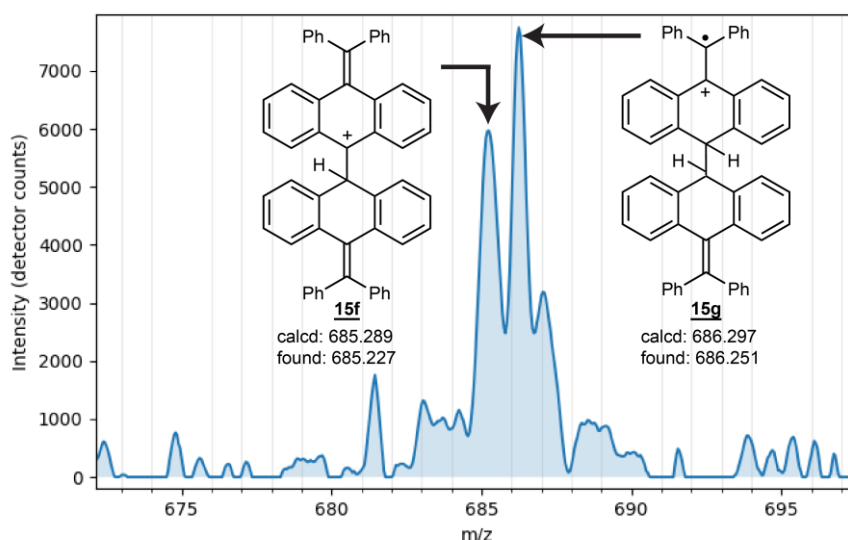


Figure S118. Mass spectrum of reaction crude containing the unexpected pink species.

4.14. Control measurements of the water content in E1 reaction mixtures

To control for possible acquisition of additional water from ambient air by acetonitrile used as solvent in E1 reaction, we used the IR-spectroscopic method of determination of trace water in acetonitrile reported by Ludvik *et al.*¹⁰⁴. Taking measurements with Perkin Elmer “Spectrum Two” FTIR spectrometer and liquid transmission cell (PIKE Technologies, SmartSeal 165-3003), we calibrated the absorbance at the water absorption peak with respect to water concentration of acetonitrile (**Figure S119b**). The fringes in the raw spectrum (**Figure S119a**) appear due to slight difference between the refractive index of acetonitrile-water mixture and the refractive index of the pure acetonitrile, which changes the optical path length and therefore the transmission spectrum of the Fabri-Perot resonator formed by the inner walls of the cell. These fringes were filtered out by a digital notch filter with center frequency $f_0 = 0.031676$ cm and Q-factor equal to unity. Filtered spectrum is shown by red curve in **Figure S119a**.

Preparation of 54 conditions of E1 reaction takes between 23 and 36 minutes. Accordingly, we have measured the amounts of water in the 2 mL vials containing 0.5 mL of acetonitrile (but no other E1 reaction substrates other than water) that were exposed to ambient 85% humidity for 23 and 36 minutes (absorption spectra are shown in **Figure S120**). The water concentrations were 196 mM (23 min) and 266 mM (36 min), with the average equal to 231 mM.

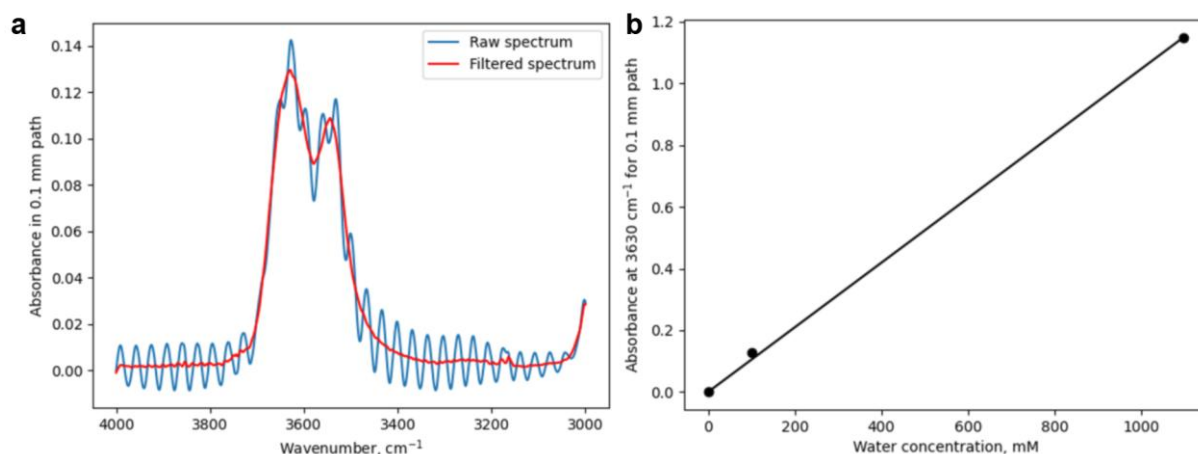


Figure S119. a, Fourier-transform infrared (FTIR) absorption spectrum of acetonitrile with 100 mM of water, in reference to the pure acetonitrile spectrum. Two IR absorption bands of water are visible at 3545 cm^{-1} and 3630 cm^{-1} . Measurements are performed in transmission cell with 0.1 mm path length. The fringes in the raw spectrum appear due to slight difference between the refractive index of acetonitrile-water mixture and the refractive index of the pure acetonitrile, which changes the optical path length and therefore the transmission spectrum of the Fabri-Perot resonator formed by the inner walls of the cell. These fringes were filtered out by a digital notch filter with center frequency $f_0 = 0.031676$ cm and Q-factor equal to unity. The filtered spectrum is shown by the red curve. **b**, Calibration curve for measuring water concentration in acetonitrile by the absorbance at 3640 cm^{-1} peak.

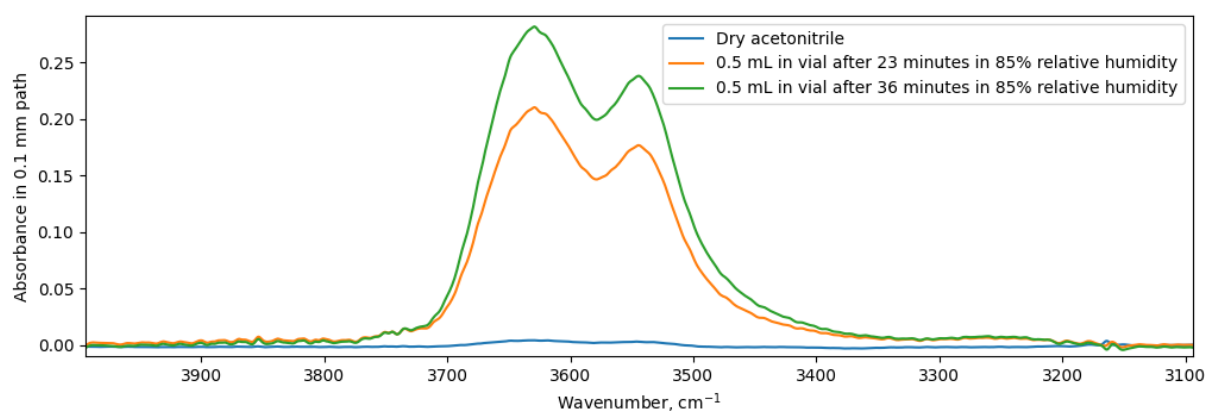


Figure S120. IR absorption spectra of dry acetonitrile (blue), as well as of 0.5 mL of acetonitrile with water in a 2 mL vial after indicated amount of time (23 min, orange and 36 min, green) in ambient 85% relative humidity.

4.15. Raw data of yield maps

4.15.1. E1 reaction

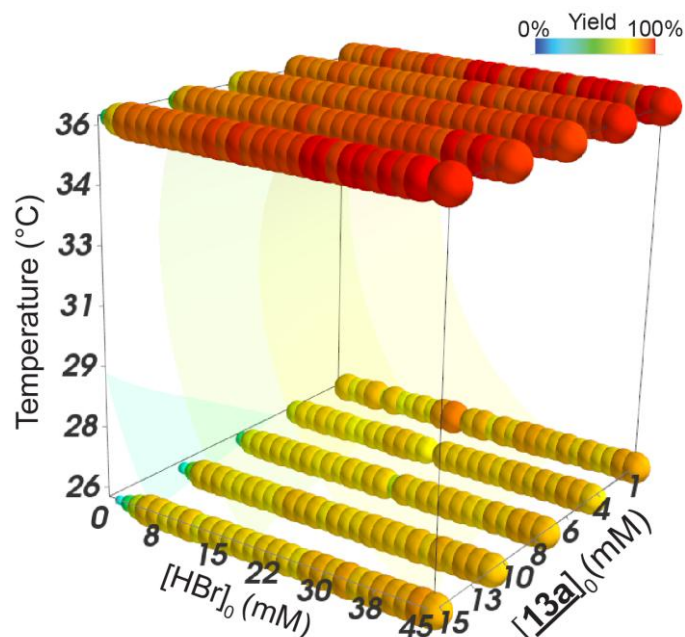


Figure S121. The yield distributions of product **13b**. The full, “saturated” cube of 310 E1 conditions (two temperatures, 26 and 36 °C; five initial substrate concentrations, 1.5–15.0 mM; 31 HBr concentrations, 0.15–45.0 mM). Marker colors and sizes are proportional to yields. Surfaces of equal yield are at 40%, 60%, and 80%.

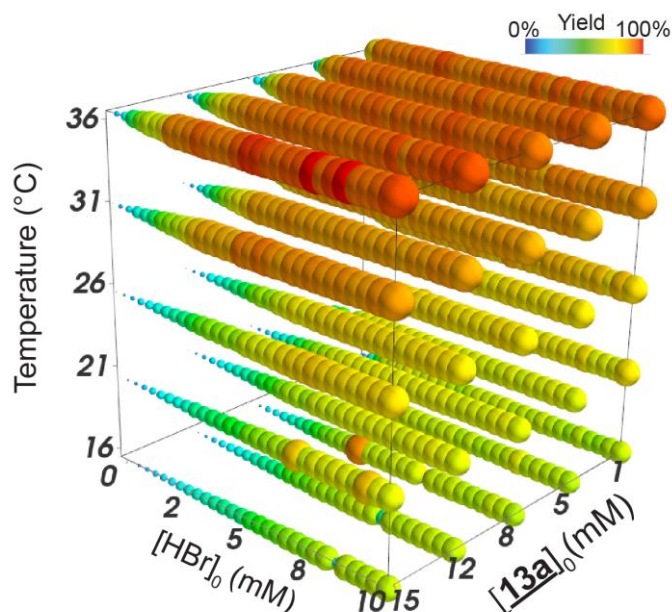


Figure S122. The yield distributions of product **13b**. The full cube of 775 E1 conditions (five temperatures, 16 to 36 °C; five initial substrate concentrations, 1.5–15.0 mM; 31 HBr concentrations, 0.1–10.0 mM). Marker colors and sizes are proportional to yields.

4.15.2. S_N1 reaction in Fig. 2

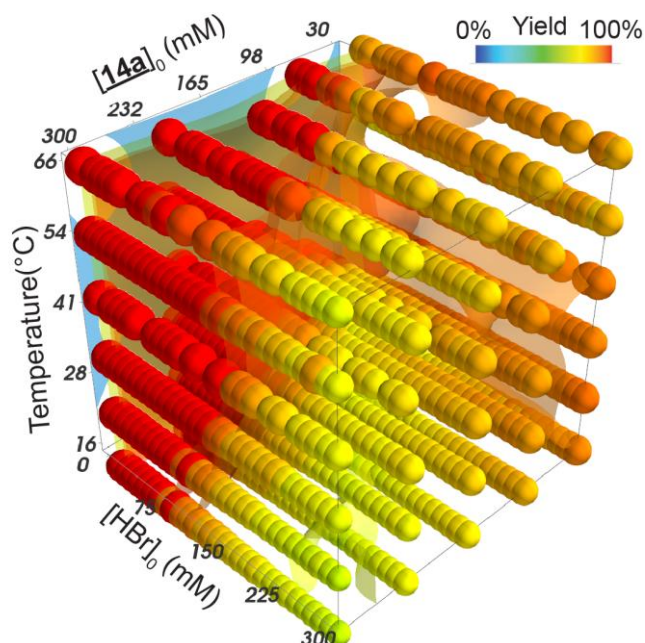


Figure S123. Reaction yield distribution from 930 raw datapoints of S_N1 reaction in **Fig. 2**. To reproduce this plot in an interactive mode, run the script `robowski/visualize_results/examples/plot_3d_map_for_simpleSN1_reaction.py` in the associated code repository. To rerun the spectral unmixing on the raw experimental spectra prior to plotting, run the `robowski/uv_vis_absorption_spectroscopy/examples/simple_SN1/process_simple_SN1_2023-12-11-run01.py` script.

4.15.3. S_N1 reaction in Fig. 3

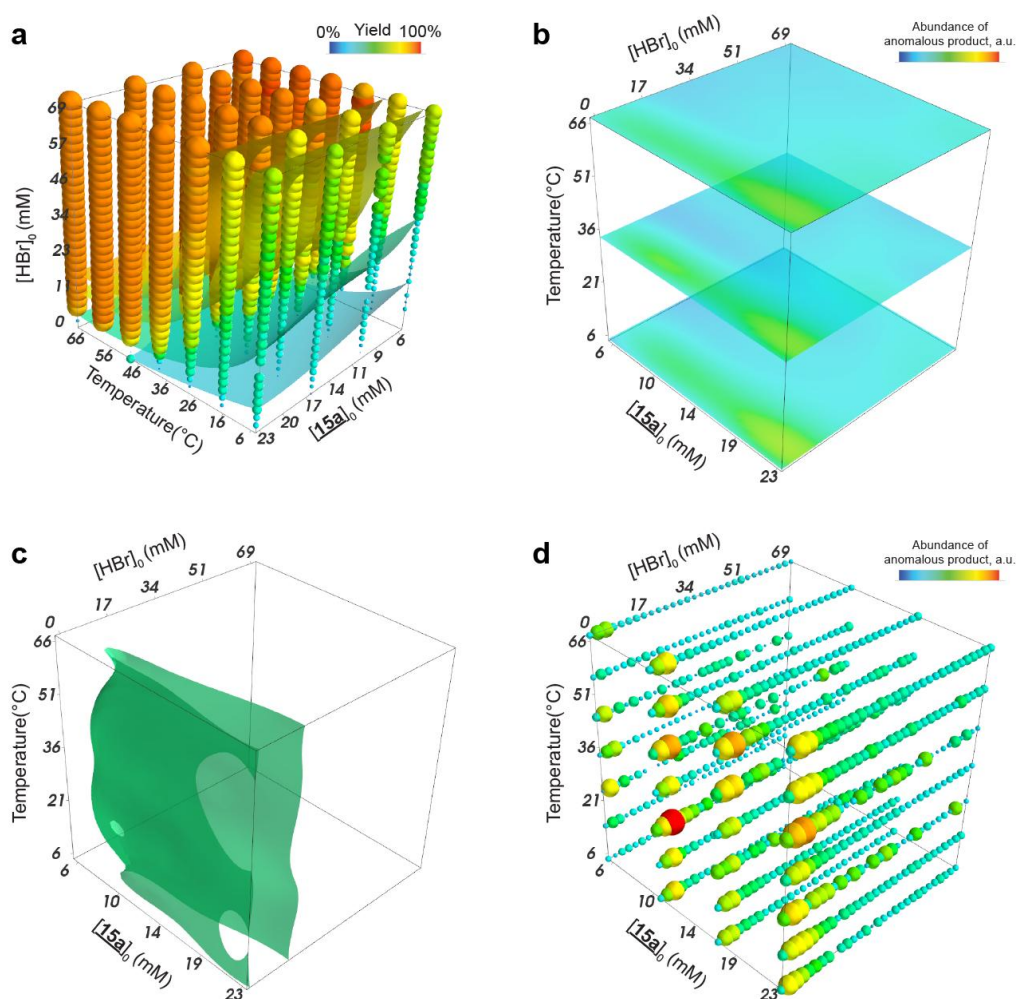


Figure S124. **a**, Reaction conversion distribution from 1085 raw datapoints of S_N1 reaction in **Fig. 3**. To reproduce this plot in an interactive mode, run the script `robowski/visualize_results/examples/plot_3d_map_for_SN1_reaction.py` in the associated code repository. On the line 130 of this script, you should set the parameter `data_for_spheres='raw'` for showing all the raw data points. To rerun the spectral unmixing on the raw experimental spectra prior to plotting, run the `robowski/uv_vis_absorption_spectroscopy/examples/SN1/process_simple_reactions_2023-07-05-run01.py`. **b-d**, Map of the conditions in which the anomalous outcome was observed. Due to the pure substance being impossible to isolate, neither the accurate spectrum nor its calibration with respect to concentrations are available for the anomalous product. In the absence of more reliable methods, the abundance of anomalous product here is quantified as the difference of absorbances at wavelengths 510 nm and 680 nm. The approximate shape of the relevant region in the condition space is visualized by means of planar cross-sections (**b**) and an isosurface (surface of equal abundance) (**c**) of the interpolated and smoothed abundance map, whose raw data points are shown in (**d**).

4.15.4. Ugi-type multicomponent reaction

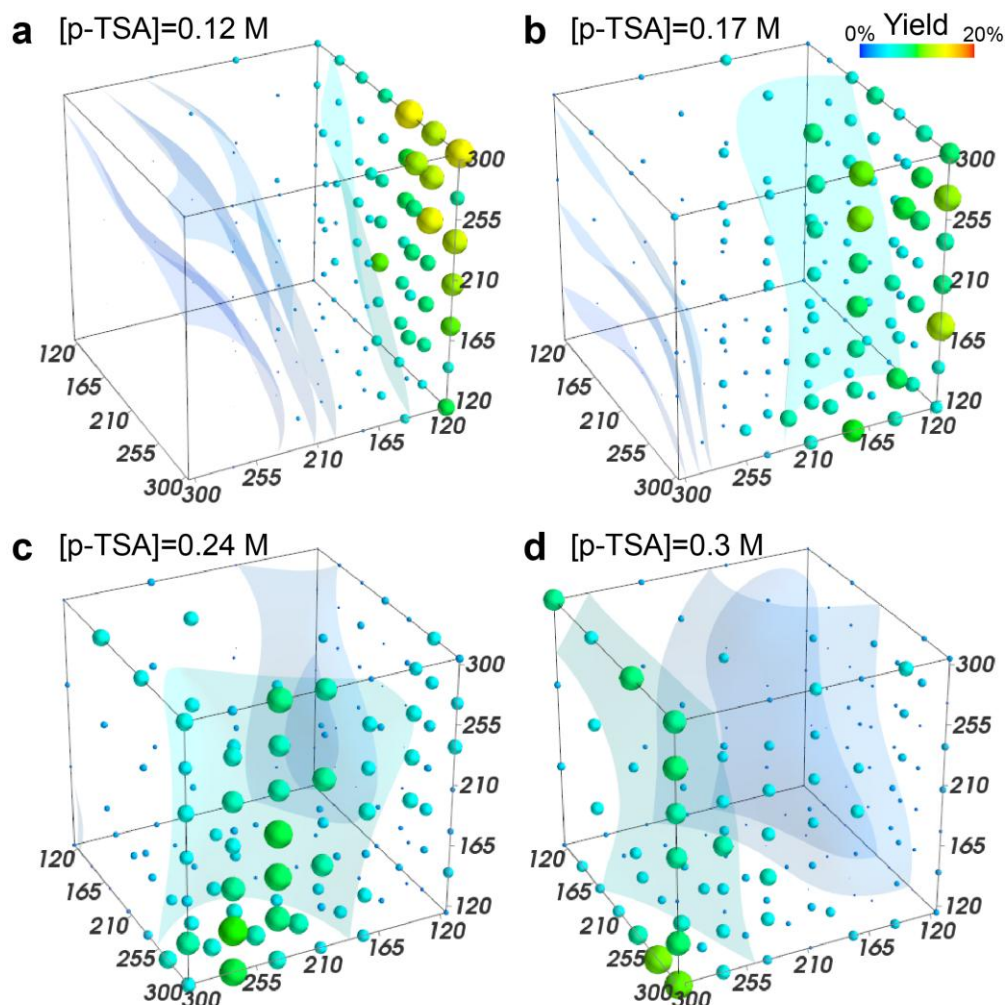


Figure S125. The yield distributions of product **16e** for various concentrations of the p-TSA (indicated in each panel). Four out of 21 experimentally sampled concentrations of p-TSA are displayed here. To reproduce this plot in an interactive mode, run the script `robowski/visualize_results/examples/view_4d_data_for_ugi_reaction.py` in the associated code repository. Selecting the p-TSA concentration (choosing which of the four panels to show) is controlled by the parameter `ptsa_target` on line 47: set `ptsa_target = ptsa_targets[1]` for showing panel **a**, or `ptsa_target = ptsa_targets[2]` for panel **b**, and so on. To rerun the spectral unmixing on the raw experimental spectra prior to plotting, run the `robowski/uv_vis_absorption_spectroscopy/examples/Ugi_reaction/process_ugi_from_2023_06_19_run01_to_2023_06_28_run03.py`.

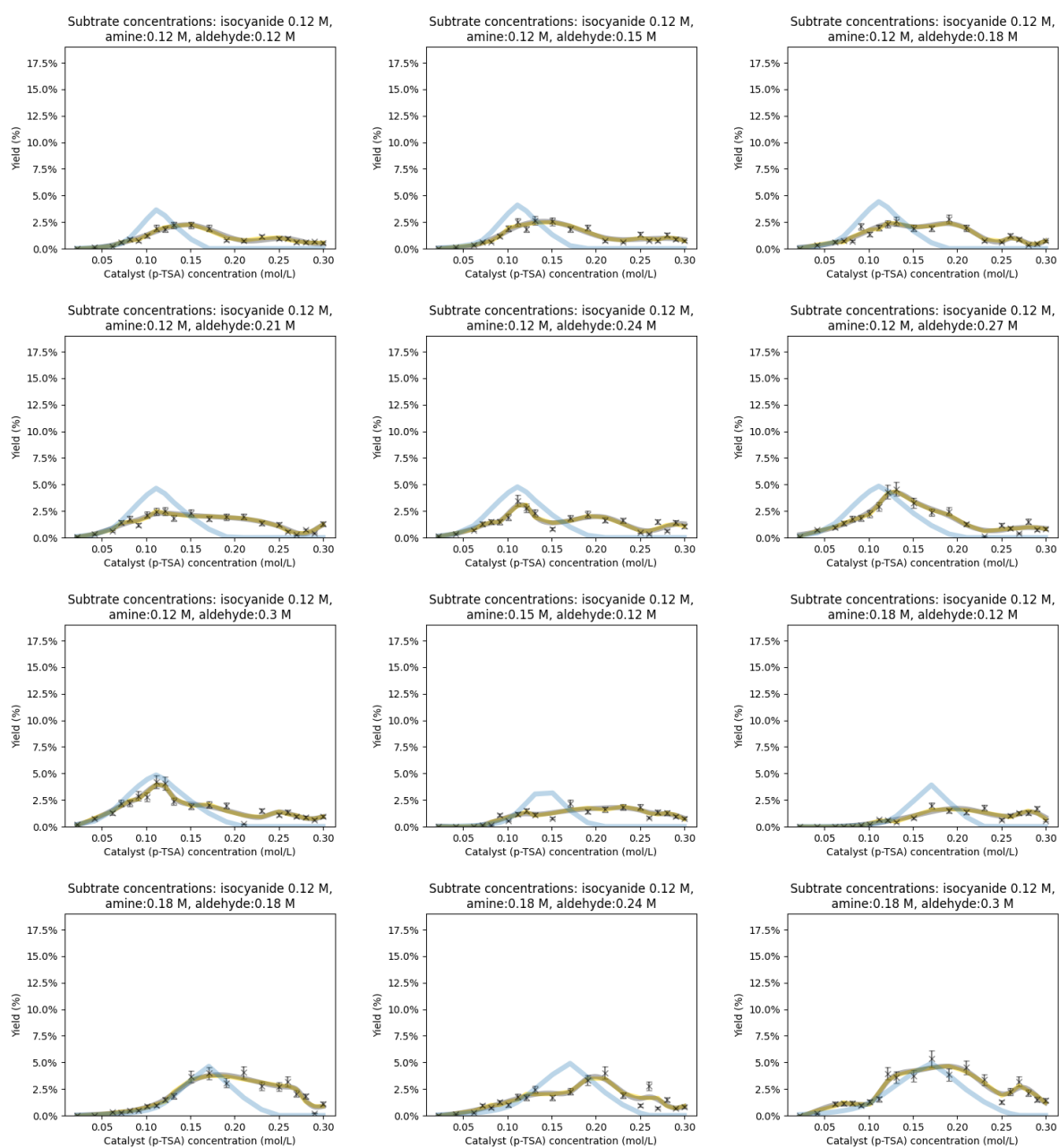


Figure S126. Yield of Ugi multicomponent reaction. Raw experimental data is shown by black “x” markers. Errorbars correspond to 14.5% relative error. Yellow curve is the result of weighted Savitsky-Golay filtering with degree 2 and window length equal to 5 points. Grey curve is the weighted B-spline. Blue curve is the prediction of the best-fit kinetics model described in **Section 5.3**. Starting concentrations of substrates are indicated above each plot.

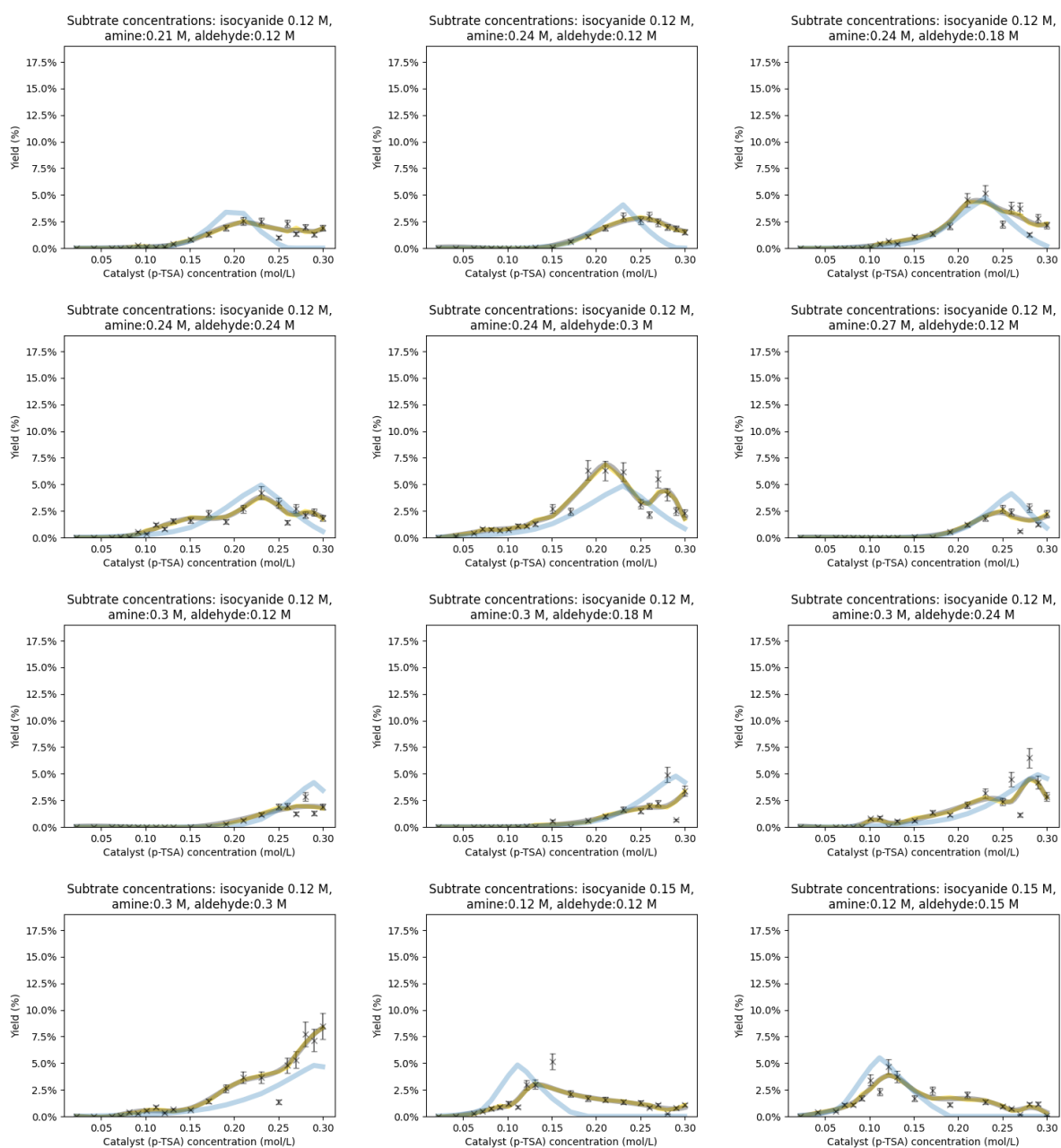


Figure S127. Yield of Ugi multicomponent reaction. Raw experimental data is shown by black “x” markers. Errorbars correspond to 14.5% relative error. Yellow curve is the result of weighted Savitsky-Golay filtering with degree 2 and window length equal to 5 points. Grey curve is the weighted B-spline. Blue curve is the prediction of the best-fit kinetics model described in **Section 5.3**. Starting concentrations of substrates are indicated above each plot.

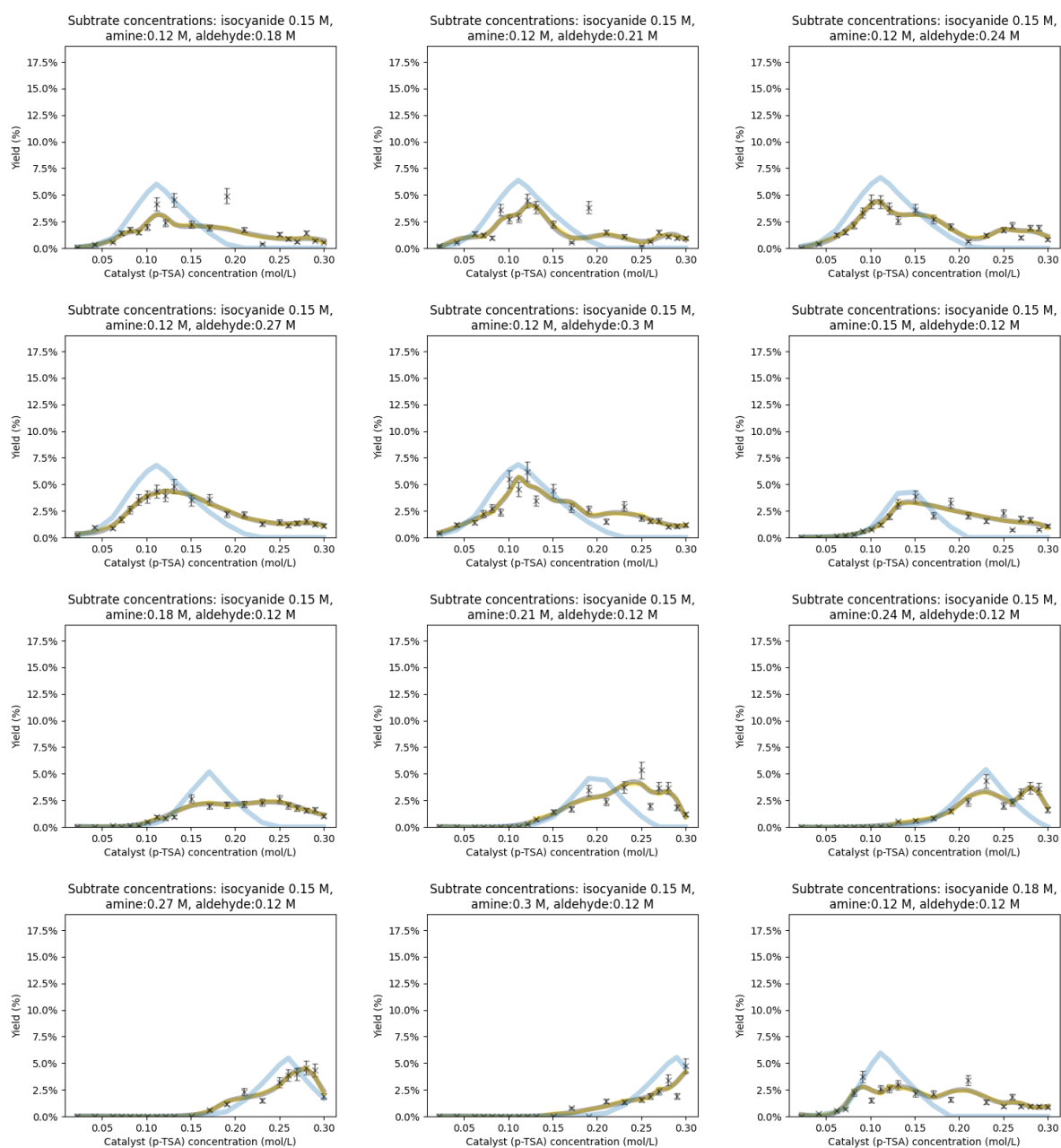


Figure S128. Yield of Ugi multicomponent reaction. Raw experimental data is shown by black “x” markers. Errorbars correspond to 14.5% relative error. Yellow curve is the result of weighted Savitsky-Golay filtering with degree 2 and window length equal to 5 points. Grey curve is the weighted B-spline. Blue curve is the prediction of the best-fit kinetics model described in **Section 5.3**. Starting concentrations of substrates are indicated above each plot.

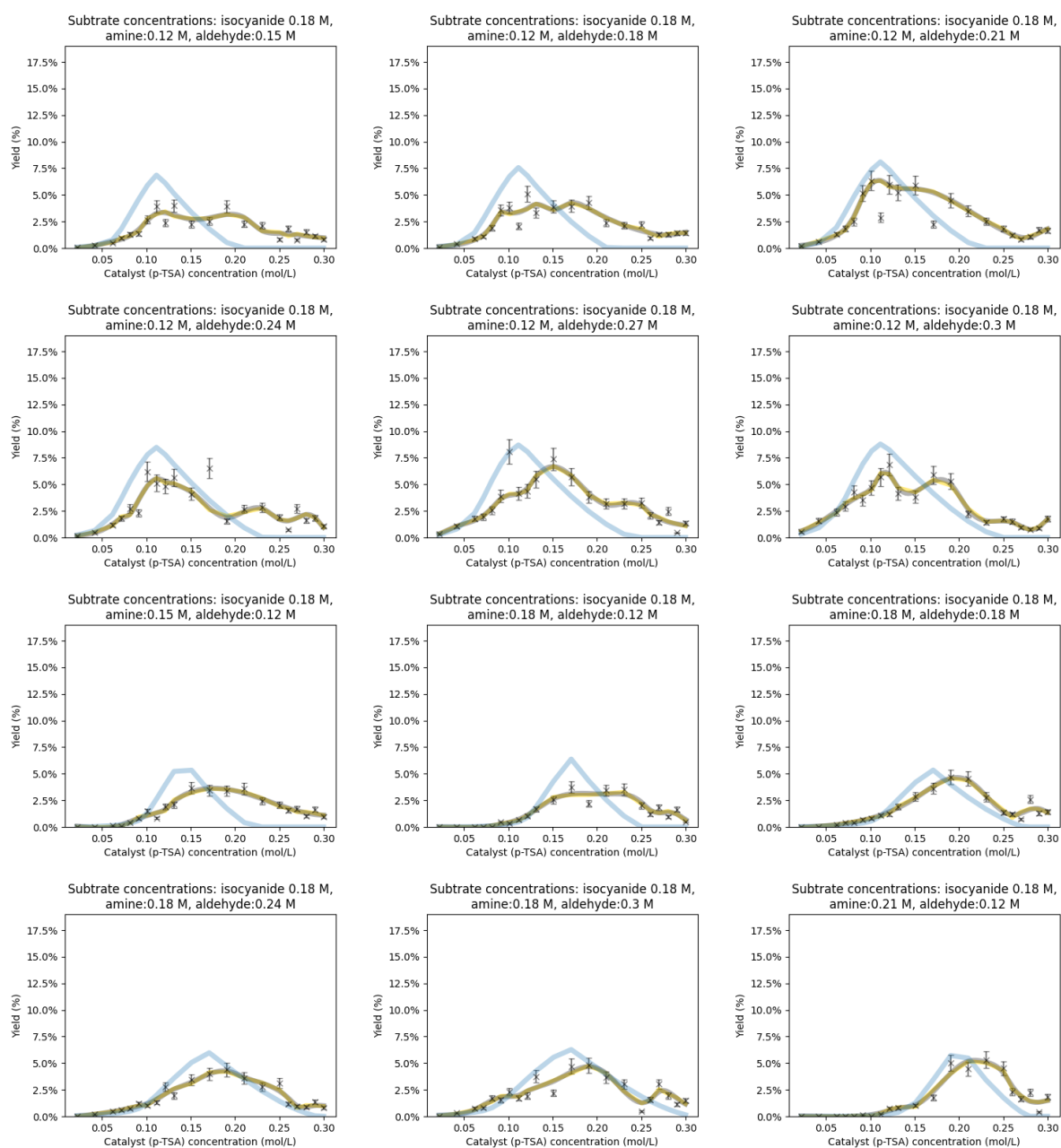


Figure S129. Yield of Ugi multicomponent reaction. Raw experimental data is shown by black “x” markers. Errorbars correspond to 14.5% relative error. Yellow curve is the result of weighted Savitsky-Golay filtering with degree 2 and window length equal to 5 points. Grey curve is the weighted B-spline. Blue curve is the prediction of the best-fit kinetics model described in **Section 5.3**. Starting concentrations of substrates are indicated above each plot.

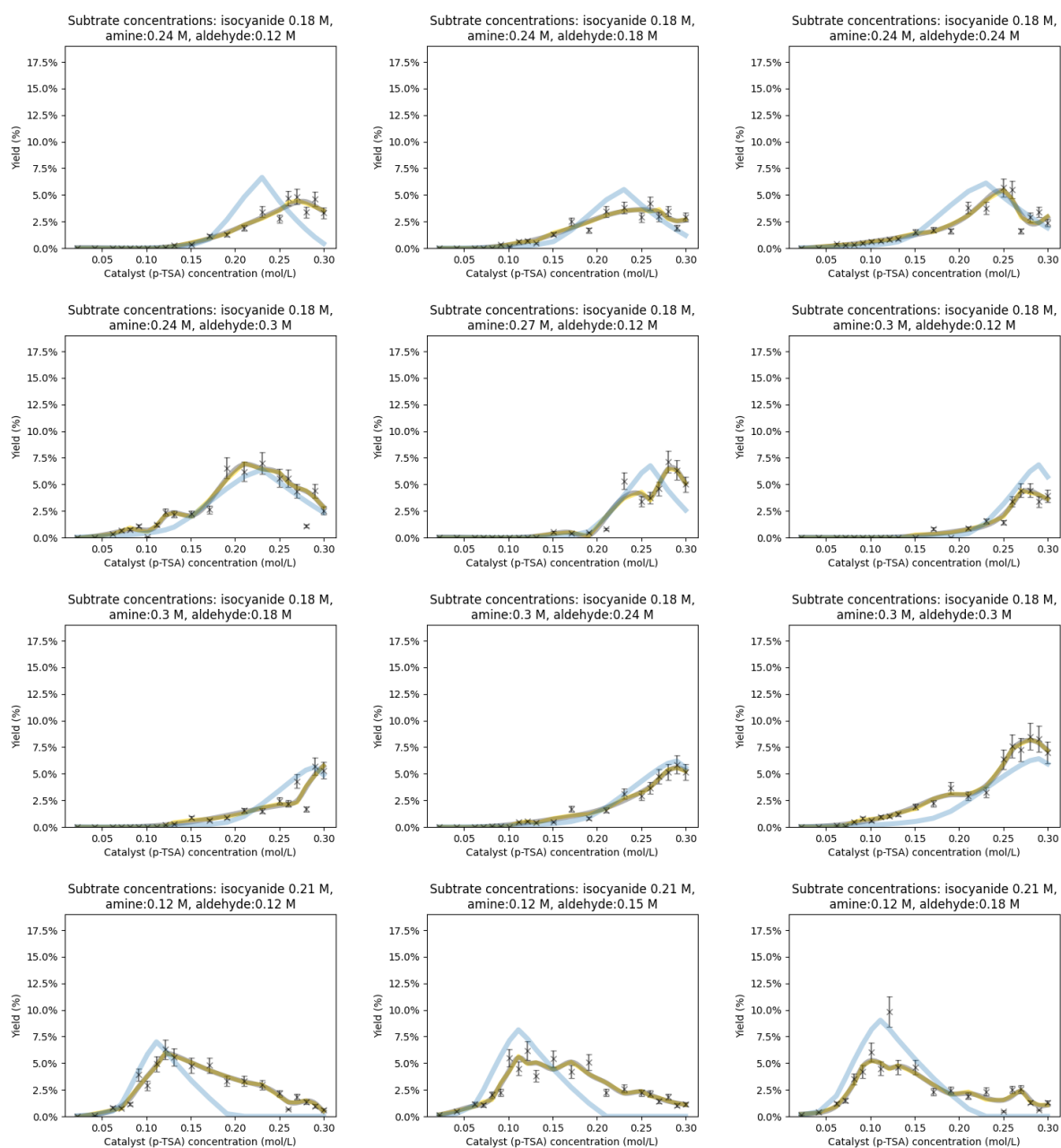


Figure S130. Yield of Ugi multicomponent reaction. Raw experimental data is shown by black “x” markers. Errorbars correspond to 14.5% relative error. Yellow curve is the result of weighted Savitsky-Golay filtering with degree 2 and window length equal to 5 points. Grey curve is the weighted B-spline. Blue curve is the prediction of the best-fit kinetics model described in **Section 5.3**. Starting concentrations of substrates are indicated above each plot.

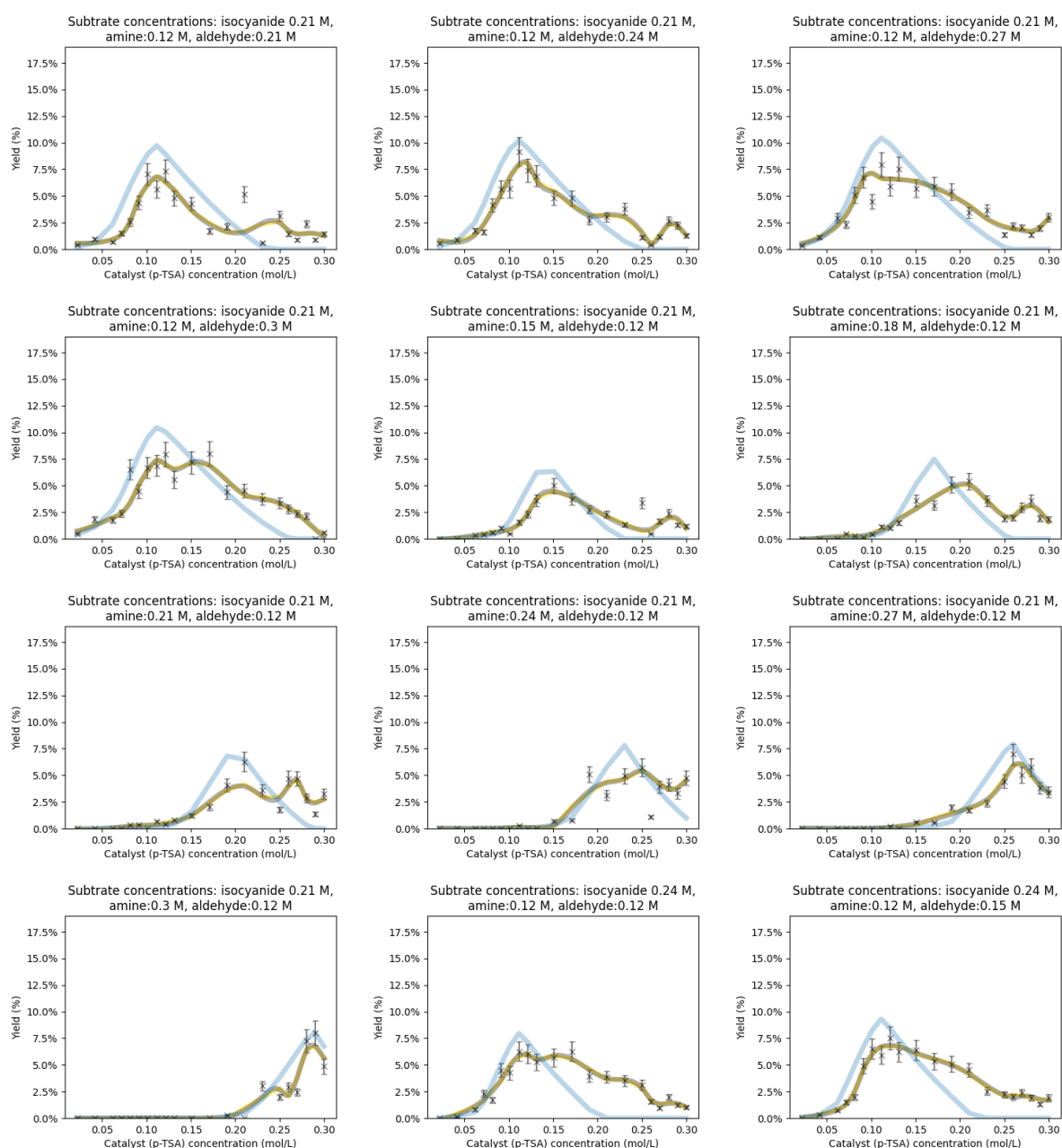


Figure S131. Yield of Ugi multicomponent reaction. Raw experimental data is shown by black “x” markers. Errorbars correspond to 14.5% relative error. Yellow curve is the result of weighted Savitsky-Golay filtering with degree 2 and window length equal to 5 points. Grey curve is the weighted B-spline. Blue curve is the prediction of the best-fit kinetics model described in **Section 5.3**. Starting concentrations of substrates are indicated above each plot.

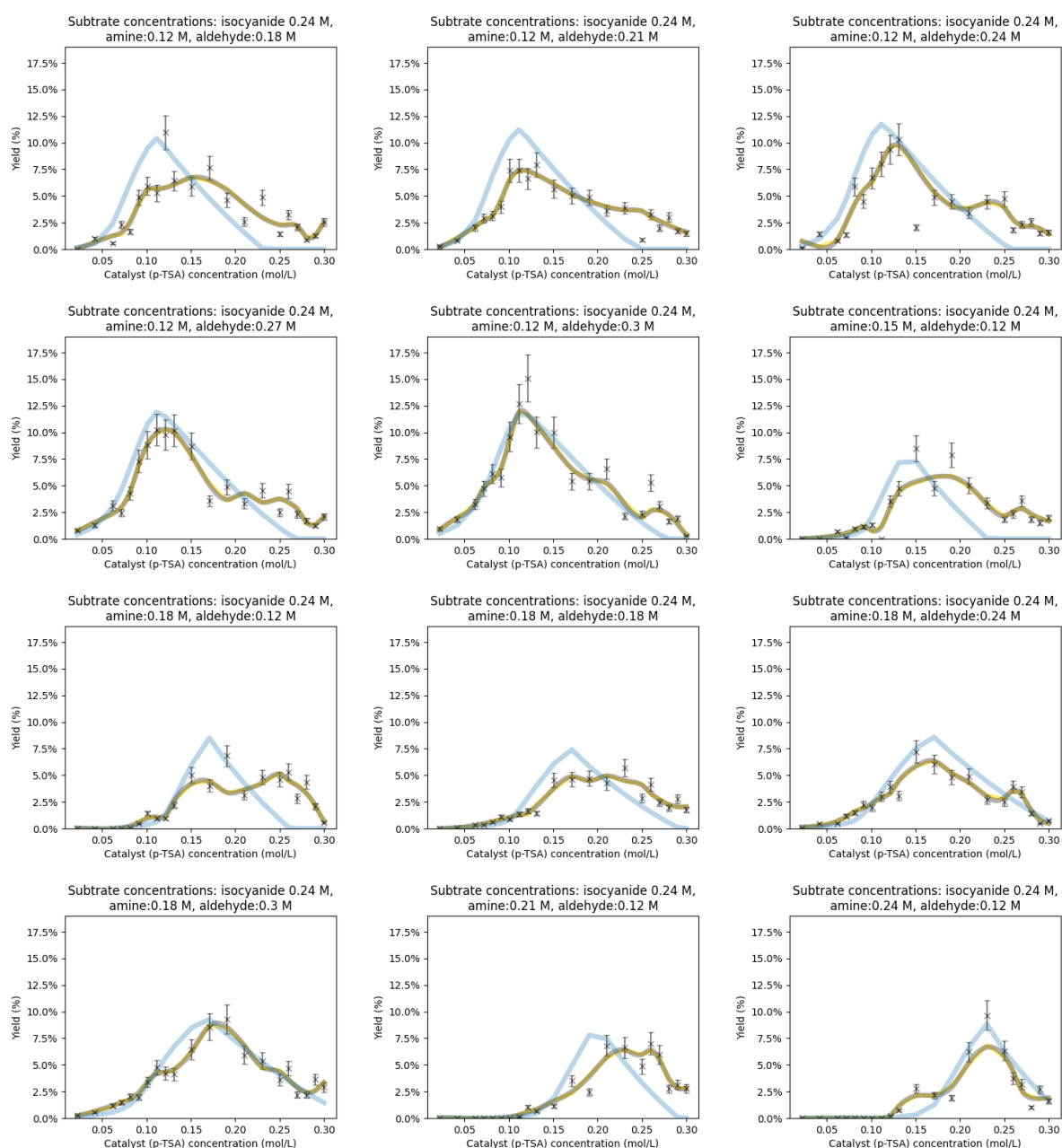


Figure S132. Yield of Ugi multicomponent reaction. Raw experimental data is shown by black “x” markers. Errorbars correspond to 14.5% relative error. Yellow curve is the result of weighted Savitsky-Golay filtering with degree 2 and window length equal to 5 points. Grey curve is the weighted B-spline. Blue curve is the prediction of the best-fit kinetics model described in **Section 5.3**. Starting concentrations of substrates are indicated above each plot.

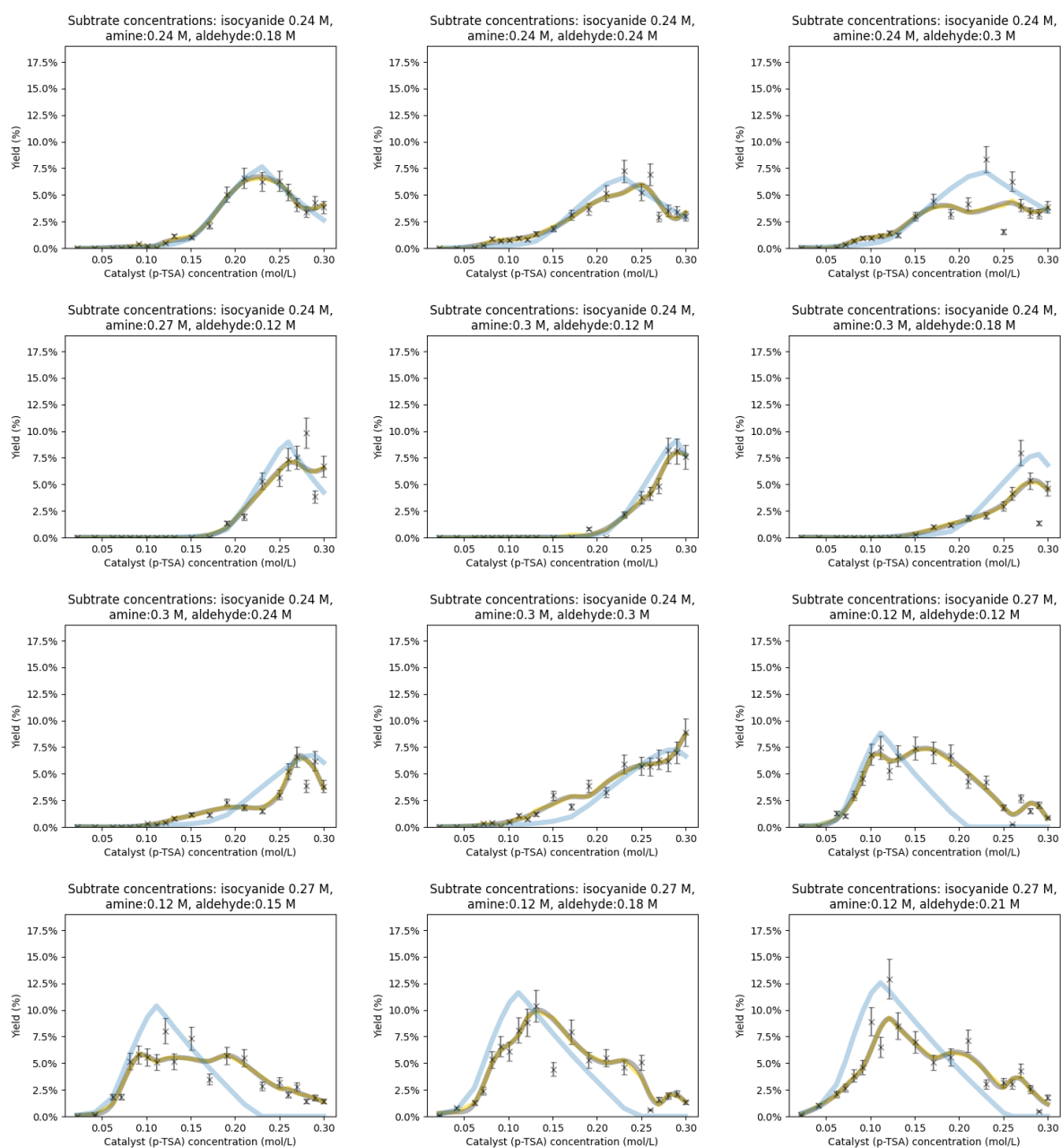


Figure S133. Yield of Ugi multicomponent reaction. Raw experimental data is shown by black “x” markers. Errorbars correspond to 14.5% relative error. Yellow curve is the result of weighted Savitsky-Golay filtering with degree 2 and window length equal to 5 points. Grey curve is the weighted B-spline. Blue curve is the prediction of the best-fit kinetics model described in **Section 5.3**. Starting concentrations of substrates are indicated above each plot.

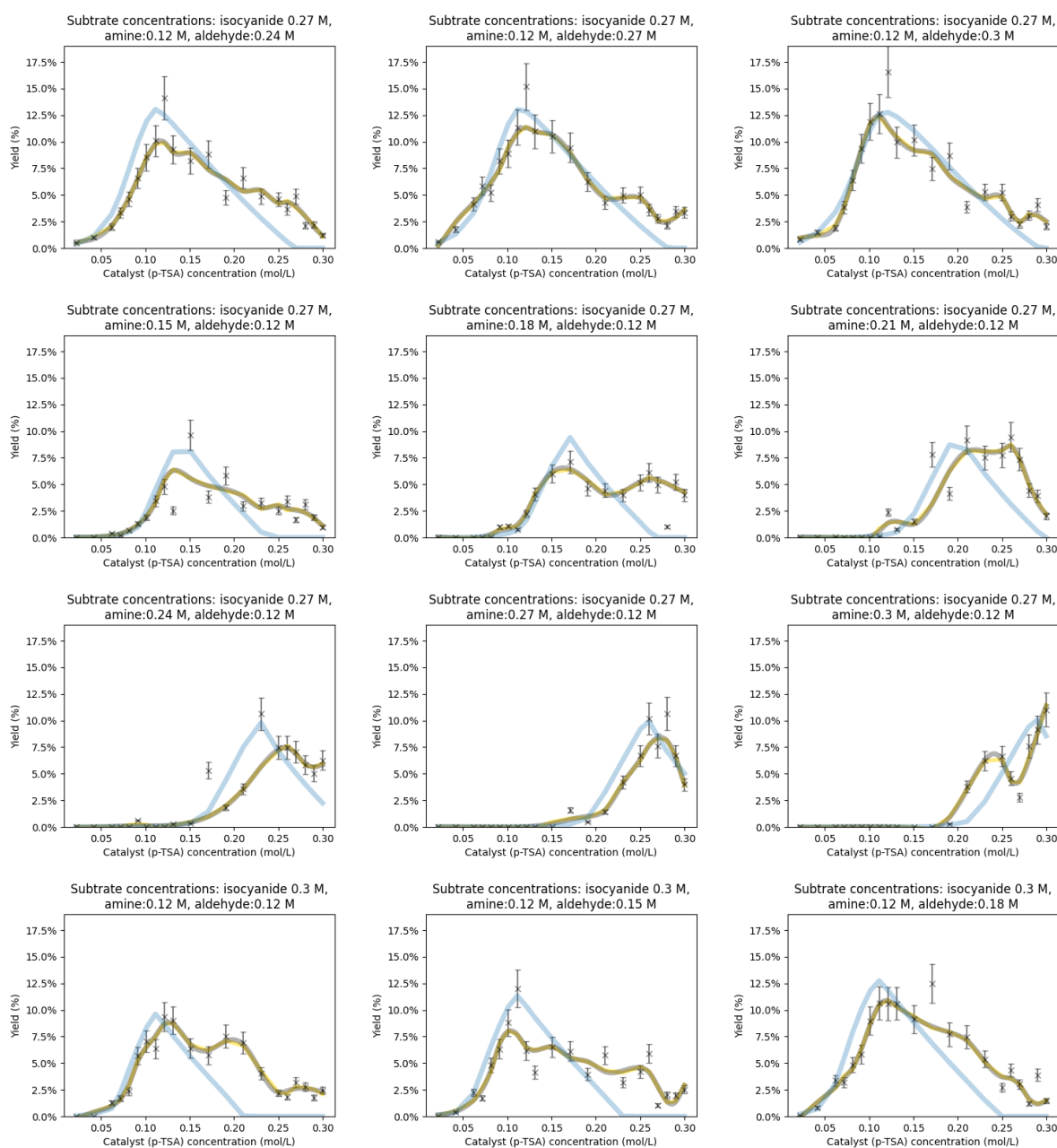


Figure S134. Yield of Ugi multicomponent reaction. Raw experimental data is shown by black “x” markers. Errorbars correspond to 14.5% relative error. Yellow curve is the result of weighted Savitsky-Golay filtering with degree 2 and window length equal to 5 points. Grey curve is the weighted B-spline. Blue curve is the prediction of the best-fit kinetics model described in **Section 5.3**. Starting concentrations of substrates are indicated above each plot.

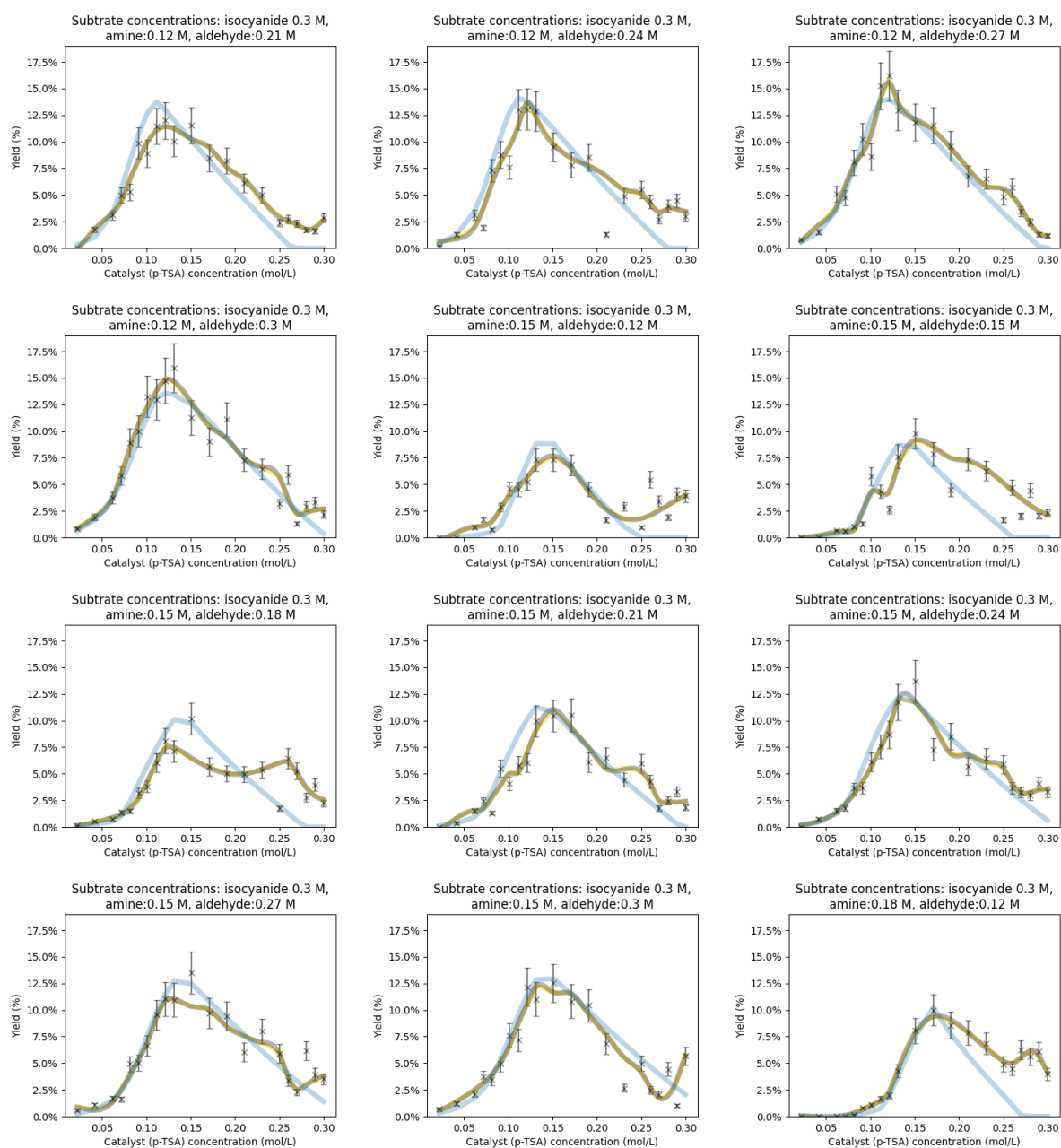


Figure S135. Yield of Ugi multicomponent reaction. Raw experimental data is shown by black “x” markers. Errorbars correspond to 14.5% relative error. Yellow curve is the result of weighted Savitsky-Golay filtering with degree 2 and window length equal to 5 points. Grey curve is the weighted B-spline. Blue curve is the prediction of the best-fit kinetics model described in **Section 5.3**. Starting concentrations of substrates are indicated above each plot.

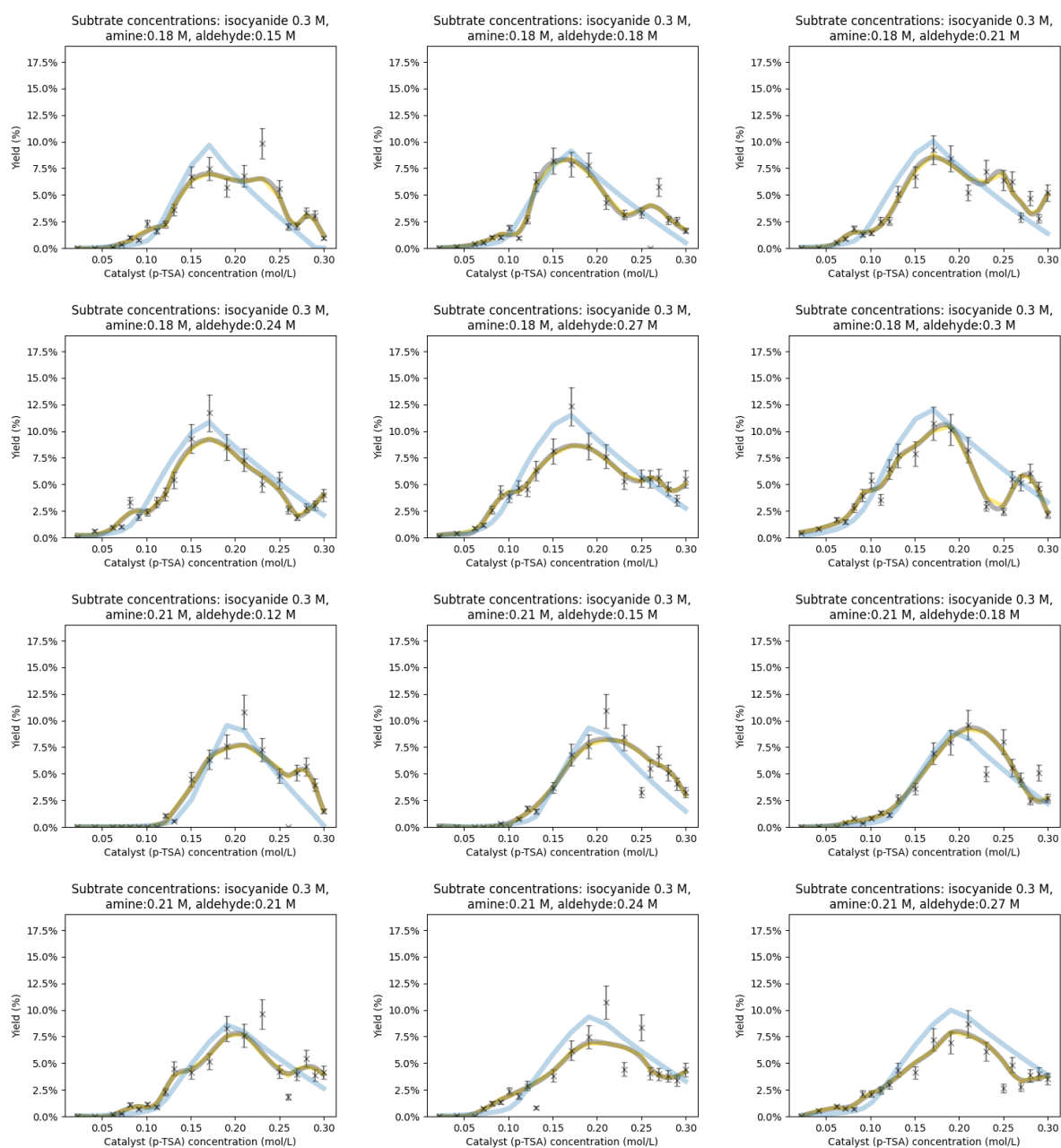


Figure S136. Yield of Ugi multicomponent reaction. Raw experimental data is shown by black “x” markers. Errorbars correspond to 14.5% relative error. Yellow curve is the result of weighted Savitsky-Golay filtering with degree 2 and window length equal to 5 points. Grey curve is the weighted B-spline. Blue curve is the prediction of the best-fit kinetics model described in **Section 5.3**. Starting concentrations of substrates are indicated above each plot.

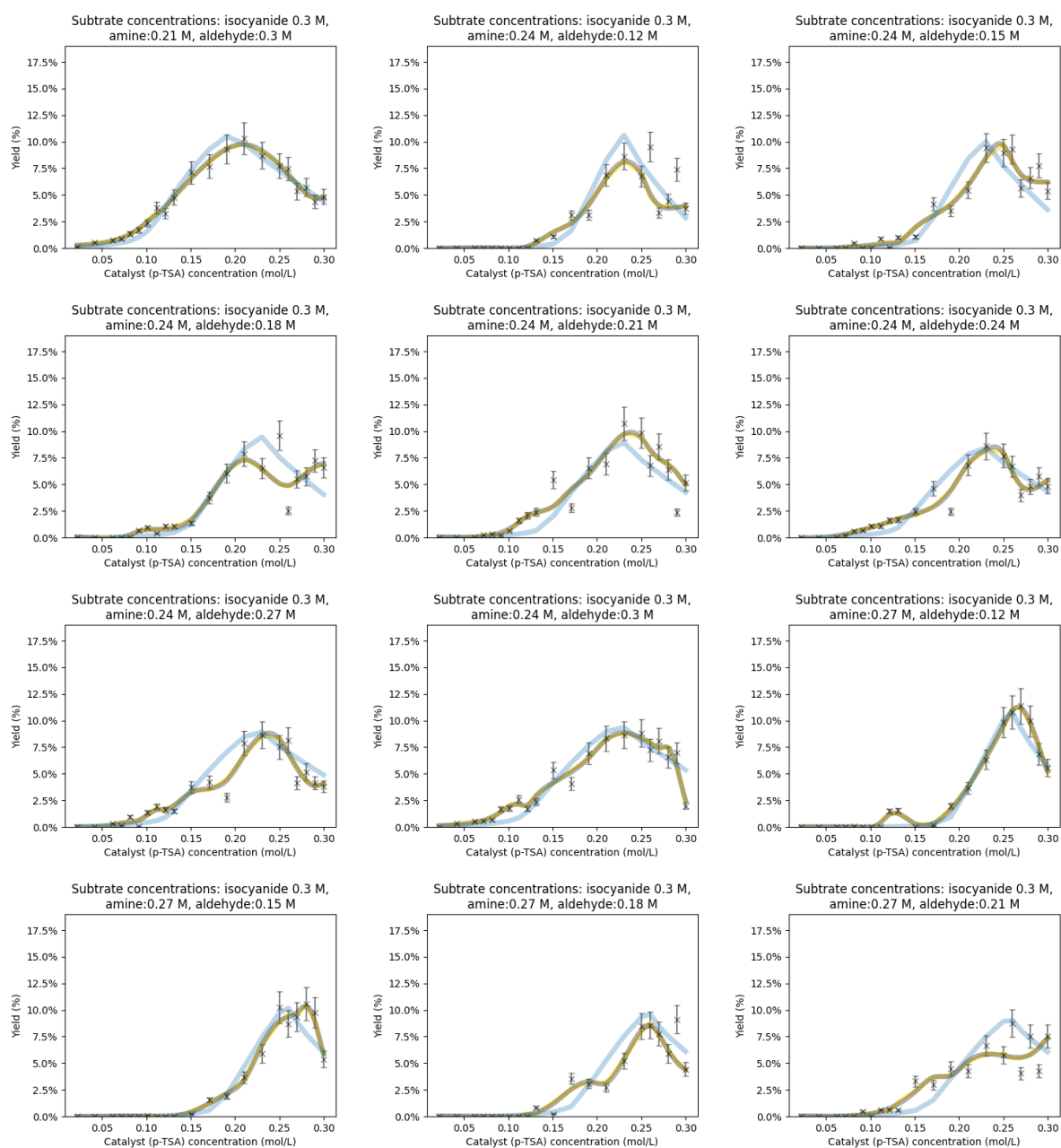


Figure S137. Yield of Ugi multicomponent reaction. Raw experimental data is shown by black “x” markers. Errorbars correspond to 14.5% relative error. Yellow curve is the result of weighted Savitsky-Golay filtering with degree 2 and window length equal to 5 points. Grey curve is the weighted B-spline. Blue curve is the prediction of the best-fit kinetics model described in **Section 5.3**. Starting concentrations of substrates are indicated above each plot.

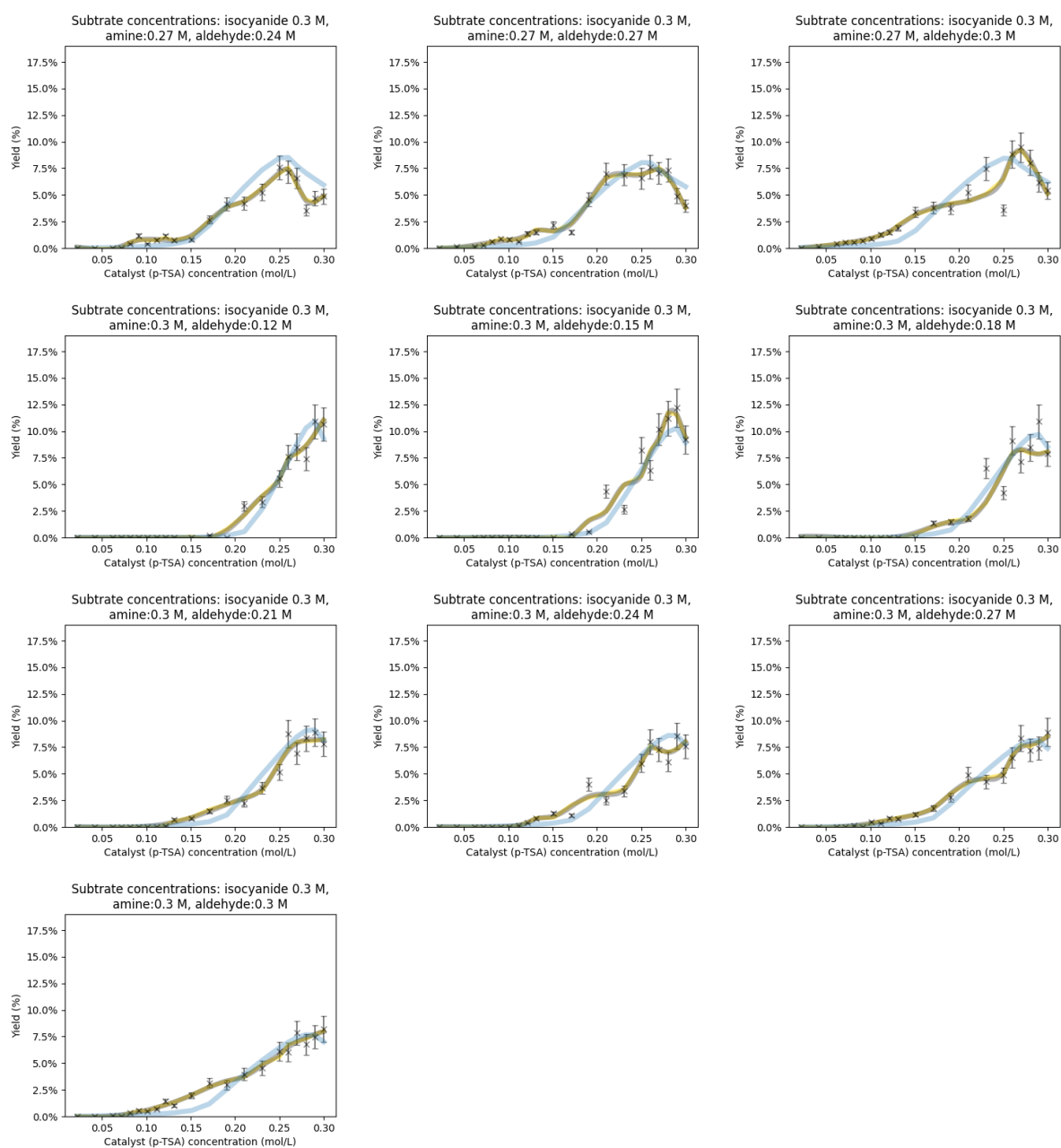


Figure S138. Yield of Ugi multicomponent reaction. Raw experimental data is shown by black "x" markers. Errorbars correspond to 14.5% relative error. Yellow curve is the result of weighted Savitsky-Golay filtering with degree 2 and window length equal to 5 points. Grey curve is the weighted B-spline. Blue curve is the prediction of the best-fit kinetics model described in Section 5.3. Starting concentrations of substrates are indicated above each plot.

4.15.5. Hantzsch reaction

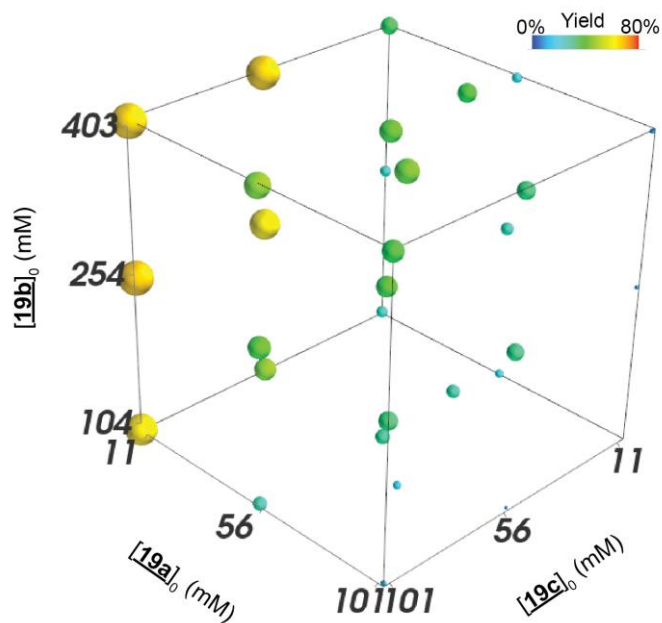


Figure S139. The yield distributions of product **19e**. The full, cube of 31 different reaction conditions at 80 °C, average from 71 experimental HPLC points.

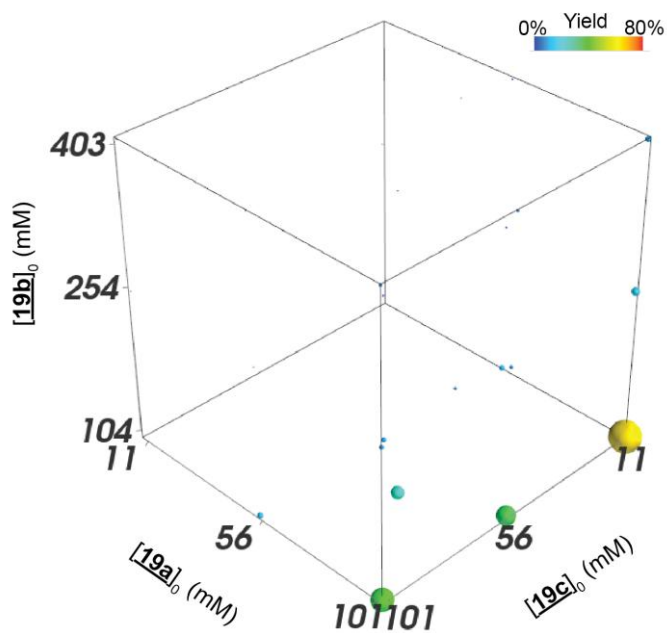


Figure S140. The yield distributions of product **19d**. The full, cube of 31 different reaction conditions at 80 °C, average from 71 experimental HPLC points.

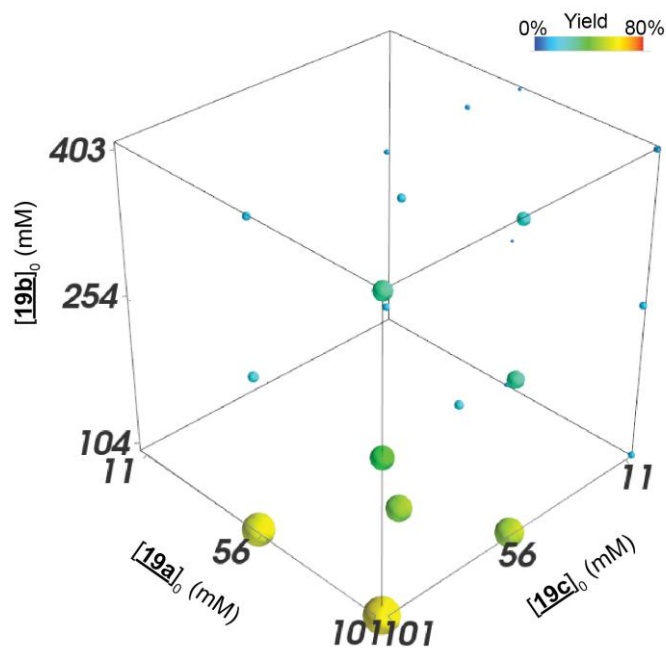


Figure S141. The yield distributions of product **19k**. The full, cube of 31 different reaction conditions at 80 °C, average from 71 experimental HPLC points.

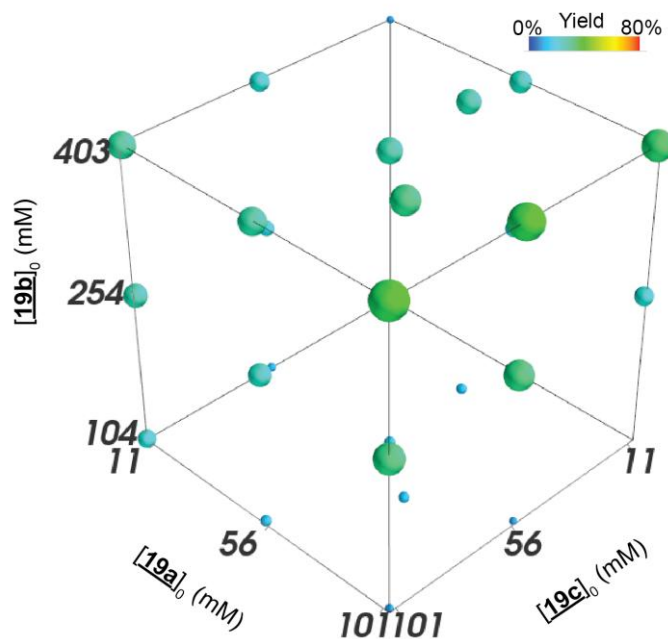


Figure S142. The yield distributions of product **19e**. The full, cube of 31 different reaction conditions at 26 °C, average from 71 experimental HPLC points.

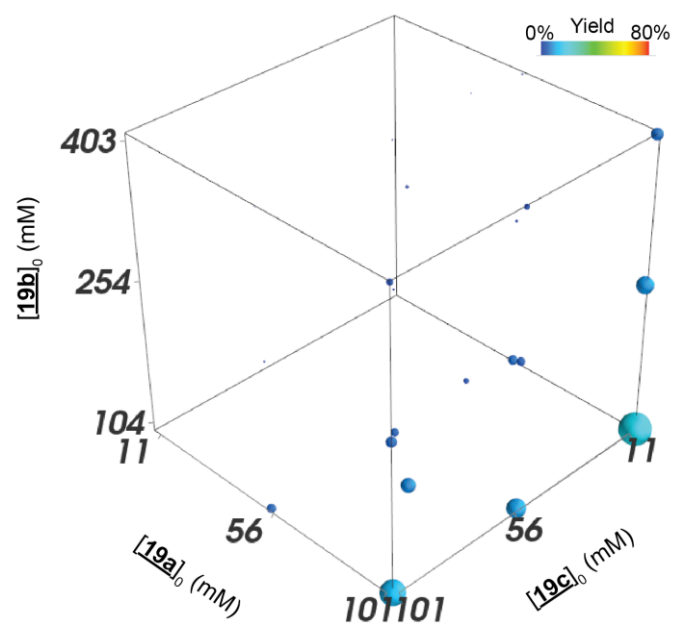
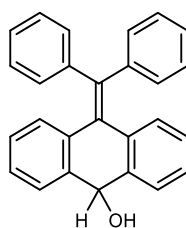


Figure S143. The yield distributions of product **19d**. The full, cube of 31 different reaction conditions at 26 °C, average from 71 experimental HPLC points.

4.16. Crystallographic data

To characterize compounds **15b** from main text Figure 3, Suitable single crystals were selected, and the measurements were performed on a Bruker APEX-II CCD diffractometer. The crystal was kept at 223.15 K during data collection. Using Olex2, the structure was solved with the olex2.solve structure solution program using Charge Flipping and refined with the olex2.refine refinement package using Levenberg–Marquardt minimization. CCDC 2418839 (**15b**), 2418838 (**15b**·C₃H₇NO) contain the supplementary crystallographic data for this paper. These data are provided free of charge by The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif.



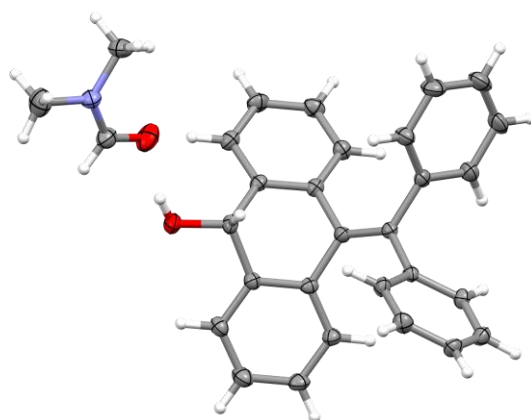
15b

Table S14. Crystal data and structure refinement for compound **15b**·C₃H₇NO.

Identification code	SN1Br01DMF
Empirical formula	C ₃₀ H ₂₇ NO ₂
Formula weight	433.554
Temperature/K	223.15
Crystal system	monoclinic
Space group	Cc
a/Å	9.139(9)
b/Å	24.15(3)
c/Å	10.564(12)
α /°	90
β /°	98.25(4)
γ /°	90
Volume/Å ³	2307(4)
Z	4
ρ_{calc} /g/cm ³	1.248
μ /mm ⁻¹	0.077
F(000)	920.5
Crystal size/mm ³	0.156 × 0.126 × 0.047
Radiation	Mo K α (λ = 0.71073)
2 Θ range for data collection/°	4.8 to 56.72
Index ranges	−12 ≤ h ≤ 10, −27 ≤ k ≤ 32, −14 ≤ l ≤ 14
Reflections collected	11829
Independent reflections	4798 [R_{int} = 0.0701, R_{sigma} = 0.0965]
Data/restraints/parameters	4798/2/301
Goodness-of-fit on F ²	0.957
Final R indexes [$I \geq 2\sigma(I)$]	R_1 = 0.0602, wR_2 = 0.1450
Final R indexes [all data]	R_1 = 0.1194, wR_2 = 0.1927
Largest diff. peak/hole / e Å ⁻³	0.29/−0.31
Flack parameter	−0.3(17)

Table S15. Crystal data and structure refinement for compound **15b**.

Identification code	SN1Br01AW
Empirical formula	C ₂₇ H ₂₀ O
Formula weight	360.459
Temperature/K	223.15
Crystal system	triclinic
Space group	P-1
a/Å	9.758(2)
b/Å	13.641(3)
c/Å	15.596(3)
α /°	72.323(7)
β /°	89.941(7)
γ /°	78.283(7)
Volume/Å ³	1932.7(7)
Z	4
$\rho_{\text{calc}}/\text{cm}^3$	1.239
μ/mm^{-1}	0.074
F(000)	760.4
Crystal size/mm ³	0.125 × 0.114 × 0.04
Radiation	Mo K α (λ = 0.71073)
2 Θ range for data collection/°	3.52 to 56.8
Index ranges	−13 ≤ h ≤ 13, −18 ≤ k ≤ 18, −20 ≤ l ≤ 20
Reflections collected	73102
Independent reflections	9634 [R_{int} = 0.1330, R_{sigma} = 0.1018]
Data/restraints/parameters	9634/0/507
Goodness-of-fit on F ²	1.026
Final R indexes [$I > 2\sigma(I)$]	R_1 = 0.0592, wR_2 = 0.1471
Final R indexes [all data]	R_1 = 0.1535, wR_2 = 0.2119
Largest diff. peak/hole / e Å ^{−3}	0.47/−0.50

**Figure S144.** Crystal structure of **15b**·C₃H₇NO (asymmetric unit, thermal ellipsoids are shown at 30% probability level).

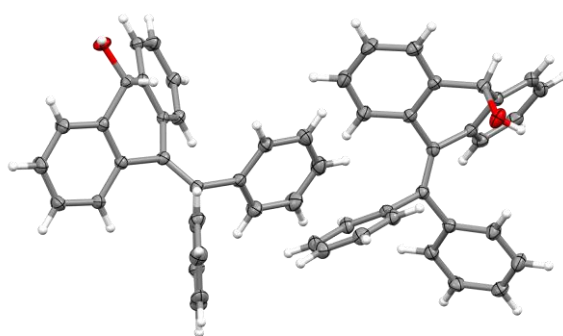


Figure S145. Crystal structure of **15b** (asymmetric unit, thermal ellipsoids are shown at 30% probability level).

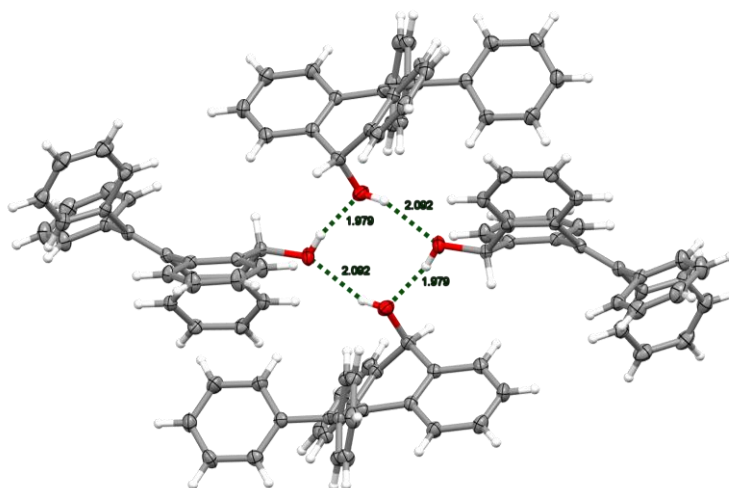


Figure S146. Supramolecular hydrogen-bonded tetramer of **15b** in the solid state (thermal ellipsoids are shown at the 30% probability level, O $\cdots$ H distances are given in Å).

4.17. Reaction mass efficiency for the Hantzsch reaction

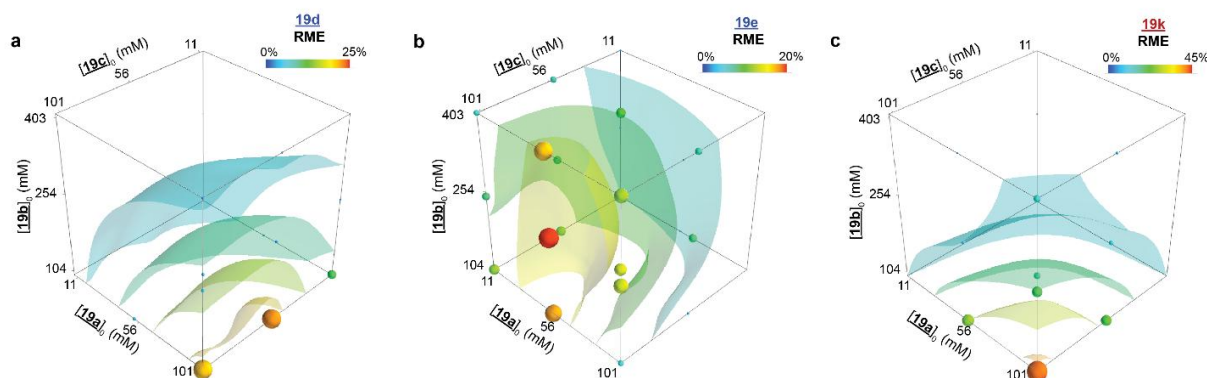


Figure S147 Space of reaction mass efficiency (RME) for the Hantzsch reaction. a-c, For a particular product, its RME is defined as the mass of product divided by the sum of masses of substrates multiplied by 100%¹⁰⁵. The maps shown here are for the Hantzsch reaction at 80 °C: **a**, Hantzsch ester product **19d** (maximum of RME ~ 23%); **b**, product **19e** (maximum of ~ 20%) and **c**, product **19k** (maximum of ~ 42%). Such maps help visualize regions that are more “practical” in the sense that a product of interest forms in appreciable yield but without impractical excess of the reagents. For example, whereas Petrenko-Kritschenko product **19e** can form in up to 67% yield, the corresponding conditions require very high excess of substrate **19b** (11 mM, 256 mM, and 100 mM; ratio ~ 1:23.2:9) corresponding to the RME of only 8.2%. A still decent yield of 42.1% can be obtained for more reasonable excess of **19b** (56 mM, 254 mM, and 100 mM, ratio ~1:4.5:2) for which the RME is 19.8%.

5. Kinetic models and fitting them to the data

5.1. Case of S_N1 reaction (main text Figure 2b)

In our experiments above 26 °C, reaction has been run to completion: the thermodynamic equilibrium between the substrates (HBr and alcohol **14a**) and products (water and product **14b**) was reached. Therefore, the final concentrations of the species obey the equilibrium relation

$$\frac{[\mathbf{14b}]_t[\text{H}_2\text{O}]_t}{[\text{HBr}]_t[\mathbf{14a}]_t} \cdot \frac{\gamma_{\mathbf{14b}}\gamma_{\text{H}_2\text{O}}}{\gamma_{\text{HBr}}\gamma_{\mathbf{14a}}} = K(T) \quad (7)$$

Where square brackets $[\cdot]_t$ denote the concentrations of species at reaction time t , the activity coefficients in dioxane are denoted by γ with a subscript indicating the species, and $K(T)$ is the equilibrium constant, which depends on the absolute temperature T of the experiment.

The activity coefficients $\gamma_{\mathbf{14a}}$ and $\gamma_{\mathbf{14b}}$ in dioxane are taken to be equal to unity, in agreement with values computed for **14a** and **14b** with a computational chemistry method based on machine learning and Dortmund Data Bank of activity coefficients (Hard-constraint Neural Network for Activity coefficient prediction, HANNA¹⁰⁶). We also assume that $\gamma_{\text{HBr}} = 1$ in dioxane. We employ published measurements of the dependence of $\gamma_{\text{H}_2\text{O}}$ on the molar fraction of water, $[\text{H}_2\text{O}]$ ^{107,108}. To be able to evaluate $\gamma_{\text{H}_2\text{O}}$ for any given concentration $[\text{H}_2\text{O}]$, we interpolate this published dependence using a sum of a second-degree polynomial and a decaying exponential function. Applying the mass conservation law to equation (7), we arrive at the equation

$$\frac{[\mathbf{14b}]_t([\text{H}_2\text{O}]_0 + [\mathbf{14b}]_t)}{([\text{HBr}]_0 - [\mathbf{14b}]_t)([\mathbf{14a}]_0 - [\mathbf{14b}]_t)} \cdot \gamma_{\text{H}_2\text{O}} = K(T) \quad (8)$$

which can be solved with respect to unknown final concentration $[\mathbf{14b}]_t$ of the product, given the known initial concentrations of all species and the equilibrium constant $K(T)$. Note that the starting concentration of water is due to water being present in the stock solution of HBr, and the specific relationship is $[\text{H}_2\text{O}]_0 = 4.864[\text{HBr}]_0$. Equation (8) is the model of product concentration $[\mathbf{14b}]_t$, where $K(T)$ is treated as the a parameter at a given temperature. We fit this model to all the experimental data collected at a given temperature. Fitting was performed using Trust Region Reflective algorithm⁷⁵ with 2-point method of computing the Jacobian. For the optimization to be robust against outliers, it was performed using a “soft L1”¹⁰⁹ loss function with a “scale” parameter equal to 0.05.

Comparison of best-fit model at each temperature is shown in **Figure S148**. This gives us the temperature dependence $K(T)$ of best-fitted equilibrium constants. The van’t Hoff plot of $\ln K(T)$ vs. $1/T$ is then used to obtain the ΔS and ΔH of the reaction, as shown in **Figure S148**.

The point at 56 °C is an outlier – perhaps because of the inapplicability of room-temperature data on activity coefficients to temperatures 56 °C and above, or due to temperature dependence of ΔS and ΔH which we presently ignore. Note also a larger discrepancy between model and data for the lowest temperature and lowest starting concentration of alcohol (**Figure S148a**, blue and orange curves) – given that these are conditions with slowest expected reaction rates, the equilibrium may not have been fully reached, hence a mismatch with our equilibrium-based model. As described in the next Section 5.2 using E1 reaction as an example, moving away from the equilibrium model to explicit modelling of the kinetics substantially increases the complexity of the analysis.

Considering the 56 °C point an outlier, the best-fit line gives $\Delta S = -(130.8 \pm 4.5) \text{ J/(mol K)}$, $\Delta H = -(30.70 \pm 1.37) \text{ kJ/mol}$. Standard errors are based on propagating 4% absolute standard error of the **14b** yield with respect to **14a** by means of the standard linear approximation method for confidence intervals of model parameters^{77–79}. To compare these thermodynamic parameters to those of other S_N1 reactions, we computed ΔH of S_N1 reaction between *tert*-butyl alcohol and HBr from enthalpies of formation compiled by NIST (*tert*-butyl alcohol¹¹⁰, HBr¹¹¹, *tert*-butyl bromide¹¹², water¹¹¹) and obtained $\Delta H = -23.49 \text{ kJ/mol}$, which is not that far from our value $-(30.70 \pm 1.37) \text{ kJ/mol}$. We were unable to do a similar comparison for ΔS because we did not find reported values of entropy for *tert*-butyl bromide. Quantum thermochemistry calculation for our specific S_N1 reaction yields $\Delta H = -24.31 \text{ kJ/mol}$, as discussed in Section 6.1. Note also that, in contrast to ΔH , measuring ΔS by fitting our model is inherently unreliable due to the absence of published data of activity coefficients for HBr in dioxane-water mixtures and their temperature dependence, as well as of the temperature dependence of the water activity in dioxane.

To reproduce the fitting of kinetic model (plots **a-e** in **Figure S148**), run the Python script “*robowski/kinetics_models/simpleSN1_kinetics/fit_kinetics_for_simpleSN1_reaction.py*” in the of the code repository associated with this article. After the script completes calculations, you can reproduce the Van’t Hoff in **Figure S148f** by running the Python script “*robowski/kinetics_models/simpleSN1_kinetics/plot_kinetics_of_simpleSN1_2023-11_2023-12.py*”. If you wish to recompute the spectral unmixing for this reaction from raw data, run the Python script “*robowski/uv_vis_absorption_spectroscopy/examples/simple_SN1/process_simple_SN1_2023-12-11-run01.py*” prior to fitting the kinetics models.

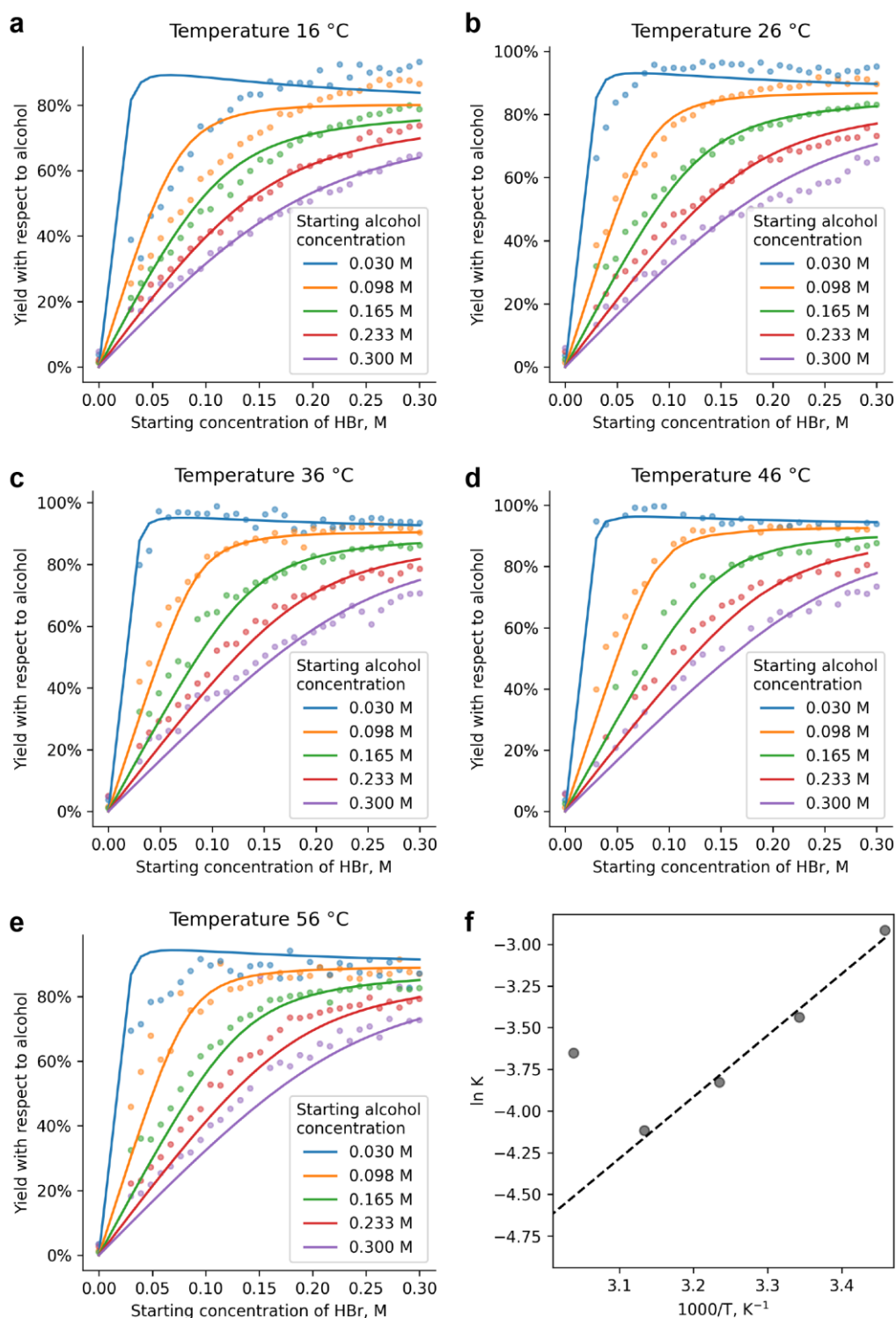


Figure S148. a-e, Experimental data (points) showing yield of **14b** with respect to alcohol **14a**, in the S_N1 reaction (scheme in main text **Figure 2b**) in comparison to the kinetic model (curves) fitted to the data. Temperatures are indicated above the plots. f, Van't Hoff plot of logarithm of the best-fit values of equilibrium constant (circles) against the inverse temperature ($1000/T$). Dashed line is fitted to all but the leftmost point (56 °C) in order to evaluate the entropy and enthalpy of the reaction.

5.2. Case of E1 reaction (main text Figure 2a)

5.2.1. Kinetic model, fitting procedure and the bootstrapping analysis

Overall equilibrium constant of the reaction is

$$K_{\text{net}} = \frac{[\mathbf{13b}][\text{H}_2\text{O}]}{[\mathbf{13a}]} \cdot \frac{\gamma_{\mathbf{13b}}\gamma_{\text{H}_2\text{O}}}{\gamma_{\mathbf{13a}}} = e^{-\frac{\Delta H_{\text{net}} - T\Delta S_{\text{net}}}{RT}} \quad (9)$$

where ΔH_{net} and ΔS_{net} are the net enthalpy and entropy of the E1 reaction. The activity coefficients $\gamma_{\mathbf{13a}}$, $\gamma_{\mathbf{13b}}$, and $\gamma_{\text{H}_2\text{O}}$ of substrate, product, and water in acetonitrile depend on the temperature, and for each of the five temperatures used in our experiments we have calculated these activity coefficients using a computation chemistry method based on machine learning and Dortmund Data Bank of activity coefficients (Hard-constraint Neural Network for Activity coefficient prediction, HANNA¹⁰⁶). The calculation results are listed in **Table S16**.

Table S16. Activity coefficients (in acetonitrile) of water, substrate **13a**, and product **13b** of E1 reaction. These coefficients were calculated using Hard-constraint Neural Network for Activity coefficient prediction (HANNA, Ref.¹⁰⁶).

Temperature, °C	$\gamma_{\mathbf{13a}}$	$\gamma_{\mathbf{13b}}$	$\gamma_{\text{H}_2\text{O}}$
16	52.9169	6.9626	10.8702
21	46.4955	6.7242	10.3507
26	40.6853	6.4882	9.8894
31	35.4860	6.2585	9.4751
36	30.8798	6.0386	9.0979

The “effective” equilibrium constant $\widetilde{K}_{\text{net}}$ is

$$\widetilde{K}_{\text{net}} = \frac{[\mathbf{13b}][\text{H}_2\text{O}]}{[\mathbf{13a}]} = \frac{\gamma_{\mathbf{13a}}}{\gamma_{\mathbf{13b}}\gamma_{\text{H}_2\text{O}}} e^{-\frac{\Delta H_{\text{net}} - T\Delta S_{\text{net}}}{RT}} \quad (10)$$

On the other hand, it can be expressed via constants of reaction steps

$$\widetilde{K}_{\text{net}} = \frac{K a_{\mathbf{13bH}^+}}{K a_{\mathbf{13aH}^+}} \cdot \frac{k_{E1}}{k_{-E1}} \quad (11)$$

where Ka_{13aH^+} and Ka_{13bH^+} are acid dissociation constants of protonated substrate and protonated product, k_{E1} and k_{-E1} are the reaction rate constants for forward reaction (dehydration of protonated substrate) and reverse reaction (hydration of protonated product). We can express Ka_{13aH^+} through respective pKa, which is assumed to have a standard form of temperature dependence:

$$pKa_{13aH^+} = \frac{\widetilde{\Delta S}_{pKa}}{R} - \frac{1}{T} \cdot \frac{\widetilde{\Delta H}_{pKa}}{R} \quad (12)$$

where $\widetilde{\Delta S}_{pKa}/R$ and $\widetilde{\Delta H}_{pKa}/R$ are entropic and enthalpic coefficients of pKa temperature dependence.

Equilibrium constant k_{E1}/k_{-E1} of dehydration obeys the thermodynamic relations

$$\frac{k_{E1}}{k_{-E1}} = e^{-\frac{\widetilde{\Delta H}_{dehyd} - T\widetilde{\Delta S}_{dehyd}}{RT}} \quad (13)$$

where k_{E1} and k_{-E1} are the reaction rate constants for forward reaction (dehydration) and reverse reaction (hydration), $\widetilde{\Delta H}_{dehyd}$ and $\widetilde{\Delta S}_{dehyd}$ are the effective enthalpy and entropy of dehydration.

The pKa of the protonated product can be obtained by combining the equations (11), (12), and (13):

$$pKa_{13bH^+} = pKa_{13aH^+} - \log_{10} \left(\widetilde{K}_{net} \frac{k_{-E1}}{k_{E1}} \right) \quad (14)$$

Second-order reverse reaction rate constant obeys the Eyring-Evans-Polanyi equation:

$$k_{-E1} = \frac{\kappa_- k_B T}{h} e^{-\frac{\Delta H^\ddagger - T\Delta S^\ddagger}{RT}} \quad (15)$$

where κ_- is the transmission coefficient, k_B is the Boltzmann constant, h is the Planck's constant, $\Delta S^\ddagger$ is the entropy of activation, $\Delta H^\ddagger$ is the enthalpy of activation. The first-order reaction rate constant k_{E1} of forward reaction can be obtained by combining equations (13) and (15):

$$k_{E1} = \frac{\kappa_- k_B T}{h} e^{-\frac{\Delta H^\ddagger - T\Delta S^\ddagger}{RT}} e^{-\frac{\widetilde{\Delta H}_{dehyd} - T\widetilde{\Delta S}_{dehyd}}{RT}} \quad (16)$$

The net reaction rate is the sum of the forward reaction rate and backward reaction rate:

$$\frac{d[\underline{\mathbf{13b}}]}{dt} = -\frac{d[\underline{\mathbf{13a}}]}{dt} = k_{E1}[\underline{\mathbf{13aH}}^+] - k_{-E1}[\underline{\mathbf{13bH}}^+][\text{H}_2\text{O}] \quad (17)$$

where the instantaneous concentrations of protonated substrate $[\underline{\mathbf{13aH}}^+]$ and protonated product $[\underline{\mathbf{13bH}}^+]$ are calculated by solving the proton exchange equilibrium as described in Section 5.3.3 at every evaluation of the right-hand side of the differential equation (17) (i.e., at every time step of the ODE integration algorithm). The $\text{p}K_{a_{\underline{\mathbf{13aH}}^+}}$ and $\text{p}K_{a_{\underline{\mathbf{13bH}}^+}}$ values required for this calculation are given by formulas (12) and (14). Autoprotolysis constant $\text{p}K_{ap}$ for water-acetonitrile mixture was taken to be equal to reported value for pure acetonitrile ¹¹³: $\text{p}K_{ap} = 33.58$. See also Table 6 in Ref. ¹¹⁴ for dependence of $\text{p}K_{ap}$ on the water content in binary water-acetonitrile mixtures. The $\text{p}K_a$ of HBr in water-acetonitrile mixture was modelled by a standard form of $\text{p}K_a$ temperature dependence:

$$\text{p}K_{a_{\text{HBr}}} = 6.6 - \frac{B_{\text{HBr}}}{298.15 \text{ }^\circ\text{K}} + \frac{B_{\text{HBr}}}{T} \quad (18)$$

where B_{HBr} is a free parameter (enthalpic coefficient) and we take advantage of the literature value (Ref. ¹¹⁵) of $\text{p}K_a$ of HBr in pure acetonitrile reported to be 6.6 at 25 °C (298.15 °K).

Differential equation (17) was integrated from zero time to the 4 hours mark (experimental reaction time) with explicit Runge-Kutta method of order 5(4) ¹¹⁶ having the first time step set to 36 microseconds and the maximum time step limited to 1 hour. The absolute tolerance was set to 10^{-8} , and the relative one to 10^{-7} .

Fitting the model to the data was performed with Trust Region Reflective algorithm ⁷⁵ with 2-point method of computing the Jacobian. In the 2-point method, the relative steps for differentiation were individually specified for each parameters as listed in **Table S17** and the problem was normalized to assist convergence (rescaling factors are also listed in **Table S17**). For optimization to be robust against outliers, it was performed using a “soft L1”¹⁰⁹ loss function with a “scale” parameter equal to 0.1.

Table S17. Jacobian steps and normalization factors for optimization of the kinetic model of E1 reaction

Kinetic model's parameter	Relative step for evaluating the Jacobian (diff_step in SciPy)	Rescaling factor (x_scale in SciPy)
$\widetilde{\Delta S}_{pKa}/R$	0.002	45
$\widetilde{\Delta H}_{pKa}/R$	0.003	10
ΔS_{net}	0.003	250
ΔH_{net}	0.003	80
$\widetilde{\Delta S}_{\text{dehyd}}/R$	0.001	20
$\widetilde{\Delta S}_{\text{dehyd}}/R$	0.001	5
$\kappa_{-}k_B/h$	0.001	10^{-6}
$\Delta S^{\ddagger}/R$	0.02	5
$\Delta H^{\ddagger}/R$	0.02	3
$B_{\text{HBr}}/1000$	0.01	180

If we exclude the reverse reaction from the model of kinetics, the model fails to match the data. For illustration, the best-fit irreversible-reaction model for 16 °C is shown against the data in **Figure S149**.

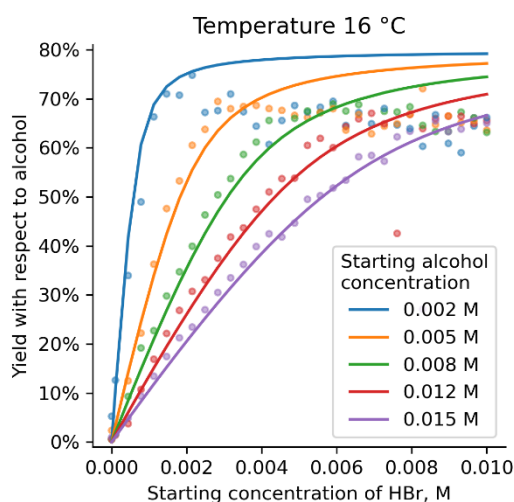


Figure S149. Yield values predicted by the best-fit irreversible-reaction model (curves) for 16 °C and various conditions (starting concentrations of HBr, horizontal axis, and of alcohol, shown by color) are compared to the experimental data (points). Note the mismatch between curves and the points at the upper part of the plot.

Uncertainties of the fitted parameters can be estimated by applying the “bootstrap” method¹¹⁷. This can be done either by resampling the experimental data points themselves (this is known as “resampling the vectors”), or by resampling the residuals. The former version is more appropriate for our case: as opposed to resampling the residuals, resampling the vectors may cause overestimation of the uncertainties, but, crucially, does not require us to make assumptions about homo- vs. heteroscedasticity and about our model being the perfect (up to experimental noise) description of the data^{117,118}.

When we fit the model to data at all temperatures, the model follows this data closely at intermediate temperatures, but systematic mismatch (model underestimates the yields) becomes higher towards the edges of the temperature range. This can be seen in **Figure S150**, which shows the best-fit model (curves) compared to the data (points): the curves are following the points much closer in **Figure S150b,c,d** than in **Figure S150a** or **Figure S150e**. This indicates that the model does not describe the temperature dependence accurately in the full temperature range, likely because the true dependences of kinetic parameters on temperature do not precisely follow the standard temperature dependencies (equations (50) and (13)) and Eyring equation (15). In other words, the model is underfitting the data. It seems to us that, at present, there is not enough data to confidently distinguish between some more complex models of temperature dependences for some or all of the kinetic parameters, so we do not attempt fitting any more complex models.

Nevertheless, perhaps at least some of the parameters do follow standard temperature dependencies and Eyring equation in a narrower range of temperatures. Thus, we compared the model that was best-fit to data at all temperatures versus the model fitted only to data below 31 °C. As shown in **Table S18**, reducing the temperature range of data did not significantly affect the best-fit values of the overall entropy and enthalpy of the reaction, whereas the best-fit values of some other parameters have shifted appreciably: we conclude that, for instance, the values of κ_- and B_{HBr} are not constrained by the current experimental data and therefore their best-fit values cannot be viewed as correct.

Fundamentally, the kinetic rate constants can only be evaluated if experimental yields do not always reach equilibrium. To illustrate how far the yield is from its equilibrium value, we plotted the fraction of the two at every tested condition (**Figure S154**).

Table S18. Parameters of the best-fit kinetic model of E1 reaction

Kinetic model's parameter	Values best-fit to all data	Values best-fit to data collected below 31 °C	Values calculated <i>ab initio</i> with quantum thermochemistry	Units
$\widetilde{\Delta S}_{\text{pKa}}/R$	$33.4^{+3.8}_{-2.0}$	$40.37^{+1.25}_{-1.63}$		1
$\widetilde{\Delta H}_{\text{pKa}}/R$	$5.36^{+1.06}_{-1.27}$	$3.92^{+0.38}_{-0.48}$		K
ΔS_{net}	$273.8^{+10.8}_{-4.9}$	$264.2^{+16.5}_{-20}$	167 ± 40	J mol ⁻¹ K ⁻¹
ΔH_{net}	$81.98^{+3.52}_{-1.64}$	$79.6^{+4.7}_{-6.2}$	37.8 ± 12.5	J mol ⁻¹
$\widetilde{\Delta S}_{\text{dehyd}}/R$	$24.64^{+0.26}_{-0.35}$	$19.15^{+0.41}_{-0.30}$		1
$\widetilde{\Delta H}_{\text{dehyd}}/R$	$7.216^{+0.088}_{-0.107}$	$5.633^{+0.125}_{-0.090}$		K
$\kappa_- k_B/h$	$3.00^{+0.25}_{-0.19}$	$0.539^{+0.0154}_{-0.019} \cdot 10^{-5}$		L K mol ⁻¹ hour ⁻¹
$\Delta S^\ddagger/R$	$-4.659^{+0.114}_{-0.112}$	$-3.494^{+0.114}_{-0.209}$		1
$\Delta H^\ddagger/R$	$-2.54^{+0.055}_{-0.042}$	$-2.728^{+0.037}_{-0.063}$		K
B_{HBr}	$0.063080^{+0.0000065}_{-0.0000113}$	$162.7^{+8.1}_{-24.8}$		K

Table S19. E1 reaction of dehydration of alcohols. Entropy and enthalpy of the reaction reported in the literature in comparison to the values estimated in this work.

Alcohol	Notes	ΔH , kJ/mol	ΔS , J/(mol K)	Ref
13a (our experiment)	5 temperatures	$81.99^{+3.52}_{-1.64}$	$273.8^{+10.8}_{-4.9}$	this work
13a (our experiment)	3 temperatures	$79.6^{+4.6}_{-6.2}$	$264.2^{+16.52}_{-20.0}$	this work
13a (DFT calculation)	vacuum	37.8 ± 12.5	167 ± 40	this work
<i>tert</i> -butyl alcohol		26 ± 9	60 ± 30	119
<i>tert</i> -butyl alcohol		27.3 - 33.5	87	120
<i>tert</i> -butyl alcohol		26.5		121
<i>tert</i> -butyl alcohol		39 ± 1	97.4 ± 3.6	122
<i>tert</i> -butyl alcohol		44.3		123,124
1-propanol		37.4	72.15	125–129
Cyclohexanol		28.4 ± 8.2	80.68	128,130–133

Table S20. E1 reaction of dehydration of alcohols. Activation energy of the reaction reported in the literature in comparison to the values estimated in this work.

Alcohol	Notes	$\Delta H^\ddagger$, kJ/mol	Ref
13a (our experiment)	5 temperatures	$21.10^{+0.33}_{-0.49}$	this work
13a (our experiment)	3 temperatures	$22.68^{+0.33}_{-0.50}$	this work
<i>tert</i> -butyl alcohol	“model 1”	24 ± 7	119
<i>tert</i> -butyl alcohol	“model 2”	18 ± 6	119
<i>tert</i> -butyl alcohol	“model 3”	27 ± 6	119

O’Ferrall and coworkers determined experimentally the equilibrium constants for the dehydration ($K_{\text{de}} = [\text{alkene}]/[\text{alcohol}]$) of alcohols in water at 25 °C¹³⁴. For the 18 aliphatic, cyclic, benzylic and allylic alcohols tested, K_{de} ranges from 1.6×10^{-5} to 22 (**Table S21**). We calculated the corresponding values of free energy change of dehydration (ΔG) that turned out to range from 27 to -7.7 . As pointed out by the authors, the dehydration equilibrium depends subtly on the structure of the alcohol (open-chain vs. cyclic, and saturated vs. allylic vs. benzylic) and hence also on the structure of the alkene formed. Therefore, there exists no narrow, typical range of ΔG for the dehydration of alcohols. Probably, the same is also true for the ΔH and ΔS values. Consequently, our method is not immediately invalidated by the apparent discrepancies of ΔH and ΔS determined for our alcohol **13a** relative to the data for other alcohols available in the literature gathered in **Table S20**. As a crude comparison, the ΔG value obtained for **13a** is 0.357 kJ/mol (five temperatures, in DMF–water mixture), and is closest to that of isobutyl alcohol (0.46 kJ/mol in water) among the 18 alcohols investigated by O’Ferrall and coworkers.

Table S21. Equilibrium constants for the dehydration ($K_{\text{de}} = [\text{alkene}]/[\text{alcohol}]$) of alcohols in water at 25 °C extracted from ref¹³⁴ and the corresponding ΔG calculated by $\Delta G = -RT \ln(K_{\text{de}})$ for $T = 298.15$ K.

K_{de}	K_{de}	ΔG (kJ/mol)
ethanol	1.6×10^{-5}	27
1-butanol	1.6×10^{-5}	27
2-butanol	1.6×10^{-5}	27
1-propanol	5.9×10^{-5}	24
<i>tert</i> -butyl alcohol	1.1×10^{-4}	23
2-butanol	1×10^{-3}	17
3-pentanol	1.1×10^{-3}	17
cyclopentanol	6.6×10^{-3}	12
2-propanol	1.1×10^{-2}	11
1-phenylethan-1-ol	2.5×10^{-2}	9.1
cyclohexanol	3.2×10^{-2}	8.5
but-3-en-2-ol	7×10^{-2}	6.6
cyclohex-2-en-1-ol	0.1	5.7
isobutyl alcohol	0.83	0.46
1-phenylpropan-1-ol	2.3	-2.1
cyclopent-2-en-1-ol	3.5	-3.1
pent-1-en-3-ol	6.0	-4.4
1-indan-1-ol	10	-5.7
1-tetralol	22	-7.7

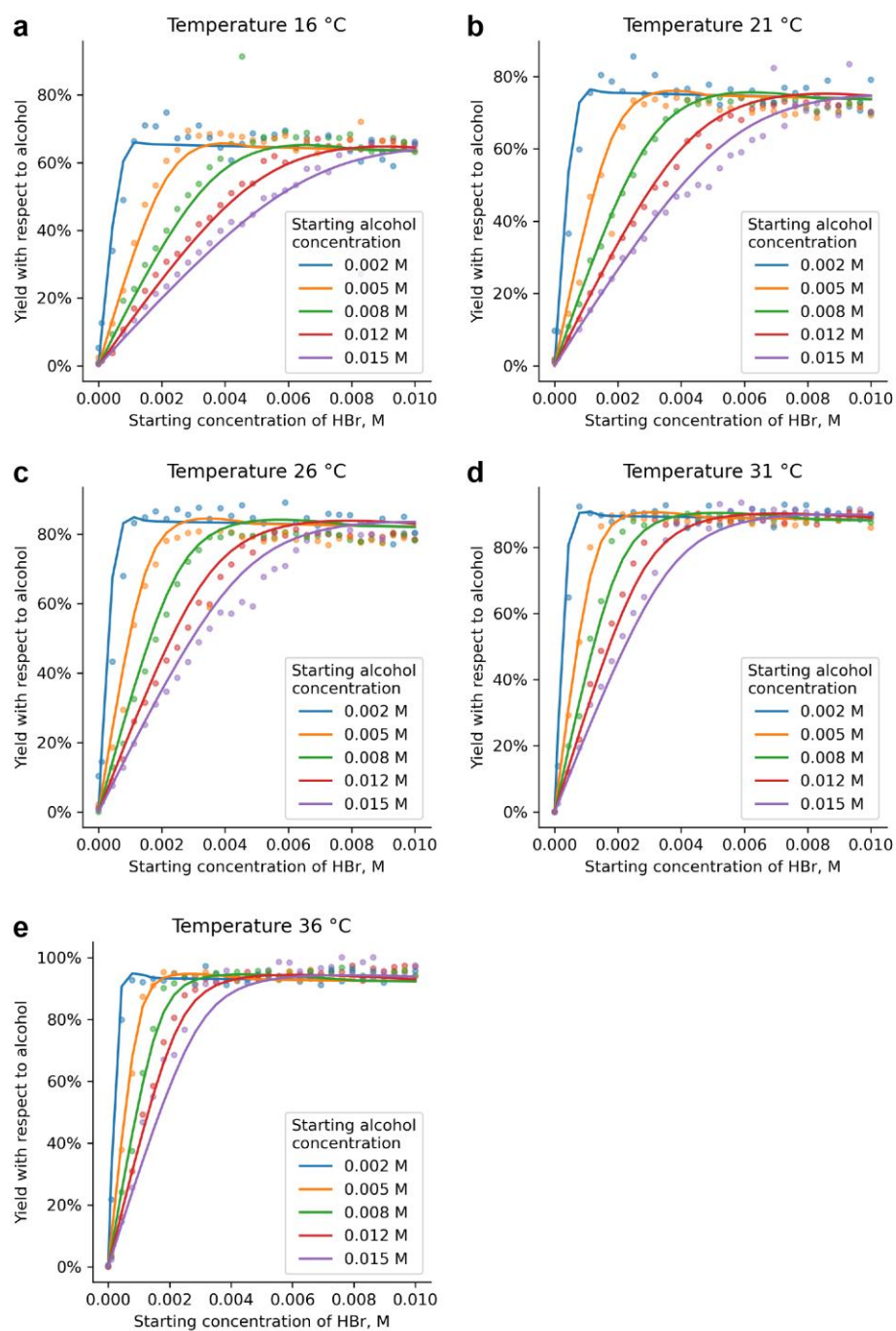


Figure S150. Yields predicted by the kinetics model (curves) best-fitted to the experimental data (circles). A single model shown here has been fitted to experimental data for all the temperatures.

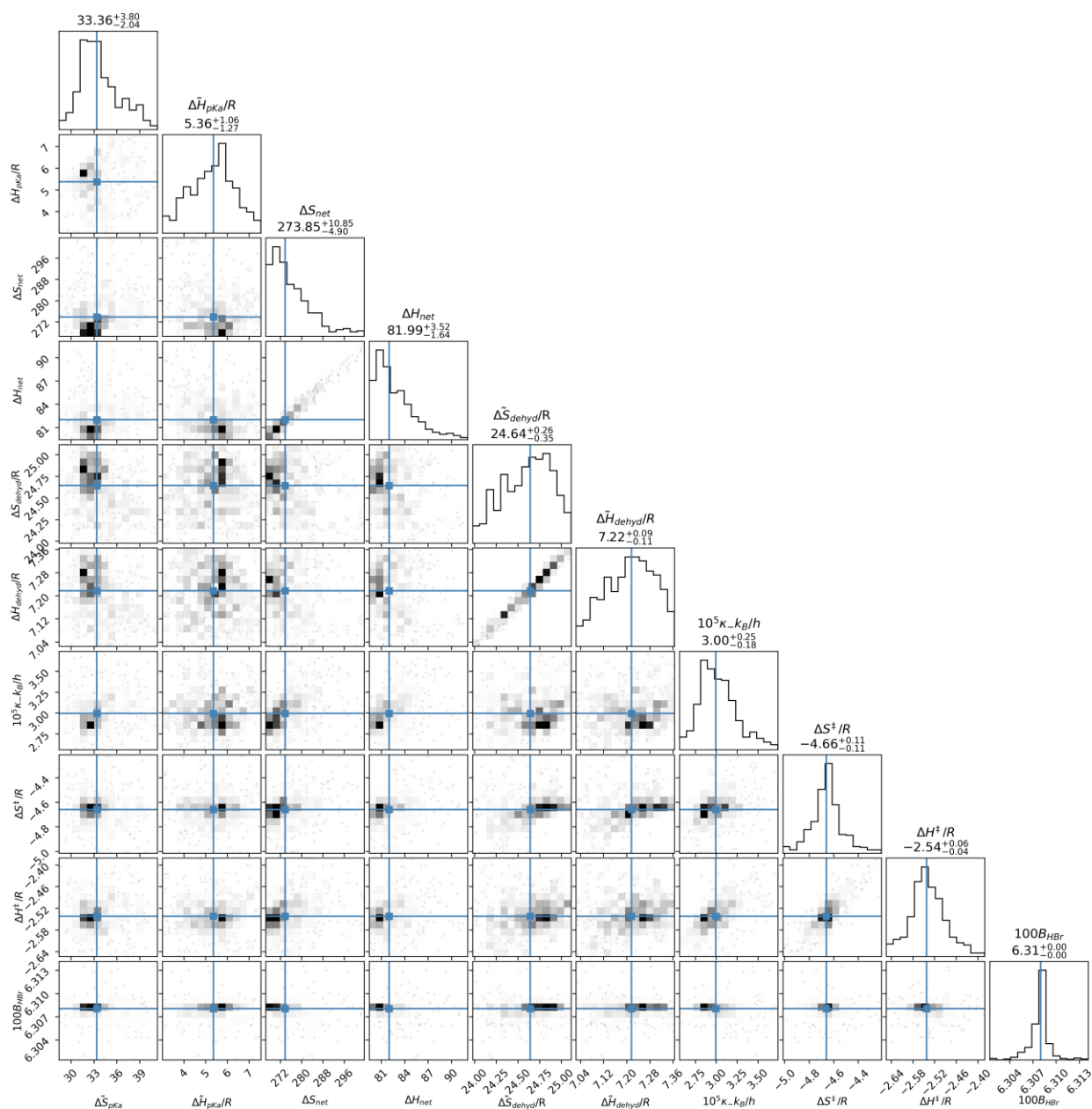


Figure S151. “Corner plot” showing bootstrapping analysis for probability distributions of model parameters of E1 reaction kinetics model. All the experimental data points were resampled in the course of bootstrapping. Number of bootstrapping iterations is 434, which took 140 hours on 40 CPU cores (Intel Xeon Gold 6230 CPU) in parallel.

At each bootstrapping iteration, the initial guesses for the overall ΔS and ΔH were equal to the respective best-fit values for the experimental sample ($\Delta S_0 = 268.463787$ J/(mol K) and $\Delta H_0 = 80.8157922$ J/mol) multiplied by random numbers sampled from uniform probability distribution between 0.85 and 1.15. The initial guesses for the other parameters were equal to the respective best-fit values obtained for the experimental sample.

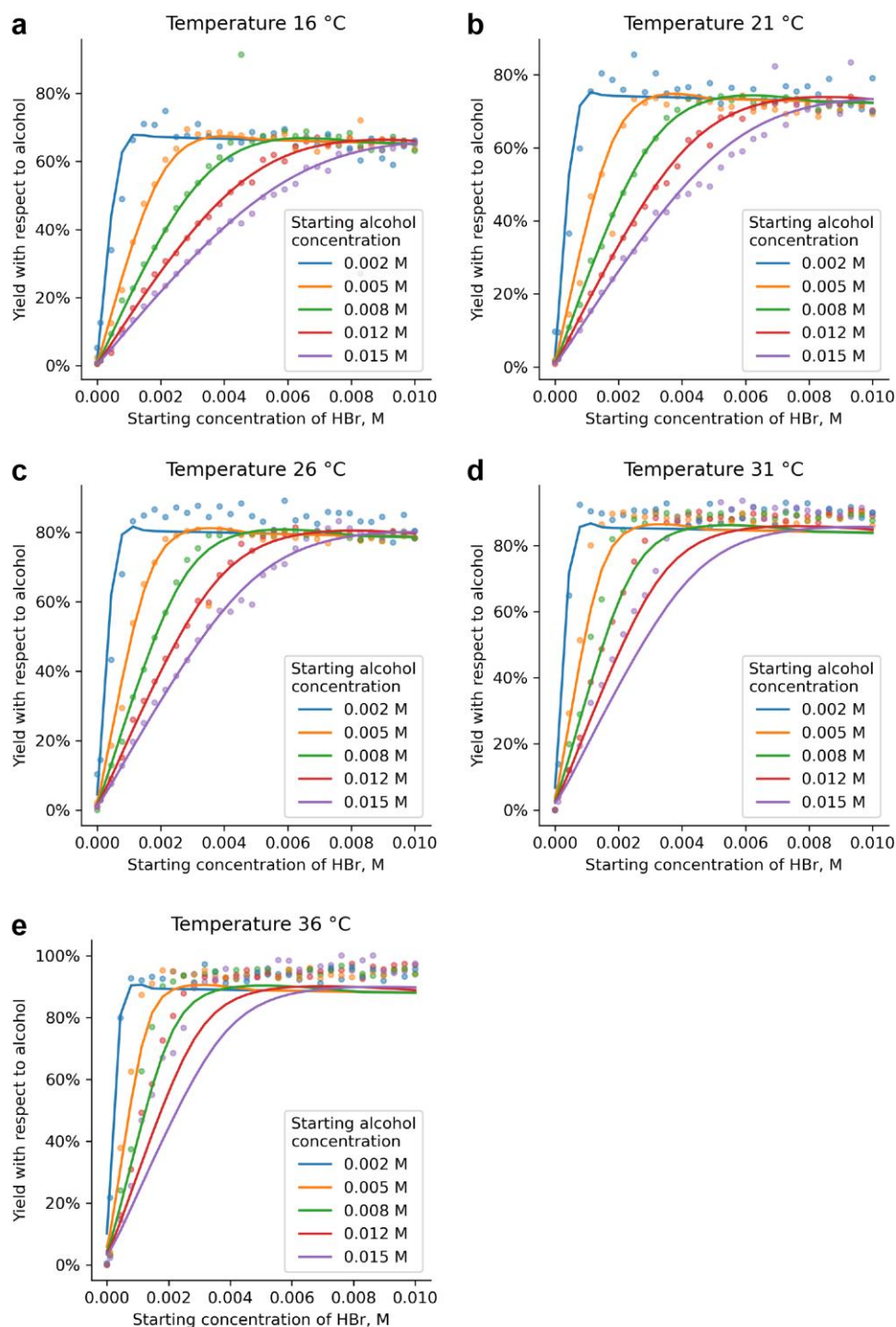


Figure S152. Yields predicted by the kinetics model (curves) best-fitted to the experimental data (circles). Temperature is indicated above each plot. A single model shown here has been fitted to experimental data for the temperatures below 31 °C only (panels **a-c**). Comparison of the model to experimental data for 31 °C (**d**) and 36 °C (**e**) is shown despite these data not being used for optimization of the model's parameters.

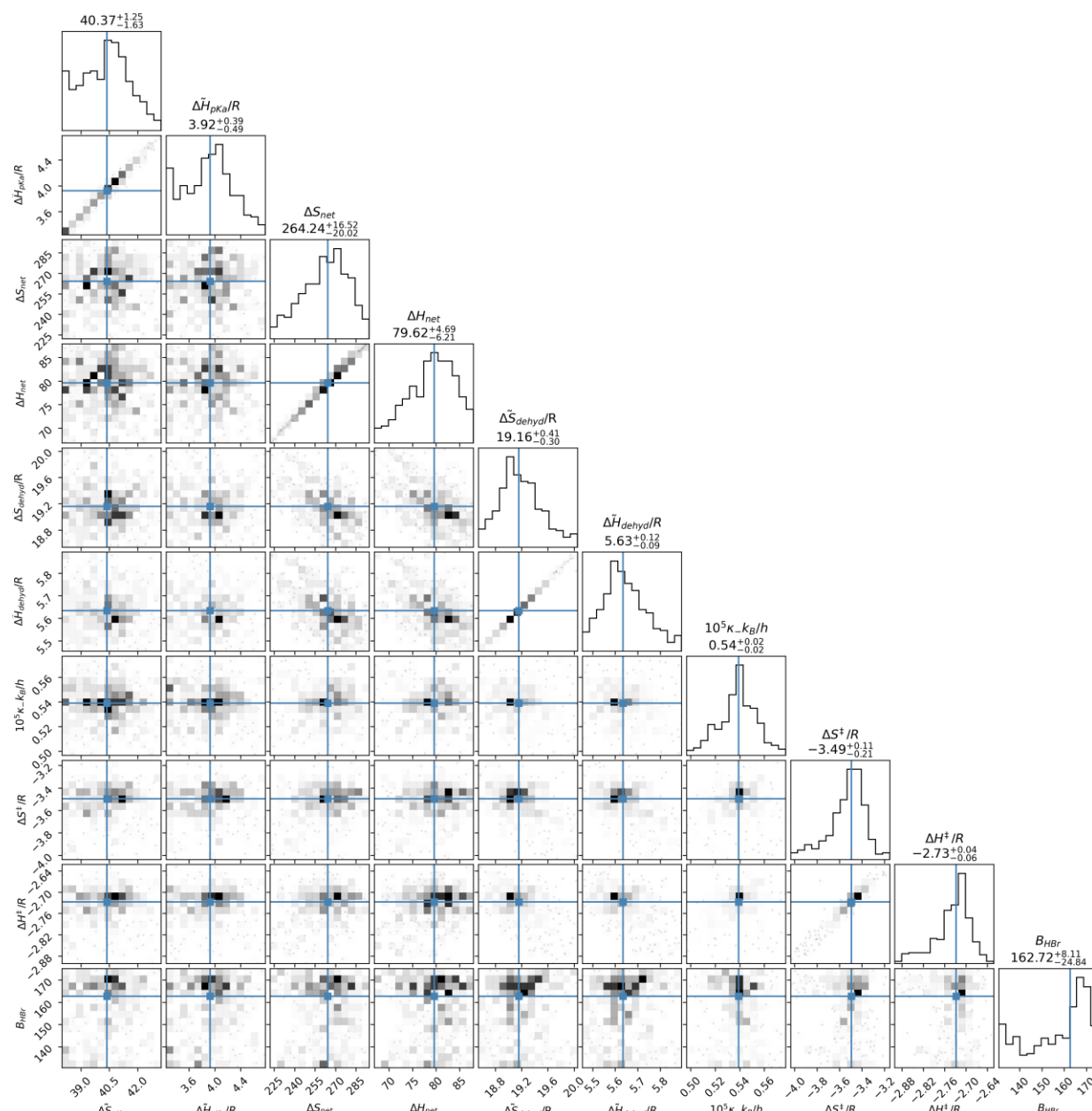


Figure S153. “Corner plot” showing bootstrapping analysis for probability distributions of model parameters of E1 reaction kinetics model fitted to data below 31 °C. Only the experimental data points for temperatures below 31 °C were used for best-fitting and for resampling in the course of bootstrapping. Number of bootstrapping iterations is 364, which took 323 hours on 40 CPU cores (Intel Xeon Gold 6230 CPU) in parallel.

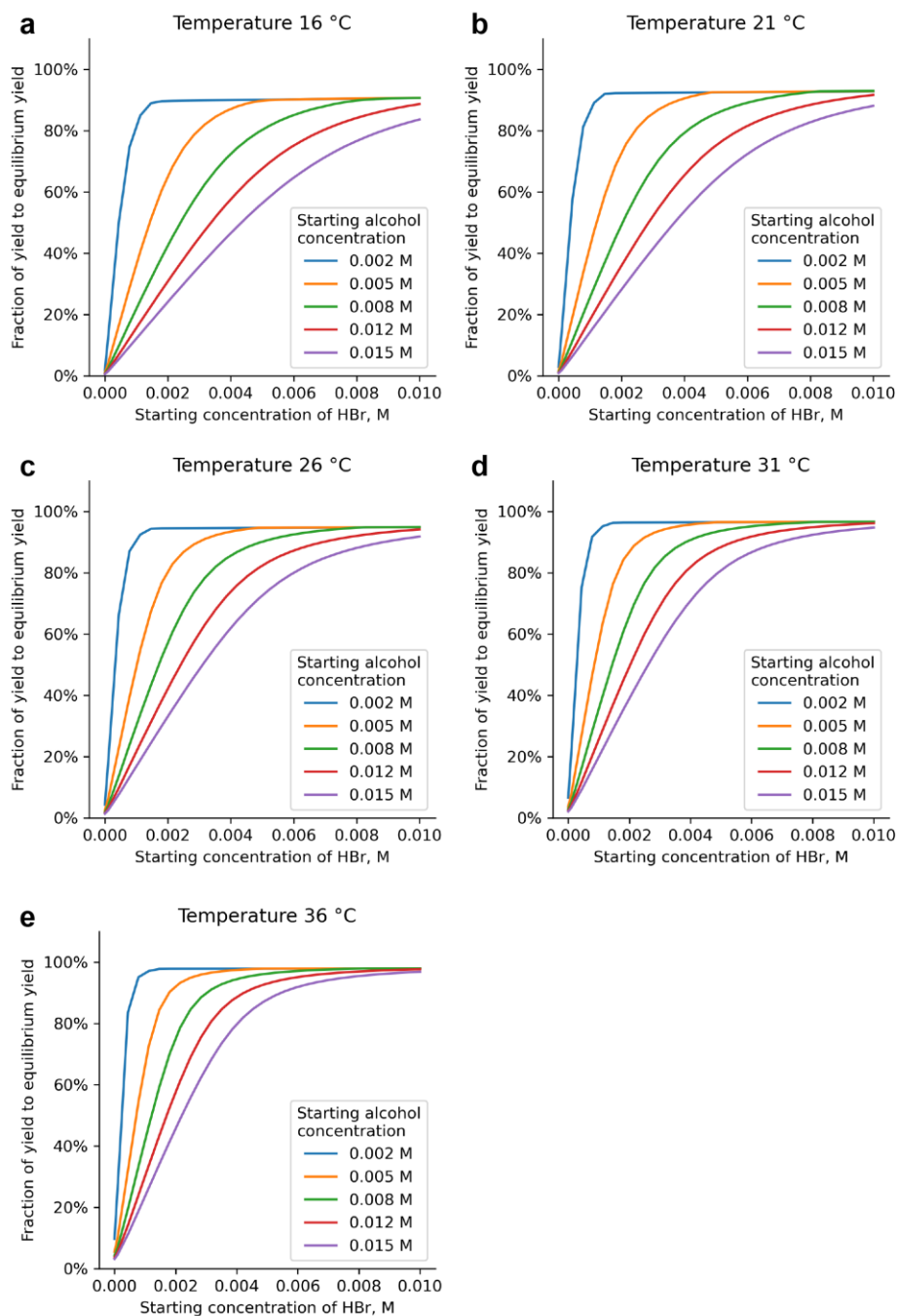


Figure S154. Ratios of yields predicted by the E1 reaction kinetics model best-fitted to the experimental data (circles in **Figure S153**) to the equilibrium yields evaluated for the same parameters of the model. Temperature is indicated above each plot. Model here (same as in **Figure S153**) has been fitted to experimental data for the temperatures below 31 °C only, and then evaluated for all temperatures at 4 hours of reaction duration (as in experiments), and at 4000 hours reaction time (approximating the equilibrium yield). The ratio of the two yields is plotted. We have validated that 4000 hours duration provides a good approximation of the equilibrium yield: the plots do not change upon further increase of the duration past that value.

5.3. Case of Ugi reaction

5.3.1. Overview of the model

Model of kinetics for Ugi reaction considered 12 reactions listed in **Table S22**, as well as equilibrium of proton exchange between 15 substance listed in Table S2. For overall scheme of the model, see **Scheme S5**. The transformation from **18d** to **16e** is represented by one single pseudo-first order rate constant (see **Scheme S6** for the underlying equilibria).

All the kinetic laws of reactions in **Table S22** were assumed to be first-order with respect to concentrations of substrates. For example, the reaction rate of reaction #2 in **Table S22** was modelled as $k_2[A][B] - k_{-2}[P_1]$, where the square brackets mean concentrations. The corresponding system of 12 ordinary differential equations was integrated by LSODA algorithm¹³⁵ with the first time step set to 36 milliseconds and the maximum time step limited to 2 hours. We used LSODA instead of the explicit Runge-Kutta method of order 5(4)¹¹⁶, as the latter was prohibitively slow for a kinetic model of this complexity. The specific implementation of LSODA we used is a SciPy wrapper⁷⁶ of the Fortran solver from ODEPACK¹³⁶. The integration proceeded from the time $t = 0$ to the time $t = 16$ hours, and the final product concentration was used to compute the final yield. Proton exchange (see the section below) was assumed to be much faster than all the 12 reactions listed in **Table S22**, and the proton exchange equilibrium was recalculated at every time step of integrating the differential equations. This constitutes the kinetics model. As a side note, given the 36 substances (including all species involved in proton exchange equilibria) and 26 reactions, the present reaction network has the “density of chemical space”, as defined by G. Restrepo¹³⁷, equal to $26/(3^{36} - 2^{37} + 1) \approx 1.7 \cdot 10^{-16}$, where $3^{36} - 2^{37} + 1$ is the theoretical upper bound¹³⁸ of the possible number of reactions among 36 substances.

The reaction rate constants and the pK_a values of the kinetics model were optimized to minimize the discrepancy between the yields predicted by the model and the yields obtained experimentally at various points of the condition space. For this optimization, we used Trust Region Reflective algorithm⁷⁵ with 2-point method of computing the Jacobian. To save computational time, we chose a subset of the experimentally probed points of the condition space to include all the prominent feature of the yield map. This subset is described in the **Table S32**.

Table S22. List of reactions and their forward and backward reaction rates corresponding to Scheme S5.

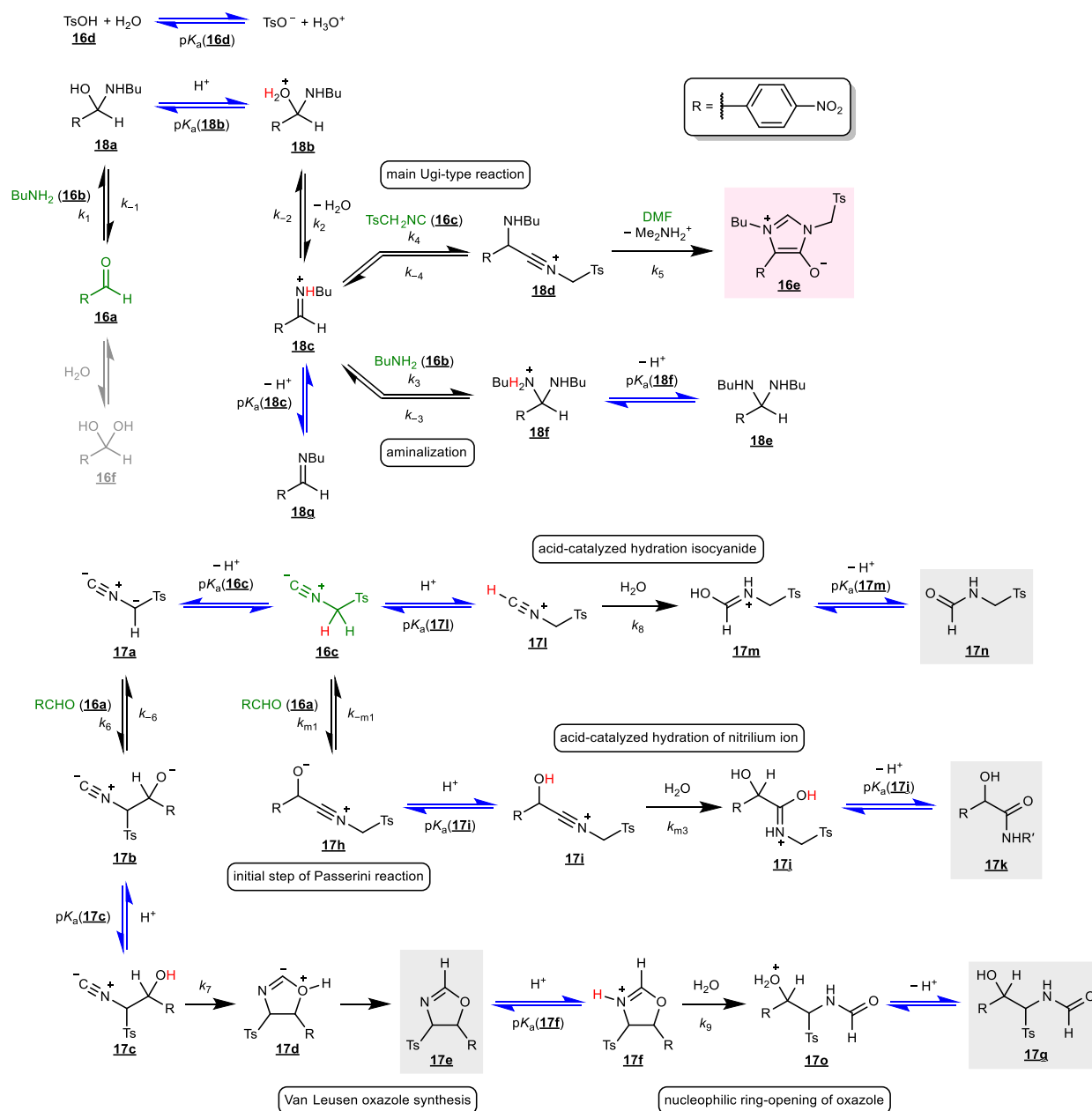
Reaction index	Reaction	Notation of kinetic rate constant for forward reaction	Notation of kinetic rate constant for backward reaction (if applicable)
1	<u>16a</u> + H ₂ O $\rightleftharpoons$ <u>16f</u>	$k_a K_{eq_a}$	k_a
2	<u>16a</u> + <u>16b</u> $\rightleftharpoons$ <u>18a</u>	k_1	k_{-1}
3	<u>18b</u> $\rightleftharpoons$ <u>18c</u> + H ₂ O	k_2	k_{-2}
4	<u>18b</u> + <u>16b</u> $\rightleftharpoons$ <u>18f</u>	k_3	k_{-3}
5	<u>18b</u> + <u>16c</u> $\rightleftharpoons$ <u>18d</u>	k_4	k_{-4}
6	<u>18d</u> + DMF $\rightleftharpoons$ <u>16e</u> + Me ₂ NH ₂	k_5	k_{-5}
7	<u>17a</u> + <u>16a</u> $\rightleftharpoons$ <u>17b</u>	k_6	k_{-6}
8	<u>17c</u> $\rightarrow$ <u>17e</u>	k_7	
9	<u>17l</u> + H ₂ O $\rightarrow$ <u>17m</u>	k_8	
10	<u>17f</u> + H ₂ O $\rightarrow$ <u>17o</u>	k_9	
11	<u>16c</u> + <u>16a</u> $\rightleftharpoons$ <u>17h</u>	k_{M1}	k_{-M1}
12	<u>17i</u> + H ₂ O $\rightarrow$ <u>17j</u>	k_{M3}	

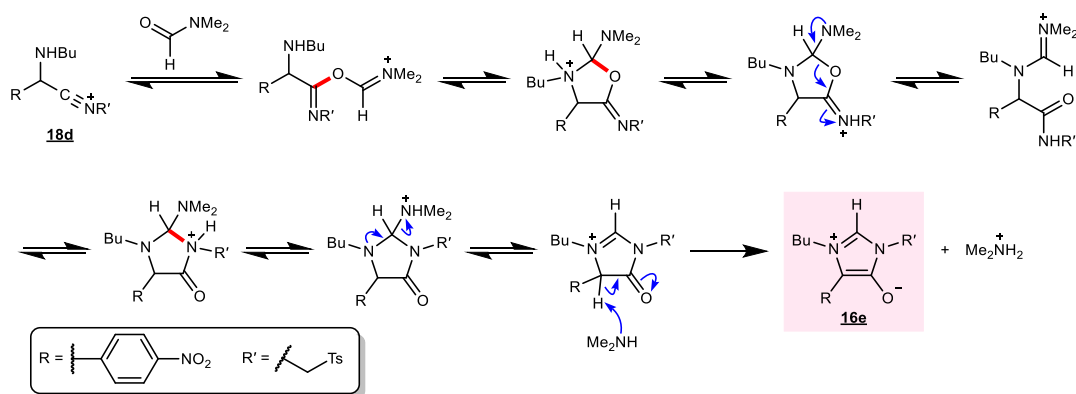
Table S23. Proton exchange equilibria and their respective pK_a values. For discussion of approximate literature values of pK_a see Section 5.3.3

Proton exchange equilibrium	pKa variable name in the code	pKa, values from the literature	pKa, best-fit value
$\text{TsOH} + \text{H}_2\text{O} \rightleftharpoons \text{TsO}^- + \text{H}_3\text{O}^+$	pKa_ptsa	-1.7	-2.49
$\text{BuNH}_3^+ + \text{H}_2\text{O} \rightleftharpoons \text{BuNH}_2 + \text{H}_3\text{O}^+$	pKa_BuNH ₃	12.23	11.15
$\text{DMFH}^+ + \text{H}_2\text{O} \rightleftharpoons \text{DMF} + \text{H}_3\text{O}^+$	pKa_DMFH	-7.2	-7.05
$\text{18b} + \text{H}_2\text{O} \rightleftharpoons \text{18a} + \text{H}_3\text{O}^+$	pKa_p1	8	7.23
$\text{18c} + \text{H}_2\text{O} \rightleftharpoons \text{18g} + \text{H}_3\text{O}^+$	pKa_p2	7	8.43
$\text{18f} + \text{H}_2\text{O} \rightleftharpoons \text{18e} + \text{H}_3\text{O}^+$	pKa_p3	10	11.17
$\text{16c} + \text{H}_2\text{O} \rightleftharpoons \text{17a} + \text{H}_3\text{O}^+$	pKa_c	13.1	14.40
$\text{17l} + \text{H}_2\text{O} \rightleftharpoons \text{16c} + \text{H}_3\text{O}^+$	pKa_ch	0.86	-3.92
$\text{17c} + \text{H}_2\text{O} \rightleftharpoons \text{17b} + \text{H}_3\text{O}^+$	pKa_q1	17	16.62
$\text{17m} + \text{H}_2\text{O} \rightleftharpoons \text{17n} + \text{H}_3\text{O}^+$	pKa_q2h	0	-0.53
$\text{17i} + \text{H}_2\text{O} \rightleftharpoons \text{17h} + \text{H}_3\text{O}^+$	pKa_m1h	17	14.50
$\text{17j} + \text{H}_2\text{O} \rightleftharpoons \text{17k} + \text{H}_3\text{O}^+$	pKa_m4h	0	-0.24
$\text{Me}_2\text{NH}_2^+ + \text{H}_2\text{O} \rightleftharpoons \text{Me}_2\text{NH} + \text{H}_3\text{O}^+$	pKa_me2nh2	11.97	11.00
$\text{17f} + \text{H}_2\text{O} \rightleftharpoons \text{17e} + \text{H}_3\text{O}^+$	pKa_q4	5.9	6.71

Table S24. Subset of the condition space points used for optimizing the parameters of the kinetics model. Experimentally, the probed pTSA concentrations were not equally spaced between 0.0219 M and 0.3 M, so the experimental data was interpolated along the pTSA coordinate (see Sections 3.13 and 4.15.4) at 17 evenly spaced points for comparing the model to the data.

Color in Figure S155 and Figure S156	Starting concentration of aldehyde	Starting concentration of amine	Starting concentration of isocyanide	Starting concentrations of pTSA
Violet	0.3 M	0.12 M	0.12 M	17 values evenly spaced between 0.0219 M and 0.2916 M
Blue	0.3 M	0.12 M	0.3 M	17 values evenly spaced between 0.0219 M and 0.2916 M
Brown	0.12 M	0.3 M	0.12 M	17 values evenly spaced between 0.0219 M and 0.2916 M
Green	0.12 M	0.3 M	0.3 M	17 values evenly spaced between 0.0219 M and 0.2916 M
Red	0.21 M	0.21 M	0.3 M	17 values evenly spaced between 0.0219 M and 0.2916 M





Scheme S6 Proposed elementary steps for the conversion of nitrilium ion **18d** to product **16e** via a Mumm rearrangement, cyclization and eventual aromatization according to ref ³¹. These steps are represented by one single pseudo-first order rate constant k_5 from **18d** to **16e** in our model shown in **Scheme S5**.

5.3.2. Fitting the model

Optimization of model parameters was done in two phases.

In the Phase 1, we optimized the following smaller (reduced) model. Only the reaction rate constants $k_1, k_{-1}, k_2, k_{-2}, k_3, k_{-3}, k_4, k_{-4}, k_5, k_{m1}, k_{-m1}$ were varied, while the rest of reaction rate constants were fixed at zero – essentially removing the respective reactions from the model. Furthermore, of all the $\text{p}K_a$ values, only the $\text{p}K_a$ of butylamine, p-TSA, and the substance **18b** were varied; all the remaining $\text{p}K_a$ values were fixed at their initial guess values (**Table S23**).

The initial guess of the model parameters was produced by means of Monte-Carlo Markov Chain (MCMC) sampler ¹³⁹. The prior distributions were chosen to reflect the parameter bounds (**Table S25**), and the initial states of the MCMC walkers were picked randomly (uniformly) from the range between the parameter bounds (for kinetic rate constants) and from ± 0.6 range around the literature values for each $\text{p}K_a$. Around 200 000 samples were acquired by MCMC. The sample with the smallest mismatch between the model and the data was used as initial guess for the optimization of the small (reduced) model.

Predictions of the best-fit small (reduced) model are compared to experimental data in **Figure S155**.

In the Phase 2, the best-fit parameter values of the small (reduced) model were used as initial guess for respective parameters of the full kinetics model. The initial guess for rate constants that were not varied in the reduced model were set to small values, which, however, were large enough to influence the model's predictions: this was necessary for the Jacobian with respect to these parameters to be non-zero, otherwise the optimization algorithm would not know in which direction to vary these parameters in order to improve the match between the model and

the data. Literature values (**Table S23**) were used as the initial guess of pKa values of substances other than butylamine, p-TSA, and the substance **18b**.

The final best-fit model is compared to the experimental data in **Figure S156**, and the respective kinetic rate constants are listed in **Table S25**. For comparison of the final best-fit model predictions against all the 3234 raw experimental data points, see Movie 3 and figures in Section 4.15.4 above.

Although the main purpose of this Section was to show that the observed yield map does not contradict the kinetic model in **Scheme S5** given pKa values close to the literature ones and given *some* set of rate constants, not necessarily a unique set, it is still interesting to estimate the degree to which our data constrains some of the model's rate constants. Due to high computational costs of the kinetics model, we were unable to perform the analysis of parameter uncertainties by the same Mote Carlo Markov Chain sampling we utilized for E1 kinetic model, so instead we resorted to the standard linear approximation method of estimating the confidence intervals of model parameters^{77–79}. For this purpose, only the rate constants were allowed to vary around the best-fit values, whereas the pKa values were fixed to their best-fit values. The Jacobian estimation method was 3-point differentiation with step equal to 0.2 (`diff_step` in SciPy).

Inspection of the correlations between the rate constants (see **Figure S157** for correlation matrix) reveals strong positive correlation between k_2 and k_{-2} , which indicates that only the ratio of these two constants may be constrained, not their individual values. This is likely because these constants are high and the respective forward and backward reactions rates that are too fast to be resolved through observation of the yield at the end of 16-hours-long reaction. Similarly strong correlations (positive or negative) are observed between k_1 and k_3 , between k_4 and k_{-4} and between k_{-m1} and k_{m3} . To analyze these correlations, we perform a change of variables so that the free parameters are the expressions that are potentially constrained by the experimental data: for example, instead of k_{-2} we introduce the equilibrium constant $K_2 = k_2/k_{-2}$ as a free parameter and then calculate k_{-2} as $k_{-2} = k_2/K_2$ when evaluating the model's equations. This change of variables does not influence the model's predictions, only the output of the linear approximation method of confidence interval estimation. Present rate constants expressed through new parameters K_2, K_3, K_4, K_{m3} are:

$$\begin{aligned}k_{-2} &= k_2/K_2 \\k_{m3} &= K_{m3}/k_{-m1} \\k_3 &= k_1/K_3 \\k_{-4} &= k_4/K_4\end{aligned}$$

This change of variables resulted in a substantial decrease of the formal errors (listed in **Table S26**, cf. uncertainties in **Table S26** before the change of variables). The resulting correlation matrix (**Figure S158**) no longer has extremely high (i.e. almost equal to 1) correlations between

forward and backward reaction rate constants. There appears a high correlation between k_8 and K_{m3} , as well as between k_4 and k_9 , which makes the kinetic model underdetermined with respect to these kinetic parameters.

Finally, according to the formal error analysis, the following parameters are constrained at least to an order of magnitude: k_1 , k_{-1} , k_2 , k_2/k_{-2} , k_1/k_3 , k_4/k_{-4} , and $k_{m3}k_{-m1}$ (indicated by green background in the “Uncertainties” column of **Table S26**)

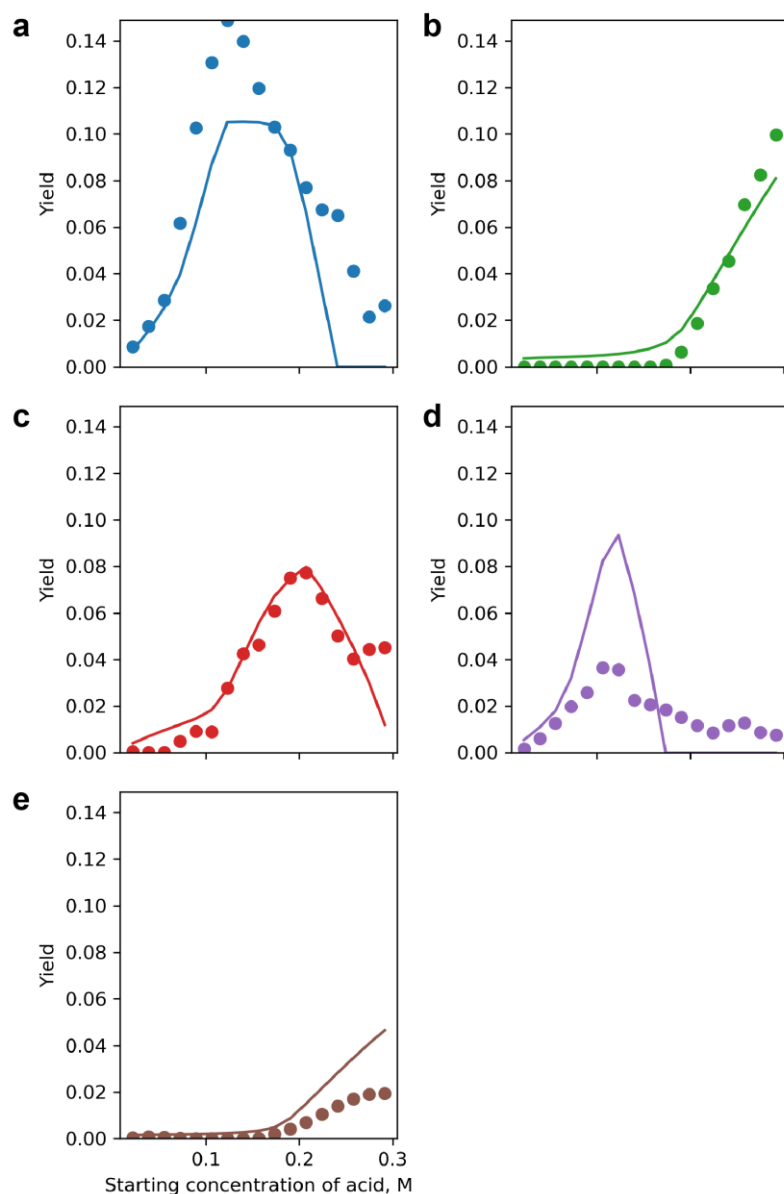


Figure S155. A smaller, reduced kinetics model (curves) fitted to the experimental data. Panels **a-e** correspond to subsets of experimental data listed in **Table S24**.

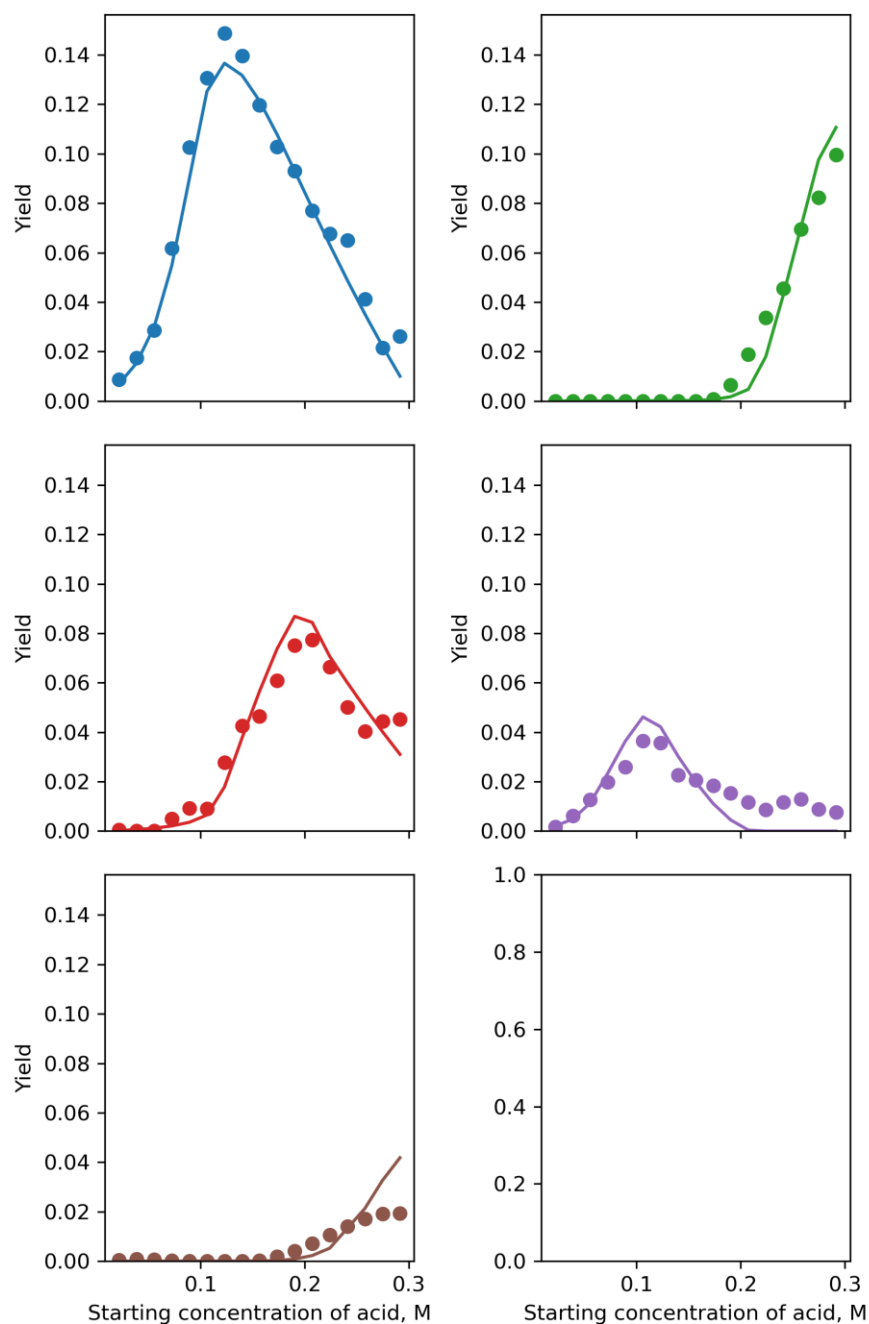


Figure S156. Full kinetics model (curves) fitted to the experimental data. Panels are arranged similarly to **Figure S155** and correspond to subsets of experimental data listed in **Table S24**—see the first column of **Table S24** for corresponding colors. Global maximum of yield in the entire condition space is the highest point in the panel **(a)**. Local maximum of yield is the highest point in the panel **(b)**. Horizontal axes are the concentrations of p-TSA.

Table S25. Best-fit values of kinetic rate constants in the model of Ugi reaction network before the change of variables. Pale green background in the “Uncertainty” column indicates parameters whose uncertainties are lower than the means, and therefore their values are constrained by the experimental data. Yellow background indicates the parameters whose uncertainties are roughly the same as the means.

Kinetic rate constant	Units	Upper bound	Initial guess for reduced model (Phase 1)	Best-fit value for reduced model (Phase 1)	Final best-fit value for the full model (Phase 2)	Uncertainty
k_1	$M^{-1}s^{-1}$	10000	175	29.48	5914	5380
k_{-1}	s^{-1}	10000	1.5155	10.06	3.38	2.24
k_2	s^{-1}	20000	97.18538	3820	13300	35000
k_{-2}	$M^{-1}s^{-1}$	10^6	2.1245	63100	150000	278000
k_3	$M^{-1}s^{-1}$	10^6	$1.1952 \cdot 10^{-5}$	130563.1	8888	3740
k_{-3}	s^{-1}	10000	$4.6 \cdot 10^{-2}$	$2.36 \cdot 10^{-6}$	0.00372	0.023
k_4	$M^{-1}s^{-1}$	10000	0.176037	241.58	25.3	979
k_{-4}	s^{-1}	10000	$1.263 \cdot 10^{-5}$	26.25	345	11400
k_5	$M^{-1}s^{-1}$	10000	7.693613	227.65	2.38	49
k_6	$M^{-1}s^{-1}$	200000	0 (fixed)	0 (fixed)	1265	62300
k_{-6}	s^{-1}	50000	0 (fixed)	0 (fixed)	248	3050000
k_7	s^{-1}	10000	0 (fixed)	0 (fixed)	137	3710000
k_8	$M^{-1}s^{-1}$	10^{13}	0 (fixed)	0 (fixed)	$8.75 \cdot 10^{11}$	2.3
k_9	$M^{-1}s^{-1}$	10000	0 (fixed)	0 (fixed)	5400	2790000
k_{M1}	$M^{-1}s^{-1}$	20000	0.0175	0.1667	0.4165	0.053
k_{-M1}	s^{-1}	20000	0	3052	11350	112000
k_{M3}	$M^{-1}s^{-1}$	10000	0 (fixed)	0 (fixed)	11238	77300

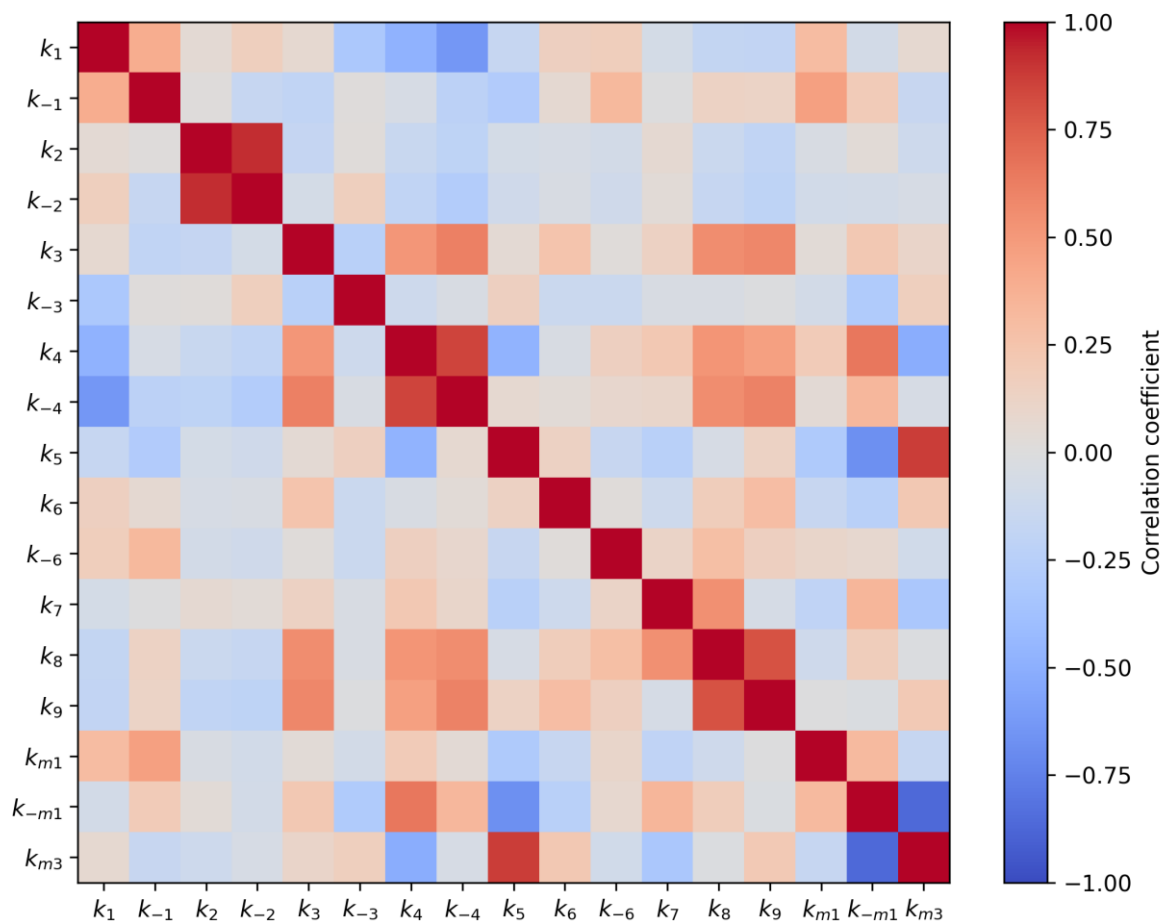


Figure S157. Matrix of correlations between the rate constants for the best-fit full kinetic model. Correlation is shown by color.

Table S26. Best-fit values of kinetic parameters after the change of variables (see text) in the model of Ugi reaction network. Pale green background in the “Uncertainty” column indicates parameters whose uncertainties are lower than the means, and therefore their values are constrained by the experimental data.

Kinetic parameter	Units	Final best-fit value for the full model (Phase 2)	Uncertainty
k_1	$M^{-1}s^{-1}$	5914	1200
k_{-1}	s^{-1}	3.4	2.5
k_2	s^{-1}	13300	11700
$K_2 = k_2/k_{-2}$	M	0.088	0.067
$K_3 = k_1/k_3$	1	0.65	0.23
k_{-3}	s^{-1}	0.0037	0.055
k_4	$M^{-1}s^{-1}$	25.3	2600
$K_4 = k_4/k_{-4}$	M^{-1}	0.063875	0.041
k_5	$M^{-1}s^{-1}$	2.38	1.75
k_6	$M^{-1}s^{-1}$	1265	73000
k_{-6}	s^{-1}	241	13300
k_7	s^{-1}	137	11400
k_8	$M^{-1}s^{-1}$	$8.75 \cdot 10^{11}$	$8.1 \cdot 10^5$
k_9	$M^{-1}s^{-1}$	5400	$5.5 \cdot 10^5$
k_{m1}	$M^{-1}s^{-1}$	0.416	0.044
k_{-m1}	s^{-1}	11350	58000
$K_{m3} = k_{m3}k_{-m1}$	$M^{-1}s^{-2}$	$1.4 \cdot 10^8$	$6.4 \cdot 10^8$

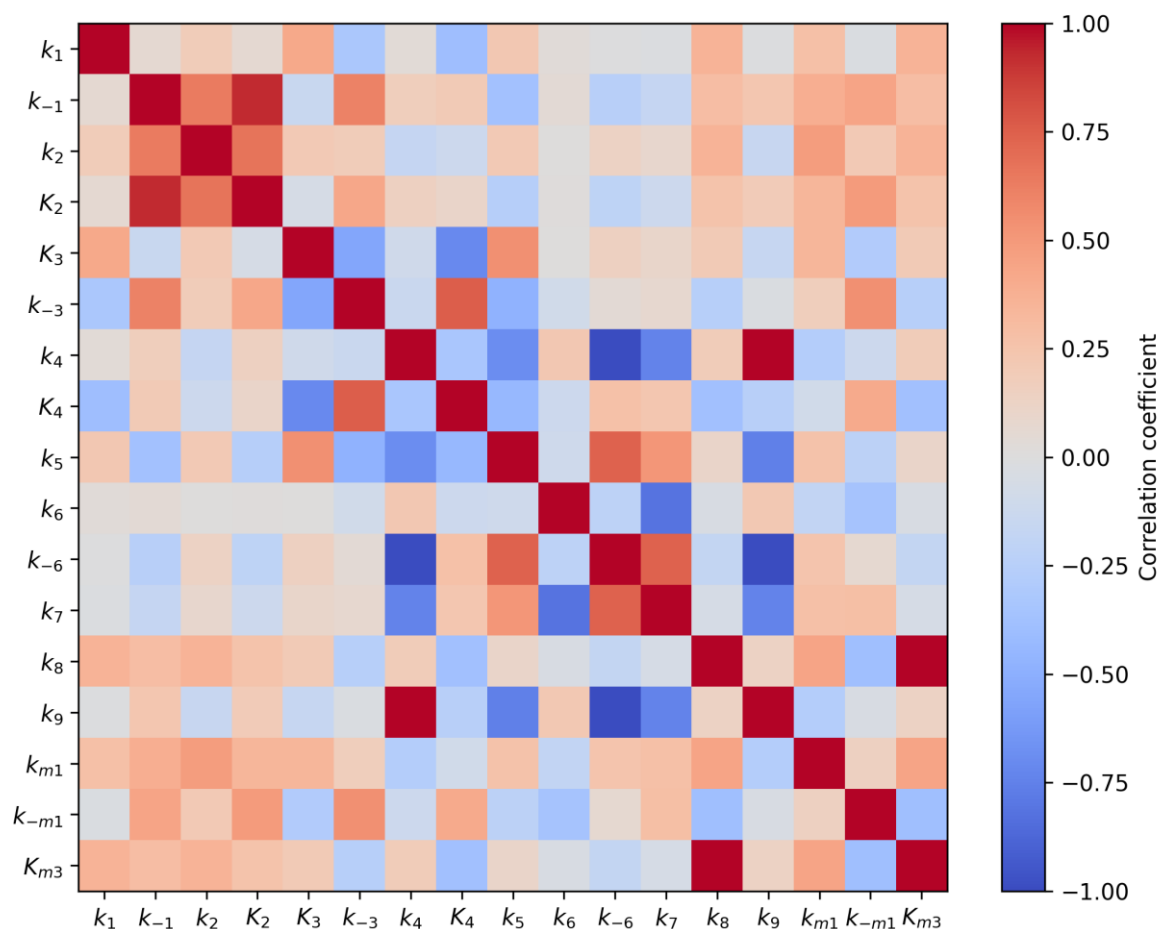


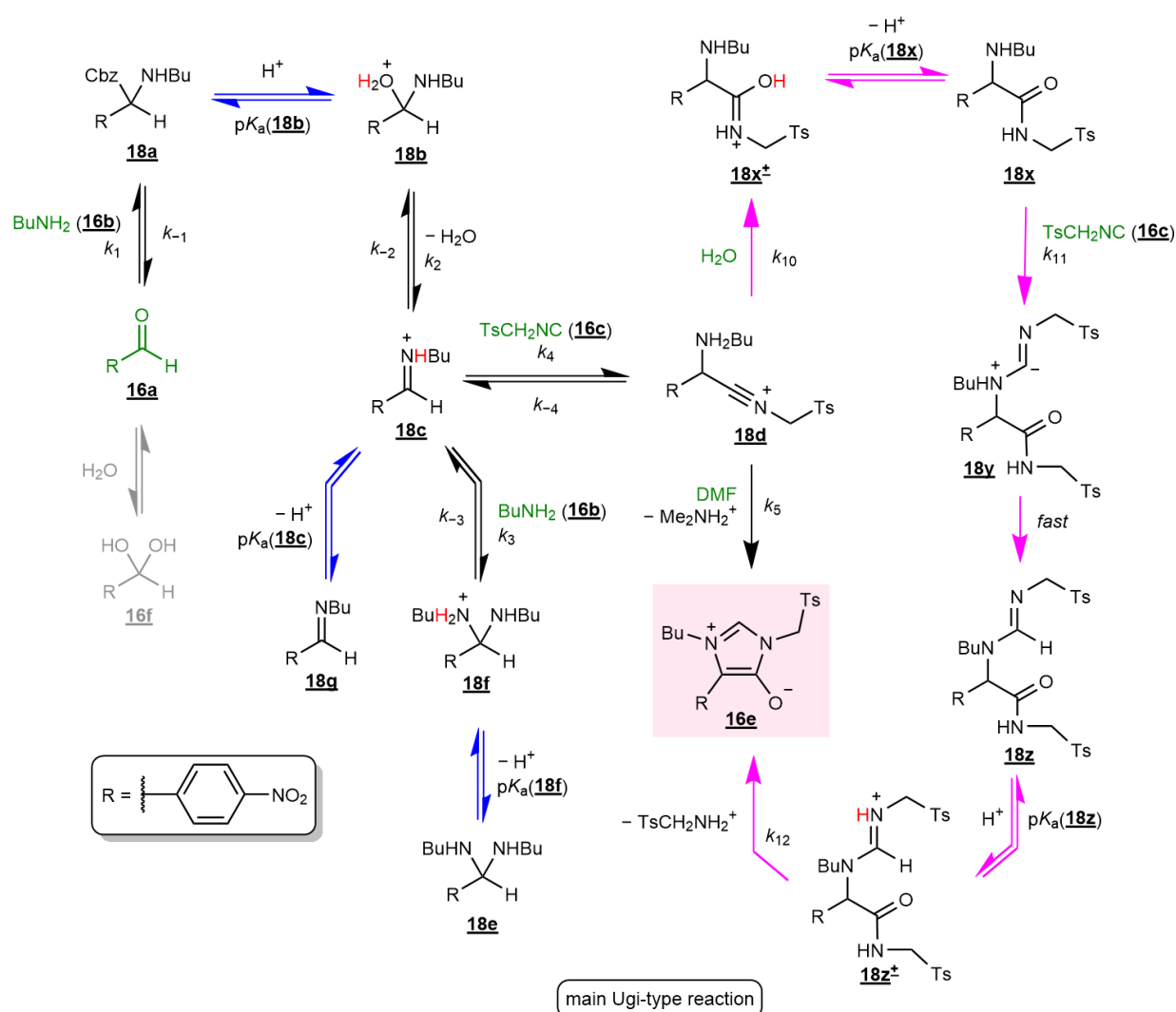
Figure S158. Matrix of correlations between the model parameters after the change of variables (see text) for the best-fit full kinetic model. Correlation is shown by color.

The fitted model follows the data quite closely, except for the slight mismatch at high acid concentrations seen at the right edges of plots in **Figure S156c,d,e**. For completeness, we tested whether adding a “hydrative” mechanism³¹, shown by pink arrows in **Scheme S7**, to the kinetics model would significantly reduce this residual mismatch. Initial guess of the pKa of **18x+** and **18z** was 0.1 and 12, respectively (**Table S28**). Initial guesses for the added mechanism’s kinetic rate constants k_{10} , k_{11} , and k_{12} were $1000 \text{ M}^{-1} \text{ s}^{-1}$, $30 \text{ M}^{-1} \text{ s}^{-1}$, and 3 s^{-1} . Initial guess for k_5 was 20% smaller than the best-fit value in the **Table S25** to give chance to the “hydrative” mechanism to make a contribution to the product yield by acting on **18d**. No significant improvement of the match between the model and data was observed upon inclusion of the “hydrative mechanism” (**Figure S159**): the root mean square discrepancy between the experimental and the modelled yield has improved from 0.74% without the “hydrative” mechanism (in units of yield) to 0.67% with the “hydrative” mechanism.

We have neglected the reversible hydration of aldehyde (reaction #1 in the **Table S22**) for the following reason. The reported product $K_{eq,a}[\text{H}_2\text{O}]$ of the equilibrium constant $K_{eq,a}[\text{H}_2\text{O}]$

and water concentration $[H_2O]$ for this reaction in pure water is 0.25^{75} . Thus, $K_{eq_a} = 0.004499\text{ M}^{-1}$. But the amount of water in our reaction can never exceed 0.6 M. Hence, the molar ratio of hydrated aldehyde to aldehyde can never exceed $K_{eq_a} \cdot (0.6\text{ M}) = 0.27\%$, which justifies the omission of this hydration reaction from the model.

As mentioned before, it is fundamentally impossible to evaluate the kinetic rate constants based on experimental yields obtained at a time point when yields no longer depend on time. In this context, it is important that, at 16 hours reaction time, the yield of Ugi reaction still keeps increasing with time, as shown in **Figure S160**, and drops at 30 hours due to product decomposition.



Scheme S7. Main section of the Ugi reaction network with *pink arrows* showing the “hydrative” mechanism³¹ incorporated into the model in addition to the classical Ugi mechanism.

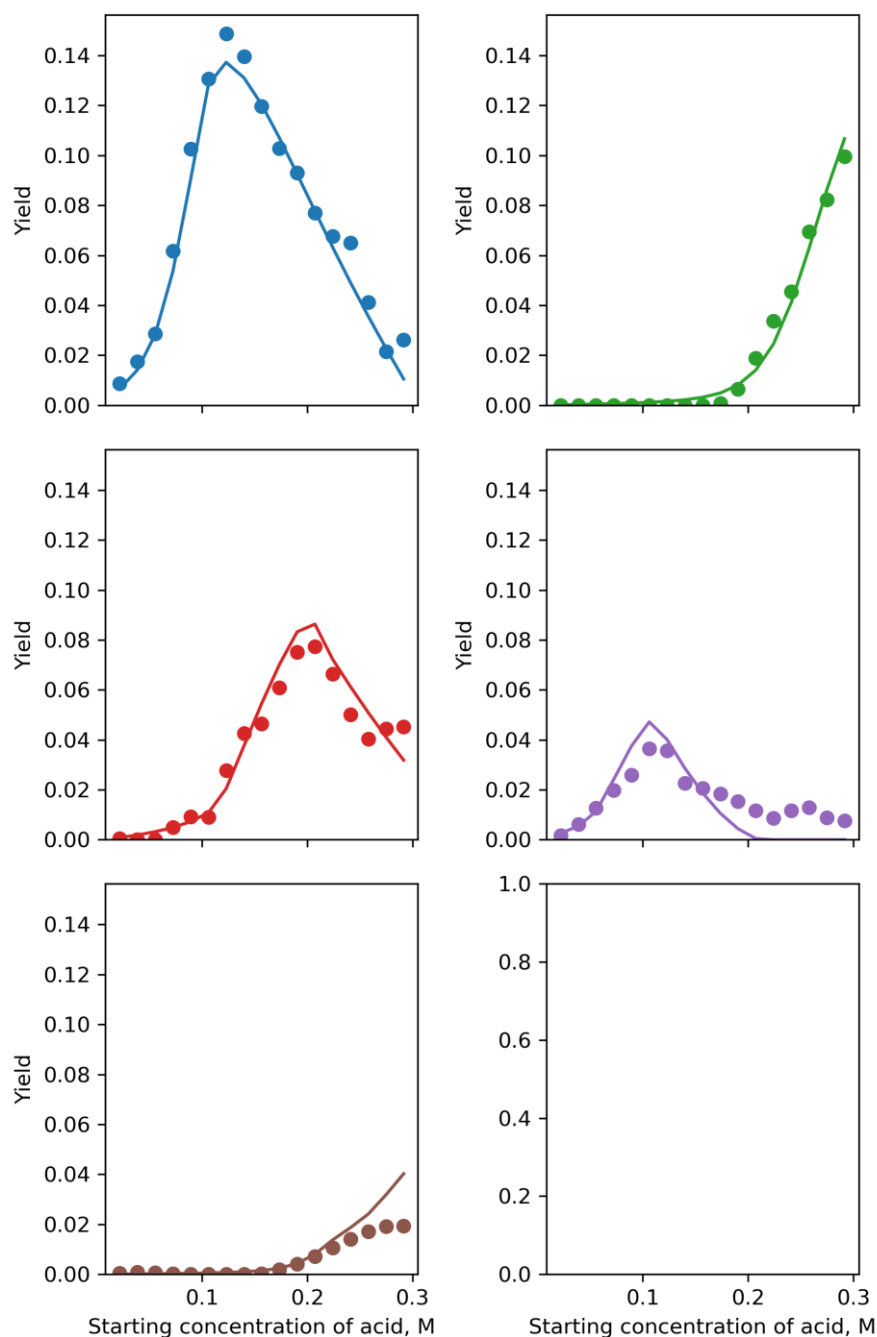


Figure S159. Full kinetics model with the additional “hydrative” mechanism³¹ (curves) fitted to the experimental data. Panels are arranged similarly to **Figure S155** and correspond to subsets of experimental data listed in **Table S24** – see the first column of **Table S24** for corresponding colors. Global maximum of yield in the entire condition space is the highest point in the panel **a**. Local maximum of yield is the highest point in the panel **b**. Horizontal axes are the concentrations of p-TSA.

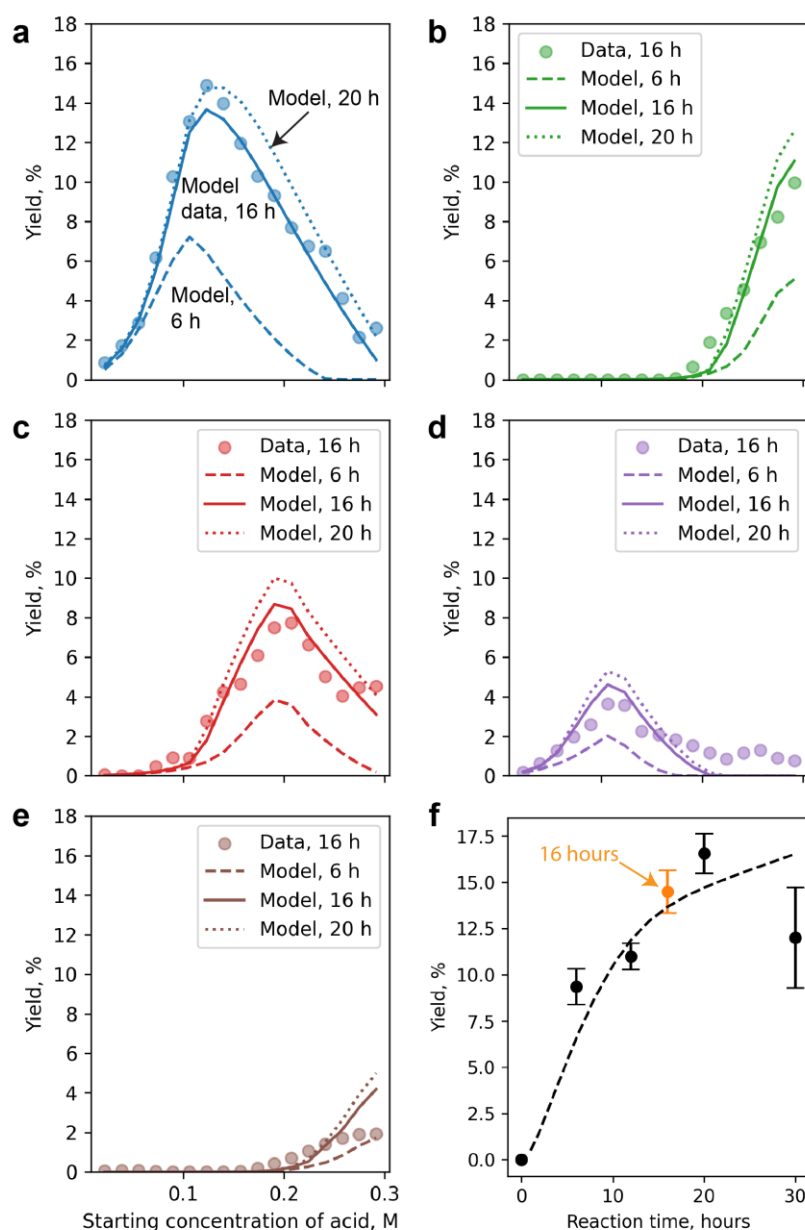


Figure S160. Time dependence of the Ugi reaction yields. **a-e**, Time dependence of yields predicted by the full Ugi reaction kinetics model without hydrative mechanism (curves) best-fitted to the experimental data at 16 hours reaction time (circles) and evaluated at reaction times 6 hours and 20 hours (indicated in the legends and on the plot in **a**). Panels are arranged similarly to **Figure S155** and correspond to subsets of experimental data listed in **Table S24** – see the first column of **Table S24** for corresponding colors. **f**, Experimental time dependence of yield at the conditions of global maximum of yield (circles). Error bars represent standard deviation of $n = 5$ experimental repetitions. Dashed curve is the full kinetics model best-fitted to the experimental data at reaction time 16 hours (same model as in panels a-e and **Figure S156**) and evaluated at different reaction times. That is, the experimental data shown by black circles in this plot were not used to fit the kinetic model represented by dashed curve. Lower experimental yield at 30 hours owes to product decomposition not modelled in our reaction network.

5.3.3. Accounting for proton exchange

It should be remembered that water was present in our starting reaction mixture, as pTSA was added as monohydrate. Hence, we write the reactions of proton exchange with respect to water as the common substance participating in proton exchange with all other substance. We could use DMF as common substance, but mathematically, it does not matter which common substance to use, as long as the equilibrium equations are consistent. On the one hand, it is reported that in mixtures of DMF and water the proton equilibrium is strongly shifted towards hydronium ions, rather than protonated DMFH⁺ ¹⁴⁰. On the other hand, it is more convenient to evaluate thus formulated equilibrium numerically by solving the zero-charge equation with respect to pH.

The total charge of the solution is the sum of charges of protonated and deprotonated species multiplied by their concentrations. For any given hypothetical value of pH, the concentrations of protonated and deprotonated species can be resolved for each row of **Table S23** like so:

$$[M_{\text{deprotonated}}] = [M]_0 * \frac{K_a}{\frac{[H_3O^+]}{[H_2O]} + K_a} \quad (19)$$

$$[M_{\text{protonated}}] = [M]_0 - [M_{\text{deprotonated}}] \quad (20)$$

where M is the substance in question (in a specific row of **Table S23**), $[M]_0$ is the net concentration of all the forms of this substance: both $M_{\text{protonated}}$ and $M_{\text{deprotonated}}$. $K_a = 10^{-pK_a}$ is the equilibrium constant, $[H_3O^+] = 10^{-pH}$ is the concentration of hydronium ions, and $[H_2O]$ is the concentration of the neutral water (which depends on $[H_3O^+]$ and the net concentration of water in the mixture in the manner similar to the two equations above).

Isocyanide can be both protonated and deprotonated. Hence, it has two values. Formulas for double ionization are derived thusly:

System of equations for species M with the net concentration $[M]_0$ is:

$$[M]_0 = [M] + [M^-] + [M^{2-}] \quad (21)$$

$$K_{a1} = \frac{[M^-][H_3O^+]}{[M][H_2O]} \quad (22)$$

$$K_{a2} = \frac{[M^{2-}][H_3O^+]}{[M^-][H_2O]} \quad (23)$$

Solving it with respect to $[M^-]$ and $[M^{2-}]$ gives:

$$[M^-] = \frac{[M]_0 K_{a1}}{\frac{[H_3O^+]}{[H_2O]} + \left(1 + K_{a2} \frac{[H_2O]}{[H_3O^+]}\right) K_{a1}} \quad (24)$$

$$[M^{2-}] = [M^-] K_{a2} \frac{[H_2O]}{[H_3O^+]} \quad (25)$$

It is important that these formulas, both for the case of single and double ionizations, though more complex than the ones normally used by practitioners, are free both from the assumption that the solute is in pure water and from the assumption that the concentration of water is much larger than the concentration of the solute. In our case, water can be produced and consumed over the course of the reaction, and its concentration can be comparable to the concentrations of other species, hence the necessity to use the more general formulas above.

Note that due to subtractions and implicit exponentiations in the formulas (19)-(20) and (24)-(25), they need to be evaluated numerically with larger floating point precision than that offered by 64-bit floating-point numbers commonly used by default in software, otherwise the rounding errors of the floating-point arithmetic used by the computer lead to significant errors. One method is to use software libraries for arbitrary-precision arithmetic (we tried 100 digits of precision and it solved the problem). We use this method for the isocyanide, as it can be both protonated and deprotonated. For the species undergoing single deprotonation only, it is possible to instead reformulate the formulas (19) and (20) in terms of logistic sigmoid function of pH and pKa values, which is often implemented numerically in a way robust against floating-point arithmetic errors (for instance, such is the implementation in Python SciPy package ⁷⁶).

Once all the concentrations of protonated and deprotonated species were computed for a given hypothetical pH, the net charge Q of the solution is obtained as

$$Q = \sum_M \left(Q_{M_{\text{protonated}}} [M_{\text{protonated}}] + Q_{M_{\text{deprotonated}}} [M_{\text{deprotonated}}] \right)$$

where summation is over species, and charges of protonated forms and deprotonated forms are $Q_{M_{\text{protonated}}}$ and $Q_{M_{\text{deprotonated}}}$.

This established the dependence of Q on pH, and therefore the equation $Q(\text{pH}) = 0$ can be solved against pH. We solve this equation numerically by the Brent's method (Chapters 3-4 in Ref. ¹⁴¹)

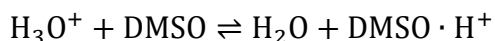
For solving the proton exchange equilibrium, we developed our own code in Python instead of using the available Python packages, since they rely on assumptions of small concentrations of solute relative to water, and/or fix the water concentration at that of pure water. Our code is available from the Github repository <https://github.com/yaroslavsobolev/robowski-maps> published along with this article. Specifically, the module implementing these calculations is located in the directory “*robowski/kinetics_models/acid_base_equilibrium.py*” of the code repository.

For the purposes of optimizing the model's parameters, the initial guess for pK_a values is listed in **Table S23**. These values were estimated from available data in the literature (**Table S27** and **Table S28**). In cases when literature values were reported for DMSO as the solvent, we used the following empirical formula for conversion between pK_a values in DMSO and DMF (Formula (9) in Ref. ¹⁴²):

$$pK_a(\text{DMF}) = 1.56 + 0.96pK_a(\text{DMSO}) \quad (26)$$

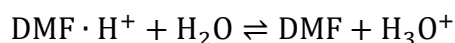
For example, pK_a of *n*-butylammonium ion in DMSO is 11.12, according to Crampton & Robotham ¹⁴³. The conversion formula above yields pK_a of *n*-butylammonium in DMF equal to $(1.56 + 0.96 \cdot 11.12) = 12.23$.

Reaction



In DMSO as solvent has $pK_a = 5.88 \pm 1.79$ as calculated by Yang *et al.* ¹⁴⁴. Conversion from DMSO solvent to DMF solvent yields $pK_a = 7.2 \pm 1.8$.

Hence, the reverse equilibrium



Is estimated to have pK_a equal to (-7.2 ± 1.8) .

Entry 4 in Table 1 in the Ref. ¹⁴⁰ gives autoprotolysis constant of water in DMF:

$$\frac{[\text{H}_3\text{O}^+][\text{OH}^-]\gamma_{\text{H}_3\text{O}^+}\gamma_{\text{OH}^-}}{\chi_{\text{H}_2\text{O}}^2} = 27.44 \pm 3.40$$

where $\gamma_{\text{H}_3\text{O}^+}$ and γ_{OH^-} are activity coefficients, which we take to be equal to 1/M, and $\chi_{\text{H}_2\text{O}}$ is the molar fraction of water.

Table S27. Literature pK_a values for species modelled in our network. The blue backgrounds indicate analogues that are structurally similar to our modelled species lacking literature pK_a data.

Species with fitted pK_a	Species for comparison	pK_a reported in literature [solvent]	Corrected pK_a in DMSO	Corrected pK_a in DMF ^[d]
TsOH (16d)	PhSO ₃ H	-2.8 [water] ¹⁴⁵	/	/
DMF·H ⁺	DMF·H ⁺	6.1 [MeCN] ¹⁴⁶	/	/
BuNH ₃ ⁺	BuNH ₃ ⁺	10.59 [water] ¹⁴⁷ 11.12 [DMSO] ¹⁴³	10.09 ^[a]	12.23
Me ₂ NH ₃ ⁺	Me ₂ NH ₃ ⁺	10.64 [water] ¹⁴⁷	10.84 ^[b]	11.97
TsCH ₂ N ⁺ ≡CH (17l)	CyN ⁺ ≡CH	0.86 [water] ¹⁴⁸	/	/
TsCH ₂ NC (16c)	TsCH ₂ NC	14 [water] ¹⁴⁹	/	/
	PhCH ₂ CN	12.0 [DMSO] ¹⁵⁰	/	13.1
oxazolinium ion 17f	2-methyl-2-oxazolinium ion	5.5 [water] ¹⁵¹	4.5 ^[c]	5.9

[a] $pK_a(\text{in DMSO}) = pK_a(\text{in water}) - 0.50$ for primary ammonium ions¹⁵²

[b] $pK_a(\text{in DMSO}) = pK_a(\text{in water}) + 0.20$ for secondary ammonium ions¹⁵²

[c] $pK_a(\text{in DMSO}) = pK_a(\text{in water}) - 1.00$ for protonated heterocycles¹⁵²

[d] Formula (26) was applied to data in Ref.¹⁴²

Table S28. Approximate pK_a values of species classes without available experimental values.

Species with fitted pK_a	Class of species	Approximate pK_a in water
17j, 17m, 18x+	Protonated amides	~0 ¹⁵³
18c	Iminium ions	5–7 ¹⁵⁴
18z+	Amidinium ions	~12 ¹⁵³

5.3.4. Interpretation of the best-fit kinetics model

In this subsection, we aim at rationalizing the occurrence of two yield maxima in the hyperspace spanning the initial concentrations (i.e. [pTSA]₀, [aldehyde]₀, [amine]₀ and [isocyanide]₀). Notably, three features of initial conditions on the yield (**I–III** described below) were observed. Below, these three features are rationalized in terms of the fitted rate constants and pK_a values of the partial network (**Scheme S8**) and can be used to rationalize presence of the global and the local yield maxima.

To start with, we consider conditions NOT fulfilling the criterion $[\text{pTSA}]_0 \approx [\text{amine}]_0$. In the limiting case of $[\text{pTSA}]_0 \gg [\text{amine}]_0$, the amine is completely protonated by pTSA (fitted $\text{p}K_a = -2.49$) to give the *n*-butylammonium cation (fitted $\text{p}K_a = 11.15$). As a result, a negligible equilibrium concentration of hemiaminal **18a** (step A) could be achieved. In the other limiting case of $[\text{pTSA}]_0 \ll [\text{amine}]_0$, the degree of protonation of hemiaminal **18a** (step B) giving **18b** ($\text{p}K_a = 7.2$) would be too minute. In both scenarios, the pre-equilibria cause a lower steady-state concentration of iminium ion **18c** (the key intermediate involved in the rate-determining step), and hence a lower yield of **16e**. Apparently, the midway between these two extremes, implying $[\text{pTSA}]_0 \approx [\text{amine}]_0$, is optimal for product formation.

III. When $[\text{aldehyde}]_0 + [\text{amine}]_0 = \text{constant}$, the final product concentration decreases gradually as $[\text{amine}]_0$ increases (Figure S162).

Here, we analyze separately how excess aldehyde or excess amine differently hampers the formation of the heterocycle **16e**. In the former case ($[\text{aldehyde}]_0 > [\text{amine}]_0$), the side reaction of concern would be the aldehyde–isocyanide reaction yielding **17h** (step F, $k_{\text{ml}} = 0.42 \text{ M}^{-1} \text{ s}^{-1}$), which is far slower than the productive aldehyde–amine reaction (step A, $k_1 = 5914 \text{ M}^{-1} \text{ s}^{-1}$). Thus, the side reaction associated with excess aldehyde can be neglected.

In the latter case ($[\text{amine}]_0 > [\text{aldehyde}]_0$), excess amine **16b** can facilely react with the iminium ion **18c** to yield the protonated aminal **18f** (step G, $k_3 = 8888 \text{ M}^{-1} \text{ s}^{-1}$). This competition is much faster than the key iminium–isocyanide reaction ($k_4 = 25.3 \text{ M}^{-1} \text{ s}^{-1}$). Consequently, the steady-state concentration of iminium ion **18c** for the productive step D is reduced.

Hence, an inspection of side reactions pinpoints that excess amine is more detrimental than excess aldehyde to the formation of product **16e** due to kinetics. In terms of thermodynamics, the same conclusion can be drawn by comparing the fitted equilibrium constants, specifically $K_F \ll K_A$ and $K_G \gg K_D$ as shown in **Scheme S8**.

To further validate the reasoning proposed above, we plotted the ratio $([\text{18f}]_{\text{final}} + [\text{18e}]_{\text{final}}) / ([\text{18c}]_{\text{final}} + [\text{18g}]_{\text{final}})$ against $[\text{amine}]_0$ (**Figure S161**). The numerator is the sum of final theoretical concentrations of aminal **18e** and its conjugate base **18f**, whereas the denominator is the sum of the final theoretical concentrations of the iminium ion **18c** and its conjugate base **18g**. It turns out that this ratio increases with $[\text{amine}]_0$, which is consistent with our proposed influence of the aminal formation (step G) on the yield of product **16e**.

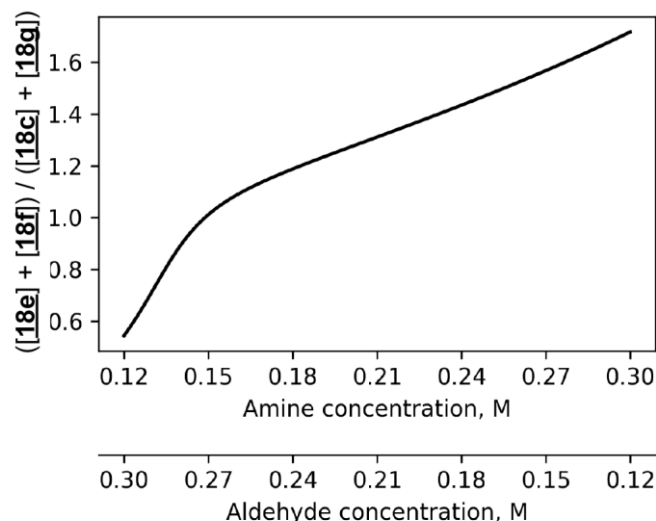


Figure S161. Ratio of concentrations of intermediates ($[18e] + [18f]$)/($[18c] + [18g]$) at the end of the reaction, according to the best-fit kinetics model of Ugi reaction network along the hyperline in the condition space given by the equations $[pTSA]_0 = [amine]_0 = ((0.42 \text{ M}) - [aldehyde]_0)$ and $[isocyanide]_0 = 0.30 \text{ M}$.

5.3.5. Occurrence of two yield maxima

Consider the case when $[aldehyde]_0 + [amine]_0 = 420 \text{ mM}$. As shown in **Figure S162b**, the plot of $[limiting \text{ reactant}]_0$ against $[amine]_0$ consists of two straight lines that meet when the amine and aldehyde are equimolar (210 mM). If $[amine]_0 < 210 \text{ mM}$, the amine is the limiting reactant. If $[amine]_0 > 210 \text{ mM}$, the aldehyde is the limiting reactant. Displayed in **Figure S162c** is the yield of product **16e** against $[amine]_0$, featuring a global maximum (at $[amine]_0 = 120 \text{ mM}$) and a local maximum (at $[aldehyde]_0 = 120 \text{ mM}$). In other words, both maxima correspond to the situation when $[limiting \text{ reactant}]_0$ is at minimum (120 mM), which is a reasonable result because yield is calculated by dividing $[product]_{final}$ by $[limiting \text{ reactant}]_0$. The relative difference in yield between the two maxima is a consequence of feature **III** explained above (i.e. $[product]_{final}$ decreases as $[amine]_0$ increases). Likewise, a yield minimum at either $[amine]_0 = 210 \text{ mM}$ (fitted curve) or 240 mM (experimental datapoint) is understandable because the value of $[limiting \text{ reactant}]_0$ is at or close to the maximum value (210 mM). For an automatically identified shortest linear hyperpath between the local and global maxima, see **Figure S163**.

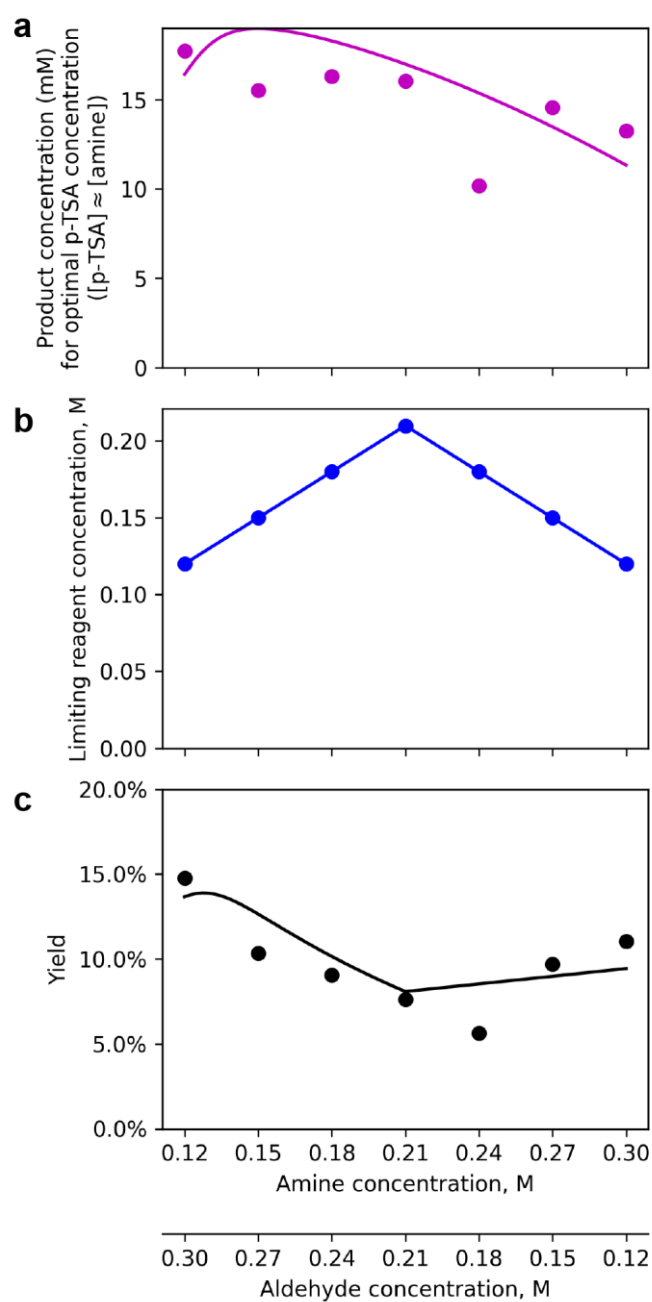


Figure S162. Behavior of the best-fit kinetics model of Ugi reaction network along the hyperline in the condition space given by the equations $[pTSA]_0 = [amine]_0 = ((0.42 \text{ M}) - [aldehyde]_0)$ and $[isocyanide]_0 = 0.30 \text{ M}$.

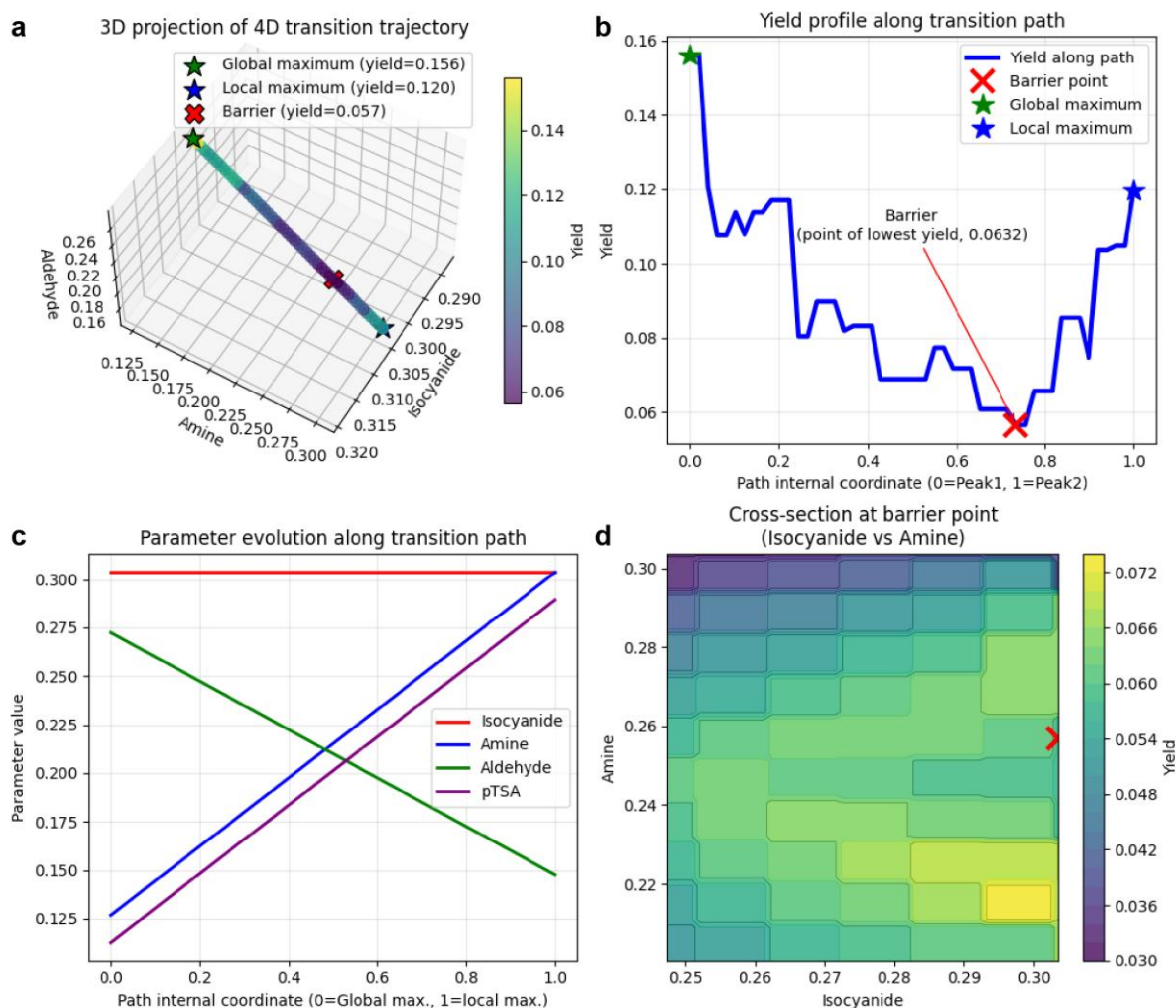


Figure S163. Shortest linear four-dimensional hyperpath connecting the local and global maxima in the experimental map of the Ugi reaction yields. **a**, 3D projection of the shortest 4D hyperpath. Color indicates yield (see colorbar). **b**, Yield plotted along the internal coordinate of the 4D hyperpath. Red cross shows the point with the lowest yield on the hyperpath. **c**, Dependence of the four coordinates on the internal coordinate of the hyperpath. **d**, Cross-section of experimental map in [amine]-[isocyanide] plane at the lowest yield point of the 4D hyperpath. Our software tool that automatically performs this analysis also computes the geometric properties of the basins of the local extrema: maximum curvature within the basin (16.68 [1/M] for the global maximum, 21.27 [1/M] for the local maximum) and the basin's volume ($6.683 \cdot 10^{-7} \text{ M}^3$ for the global maximum, $8 \cdot 10^{-6} \text{ M}^3$ for the local maximum). Basin is defined at a cutoff equal to 80% of the respective extremum. For reproducing this figure and the calculation of basin's geometry, execute the Python script "*robowski/geometric_algorithms/geometry_tools.py*" in the code repository associated with this article.

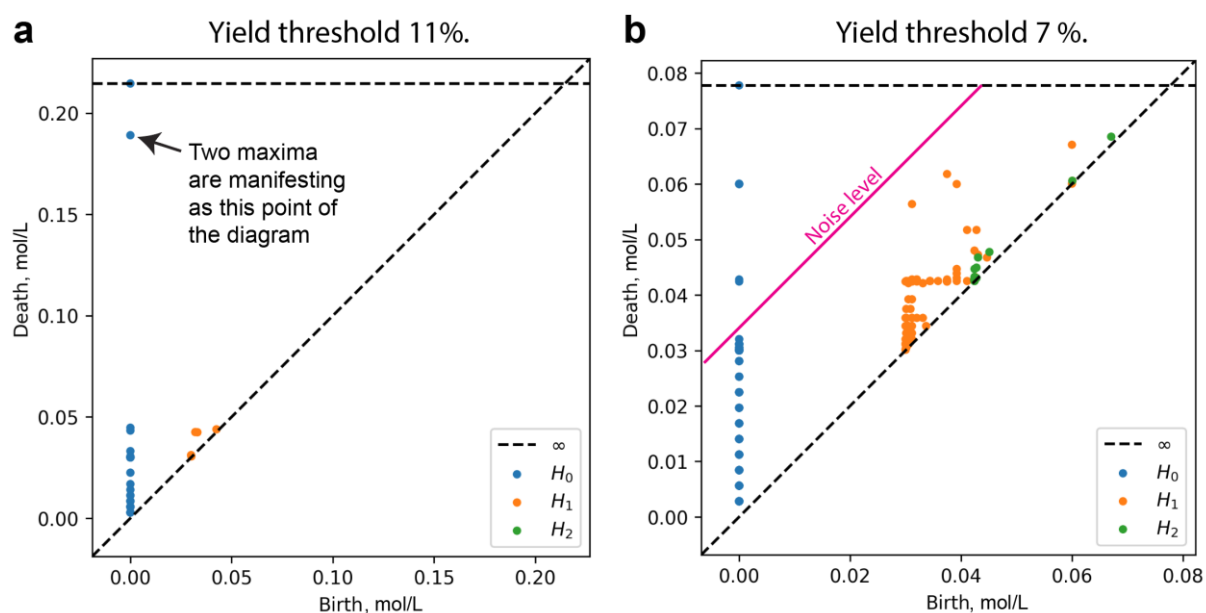


Figure S162. Persistent homology diagrams^{155,156} showing the topological features¹⁵⁷ of the Ugi reaction dataset across different scales. Each colored dot represents a topological feature that appears ("birth") and disappears ("death") as we gradually connect nearby points of the condition space. A given diagram considers only the conditions for which the reaction yield is above the selected yield threshold (11% for **a**, 7% for **b**). The horizontal axis ("Birth"), indicates the scale parameter when a feature first appears, while the vertical axis ("Death") shows when the feature is filled in or merged with another feature. Blue dots (H_0) represent connected components – distinct clusters or groups in the data that form and later merge together. Orange dots (H_1) represent one-dimensional holes or loops – circular structures that appear in the data and later get filled in. Green dots (H_2) represent two-dimensional voids or cavities – enclosed bubble-like structures in higher-dimensional space. The diagonal dashed line indicates where birth equals death; dots close to this line represent short-lived features likely due to noise, while dots far from the diagonal indicate persistent, significant topological structures in the data. The further a point is from the diagonal, the longer that topological feature persists across scales, suggesting it represents a genuine structural property rather than random noise. The presence of two maxima of Ugi reaction yield manifests as the off-diagonal blue dot in **a** (indicated by an arrow). To reproduce these plots, execute the Python script located at "robowski/geometric_algorithms/persistent_homologies.py" in the code repository associated with this article.

6. Quantum chemical calculations

6.1. Quantum thermochemistry of E1 and S_N1 reactions

Quantum-mechanical calculations were used to obtain enthalpy $\Delta H = (37.8 \pm 12.5)$ kJ/mol and entropy $\Delta S = (167 \pm 40)$ J/(mol K) of the E1 reaction from main text **Figure 2a**, as well as enthalpy $\Delta H = (-24.3 \pm 12.5)$ kJ/mol and entropy $\Delta S = (9.6 \pm 40)$ J/(mol K) for the S_N1 reaction from main text **Figure 2b**.

To perform these calculations, we used density-functional theory with spin-component-scaled double-hybrid revDSD-PBEP86 functional¹⁵⁸ with D3(BJ) dispersion term¹⁵⁹ having Becke–Johnson damping¹⁶⁰. Prefix ‘rev’ indicates that the parameters of this functional were revised by Golokesh, Sylvetsky & Martin¹⁵⁸ to achieve the best possible performance on a very large benchmark (GMTKN55¹⁶¹) of various interaction energies and reaction thermochemistry. Relevant to our present study, the weighted mean absolute deviation of the predicted energies from the experimental values of the thermochemistry subset of GMNTEK55 benchmark was reported to be 2.5 kJ/mol¹⁵⁸. However, the GMTKN55 benchmark uses the wavefunction-based calculations as the “ground truth”, so in addition to the 2.5 kJ/mol RMS discrepancy between the double-hybrid approximation and that “ground truth”, there is additional discrepancy between the “ground truth” of quantum-mechanical calculations and the experimentally measured values. Dedicated benchmarks have arrived at roughly the same root-mean squared error when comparing DFT energies to experiments^{162,163}: 3 kcal/mol = 12.5 kJ/mol, and this is the value we are going to use as uncertainty of our ΔH calculations. For uncertainty of ΔS , this error value is divided by the ambient temperature 298.15 K and becomes 40 J/(mol K).

Before running geometry optimization with the revDSD-PBEP86 functional, the lowest energy conformer of each molecule was first optimized with a less accurate but much faster self-consistent tight-binding quantum chemical method (xTB)¹⁶⁴ in its version accounting for second-order density fluctuations (GFN2-xTB)¹⁶⁵. In all our DFT calculations, we used the triple- ζ def2-TZVPP basis set by Weigend & Ahlrichs^{166,167} and the Gaussian16 software, for which the settings (usually called “route”, or “card”) of our calculations were set as follows:

```
# opt freq=noraman dsdpbep86/def2TZVPP
iop(3/125=0079905785,3/78=0429604296,3/76=0310006900,3/74=1004)
# empiricaldispersion=gd3bj iop(3/174=0437700,3/175=-1,3/176=0,3/177=-
1,3/178=5500000)
```

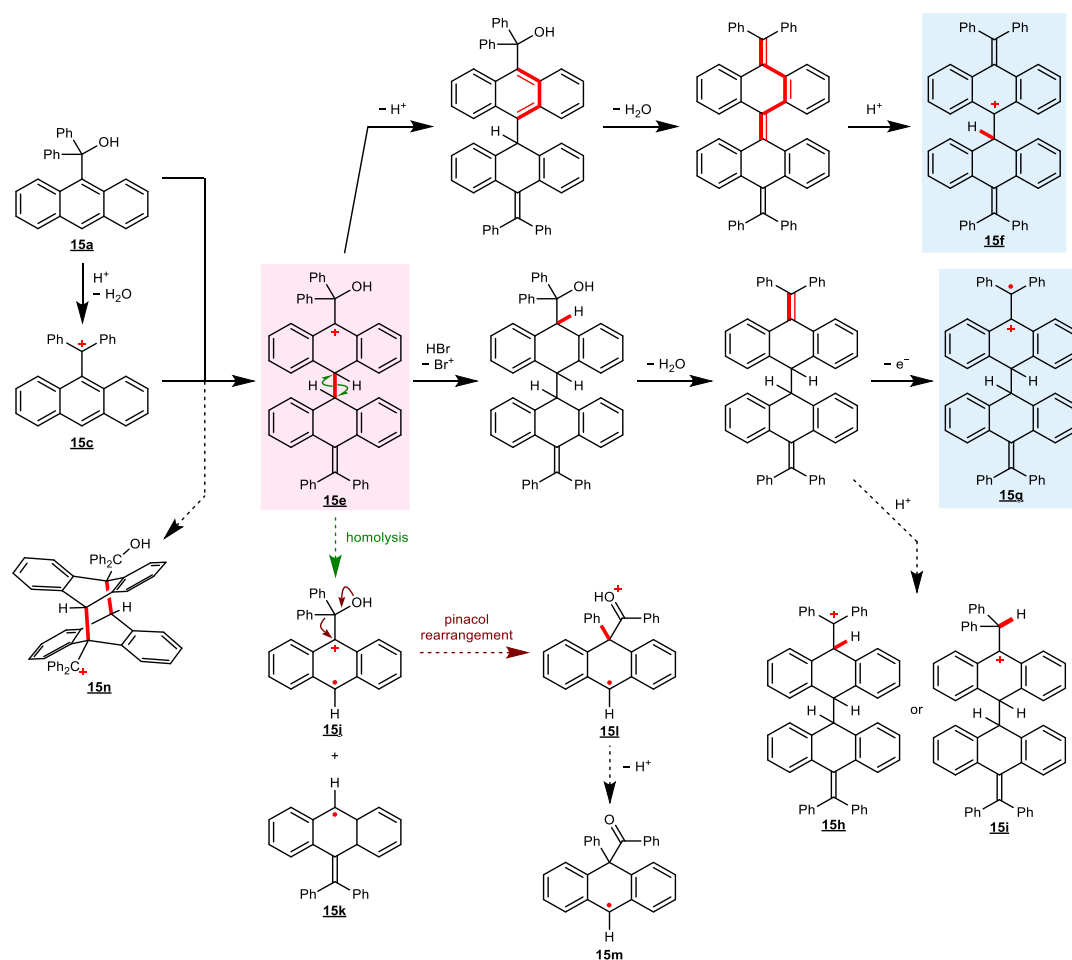
6.2. Theoretical light absorption spectra for identification of product **15e**

The theoretical UV-Vis absorption spectra for all the candidate species related to the identification of the anomalous product **15e** were calculated as follows.

First, we optimized the molecular geometry of the ground state with long-range corrected hybrid density functional ω B97X-D¹⁶⁸ and cc-pVDZ basis set¹⁶⁹. IEF-PCM solvation model¹⁷⁰ for DMF was used to account for solvation. For the optimized geometry, we then performed a time-domain density functional (TD-DFT) calculation with the same functional, basis set, and solvation model to obtain 60 excited states closest to the ground state in terms of energy. It was established in benchmark studies¹⁷¹ that the thus computed transition frequencies should be divided by a functional-dependent correction factor, 1.14 in the case of ω B97X-D functional.

Using the computed oscillator strengths and corrected frequencies for each transition between the ground and an excited state, corrected frequencies of transitions, and the typical spectral linewidth value 0.22 eV reported in the benchmark studies¹⁷¹, we computed the respective spectral band of molar extinction coefficient for every transition using the formulas recommended by Gaussian Inc.¹⁷². Summation of thus computed bands from all the excited states gives the overall spectrum of molar extinction coefficient. For the purposes of comparing it to the experimental absorbance spectrum, the computed molar extinction coefficient spectrum was multiplied by the optical path length (1 cm) of the cuvette used in experiments and by such a molar concentration (indicated in captions of **Figure S164**, **Figure S165**, **Figure S166**) so as to make absorbance comparable in magnitude to the experimental spectrum.

To estimate the confidence intervals of this theoretical calculation, we used the root-mean squared error of TD-DFT calculation of transition frequency, evaluated by Ali *et al.*¹⁷³ to be 0.086 eV. Accordingly, for each species we sample 1000 theoretical spectra, each computed as above but with each transition frequency shifted by a normally distributed random shift with standard deviation 0.086 eV. The grey bands in the figures below correspond to 16% and 84% percentiles (the common “1 σ percentiles”) of absorbance among these 1000 sampled spectra at a given wavelength.



Scheme S9. Mechanistic pathways from alcohol **15a** and HBr toward plausible (radical)-cationic species that may absorb at around 520 nm (causing the pink coloration). The theoretical absorption spectra for species **15e**, **15f**, **15h**, **15i**, **15k**, **15k**, **15l**, **15m** and **15n** were computed by TD-DFT (see below). Species **15f** and **15g** (on blue backgrounds) were observed by APCI-MS (see section 4.13). The reaction between **15a** and its own benzhydrylic carbocation **15c** can give **15e** or the cage **15n**. An overall dehydration of **15e** can yield **15f**. The formation of radical cation **15g**, **15h** and **15i** necessitates the reduction of cation **15e** by HBr.¹⁷⁴ The homolytic cleavage of **15e** (green arrow) can lead to the radical cation **15i** and the radical **15k**, and the pinacol rearrangement of the former (brown arrow) leading to **15l** and **15m** is conceivable. The TD-DFT results indicate that species **15e** (on pink background) is highly probable for explaining the origin of the pink coloration (**Figure S164a**).

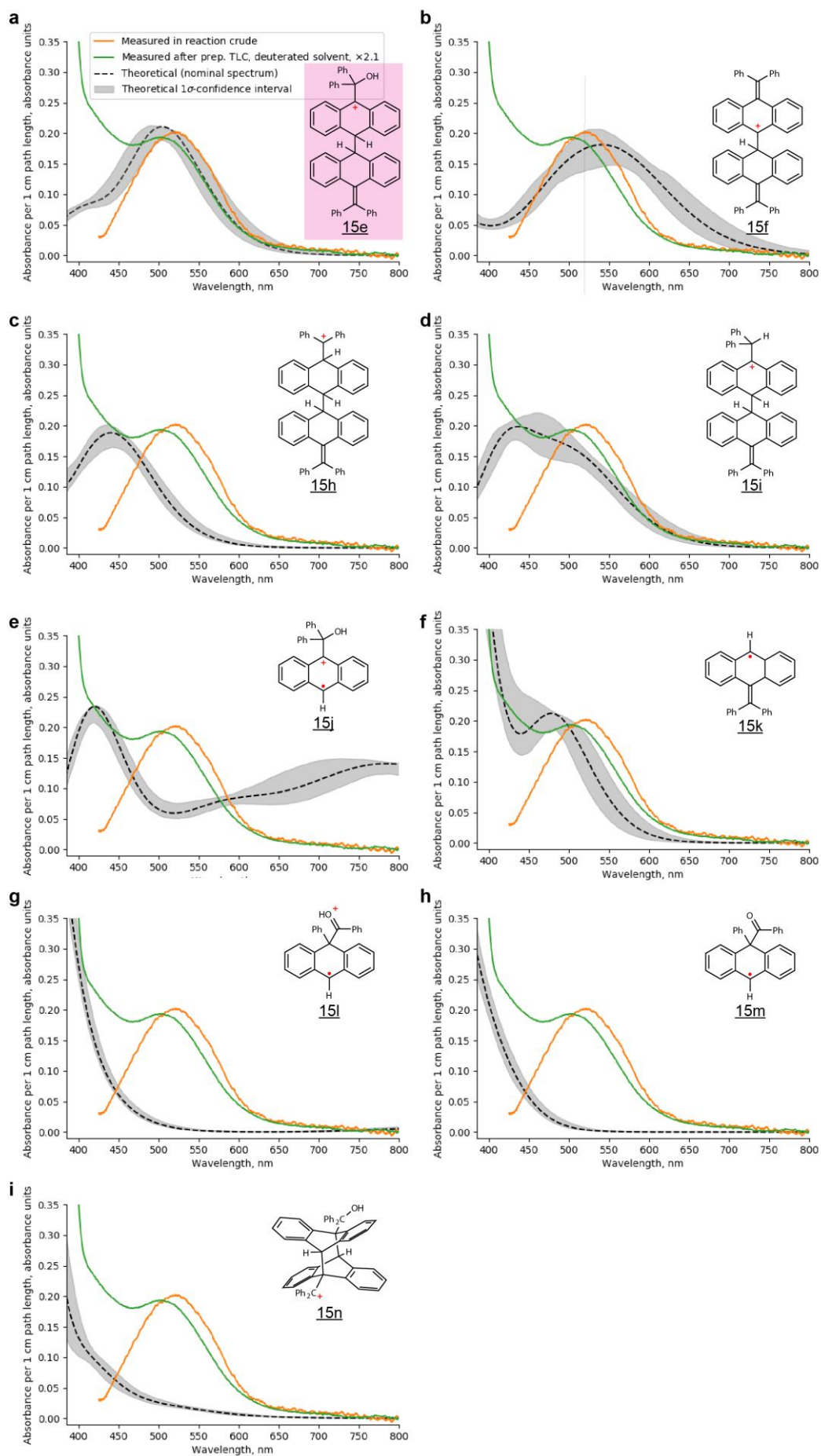


Figure S164. Comparison of visible light absorption spectra of the unexpected pink species detected in a crude reaction mixture (**Fig. 3e**) to the theoretical absorption spectra of various mechanistically relevant species (see **Scheme S9** for the proposed pathways), among which carbocation **15e** (panel A) is the most probable candidate due to the closest resemblance of the experimental and calculated spectra. Orange curve is the absorption band of the unexpected compound measured in the reaction crude: it was obtained by subtracting the spectrum of the reaction crude prepared with conditions not producing the unexpected product (**Fig. 3e**, black dashed curve) from the spectrum of the reaction crude prepared with a condition that resulted in the presence of the absorption band of the unexpected product (**Fig. 3e**, pink curve). The green curve is the absorption spectrum of the unexpected compound separated from the crude by preparative thin layer chromatography (TLC) and dissolved in deuterated solvent. To aid the comparison between the spectra, the absorbance of the latter spectrum was multiplied by a factor 2.1. Black dashed line is the theoretical light absorption spectrum computed using time-domain density functional theory (TD-DFT) for the species whose structure is shown in the upper right corner of the panel. In the chemical structures, radical is indicated by the red dot, whereas the positive charge is indicated by red plus inside a red circle. Grey band is the 68% confidence interval of the theoretical absorption spectrum, whereas the black dashed line is the most likely theoretical spectrum calculations and estimation of their uncertainty. Theoretical absorbance was computed for the optical path length (1 cm) used in the measurements and assuming the following molarities: 1.23 nM (**a**), 1.2 nM (**b**), 1.4 nM (**c**), 1.45 nM (**d**), 2.6 nM (**e**), 4 nM (**f**), 2.5 nM (**g**), 2.5 nM (**h**), 2.5 nM (**i**).

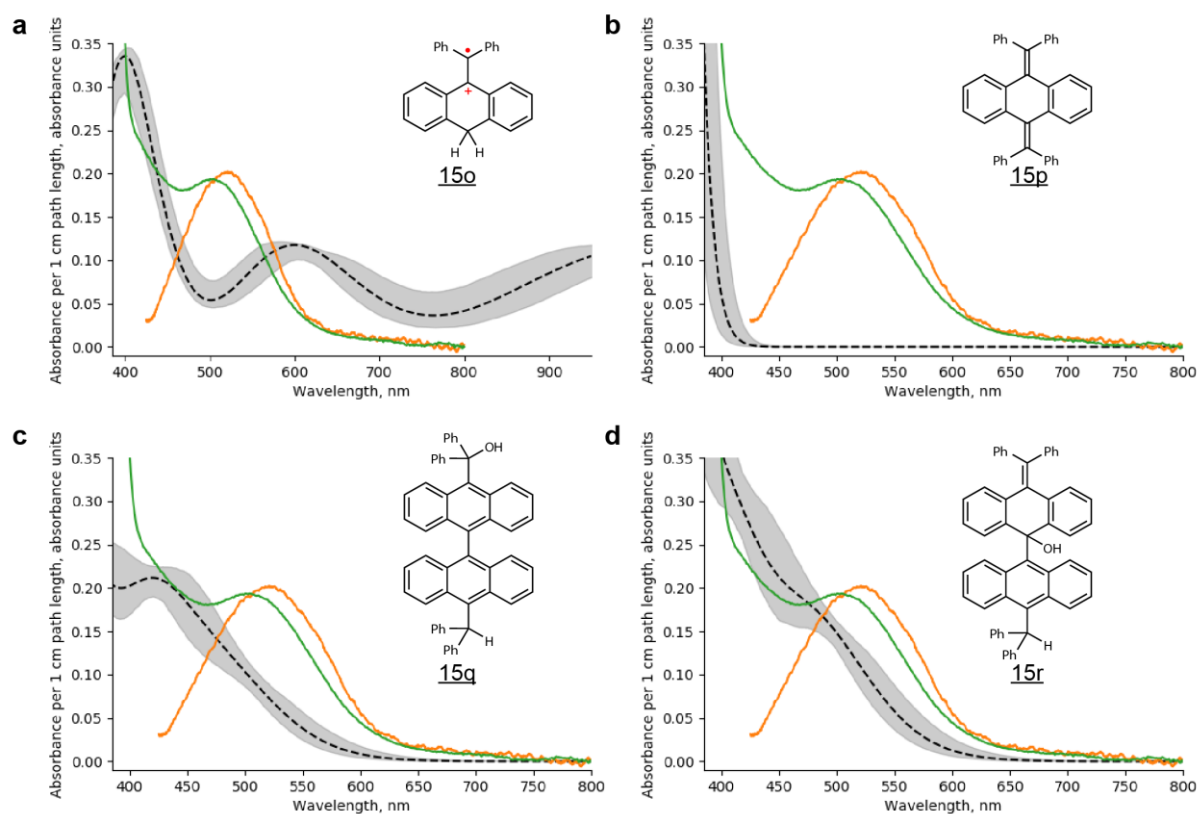


Figure S165. Comparison of visible light absorption spectra of the unexpected pink spices detected in a crude reaction mixture (**Fig. 3e**) to the theoretical absorption spectra of various possible species. Species from panel A, B, C and D are derivatives of the spices from **Figure S164**. Same as **Figure S164**, but showing theoretical spectra for a different set of species. Theoretical absorbance was computed for the optical path length (1 cm) used in the measurements and assuming the following molarities: 1.6 nM (**a**), 100 nM (**b**), 1 nM (**c**), 2.5 nM (**d**).

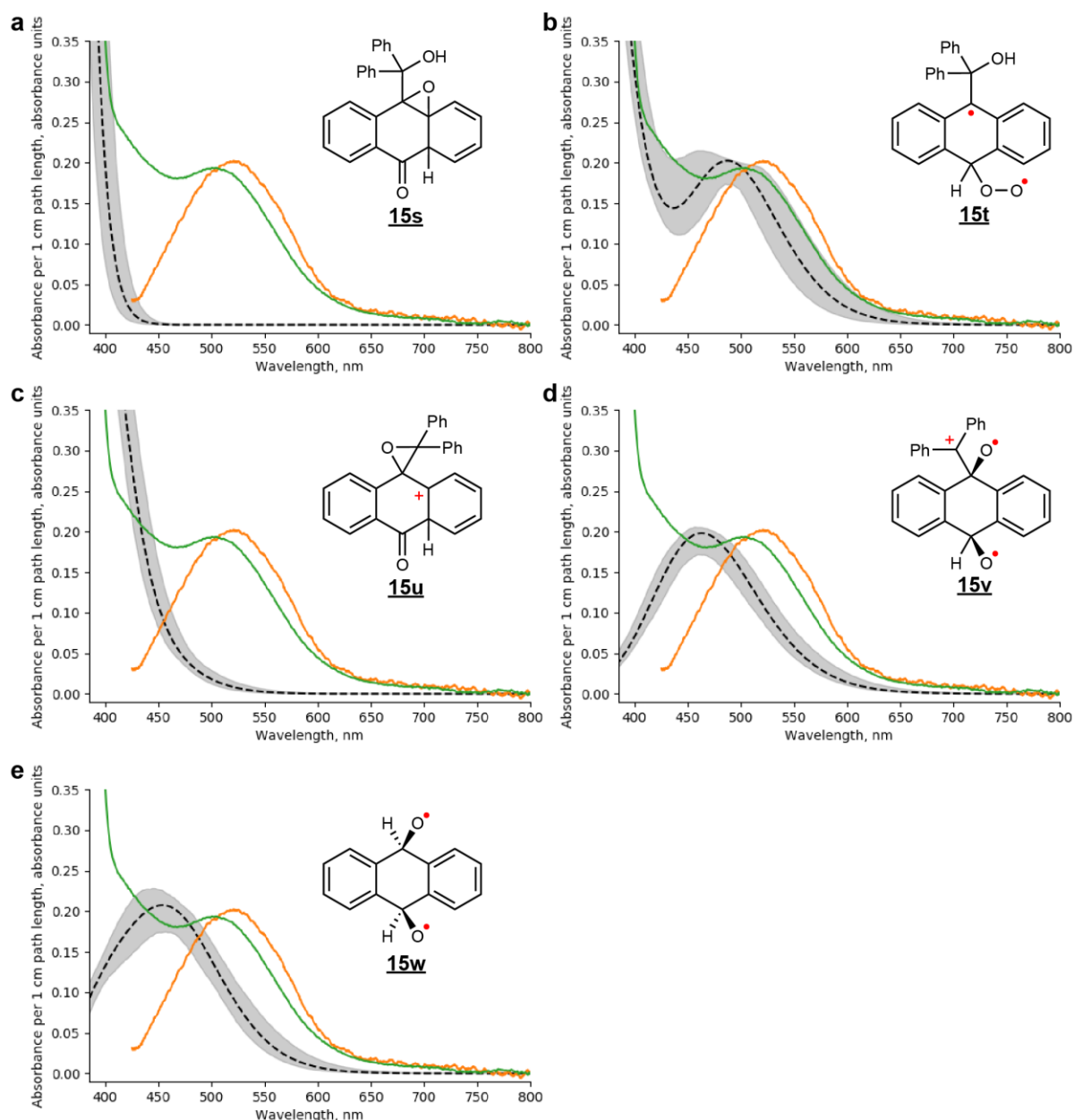


Figure S166. Comparison of visible light absorption spectra of the unexpected pink species detected in a crude reaction mixture (**Fig. 3e**) to the theoretical absorption spectra of various possible species. The photooxygenation of anthracene involves biradical intermediates (such as **15t**, **15v** and **15w**)^{25,175} leading to the formation of endoperoxides^{176,177} and anthraquinone. Compounds **15s** and **15u** are converged structures from plausible biradical species. Same as **Figure S164**, but showing theoretical spectra for a different set of species. Theoretical absorbance was computed for the optical path length (1 cm) used in the measurements and assuming the following molarities: 10 μ M (**a**), 11 nM (**b**), 6 nM (**c**), 1 nM (**d**), 12 nM (**e**).

7. Theoretical perspective on the “slopes” of yield maps

7.1. Overview

As narrated in the main text, the experimental yield distributions over condition-temperature hyperspaces are changing very gradually and, colloquially put, they appear “smooth”. This can be restated by saying that the absolute values of derivatives of the yield with respect to its argument (coordinates of the condition space) are small. There are some caveats on this path, however. The yield is the ratio of final concentration of some substance by the concentration of the limiting reactant. The denominator, the limiting reactant concentration, has a discontinuity of derivative at the “watershed” manifold where one limiting reactant switches to another (dashed line in **Figure S167**), and therefore the derivative of yield does not exist at this manifold. One may say that smoothness of the yield is undefined at the “watershed”, or that yield map is infinitely sharp there. This is one fundamental obstacle to systematically studying the “smoothness” of the yield itself. Another obstacle is that derivatives of the yield (which is dimensionless) by the coordinates (which have dimensions of, e.g. concentrations of reactants) is a dimensional quantity. Hence, its value will not be invariant with respect to the arbitrary choice of concentration units, and therefore such a derivative cannot be used as a measure of “smoothness” with any degree of generality.

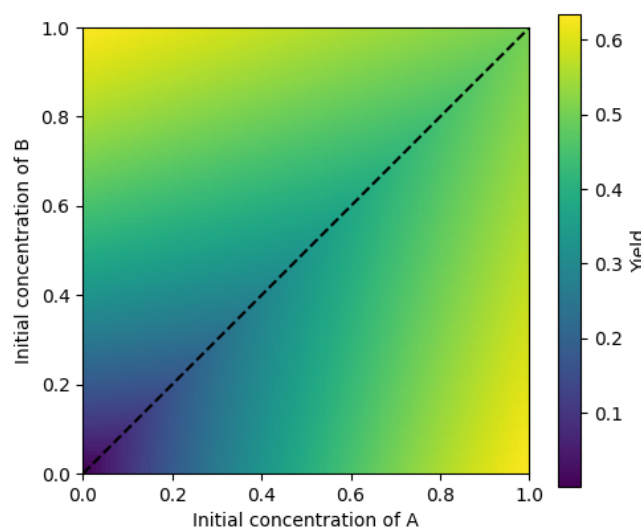


Figure S167. Yield (color scale) of product C with respect of limiting reactant at different starting concentrations $[A]$ (horizontal axis) and $[B]$ (vertical axis) for a reaction $A + B \rightarrow C$ with rate law $dC/dt = [A][B]$ and reaction time $t = 1$. Dashed line is the manifold where limiting reactant switches from A to B .

To overcome these basic problems, we consider the maps of product’s final concentration, rather than maps of yield. The experimental maps of final concentrations also look smooth, but the respective partial derivatives of the final concentrations by the initial concentrations, $D_{ij} = \partial C_i(t)/\partial C_j(0)$, are (i) dimensionless and (ii) do not suffer from the “limiting reactant watershed”

discontinuity because they are not defined in terms of the limiting reactant. This allows us to define “smoothness” via the maximum of the absolute values of these derivatives: the lower the maximum, the smoother the map is. We leave the smoothness with respect to other controlling parameters (temperature, catalyst concentration) outside of the scope of the present study.

First, we evaluate thus defined partial derivatives for our experimental data sets. The highest experimentally observed magnitudes of these derivatives are listed in the third column of the **Table S29**. The histograms of the observed values are shown in **Figure S168**. For all four reactions that could be quantified solely with UV-Vis detection method (i.e., with exception of Hantzsch), the distributions are skewed to the small values and most of the observed derivatives are below or comparable to unity. Note that any experimental noise always increases the apparent derivatives. The highest observed derivative value 22.226 is for E1 reaction and may be caused by outliers, with other high values in the distribution’s tail (around 10 in **Figure S168a**) corresponding to a sharp increase of yield with HBr concentration at very small values of the latter (compare with the bottom-left corners of plots in **Figure S150**).

Table S29. Statistics of experimentally observed derivatives $|D_{ij}| = |\partial C_i(t)/\partial C_j(0)|$ of the final concentration of product with respect to the initial concentrations of substrates. Median is evaluated across all conditions and product-substrate pairs.

Reaction	Data preprocessing	Highest observed derivative	Median of the observed derivatives
E1 (Figure 2a)	Raw data is used	22.226	0.674
S _N 1 (Figure 2b)	Raw data is used	5.734	0.441
S _N 1 with anomaly (Figure 3)	Raw data is used	5.420	0.338
Ugi (Figure 4)	Using denoised data	0.616	0.020

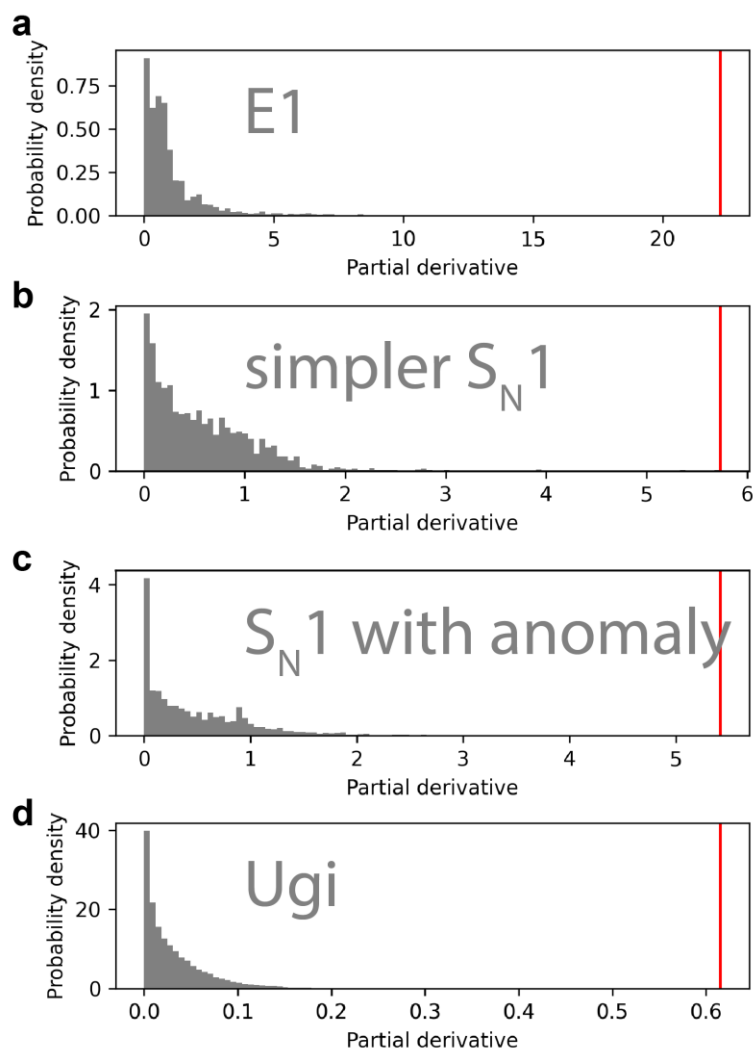


Figure S168. Histograms (across conditions and product-substrate pairs) of experimentally observed absolute values of partial derivatives, $|D_{ij}| = |\partial C_i(t)/\partial C_j(0)|$, of the final concentration of product with respect to the initial concentration of a substrate. The reactions are: (a) E1 reaction depicted in main text **Figure 2a**, (b) S_N1 reaction depicted in main text **Figure 2b**, (c), S_N1 reaction depicted in main text **Figure 3**, (d) Ugi reaction depicted in main text **Figure 4**. Red vertical lines in show the highest observed value of the derivative for a given reaction.

In the following Sections 7.2 and 7.3, we derive two upper bounds for the magnitudes of the partial derivatives of the final concentrations of substances participating in the reaction with respect to the initial concentrations of same or different substances – one bound for the case of first-order kinetic laws, and another – for the case of second-order kinetic laws. We also derive an inequality (46) that can be used to calculate a similar upper bound for any specific kinetic

law (not necessarily of a certain order) and for any kinetic graph. The first two bounds are formulated as the following two theorems.

Theorem 1. (*First-order kinetics*). Let there be n substances potentially present in the reaction mixture. Let all the reactions be unimolecular and obey first-order kinetic laws. Let $C(t)$ be the time-dependent vector of mass concentrations and $\dot{C}(t) = KC$ is the vector of time derivatives of concentrations, where K is the constant kinetic matrix with elements k_{ij} . Then, at time t , the absolute value $|D_{ij}| = |\partial C_i(t)/\partial C_j(0)|$ of partial derivative of the final concentration $C_i(t)$ of any i -th substance by the initial concentration $C_j(0)$ of same or any other (j -th) substance cannot exceed $1 + (e^{tn \max_{i,j} |k_{ij}|} - 1)/n$, where $\max_{i,j} |k_{ij}|$ is the largest absolute value of the kinetic matrix element. If, additionally, the kinetic graph is strongly connected (i.e., corresponds to a mechanism of one reaction rather than two or more independent reactions taking place in reaction mixture), then $|D_{ij}| \leq 1$.

Theorem 2. (*Second-order kinetics*) Let there be n substances potentially present in the reaction mixture, and all the reactions are bimolecular and obey second-order kinetic laws described by kinetic rate constants k_{ijm} such that the mass concentrations C_i evolve in time according to ordinary differential equations

$$\frac{dC_i}{dt} = \frac{1}{2} \sum_{m \neq j} k_{ijm} C_j C_m + \sum_{m=j} k_{ijm} C_j C_m$$

Then, at time t , the absolute value $|D_{ij}|$ of partial derivative of the final concentration of any substance by the initial concentration of same or any other substance cannot exceed $e^{n(n+1)C_{\max}k_{\max}t}$, where $k_{\max}$ is the largest absolute value of the k_{ijm} , and $C_{\max}$ is the largest possible value of any concentration. Note that the stoichiometry of the reactions and mass balance allow one to evaluate $C_{\max}$ based on the highest value of the initial concentration among all substances.

We now proceed to the definitions and proofs.

For kinetic equations

$$\begin{aligned} \frac{dC}{dt} &= F(C) \\ C(t=0) &= C_0 \end{aligned} \tag{27}$$

where C is the time-dependent n -length vector of mass concentrations, and $F(C)$ is the (possibly non-linear) function $\mathbb{R}^n \rightarrow \mathbb{R}^n$ describing the kinetics. The so-called “sensitivity

matrix” $\mathbf{D}$ of partial derivatives of the solution with respect to the initial condition vector C_0 is defined as

$$\mathbf{D} = \frac{\partial C}{\partial C_0} \quad (28)$$

and obeys a time-dependent linear ordinary differential equation that follows from the chain rule and the interchange property of derivative¹⁷⁸

$$\frac{d}{dt} \left(\frac{\partial C}{\partial C_0} \right) = \frac{\partial F}{\partial C}(t, X) \frac{\partial C}{\partial C_0} \quad (29)$$

By definition, matrix $\frac{\partial F}{\partial C}(t, X)$ is the Jacobian matrix $\mathbf{J}_F$ of the function $F(C)$. Assuming no explicit dependence of kinetics $F(C)$ on time t , Jacobian $\mathbf{J}_F$ will not depend on t explicitly, but only on the solution $C(t)$, at which the $\mathbf{D}$ is being evaluated. The equation (29) can be rewritten as

$$\frac{d}{dt} \mathbf{D} = \mathbf{J}_F \mathbf{D} \quad (30)$$

and is accompanied by the initial condition $\mathbf{D}(t = 0) = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix.

Since a given column of $\mathbf{J}_F \mathbf{D}$ depends only on one respective column of $\mathbf{D}$, this expression describes n independent systems of time-dependent linear ODEs – one system for each column of $\mathbf{D}$.

7.2. Case of first-order-kinetics networks

Formalism of reaction networks with first-order kinetic laws discussed in this sections is not only applicable to systems with strictly first-order kinetic laws, but also serves as useful approximation of reaction networks that have non-linear kinetic laws. As discussed by Gorban, Radulescu & Zinoviev¹⁷⁹, the nature of versatility of such linear approximations originates from several sources:

1. Practically important networks with often contain subnetworks whose kinetic laws are of first order.
2. Bimolecular reactions obey pseudo-first order kinetic laws if one of the two reactants is in substantial excess over another, and therefore its concentration does not change appreciably over the course of the reaction.
3. Behavior of a network with complex kinetic laws can sometimes be approximated as sequential switching between networks having pseudo-first order kinetic laws^{180,181}.

If only the first-order reactions are considered, and the stoichiometry is trivial (one molecule of reactant always produces one molecule of the product), the right-hand side $F(C)$ of equation (27) is simply the product of the so-called “kinetic matrix” $\mathbf{K}$ by the vector of concentrations C : $F(C) = \mathbf{K}C$. Elements of the kinetic matrix are comprised of first-order kinetic rate constants. The solution of (2) can be expressed via matrix exponential:

$$C(t) = e^{tK}C(0) \quad (31)$$

and the matrix $\mathbf{D}$ becomes simply $\mathbf{D} = e^{tK}$

For our purposes, we will utilize the so-called “max norm” $\|\cdot\|_{\max}$, which is $\sqrt{mn}$ times the maximum absolute value of the elements a_{ij} of a given $m \times n$ matrix $\mathbf{A}$:

$$\|\mathbf{A}\|_{\max} = \sqrt{mn} \max_{i,j} |a_{ij}| \quad (32)$$

For square, $n \times n$ matrices such as $\mathbf{D} = e^{tK}$, the max norm is $n \max_{i,j} |D_{ij}|$, where D_{ij} are the elements of $\mathbf{D}$. From the definition of the general matrix norm $\|\cdot\|$ and from the fact that max norm of $n \times n$ identity matrix is equal to n , it follows that for any $t > 0$ and for any matrix $\mathbf{K}$,

$$n - 1 + e^{-t\|\mathbf{K}\|_{\max}} \leq \|e^{tK}\|_{\max} \leq n - 1 + e^{t\|\mathbf{K}\|_{\max}} \quad (33)$$

Substituting the max norm definition into (33) and dividing all parts by n produces the following bounds for the absolute values of the elements D_{ij} of the matrix $\mathbf{D}$:

$$1 + \frac{e^{tn \max_{i,j} |k_{ij}|} - 1}{n} \leq \max_{i,j} |D_{ij}| \leq 1 + \frac{e^{tn \max_{i,j} |k_{ij}|} - 1}{n} \quad (34)$$

Practically, common reactions are allowed to run for such a duration t that their yield becomes appreciable, and therefore $t \max_{i,j} |k_{ij}|$ is comparable to unity, unless equilibria with very fast equilibration (and therefore high magnitudes of associated kinetic rates k_{ij}) are present and have strong impact on reactivity. If $t \cdot \max_{i,j} |k_{ij}| \approx 1$, then

$$\max_{i,j} |D_{ij}| \leq 1 + \frac{e^{tn \max_{i,j} |k_{ij}|} - 1}{n} \approx 1 + \frac{e^n - 1}{n} \quad (35)$$

is the upper bound for the derivatives D_{ij} of the final concentration of any reacting substance with respect to the initial concentration of any other (or same) reacting substance. For example, for three substances ($n = 3$) this upper bound is $1 + (e^3 - 1)/3 \approx 7.36$. Notably, at $t = 0$, the matrix $\mathbf{D}$ is the identity matrix, and left and right bounds in (34) become tight: equal to unity, as expected.

The upper bound (34), however, can be improved upon if the kinetic graph is strongly connected (that is, cannot be split into non-interacting subsets of reactions). For strongly connected kinetic graphs and two solutions $C_A(t)$ and $C_B(t)$ computed for different initial conditions, it has been shown^{182–185} that the following expression (“pseudo-Lyapunov function”) monotonously decreases with time:

$$L_{AB}(t) = \sum_{i=1}^n |C_{A,i}(t) - C_{B,i}(t)| \quad (36)$$

Where $C_{A,i}(t)$ and $C_{B,i}(t)$ are the concentrations of i -th substance in solutions $C_A(t)$ and $C_B(t)$, respectively. Then the sum of the absolute values of partial derivatives

$$\sum_{i=1}^n |D_{ij}| = \lim_{C_{A,j}(0) \rightarrow C_{B,j}(0)} \frac{\sum_{i=1}^n |C_{A,i}(t) - C_{B,i}(t)|}{|C_{A,j}(0) - C_{B,j}(0)|} = \lim_{C_{A,j}(0) \rightarrow C_{B,j}(0)} \frac{L_{AB}(t)}{|C_{A,j}(0) - C_{B,j}(0)|}$$

monotonously decreases with time as well. At $t = 0$, matrix $\mathbf{D}$ is the identity matrix and therefore $\sum_{i=1}^n |D_{ij}| = 1$ for any j at $t = 0$. As a consequence of monotonicity, for $t \geq 0$ and any j we have

$$\sum_{i=1}^n |D_{ij}| \leq 1 \quad (37)$$

This also gives us an upper bound on the largest $|D_{ij}|$ at any time $t \geq 0$:

$$|D_{ij}| \leq \sum_{i=1}^n |D_{ij}| \leq 1$$

This proves **Theorem 1**.

7.3. Case of higher-order-kinetics networks

Next, we will employ the following known theorem (p. 132, Theorem 8.2 of Ref. ¹⁸⁶):

Theorem 3: For the linear system $\dot{x} = \mathbf{A}(t)x$, $t \geq t_0$ denote the largest and smallest point-wise eigenvalues of $\mathbf{A}^T(t) + \mathbf{A}(t)$ by $\lambda_{\max}(t)$ and $\lambda_{\min}(t)$. Then for any t_0 and $x(t_0)$ the solution $x(t)$ satisfies

$$\|x(t_0)\|_I e^{\frac{1}{2} \int_{t_0}^t \lambda_{\min}(\tau) d\tau} \leq \|x(t)\|_I \leq \|x(t_0)\|_I e^{\frac{1}{2} \int_{t_0}^t \lambda_{\max}(\tau) d\tau}, t > t_0 \quad (38)$$

where $\|\cdot\|_I$ is the Euclidean matrix norm.

Since the maximum of a function on an interval is greater than or equal to its average value, we get

$$\int_{t_0}^t \lambda_{\max}(\tau) d\tau \leq (t - t_0) \max_{t_0 < \tau < t} \lambda_{\max}(\tau) \quad (39)$$

from which it follows that

$$e^{\frac{1}{2} \int_{t_0}^t \lambda_{\max}(\tau) d\tau} \leq e^{\frac{1}{2}(t-t_0) \max_{t_0 < \tau < t} \lambda_{\max}(\tau)} \quad (40)$$

Using this inequality together with the **Theorem 3**, and taking $t_0 = 0$ and $\mathbf{A} = \mathbf{J}_F^T + \mathbf{J}_F$, we thus obtain an inequality for any given row D_i of the sensitivity matrix $\mathbf{D}$:

$$\|D_i(t)\|_I \leq \|D_i(0)\|_I e^{\frac{1}{2} \int_0^t \lambda_{\max}(\tau) d\tau} \leq \|D_i(0)\|_I e^{\frac{1}{2} t \max_{0 < \tau < t} \lambda_{\max}(\tau)} \quad (41)$$

Obviously, the sensitivity matrix at $t = 0$ is simply equal to the identity matrix, and so the Euclidean norm of any row D_i at $t = 0$ is equal to unity: $\|D_i(0)\|_I = 1$. Hence,

$$\|D_i(t)\|_I \leq e^{\frac{1}{2} t \max_{0 < \tau < t} \lambda_{\max}(\tau)} \quad (42)$$

To obtain an upper bound for $\lambda_{\max}(t)$, as well as for its maximum $\max_{t_0 < \tau < t} \lambda_{\max}(\tau)$ over the duration of the reaction, we recall that any eigenvalue λ of a matrix, say, of $\mathbf{J}_F^T + \mathbf{J}_F$, obeys the following inequality

$$|\lambda| \leq \|\mathbf{J}_F^T + \mathbf{J}_F\| \quad (43)$$

where $\|\cdot\|$ is any sub-multiplicative matrix norm. This inequality follows from the definition of sub-multiplicative matrix norm and the definition of the eigenvalues.

Since $\mathbf{J}_F^T + \mathbf{J}_F$ is an $n \times n$ matrix, application of the “max norm” (32) gives us the following upper bound on the largest eigenvalue of the matrix $\mathbf{J}_F^T + \mathbf{J}_F$

$$\lambda_{\max}(t) \leq \|\mathbf{J}_F^T + \mathbf{J}_F\|_{\max} = n \max_{i,j} \left| \frac{\partial F_i}{\partial C_j} + \frac{\partial F_j}{\partial C_i} \right| \quad (44)$$

Substituting into (42), we arrive at

$$\|D_i(t)\|_I \leq e^{(tn/2) \max_{t_0 < \tau < t} \max_{i,j} \left| \frac{\partial F_i}{\partial C_j} + \frac{\partial F_j}{\partial C_i} \right|} \quad (45)$$

Because the absolute magnitude of any component of a vector is always smaller than vector’s Euclidean norm, inequality (45) gives us an upper bound on the absolute magnitude $|D_{ij}(t)|$ of the elements of the matrix $\mathbf{D}$,

$$|D_{ij}(t)| \leq \max_i \|D_i(t)\|_I \leq e^{(tn/2) \max_{t_0 < \tau < t} \max_{i,j} \left| \frac{\partial F_i}{\partial C_j} + \frac{\partial F_j}{\partial C_i} \right|} \quad (46)$$

Note that there is a second way to arrive at a very similar inequality that is obtained from (42) and (43) by additionally using $\|\mathbf{J}_F^T + \mathbf{J}_F\| \leq 2\|\mathbf{J}_F\|$:

$$|D_{ij}(t)| \leq e^{t \max_{t_0 < \tau < t} \max_C \|\mathbf{J}_F\|} \quad (47)$$

where $\max_C$ of the Jacobian norm is taken over the permitted space of concentration vectors $\mathbf{C}$. Inequality (47) can be derived from the following Theorem 4 (Page 296, Theorem 7.17 in Ref.

¹⁸⁷), which itself is a corollary of the Grönwall's inequality, the Mean Value Theorem, and the definition of Lipschitz constant:

Theorem 4. Let $F(C)$ be defined on an open set U in $\mathbb{R}^n$, and assume that $F(C)$ has Lipschitz constant L in the variables C on U . That is, for all C_1 and C_2 on U ,

$$|F(C_1) - F(C_2)| \leq L|C_1 - C_2|$$

Let $C_1(t)$ and $C_2(t)$ be solutions of (2). Then

$$|C_1(t) - C_2(t)| \leq |C_1(0) - C_2(0)|e^{tL}$$

For illustration of inequality (46), let us consider an example, in which the components of vector function $F(C)$ have the form describing bimolecular reactions with reaction rates proportional to the product of the concentrations of two reacting molecules. In other words, the reaction order is unity with respect to each reactant if the two reacting molecules are different, and equal to two if the two reacting molecules are the same. Namely, the concentration vector obeys the system of ordinary differential equations

$$\frac{dC_i}{dt} = F_i(C) = \frac{1}{2} \sum_{m \neq j} k_{ijm} C_j C_m + \sum_{m=j} k_{ijm} C_j C_m \quad (48)$$

which describes both the reactions that consume the i -th substance (in which case the respective constants k_{ijm} are negative, and i is equal to either j or m) and the reactions that produce the i -th substance (in which case the respective constants k_{ijm} are positive, and $i \neq j, i \neq m$).

Though the following consideration is not used in the proof of Theorem 2, it is worth noting that not all the k_{ijm} elements are independent from each other. Firstly, $k_{ijm} = k_{imj}$. Secondly, due to mass balance, if there is a term $k_{ijm} C_j C_m$ in the expression for i -th component's $F_i(C)$, it describes consumption of j -th substance and m -th substance to produce i -th substance – therefore there will be an opposite-sign contribution ($-k_{ijm} C_j C_m$) to the term $k_{jjm} C_j C_m$ of j -th component's $F_j(C)$ to account for the consumption of the j -th component. In other words, $\sum_i k_{ijm} = 0$ for any j and m .

For thus postulated $F(C)$, we have

$$\frac{\partial F_i}{\partial C_j} = \sum_{m \neq j} k_{ijm} C_m + 2 \sum_{m=j} k_{ijm} C_m \quad (49)$$

In order to apply the inequality (46), we are interested in

$$\left| \frac{\partial F_i}{\partial C_j} + \frac{\partial F_j}{\partial C_i} \right| = \left| \sum_{m \neq j} k_{ijm} C_m + \sum_{m \neq i} k_{jim} C_m + 2 \sum_{m=j} k_{ijm} C_m + 2 \sum_{m=i} k_{jim} C_m \right| \quad (50)$$

The first two sums have $(n - 1)$ terms each, the last two sums have one term each. We introduce the highest absolute kinetic rate constant $k_{\max} = \max_{i,j,m} |k_{ijm}|$ and the maximum concentration $C_{\max} = \max_{t_0 < \tau < t} \max_m C_m(\tau)$, which is across all substances and across time: between the time t_0 (start of the reaction) and the time t (end of the reaction). We obtain an upper bound on (50):

$$\begin{aligned} \max_{t_0 < \tau < t} \max_{i,j} \left| \frac{\partial F_i}{\partial C_j} + \frac{\partial F_j}{\partial C_i} \right| &\leq 2(n - 1)C_{\max}k_{\max} + 4C_{\max}k_{\max} \\ &= 2(n + 1)C_{\max}k_{\max} \end{aligned} \quad (51)$$

And, finally, by substituting (51) into (46), we find that the absolute magnitude of the partial derivative of any final concentration with respect to any initial one is bound as follows:

$$|D_{ij}(t)| \leq e^{n(n+1)C_{\max}k_{\max}t} \quad (52)$$

This proves **Theorem 2**. Assuming simple stoichiometry, where one molecule of a product is produced from one molecule of each of the two substrates, $C_{\max}$ cannot exceed the initial concentration of the most abundant substance.

The upper bound (52) does not seem very tight in the sense that its value is too large to be relevant to our experimental observations: commonly, reactions have appreciable yield by the end of the reaction time, and so the product $C_{\max}k_{\max}t$ is comparable to unity, which translates to the magnitudes of derivatives $|D_{ij}(t)|$ being limited by approximately $e^{3(3+1)} = e^{12} \approx 162754$ for $n = 3$ substances. It is interesting, however, that the theoretical upper bound does exist even for this rich family of kinetic graphs.

We review this bound for kinetic graphs further specialized by the following constraints.

7.4. Case of constrained second-order-kinetics network

Suppose that in a reaction network with second-order rate laws (48) each substance can react with no more than one other substance. We then use the inequality (47) and apply the matrix norm $\|\cdot\|_1$ induced by the vector p -norm with $p = 1$. This norm is the maximum row sum of the absolute values of matrix elements:

$$\|A\|_1 = \max_j \sum_i |a_{ij}|$$

The partial derivative $\frac{\partial F_i}{\partial C_j}$ is negative if and only if i -th substance reacts with j -th substance. But by construction, in this section, for every j there is only one value of i for which this is true – let's denote this index value as $i_{<0}(j)$ and therefore there will be at most one negative $\frac{\partial F_i}{\partial C_j}$ for

a given j . Additionally, due to mass balance in the system, the sum of concentrations never changes with time:

$$\sum_i \frac{dC_i}{dt} = 0$$

As a consequence,

$$\sum_i \frac{\partial F_i}{\partial C_j} = \frac{\partial}{\partial C_j} \left(\sum_i F_i \right) = \frac{\partial}{\partial C_j} \left(\sum_i \frac{dC_i}{dt} \right) = 0$$

On the other hand, the sum $\sum_i \left| \frac{\partial F_i}{\partial C_j} \right|$ differs from the sum $\sum_i \frac{\partial F_i}{\partial C_j}$ by twice the sum of the negative terms (and there is, at most, one negative term, as we discussed above):

$$\sum_i \left| \frac{\partial F_i}{\partial C_j} \right| \leq \sum_i \frac{\partial F_i}{\partial C_j} + 2 \left| \frac{\partial F_{i < 0(j)}}{\partial C_j} \right| = 2 \left| \frac{\partial F_{i < 0(j)}}{\partial C_j} \right|$$

Now we have two possibilities. If $i_{<0}(j) = j$, then it follows from the formula (49) that

$$\left| \frac{\partial F_{i < 0(j)}}{\partial C_j} \right| = |2k_{jjj}C_j| \leq 2k_{\max}C_{\max}$$

Otherwise,

$$\left| \frac{\partial F_{i < 0(j)}}{\partial C_j} \right| = |k_{ij i_{<0}(j)} C_{j i_{<0}(j)}| \leq k_{\max}C_{\max}$$

In both cases,

$$\sum_i \left| \frac{\partial F_i}{\partial C_j} \right| \leq 2 \left| \frac{\partial F_{i < 0(j)}}{\partial C_j} \right| \leq 4k_{\max}C_{\max}$$

And therefore $\|J_F\|_1 \leq 4k_{\max}C_{\max}$. After substituting it into the inequality (47), we obtain the final upper bound:

$$|D_{ij}(t)| \leq e^{4C_{\max}k_{\max}t}$$

In practice, if a reaction has appreciable yield by the time t , then $C_{\max}k_{\max}t \approx 1$, and we get $|D_{ij}(t)| \leq e^4 \approx 54.6$.

Upper bounds for higher kinetic orders and more complicated kinetic laws can be obtained similarly by substituting the appropriate function $F(C)$ into inequality (46). In cases with more *a priori* knowledge about the structure of the Jacobian matrix of $F(C)$, tighter bounds may perhaps be obtained by using variants of the matrix norm other than the “max norm” we used here for generality. If there are further *a priori* constraints on the time evolution of the Jacobian matrix, then omission of the inequality (40) may lead to tighter bounds still.

7.5. Numerical exploration of reaction networks' smoothness

To roughly gauge how tight are the analytical bounds described in the preceding sections, we generated reaction networks randomly in terms of both connectivity and rate constants, integrated each network numerically, and computed all the partial derivatives D_{ij} . We were also curious how the maximum $|D_{ij}|$ depends on the topological properties of the network's graph and on time of the reaction. We did not intend this brief numerical exploration to be exhaustive in any sense, but to give us a rough idea of what orders of magnitudes of $|D_{ij}|$ are possible in a few specific families of reaction networks.

Though we do plot histograms of the highest (over $\{i, j\}$ pairs, time, and initial conditions) $|D_{ij}|$ sampled over randomly generated reaction systems, it should be noted that the respective probability densities are not guaranteed to have a meaning except in the context of a specific algorithm we use for generating the random kinetic networks and sampling the random kinetic rate constants. In other words, presented histograms of the highest $|D_{ij}|$ may or may not reflect the probabilities of encountering certain $|D_{ij}|$ in a "random" reaction system in the real world. Like in the famous Bertrand paradox, there is no clear way to define probability on the space of reaction networks.

Unless otherwise noted, the kinetic rate constant for each individual reaction is a random real number sampled from a uniform probability distribution on the interval $[0.001, 1]$. This in itself does not constrain the constants k_{ijm} to the same interval, because a given k_{ijm} may contain contributions from more than one reaction. All the elements k_{ijm} and the elements of the matrix $\mathbf{K}$ (which is responsible for first order reactions) were normalized by the highest absolute value of their elements (maximum is taken over k_{ijm} and $\mathbf{K}$ together). This normalization is equivalent to rescaling the time interval of the numerical integration and allows us to fix said time interval to $[0, 10^{3.5}]$ and be sure that the final steady state is reached by the time $t = 10^{3.5}$. At the same time, we can be sure that no kinetic rate becomes unexpectedly high and the minimum time step (10^{-6}) of the integration algorithm is sufficient to treat the highest kinetic rate in the system. Integration of ordinary differential equations was performed using the explicit Runge-Kutta method of order 5(4) (Ref. ¹⁸⁸) with first step 10^{-6} time units, maximum step 100 time units, absolute tolerance 10^{-8} , relative tolerance 10^{-7} . Concentrations were evaluated at 200 time points spaced evenly on a log scale between 10^{-3} and $10^{3.5}$ time units. Partial derivatives D_{ij} were evaluated for a given initial concentrations vector using two-point finite-difference scheme with the relative step equal to 0.001. One hundred sets of initial concentrations were tried for every kinetic graph or every set of kinetic rate constants, and the highest $|D_{ij}|$ across all the tried initial concentration sets, all time points, and all $\{i, j\}$ pairs was recorded. Initial concentrations were first sampled from a uniform probability distribution on $[0.01, 1]$ and then divided by their sum. This way, the sum of initial concentrations is equal to

unity (and, obviously, the sum is conserved at all times). This normalization is not repeated after one of the initial concentrations is incremented in the two-point finite-difference scheme of computing D_{ij} . Therefore, the sum of concentrations within the finite-difference scheme can be higher than 1, but does not exceed 1.001, because the relative step of the differentiation scheme is 0.001. The lower limit 0.01 of initial concentrations was chosen with two nuances in mind: first, if one of the two reactants is in a large excess, a second-order reaction behaves as a pseudo-first order reaction, so there is no point in modelling it as second-order reaction; second, small steps in finite-difference scheme result in large numerical errors of $|D_{ij}|$. The GitHub repository released along with this article contains all Python code we used for generation of reaction network, integrating them in time and analyzing the resulting $|D_{ij}|$.

The properties that we optionally enforce when generating reaction systems are as follows. The first restriction that we optionally enforce is that a substance can be a reactant in no more than one bimolecular reaction. This condition (“single use”) corresponds to the first column of the **Table S30** and **Table S31**. In our algorithm for generating a reaction system, the reactions are added one by one until either their total number is 20 or until it is impossible to satisfy the requirements (e.g., if “single use” is enforced but there are no more unused reactants left in the system).

Two nodes A and B of the kinetic graph are said to be “connected in a weakly ergodic way”¹⁷⁹ if there exists a node that can be reached both from A and B by following the reactions’ arrows. Graphs is said to be “weakly ergodic”¹⁷⁹ if every pair of its nodes is connected in a weakly ergodic way. When we attempt to randomly generate a weakly ergodic graphs in cases A3-A5 described below, the second-order reactions are added to the system one by one, and each next reaction preferably chooses a pair of reactants that are not yet connected in a weakly ergodic way (if such disconnected pairs are present). If the graph is already weakly ergodic, two substrates for each next reaction are chosen at random. After 20 reactions are added, we check whether the resulting kinetic graph turned out to be weakly ergodic.

Cycles in the kinetic graph can be prevented by disallowing the products of each next added reaction to be one of the “graph ancestors” of this reaction’s two substrates. Finally, the number of products of each bimolecular reaction can be restricted. If it is not restricted, the number of products of each reaction is chosen at random from 1 to the maximum possible number of products (which depends on whether we are disallowing the cycles, for instance). In our calculations, we keep all molar masses equal to 1 for simplicity, and therefore in a given reaction the total consumed mass of the reactants is distributed evenly among the masses of all the products.

The first group of cases (**Table S30**) are systems containing 10 substances and up to 20 bimolecular second-order reactions chosen at random. When cycles are disallowed (**Cases A1-A3**), the highest derivatives are below 10, and the median (over random reaction systems) of the highest derivative is between 2.2 and 2.8. When we allow the cycles (**Cases A4, A5**), the

highest derivatives increase to 47.2 in “single use” case, and to 3051 if we allow multiple reactions per reactant. Peculiarly, in the latter case **A5 (Figure S17e)** the median stays around 1. We do not understand why this is the case, but maybe it is because allowing multiple reactions per reactant often produces a tightly connected graph, where most pairs of substances are connected by a reaction.

Table S30. Properties and the resulting $|D_{ij}|$ for randomly generated kinetic graphs containing only second-order bimolecular reactions (Cases A1-A5).

Case	No more than one reaction per reactant (“single use”)	Weakly ergodic	Cycles are allowed	Number of products of each reaction	Maximum of $ D_{ij} $ over networks and initial conditions	Median (over networks) of maximum $ D_{ij} $ over initial conditions
A1	Yes	No	No	Unrestricted	8.940	2.780
A2	No	No	No	Unrestricted	8.021	2.574
A3	No	Yes	No	Unrestricted	7.734	2.260
A4	Yes	Yes	Yes	Unrestricted	47.249	2.647
A5	No	Yes	Yes	Unrestricted	3051.711	1.000

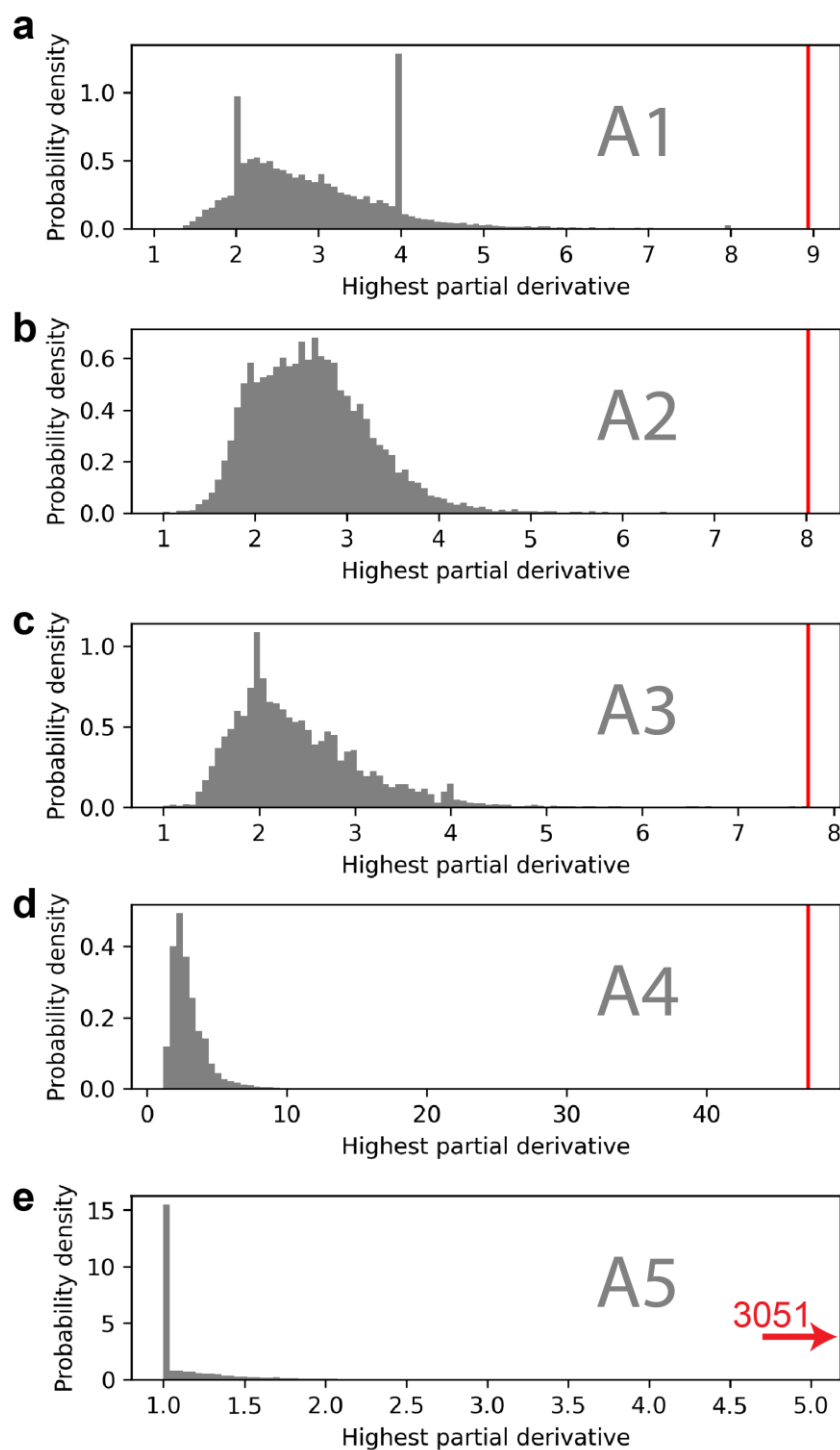


Figure S169. Histograms of randomly chosen reaction systems (Cases A1-A5, **Table S30**) in terms of the highest (across time, initial conditions, and substance pairs $\{i, j\}$) absolute value of partial derivatives, $|D_{ij}|$. Red vertical lines in **a-d** show the highest observed value of $|D_{ij}|$. In **e**, the highest value is far beyond the right edge of the histogram, because we wanted to show the distribution centered around unity more clearly.

The second group of cases, B1-B5 (**Table S31**) are reaction networks containing 10 substances and up to 20 reversible bimolecular second-order reactions chosen at random. What differentiates these cases from A1-A5 is that all reactions are reversible and have either one or two products, never more. When a forward reaction has one product, the corresponding reverse reaction obeys first-order kinetics: the reverse reaction rate is proportional to the concentration of the product. When a forward reaction has two products, the corresponding reverse reaction obeys second-order kinetics: the reverse reaction rate is proportional to the product of concentrations of the two products. Due to reversibility of reactions, the respective graph always has simple cycles of length 2 and even 4; what we mean in cases B1-B5 by “cycles are allowed” is that each next reaction being added to the system is allowed to randomly pick its products even from the ancestors of its reactants. When we attempted to generate reaction networks without “single use” property, they all turned out to be weakly ergodic – for this reason, there is no case listed in **Table S31** with “No” in both the first and the second column.

While we do see a few random networks having $|D_{ij}| > 1$, for the vast majority $|D_{ij}|$ is close to unity, as can be seen from the median (across reaction networks) listed in the last column of **Table S31** and histograms in **Figure S170**.

Table S31. Properties and the resulting $|D_{ij}|$ for randomly generated kinetic graphs containing reversible bimolecular forward reactions with either 1 or 2 products each: forward reactions are second order, reverse reactions are first- or second-order depending on the number of products (Cases B1-B5).

Case	No more than one reaction per reactant (“single use”)	Weakly ergodic	Cycles are allowed	Number of products of each reaction	Maximum of $ D_{ij} $ across networks and initial conditions	Median (across networks) of maximum $ D_{ij} $ across initial conditions
B1	Yes	No	No	1 or 2	3.648	1.000
B2	Yes	Yes	No	1 or 2	3.675	1.000
B3	No	Yes	No	1 or 2	2.691	1.000
B4	Yes	Yes	Yes	1 or 2	3.753	1.000
B5	No	Yes	Yes	1 or 2	73.201	1.000

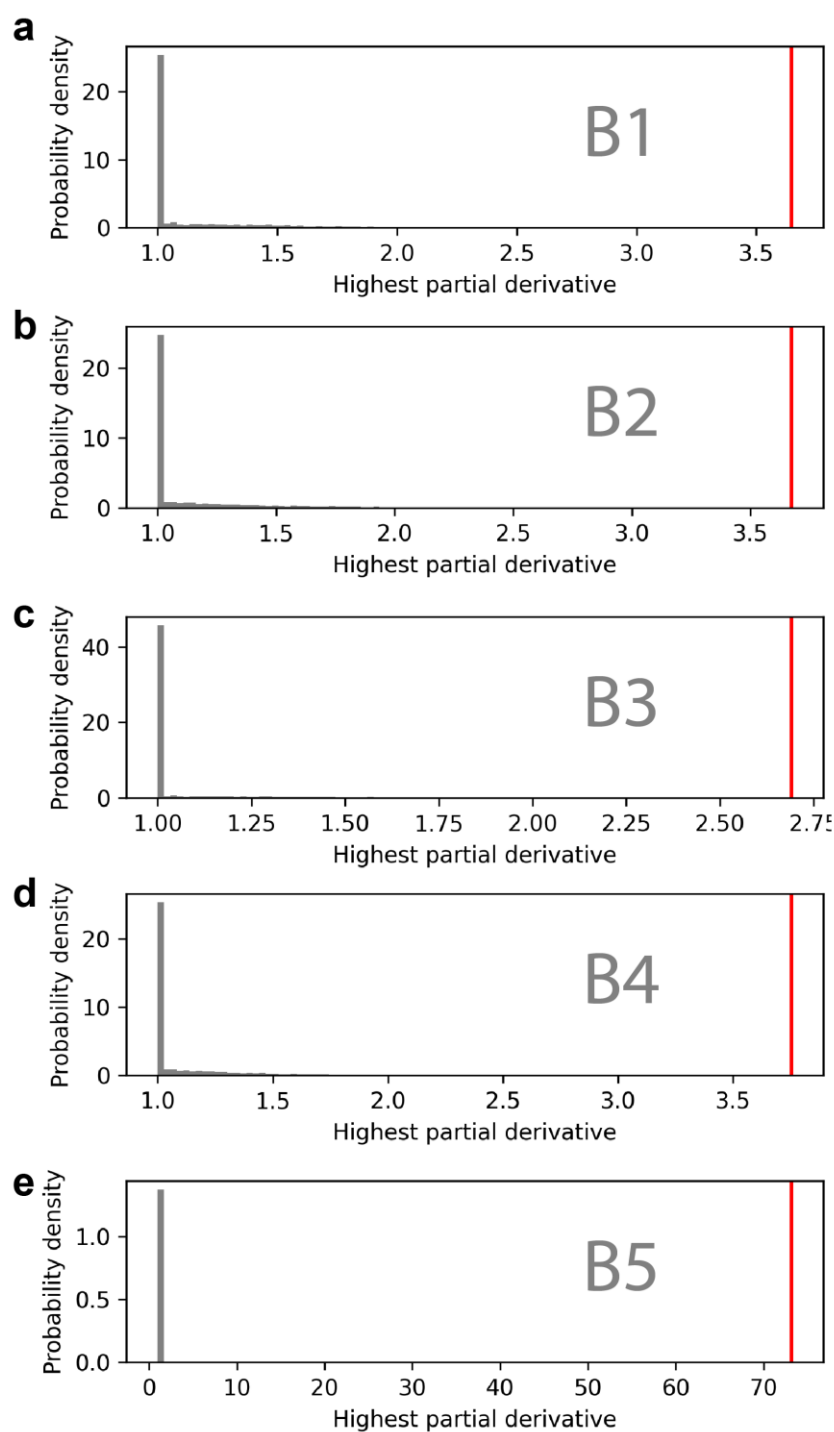


Figure S170. Histograms of randomly chosen reaction systems (Cases B1-B5, **Table S31**) in terms of the highest (across time, initial conditions, and substance pairs $\{i, j\}$) absolute value of partial derivatives, $|D_{ij}|$. Red vertical lines show the highest observed value of $|D_{ij}|$. Most of the probability density is concentrated in the first bin and only a small “tail” is stretching to the right.

Case C1-C3 (Figure S171). Two second-order reactions in sequence, accompanied by first-order side reactions. All these kinetic graphs are weakly ergodic and strongly connected because all the reactions are reversible.

Case C4: Second-order reaction is followed by first-order reaction, then by another second-order reaction. All side reactions are first-order.

The maximum values of $|D_{ij}|$ in **Table S32** deviate from unity only because of finite numerical precision: analytically, $|D_{ii}| = 1$ at $t = 0$. Hence, $|D_{ij}|$ is bound by 1 in cases **C1 – C5**.

We suspected that the apparent limit $|D_{ij}| \leq 1$ for these cases could have been due to first-order reactions overwhelming the effect of second-order reactions on the overall network dynamics, but this suspicion proved to be wrong: we reran the calculations of **Case C1** after slowing the first-order reactions ten (**Case C2**) or hundred times (**Case C3**) and it turned out that $|D_{ij}| \leq 1$ even with slowed-down first-order reactions.

Table S32. Reaction systems with specific kinetic graphs (**Figure S171**) but randomly chosen kinetic rate constants and initial concentrations.

Case	Graph	Range of first-order kinetic constants	Maximum of $ D_{ij} $ over networks and initial conditions	Median (over networks) of maximum $ D_{ij} $ over initial conditions
C1	Figure S171a	$[10^{-3}, 1]$	0.9999987	0.9997661
C2	Figure S171a	$[10^{-4}, 0.1]$	0.99999985	0.99996595
C3	Figure S171a	$[10^{-5}, 0.01]$	0.99999998	0.99999450
C4	Figure S171b	$[10^{-3}, 1]$	0.99999878	0.99982742
C5	Figure S171c	$[10^{-3}, 1]$	0.99999930	0.99987626

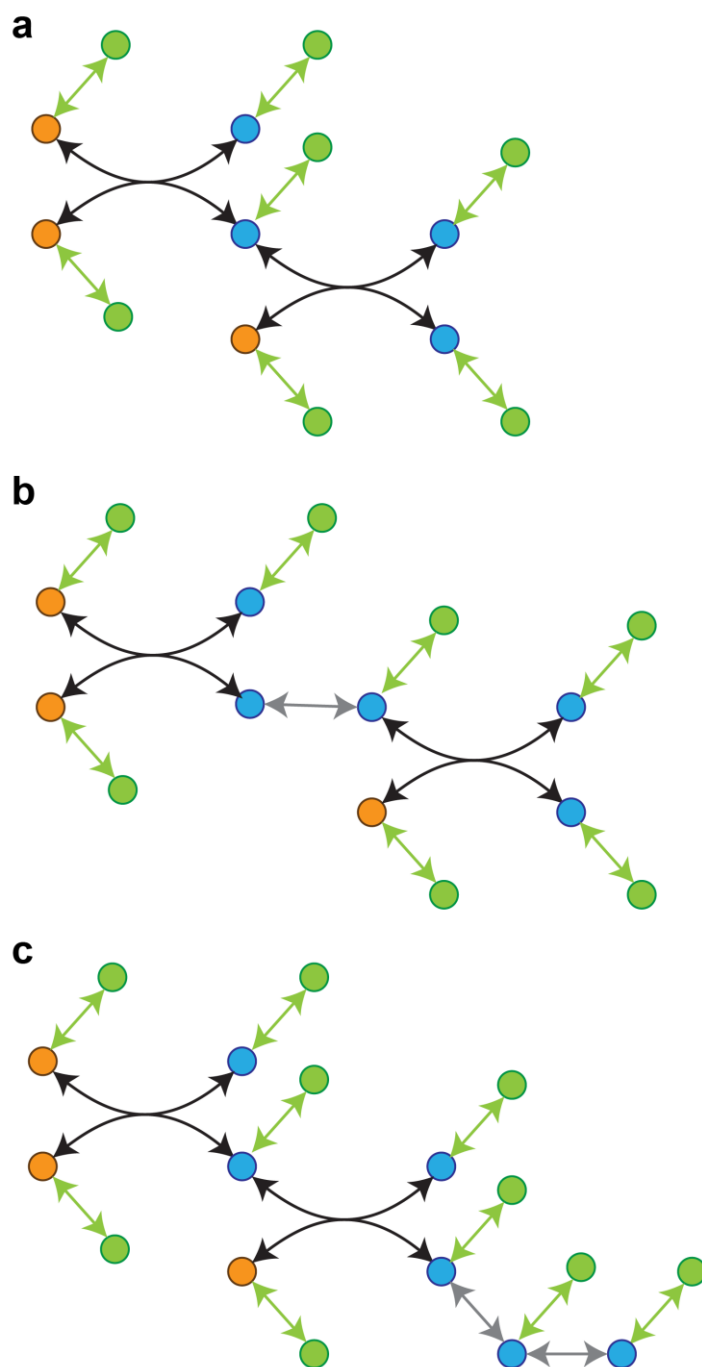


Figure S171. Kinetic graphs tested for their highest derivatives $|D_{ij}|$. Respective case labels are C1-C3 (panel **a**), C4 (panel **b**), C5 (panel **c**). Reversible second-order reactions are shown by *black arrows*. Reversible first-order reactions of the “main” pathway are shown by *grey arrows*. Additional side products (*green circles*) are produced by reversible first-order reactions (*green arrows*). We evaluate partial derivatives of final concentrations of all substances by the starting concentrations of substrates only (*orange circles*). Starting concentrations of species other than the substrates are set to zero.

7.6. Ramifications of limited slopes for the experimental sampling strategies

A practically important consequence of the $|D_{ij}|$ “slopes” being small is for the conditions’ screening and optimization campaigns – namely, hypersurfaces along k concentration variables $C_{0,j}, j = 1..k$, can be probed at sparse intervals. In particular, the property $|D_{ij}| \leq 1$ makes the function $C_i(C_{0,j})$ Lipschitz-continuous with Lipschitz constant $L = 1$, a class of functions for which a dedicated theory of sampling has been developed^{189–192}.

The mathematical theory of Lipschitz function approximation, pioneered by Sukharev¹⁸⁹ and later developed by Cooper^{189,190} and others¹⁹¹, provides rigorous bounds on how densely we need to sample such functions to achieve a desired approximation accuracy.

One possibility is to sample one’s data on a regular k -dimensional hypercubic grid in the condition space. In this case, it has been shown^{189,190} that for a Lipschitz-continuous function with constant L , to guarantee that a piecewise-linear interpolator deviates by no more than ϵ from the true function value at any point, it is sufficient to sample on a regular grid with spacing:

$$\Delta x = \frac{2\epsilon}{L\sqrt{k}}$$

In our chemical context, $L = 1$ and this means that if we want to predict product concentrations with accuracy ΔC_i , we need to sample initial substrate concentrations with spacing:

$$\Delta C_{0,j} = \frac{2\Delta C_i}{\sqrt{k}}$$

This result provides a quantitative relationship between experimental accuracy requirements and the minimum sampling density needed to achieve that accuracy.

For a concrete example, consider 2D concentration-concentration slices ($k = 2$) of the 3D space in **Figure 2a,c**. Suppose we want to predict product concentrations with accuracy $\Delta C_i = 2$ mM across the entire space. The theory tells us we need to try concentrations $C_{0,13a}$ of substrate **13a** at intervals of:

$$\Delta C_{0,13a} = \frac{2 \cdot 2 \text{ mM}}{\sqrt{2}} \approx 2.83 \text{ mM}$$

Since the substrate concentration range is 0-15 mM, this requires

$$\frac{15 \text{ mM}}{2.83 \text{ mM}} \approx 5 \text{ points}$$

In other words, sampling at 6 evenly spaced concentrations of substrate **13a** (including endpoints) would suffice for sampling on a regular grid.

While regular grids are conceptually simple and trivial to implement, sampling in more complex k -dimensional patterns can achieve the same accuracy with even fewer experiments. A pattern of points in k -dimensional space such that the distance from any point of space to the closest point (node) of the pattern does not exceed a constant ϵ is called “ k -dimensional ϵ -covering”, and ϵ is called “radius” of that covering. A k -dimensional ϵ -covering that has the least possible density of its points among all possible k -dimensional ϵ -covering with the same radius ϵ is called “*optimal* k -dimensional ϵ -covering”. Geometries of optimal k -dimensional ϵ -coverings are known and proven: for example, in 2D space, it is a hexagonal lattice, and in 3D space it is a body-centered cubic (BCC) lattice^{193,194}. In the context of Lipschitz-continuous functions, sampling the substrate concentrations at optimal k -dimensional ϵ -coverings ensures that no point of condition hyperspace is further than distance ϵ from the nearest sample. Sukharev¹⁸⁹ has shown that in this case the approximation error of Lipschitz-continuous functions does not exceed the radius ϵ of the covering.

In other words, sampling substrate concentration space at optimal k -dimensional ϵ -covering of radius ΔC_i offers the same accuracy ΔC_i of product concentration but requires fewer samples than for regular grids discussed above (e.g. 23% fewer samples with honeycomb lattice in 2D, 46% fewer samples with BCC lattice in 3D^{193,194}).

The number of samples required to achieve a predefined accuracy can be further reduced by relying on data from already sampled conditions to optimally choose each next condition to sample. For instance, Zabinsky *et al.*¹⁹¹ developed an optimal algorithm for such a sequential sampling, which, broadly speaking, consists of the following steps:

1. Start with a minimal initial sampling.
2. Construct confidence bounds based on Lipschitz-continuity constraints.
3. Identify regions of highest uncertainty.
4. Select the next sampling point to maximally reduce this uncertainty.

We conclude this overview of Lipschitz-continuous function sampling strategies with a reminder that, while these strategies are optimal for choosing the starting concentrations of reactants, the final concentrations may not be similarly Lipschitz-continuous with respect to condition parameters other than starting concentrations (e.g., temperature or pH).

8. Experimental section for the PBA-catalyzed styrene epoxidation.

8.1. PBA particle size distribution

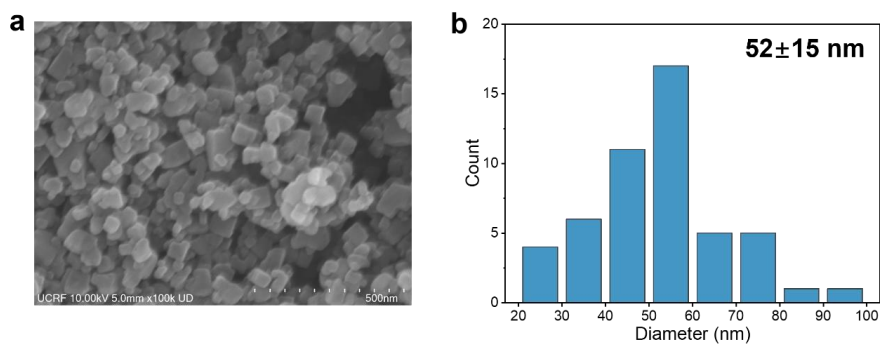


Figure S172. Typical size distribution of the PBAs. **a.** SEM image of the $\text{Mn}_{0.2}\text{Cu}_{0.8}\text{-Fe}_{0.4}\text{Co}_{0.6}$ PBA. **b.** Histogram of this PBA size distribution evidences average diameter of 52 ± 15 nm.

8.2. Grid generation and t-SNE for visualization

Based on previous reports^{44,45} and on commercial availability, two types of Metal A (Fe, Co) and five types of Metal B (Mn, Fe, Co, Ni, Cu) were selected to generate PBAs, accordingly denoted as 5+2 PBAs. To obtain a reasonably sized hyperspace ($\sim 10^3$ PBAs), a step size of 0.2 molar fraction was chosen for generating a uniform grid, in total spanning 756 different PBAs. **Figure S173** illustrates the visualization of the chemical composition space for the 5+2 PBA family. The compositional subspace of metal B, represented as a 4-dimensional hyperplane within a 5-dimensional space, is projected onto a two-dimensional plane using t-SNE for visualization purposes. The compositional subspace of metal A is incorporated as an additional dimension, along the vertical axis. Together, these form a three-dimensional representation—referred to as the PBA composition hyperspace—capturing the integrated compositional landscape of the system. In **Figure S173**, the highlighted plane stands for one slice of metal B composition subspace with fixed composition of metal A ($\text{Fe}_{0.6}\text{Co}_{0.4}$); individual spheres mark individual PBA compositions sampled and experimentally tested. The green sphere is one example of the multimetallic PBA, here with composition $\text{Mn}_{0.2}\text{Ni}_{0.4}\text{Cu}_{0.4}\text{-Fe}_{0.6}\text{Co}_{0.4}$.

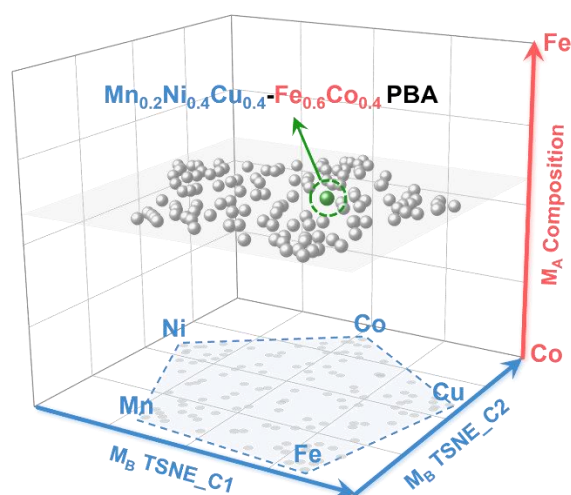


Figure S173. Illustration of the composition space of 5+2 PBAs.

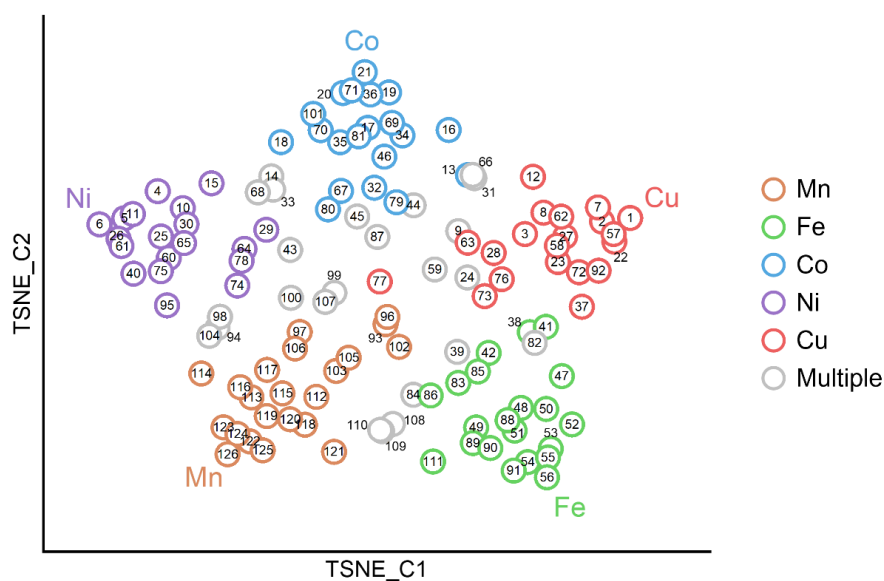


Figure S174 t-SNE projection of metal composition on M_B site of PBAs. The compositional space on M_B site (5 types of metals) is a 4-dimensional hyperplane in a 5-dimensional space, which is projected on a 2-dimensional plane for the purpose of visualization. Here, each circle represents a sampled composition on M_B site, and the number in each circle is their composition index, which is detailed in **Table S33**. Different colors of the circle represent the major metal type of each composition: e.g. composition 7 (Co_{0.2}Cu_{0.8}) has a major metal of Cu, composition 59 (Mn_{0.2}Ni_{0.4}Cu_{0.4}) has multiple major metals.

Table S33 Metal composition on M_B site of PBAs. Numbers in each cell represent the ratio of corresponding metal, e.g. composition 2 stands for Ni_{0.2}Cu_{0.8}, composition 64 stands for Mn_{0.2}Co_{0.2}Ni_{0.4}Cu_{0.2}.

index	Mn	Fe	Co	Ni	Cu	index	Mn	Fe	Co	Ni	Cu
1	0	0	0	0	1	64	0.2	0	0.2	0.4	0.2
2	0	0	0	0.2	0.8	65	0.2	0	0.2	0.6	0
3	0	0	0	0.4	0.6	66	0.2	0	0.4	0	0.4
4	0	0	0	0.6	0.4	67	0.2	0	0.4	0.2	0.2
5	0	0	0	0.8	0.2	68	0.2	0	0.4	0.4	0
6	0	0	0	1	0	69	0.2	0	0.6	0	0.2
7	0	0	0.2	0	0.8	70	0.2	0	0.6	0.2	0
8	0	0	0.2	0.2	0.6	71	0.2	0	0.8	0	0
9	0	0	0.2	0.4	0.4	72	0.2	0.2	0	0	0.6
10	0	0	0.2	0.6	0.2	73	0.2	0.2	0	0.2	0.4
11	0	0	0.2	0.8	0	74	0.2	0.2	0	0.4	0.2
12	0	0	0.4	0	0.6	75	0.2	0.2	0	0.6	0
13	0	0	0.4	0.2	0.4	76	0.2	0.2	0.2	0	0.4
14	0	0	0.4	0.4	0.2	77	0.2	0.2	0.2	0.2	0.2
15	0	0	0.4	0.6	0	78	0.2	0.2	0.2	0.4	0
16	0	0	0.6	0	0.4	79	0.2	0.2	0.4	0	0.2
17	0	0	0.6	0.2	0.2	80	0.2	0.2	0.4	0.2	0
18	0	0	0.6	0.4	0	81	0.2	0.2	0.6	0	0
19	0	0	0.8	0	0.2	82	0.2	0.4	0	0	0.4
20	0	0	0.8	0.2	0	83	0.2	0.4	0	0.2	0.2
21	0	0	1	0	0	84	0.2	0.4	0	0.4	0
22	0	0.2	0	0	0.8	85	0.2	0.4	0.2	0	0.2
23	0	0.2	0	0.2	0.6	86	0.2	0.4	0.2	0.2	0
24	0	0.2	0	0.4	0.4	87	0.2	0.4	0.4	0	0
25	0	0.2	0	0.6	0.2	88	0.2	0.6	0	0	0.2
26	0	0.2	0	0.8	0	89	0.2	0.6	0	0.2	0
27	0	0.2	0.2	0	0.6	90	0.2	0.6	0.2	0	0
28	0	0.2	0.2	0.2	0.4	91	0.2	0.8	0	0	0
29	0	0.2	0.2	0.4	0.2	92	0.4	0	0	0	0.6
30	0	0.2	0.2	0.6	0	93	0.4	0	0	0.2	0.4
31	0	0.2	0.4	0	0.4	94	0.4	0	0	0.4	0.2
32	0	0.2	0.4	0.2	0.2	95	0.4	0	0	0.6	0
33	0	0.2	0.4	0.4	0	96	0.4	0	0.2	0	0.4
34	0	0.2	0.6	0	0.2	97	0.4	0	0.2	0.2	0.2
35	0	0.2	0.6	0.2	0	98	0.4	0	0.2	0.4	0
36	0	0.2	0.8	0	0	99	0.4	0	0.4	0	0.2
37	0	0.4	0	0	0.6	100	0.4	0	0.4	0.2	0
38	0	0.4	0	0.2	0.4	101	0.4	0	0.6	0	0
39	0	0.4	0	0.4	0.2	102	0.4	0.2	0	0	0.4
40	0	0.4	0	0.6	0	103	0.4	0.2	0	0.2	0.2
41	0	0.4	0.2	0	0.4	104	0.4	0.2	0	0.4	0
42	0	0.4	0.2	0.2	0.2	105	0.4	0.2	0.2	0	0.2
43	0	0.4	0.2	0.4	0	106	0.4	0.2	0.2	0.2	0

index	Mn	Fe	Co	Ni	Cu	index	Mn	Fe	Co	Ni	Cu
44	0	0.4	0.4	0	0.2	107	0.4	0.2	0.4	0	0
45	0	0.4	0.4	0.2	0	108	0.4	0.4	0	0	0.2
46	0	0.4	0.6	0	0	109	0.4	0.4	0	0.2	0
47	0	0.6	0	0	0.4	110	0.4	0.4	0.2	0	0
48	0	0.6	0	0.2	0.2	111	0.4	0.6	0	0	0
49	0	0.6	0	0.4	0	112	0.6	0	0	0	0.4
50	0	0.6	0.2	0	0.2	113	0.6	0	0	0.2	0.2
51	0	0.6	0.2	0.2	0	114	0.6	0	0	0.4	0
52	0	0.6	0.4	0	0	115	0.6	0	0.2	0	0.2
53	0	0.8	0	0	0.2	116	0.6	0	0.2	0.2	0
54	0	0.8	0	0.2	0	117	0.6	0	0.4	0	0
55	0	0.8	0.2	0	0	118	0.6	0.2	0	0	0.2
56	0	1	0	0	0	119	0.6	0.2	0	0.2	0
57	0.2	0	0	0	0.8	120	0.6	0.2	0.2	0	0
58	0.2	0	0	0.2	0.6	121	0.6	0.4	0	0	0
59	0.2	0	0	0.4	0.4	122	0.8	0	0	0	0.2
60	0.2	0	0	0.6	0.2	123	0.8	0	0	0.2	0
61	0.2	0	0	0.8	0	124	0.8	0	0.2	0	0
62	0.2	0	0.2	0	0.6	125	0.8	0.2	0	0	0
63	0.2	0	0.2	0.2	0.4	126	1	0	0	0	0

8.3. Yield map of benzaldehyde

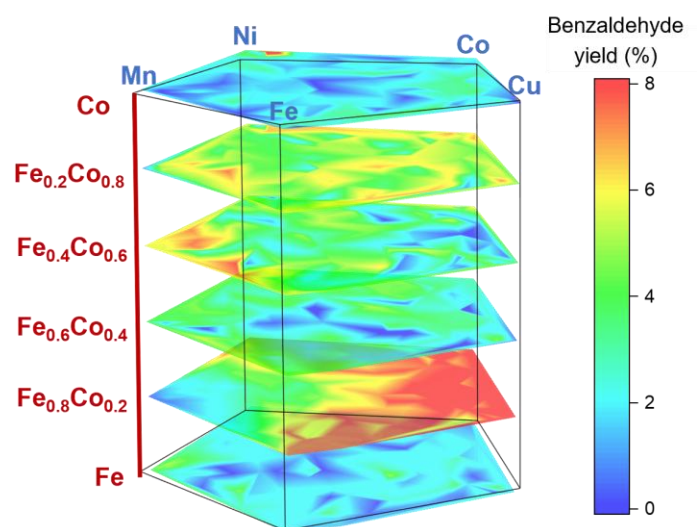


Figure S175. Yield map of benzaldehyde obtained from 756 different PBAs.

8.4. Comparison of yield and selectivity

Table S34. Comparison of yield and selectivity of styrene oxide for different PBAs and metal salts catalysts.

Catalyst	Yield of epox. (%)	Selectivity of epox. (%)	Note
Mn _{0.2} Cu _{0.8} -Fe _{0.4} Co _{0.6} PBA	69.8	98	Top PBAs in this work
Mn _{0.2} Cu _{0.8} -Fe _{0.6} Co _{0.4} PBA	67.1	94	
Co _{0.4} Ni _{0.2} Cu _{0.4} -Fe _{0.4} Co _{0.6} PBA	66.2	100	
Mn _{0.4} Fe _{0.2} Co _{0.2} Cu _{0.2} -Fe _{0.4} Co _{0.6} PBA	65.2	98	
Mn _{0.4} Cu _{0.6} -Fe _{0.8} Co _{0.2} PBA	65.1	98	
Mn ²⁺	2.3	100	Control experiments in this work
Fe ²⁺	1.0	100	
Co ²⁺	26.7	100	
Ni ²⁺	2.4	100	
Cu ²⁺	26.7	100	
Fe(CN) ₆ ³⁻	41.1	100	
Co(CN) ₆ ³⁻	4.6	100	
Without any catalyst	4.9	100	
Fe-Fe PBA	61.4	64	Reported PBAs in ref ^{44,45}
Cu-Co PBA	48.0	50	
Fe-Co PBA	30.8	35	
Cd-Co PBA	17.5	33	
Co-Co PBA	18.1	37	
Mn-Co PBA	11.3	24	
Zn-Co PBA	2.2	7	
Fe _{0.5} Cd _{0.5} -Co PBA	39.9	42	
Fe _{0.5} Co _{0.5} -Co PBA	43.2	46	
Fe _{0.5} Mn _{0.5} -Co PBA	41.4	47	
Fe _{0.5} Zn _{0.5} -Co PBA	32.4	40	
Cu _{0.5} Fe _{0.5} -Co PBA	46.1	49	
Cu _{0.5} Cd _{0.5} -Co PBA	58.9	62	
Cu _{0.5} Co _{0.5} -Co PBA	65.8	70	
Cu _{0.5} Mn _{0.5} -Co PBA	53.9	58	
Cu _{0.5} Zn _{0.5} -Co PBA	39.6	45	

8.5. NMR UV yield comparison

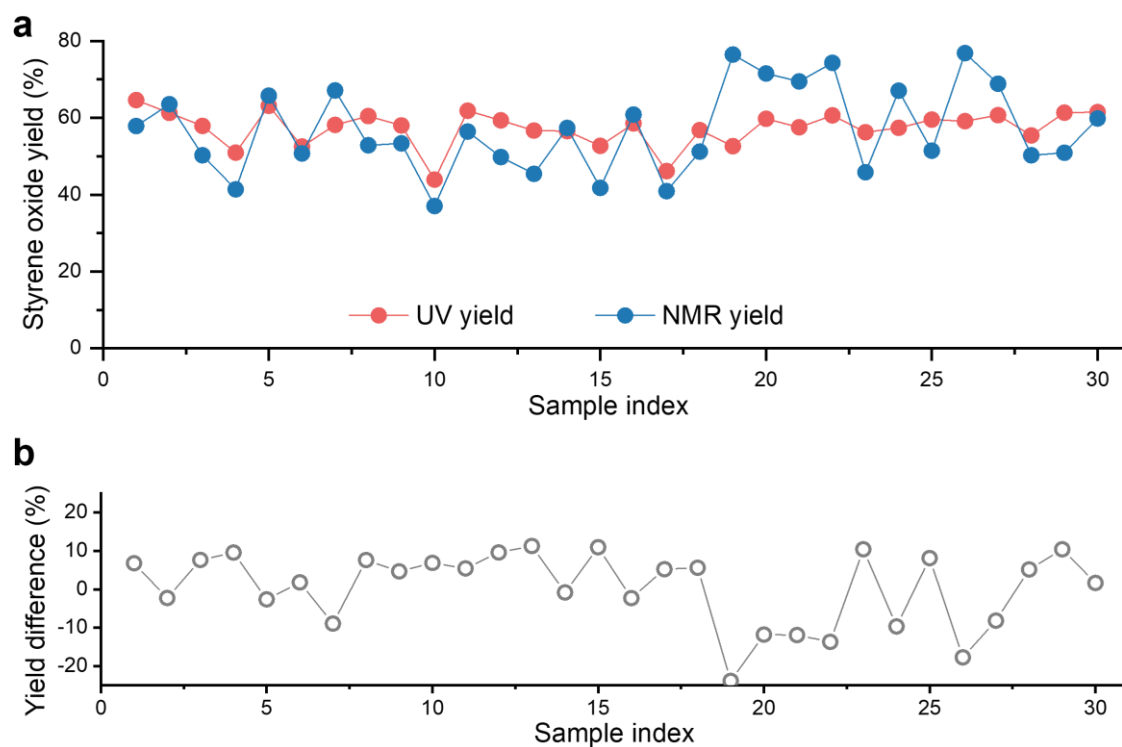


Figure S176 a. Yield comparison of styrene oxide determined by UV-Vis spectrum unmixing and manual NMR analysis for 30 selected PBAs. **b.** Yield difference between UV yield and NMR yield for 30 selected PBAs, the average difference between two methods is $0.5 \pm 9.6\%$, and the mean absolute difference between two methods is $8.1 \pm 4.9\%$.

8.6. Recycling of PBAs

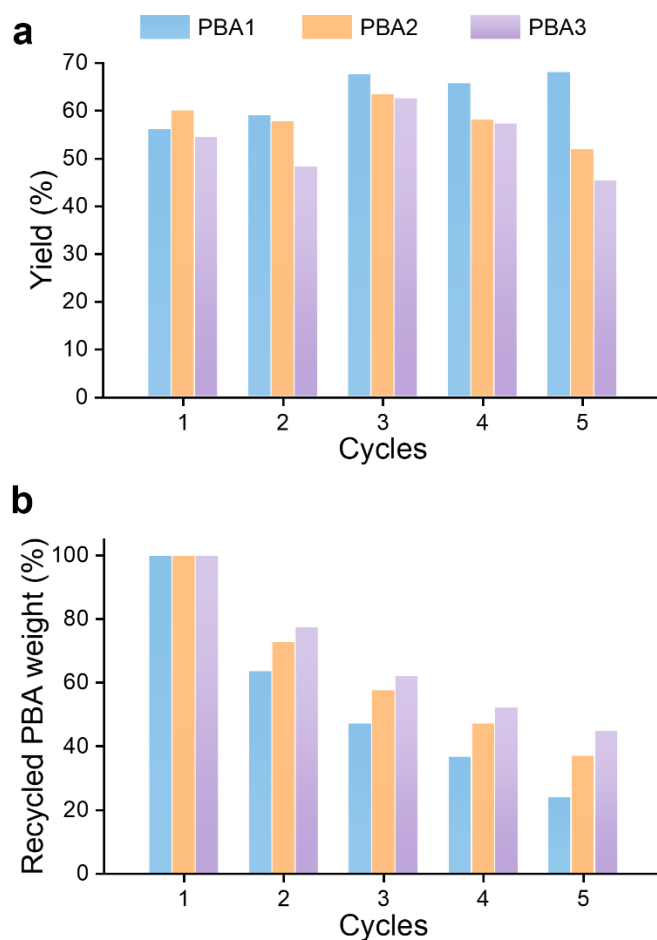
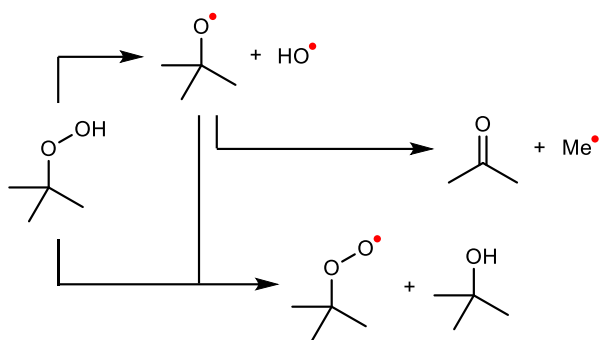


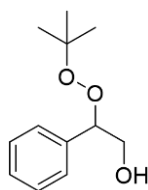
Figure S177. Three selected PBAs (PBA1: $\text{Mn}_{0.2}\text{Cu}_{0.8}\text{-Fe}_{0.4}\text{Co}_{0.6}$ PBA; PBA2: $\text{Mn}_{0.4}\text{Cu}_{0.6}\text{-Fe}_{0.8}\text{Co}_{0.2}$ PBA; PBA3: $\text{Mn}_{0.2}\text{Co}_{0.2}\text{Cu}_{0.6}\text{-Fe}_{0.4}\text{Co}_{0.6}$ PBA) with initial weight of 100 mg were used to evaluate the recycling stability of PBAs over 5 cycles of styrene oxidation. After each reaction cycle, PBAs were recollected via filtration of reaction crudes, followed by washing sequentially with acetonitrile (2 times) and deionized water (3 times). After washing, PBAs were dried at 40 °C under vacuum. The obtained PBAs powders were then used for the next reaction cycle. Notably, 10-15 mg of PBAs were lost after each recycling cycle due to separating and purification procedures. Therefore, the equivalence of all reactants, solvent, and surfactant were adjusted accordingly. **a.** Yield of styrene oxide of the 3 selected PBAs during 5 recycling cycles. **b.** The weight of PBAs used in each reaction cycle.

8.7. Generation of four radicals that react with styrene



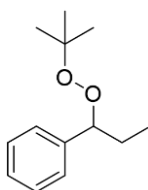
Thermal generation of radicals from *tert*-butyl hydroperoxide (*t*-BuOOH) is known and has been reported.^{195–197} The homolytic O–O bond cleavage of *t*-BuOOH gives the *tert*-butoxy (*t*-BuO•) and hydroxyl (HO•) radicals. The *t*-BuO• radical can undergo β -scission to produce acetone and the methyl ($\bullet\text{CH}_3$) radical. The *t*-BuO• radical can also abstract a hydrogen atom from *t*-BuOOH to yield *tert*-butylperoxyl (*t*-BuOO•) radical and *tert*-butanol. The four generated radicals react with styrene as illustrated in the main text **Figure 6**.

8.8. ^1H NMR data of isolated compounds



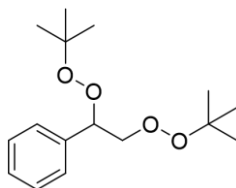
Compound 20d 2-(*tert*-butylperoxy)-2-phenylethan-1-ol matches literature description¹⁹⁸.

^1H NMR (400 MHz, chloroform-*d*): δ = 7.39–7.27 (m, 5H), 5.60 (t, J = 7.1, 5.9 Hz, 1H), 3.78 (dd, J = 16.2, 7.1 Hz, 1H), 3.21 (dd, J = 16.2, 5.9 Hz, 1H), 1.16 (s, J = 0.6 Hz, 9H).



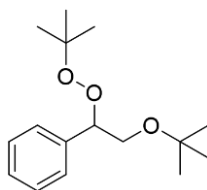
Compound 20e 1-(*tert*-butylperoxy)propylbenzene matches literature description¹⁹⁹.

^1H NMR (400 MHz, chloroform-*d*): δ = 7.36–7.27 (m, 5H), 4.75 (t, J = 6.9 Hz, 1H), 2.00–1.86 (m, 1H), 1.63–1.76 (m, 1H), 1.21 (s, 9H), 0.90 (t, J = 7.5 Hz, 3H).



Compound 20f 1,2-bis(*tert*-butylperoxy)ethylbenzene matches literature description²⁰⁰.

^1H NMR (400 MHz, chloroform-*d*): δ = 7.37–7.27 (m, 5H), 5.21 (dd, J = 7.7, 4.0 Hz, 1H), 4.23 (dd, J = 11.5, 7.7 Hz, 1H), 4.07 (dd, J = 11.5, 4.0 Hz, 1H), 1.25 (s, 9H), 1.24 (s, 9H).



Compound 20g (2-(*tert*-butoxy)-1-(*tert*-butylperoxy)ethyl)benzene matches literature description²⁰¹.

^1H NMR (400 MHz, chloroform-*d*): δ = 7.39–7.28 (m, 5H), 5.02 (dd, J = 7.4, 4.7 Hz, 1H), 3.67 (dd, J = 10.4, 7.4 Hz, 1H), 3.48 (dd, J = 10.5, 4.7 Hz, 1H), 1.26 (s, 9H), 1.15 (s, 9H).

9. Supplementary Videos

Supplementary Video 1. Liquid transfer for volume calibration, reaction preparation and UV-Vis spectrum measurement. 00:00-00:11, pipetting to analytical balance for volume calibration. 00:12-00:24, reaction preparation. 00:25-00:41 pipetting to UV-Vis spectrophotometer for spectrum acquisition.

Supplementary Video 2. Automatic measurement of UV-Vis spectra of reactions crudes. 00:26-00:16, transferring reaction crude from a vial to the spectrophotometer. 00:17-00:21, spectrum acquisition. 00:22-00:28, cleaning detection area of spectrophotometer. 00:20-00:57, exemplary measurements of another 9 samples in the same reaction plate.

Supplementary Video 3. Experimental (left) and theoretical (right) four-dimensional map of the yield of Ugi multicomponent reaction. Yield is shown by color and size of spheres. 3D space coordinates are the starting concentrations of isocyanide, amine and aldehyde. The fourth dimension is the concentration of *p*-toluenesulfonic acid (indicated above the 3D plots), which keeps increasing as the movie plays. Experimental data has been smoothed for this illustration as described in Supplementary Section 3.13. For comparison of the smoothed and raw data, see Supplementary Section 4.15.4. Theoretical yield map (right half of the video) corresponds to the kinetic model best-fit to the data as described in Supplementary Section 5.3. Same theoretical yields are shown by the blue curves in Figure S116-Figure S128 (Supplementary Section 4.13.4) for all 3234 conditions tried in experiments.

Supplementary Video 4. Illustration of the house-written HyperspaceViewer software allowing for direct visualization/analysis of 3D and 4D hyperspace data. The software features multi-isosurface rendering, interactive data point analysis and intuitive cube manipulation. The isosurfaces are generated through three sequential steps: octree generation, data interpolation, and meshing by the Marching Cubes algorithm. 00:00-00:54, visualization of 3D data; 00:55-01:48, visualization of 4D data. The application, source code and user manual of HyperspaceViewer are all deposited at DOI 10.5281/zenodo.14880579 and can also be found in the HyperspaceViewer.zip Supplementary File available from the webpage of this article.

10. References

62. Robbins, D. W. & Hartwig, J. F. A simple, multidimensional approach to high-throughput discovery of catalytic reactions. *Science* **333**, 1423–1427 (2011).
63. Troshin, K. & Hartwig, J. F. Snap deconvolution: an informatics approach to high-throughput discovery of catalytic reactions. *Science* **357**, 175–181 (2017).
64. Brereton, R. G. Introduction to multivariate calibration in analytical chemistry. *The Analyst* **125**, 2125–2154 (2000).
65. Kaspar, F. *et al.* Spectral unmixing-based reaction monitoring of transformations between nucleosides and nucleobases. *ChemBioChem* **21**, 2604–2610 (2020).
66. Keshava, N. & Mustard, J. F. Spectral unmixing. *IEEE Signal Process Mag* **19**, 44–57 (2002).
67. Berman, M. *et al.* ICE: a statistical approach to identifying endmembers in hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* **42**, 2085–2095 (2004).
68. Langergraber, G., Fleischmann, N. & Hofstädter, F. A multivariate calibration procedure for UV/VIS spectrometric quantification of organic matter and nitrate in wastewater. *Water Sci. Technol.* **47**, 63–71 (2003).
69. Andrew, K. N. & Worsfold, P. J. Comparison of multivariate calibration techniques for the quantification of model process streams using diode-array spectrophotometry. *The Analyst* **119**, 1541 (1994).
70. Durán-Merás, I., De La Peña, A. M., Espinosa-Mansilla, A. & Salinas, F. Multicomponent determination of flavour enhancers in food preparations by partial least squares and principal component regression modelling of spectrophotometric data. *The Analyst* **118**, 807–813 (1993).
71. Durbin, J. & Watson, G. S. Testing for serial correlation in least squares regression: I. *Biometrika* **37**, 409 (1950).
72. Ljung, G. M. & Box, G. E. P. On a measure of lack of fit in time series models. *Biometrika* **65**, 297–303 (1978).
73. Chen, Y. Spatial autocorrelation approaches to testing residuals from least squares regression. *PLOS ONE* **11**, e0146865 (2016).
74. Barker, L. E. & Shaw, K. M. Best (but oft-forgotten) practices: checking assumptions concerning regression residuals. *Am. J. Clin. Nutr.* **102**, 533–539 (2015).
75. Sayer, J. M. Hydration of p-nitrobenzaldehyde. *J. Org. Chem.* **40**, 2545–2547 (1975).

76. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
77. Bates, D. M. & Watts, D. G. *Nonlinear Regression Analysis and Its Applications*. (Wiley, New York, NY, 2007).
78. Rooney, W. C. & Biegler, L. T. Design for model parameter uncertainty using nonlinear confidence regions. *AIChE J.* **47**, 1794–1804 (2001).
79. Vugrin, K. W., Swiler, L. P., Roberts, R. M., Stucky-Mack, N. J. & Sullivan, S. P. Confidence region estimation techniques for nonlinear regression in groundwater flow: three case studies. *Water Resour Res* **43**, (2007).
80. Seabold, S. & Perktold, J. Statsmodels: Econometric and statistical modeling with Python. in *Proceedings of the 9th Python in Science Conference* 92–96 (2010).
81. Jamison, T. F., Lubell, W. D., Dener, J. M., Krisché, M. J. & Rapoport, H. 9-bromo-9-phenylfluorene. *Org. Synth.* **71**, 220–220 (1993).
82. Tilly, D. *et al.* On the mechanism of the directed *ortho* and remote metalation reactions of *N, N*-dialkylbiphenyl 2-carboxamides. *Org. Lett.* **12**, 68–71 (2010).
83. Nishiuchi, T. *et al.* Anthracene-attached persistent tricyclic aromatic hydrocarbon radicals. *Chem. – Asian J.* **14**, 1830–1836 (2019).
84. Santos, V. G. *et al.* The multicomponent Hantzsch reaction: comprehensive mass spectrometry monitoring using charge-tagged reagents. *Chem. – Eur. J.* **20**, 12808–12816 (2014).
85. Kohler, E. P. & Chadwell, H. M. Benzalacetophenone. *Org Synth* **2**, 1–1 (1922).
86. Chang, Z., Jing, X., He, C., Liu, X. & Duan, C. Silver clusters as robust nodes and π -activation sites for the construction of heterogeneous catalysts for the cycloaddition of propargylamines. *ACS Catal.* **8**, 1384–1391 (2018).
87. Li, D., Yin, K., Li, J. & Jia, X. CuI/iodine-mediated homocoupling reaction of terminal alkynes to 1,3-diynes. *Tetrahedron Lett.* **49**, 5918–5919 (2008).
88. Arthur Lachman. Benzophenone oxime. *Org. Synth.* **10**, 10 (1930).
89. Liu, L.-F. *et al.* Facile AlCl₃-promoted catalytic Beckmann rearrangement of ketoximes. *Synth. Commun.* **41**, 553–560 (2011).
90. Murai, M., Ogita, T. & Takai, K. Regioselective arene homologation through rhenium-catalyzed deoxygenative aromatization of 7-oxabicyclo[2.2.1]hepta-2,5-dienes. *Chem. Commun.* **55**, 2332–2335 (2019).

91. Çapan, İ., Servi, S., Dalkılıç, S. & Dalkılıç, L. K. Synthesis and anticancer evaluation of benzimidazole derivatives having norbornene/dibenzobarrelene skeletons and different functional groups. *ChemistrySelect* **5**, 14393–14398 (2020).
92. 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).
93. Alvim, H. G. O. *et al.* Facts, presumptions, and myths on the solvent-free and catalyst-free Biginelli reaction. What is catalysis for? *J. Org. Chem.* **79**, 3383–3397 (2014).
94. Companyo, X., Moyano, A. & Rios, R. A mild and convenient synthesis of 4-tosyl-4,5-dihydrooxazoles. *Lett. Org. Chem.* **6**, 293–296 (2009).
95. Khan, A. T., Parvin, T. & Choudhury, L. H. Effects of substituents in the β -position of 1,3-dicarbonyl compounds in bromodimethylsulfonium bromide-catalyzed multicomponent reactions: a facile access to functionalized piperidines. *J. Org. Chem.* **73**, 8398–8402 (2008).
96. Iglesias, E. Determination of keto–enol equilibrium constants and the kinetic study of the nitrosation reaction of β -dicarbonyl compounds. *J. Chem. Soc. Perkin Trans. 2* 431–440 (1997).
97. Roszak, R. *et al.* Systematic, computational discovery of multicomponent and one-pot reactions. *Nat. Commun.* **15**, 10285 (2024).
98. Attaby, F. A. Reactions with cyanothioacetamide derivatives: synthesis of several new pyridine and annelated pyridine derivatives. *Arch. Pharm. Res.* **13**, 342–346 (1990).
99. Elneairy, M. A. A. Cyanothioacetamide in heterocyclic synthesis: a novel synthesis of styrylpyridinethione, styrylthieno[2,3-b]pyridine, styrylpyrazolo[3,4-b]pyridine, pyrido[2',3':3,4]pyrazolo[5,1-a]pyrimidine, pyrido[3',2':4,5]thieno[3,2-d]pyrimidine and pyrido[3',2':4,5]thieno[3,2-d]-1,2,3-triazin-4-one derivatives. *Phosphorus Sulfur Silicon Relat. Elem.* **178**, 2201–2214 (2003).
100. Ia, R. A multi-component reaction to 6-aminothiouracils: synthesis, mechanistic study and antitumor activity. *Chemother. Open Access* **05**, (2016).
101. Vigante, B. *et al.* An efficient synthesis of multisubstituted 4-nitrobuta-1,3-dien-1-amines and application in cyclisation reactions. *Tetrahedron* **74**, 2596–2607 (2018).
102. Thamaraiselvi, S. & Mohan, P. S. An approach to the synthesis of new 1-phenylacridones and naphthacridones. *Z. Für Naturforschung B* **54**, 1337–1341 (1999).
103. Esquivel, R. O., Molina-Espíritu, M. & López-Rosa, S. 3D information-theoretic analysis of the simplest hydrogen abstraction reaction. *J. Phys. Chem. A* **127**, 6159–6174 (2023).

104. Ludvík, J., Hilgard, S. & Volke, J. Determination of water in acetonitrile, propionitrile, dimethylformamide and tetrahydrofuran by infrared and near-infrared spectrometry. *The Analyst* **113**, 1729–1731 (1988).
105. Constable, D. J. C., Curzons, A. D. & Cunningham, V. L. Metrics to ‘green’ chemistry— which are the best? *Green Chem* **4**, 521–527 (2002).
106. Specht, T. *et al.* HANNA: hard-constraint neural network for consistent activity coefficient prediction. *Chem. Sci.* **15**, 19777–19786 (2024).
107. Sirotkin, V. A., Solomonov, B. N., Faizullin, D. A. & Fedotov, V. D. IR spectroscopic study of the state of water in dioxane and acetonitrile: Relationship with the thermodynamic activity of water at 278–318 K. *J. Struct. Chem.* **41**, 997–1003 (2000).
108. Kogan, V., Fridman, B. & Kafarov, V. *Liquid–Vapor Equilibrium [in Russian]*. vol. 1 (Nauka, Moscow, 1966).
109. Triggs, B., McLauchlan, P. F., Hartley, R. I. & Fitzgibbon, A. W. Bundle Adjustment — A Modern Synthesis. in *Vision Algorithms: Theory and Practice* (eds. Triggs, B., Zisserman, A. & Szeliski, R.) vol. 1883 298–372 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2000).
110. Skinner, H. A. & Snelson, A. The heats of combustion of the four isomeric butyl alcohols. *Trans. Faraday Soc.* **56**, 1776 (1960).
111. Chase, M. *NIST-JANAF Thermochemical Tables*. (American chemical society, Washington, D.C, 1998).
112. Nesterova, T. & Rozhnov, A. Isomerization of isostructural monobromobutanes. *Izv Vyssh Uchebn Zaved Khim Khim Tekhnol* **17**, 556–558 (1974).
113. Barbosa, J. & Sanz-Nebot, V. Acid—base equilibria and assay of benzodiazepines in acetonitrile medium. *Talanta* **36**, 837–842 (1989).
114. Barbosa, J. & Sanz-Nebot, V. Autoprotolysis constants and standardization of the glass electrode in acetonitrile-water mixtures. Effect of solvent composition. *Anal. Chim. Acta* **244**, 183–191 (1991).
115. Kütt, A. *et al.* Equilibrium acidities of superacids. *J. Org. Chem.* **76**, 391–395 (2011).
116. Dormand, J. R. & Prince, P. J. A family of embedded Runge-Kutta formulae. *J. Comput. Appl. Math.* **6**, 19–26 (1980).
117. Freedman, D. A. Bootstrapping regression models. *Ann. Stat.* **9**, 1218–1228 (1981).
118. Fox, J. Chapter 21. Bootstrapping regression models. in *Applied regression analysis and generalized linear models* (SAGE Publications, Los Angeles, 2008).

119. Honkela, M. L., Ouni, T. & Krause, A. O. I. Thermodynamics and kinetics of the dehydration of *tert*-butyl alcohol. *Ind. Eng. Chem. Res.* **43**, 4060–4065 (2004).
120. Delion, A., Torck, B. & Hellin, M. Equilibrium constant for the liquid-phase hydration of isobutylene over ion-exchange resin. *Ind. Eng. Chem. Process Des. Dev.* **25**, 889–893 (1986).
121. Velo, E., Puigjaner, L. & Recasens, F. Inhibition by product in the liquid-phase hydration of isobutene to *tert*-butyl alcohol: kinetics and equilibrium studies. *Ind. Eng. Chem. Res.* **27**, 2224–2231 (1988).
122. Iborra, M. *et al.* Experimental study of the liquid-phase simultaneous syntheses of methyl *tert*-butyl ether (MTBE) and *tert*-butyl alcohol (TBA). *Ind. Eng. Chem. Res.* **41**, 5359–5365 (2002).
123. Eberz, W. F. & Lucas, H. J. The hydration of unsaturated compounds. II. The equilibrium between *i*-butene and *t*-butanol and the free energy of hydration of *i*-butene. *J. Am. Chem. Soc.* **56**, 1230–1234 (1934).
124. Taft, R. W., Purlee, E. L. & Riesz, P. A method for determining the distribution constant for a substance between the gas phase and a condensed phase. *J. Am. Chem. Soc.* **77**, 899–902 (1955).
125. Counsell, J. F., Lees, E. B. & Martin, J. F. Thermodynamic properties of organic oxygen compounds. Part XIX. Low-temperature heat capacity and entropy of propan-1-ol, 2-methylpropan-1-ol, and pentan-1-ol. *J Chem Soc A* 1819–1823 (1968).
126. Chao, J., Hall, K. R. & Yao, J. Thermodynamic properties of simple alkenes. *Thermochim. Acta* **64**, 285–303 (1983).
127. Mosselman, C. & Dekker, H. Enthalpies of formation of *n*-alkan-1-ols. *J. Chem. Soc. Faraday Trans. 1 Phys. Chem. Condens. Phases* **71**, 417 (1975).
128. Cox, J. D., Wagman, D. D. & Medvedev, V. A. *CODATA Key Values for Thermodynamics*. (Hemisphere Pub. Corp, New York, 1989).
129. Furuyama, S., Golden, D. M. & Benson, S. W. Thermochemistry of the gas phase equilibria $i\text{-C}_3\text{H}_7\text{I} \rightleftharpoons \text{C}_3\text{H}_6 + \text{HI}$, $n\text{-C}_3\text{H}_7\text{I} \rightleftharpoons i\text{-C}_3\text{H}_7\text{I}$, and $\text{C}_3\text{H}_6 + 2\text{HI} \rightleftharpoons \text{C}_3\text{H}_8 + \text{I}_2$. *J. Chem. Thermodyn.* **1**, 363–375 (1969).
130. Steele, W. V. *et al.* Thermodynamic properties and ideal-gas enthalpies of formation for cyclohexene, phthalan (2,5-dihydrobenzo-3,4-furan), isoxazole, octylamine, dioctylamine, trioctylamine, phenyl isocyanate, and 1,4,5,6-tetrahydropyrimidine. *J. Chem. Eng. Data* **41**, 1269–1284 (1996).

131. Adachi, K., Suga, H. & Seki, S. Phase changes in crystalline and glassy-crystalline cyclohexanol. *Bull. Chem. Soc. Jpn.* **41**, 1073–1087 (1968).
132. Haida, O., Suga, H. & Seki, S. Calorimetric study of the glassy state. XI. Plural glass transition phenomena of cyclohexene. *Bull. Chem. Soc. Jpn.* **50**, 802–809 (1977).
133. Wiberg, K. B., Wasserman, D. J., Martin, E. J. & Murcko, M. A. Enthalpies of hydration of alkenes. 3. Cycloalkenes. *J. Am. Chem. Soc.* **107**, 6019–6022 (1985).
134. Dey, J., O'Donoghue, A. C. & More O'Ferrall, R. A. Equilibrium constants for dehydration of water adducts of aromatic carbon–carbon double bonds. *J. Am. Chem. Soc.* **124**, 8561–8574 (2002).
135. Petzold, L. Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations. *SIAM J. Sci. Stat. Comput.* **4**, 136–148 (1983).
136. Hindmarsh, A. ODEPACK, a systemized collection of ODE solvers. *Sci. Comput.* (1983).
137. Restrepo, G. Spaces of mathematical chemistry. *Theory Biosci.* **143**, 237–251 (2024).
138. Restrepo, G. Chemical space: limits, evolution and modelling of an object bigger than our universal library. *Digit. Discov.* **1**, 568–585 (2022).
139. Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J. emcee: The MCMC hammer. *Publ. Astron. Soc. Pac.* **125**, 306–312 (2013).
140. Ráfols, C., Bosch, E., Rosés, M. & Asuero, A. G. Autoprotolysis in aqueous organic solvent mixtures. Water-amide and water-amine binary systems. *Anal. Chim. Acta* **302**, 355–363 (1995).
141. Brent, R. P. *Algorithms for Minimization without Derivatives*. (Prentice-Hall, Englewood Cliffs, N.J., 1973).
142. Maran, F., Celadon, D., Severin, M. G. & Vianello, E. Electrochemical determination of the pK_a of weak acids in N,N-dimethylformamide. *J. Am. Chem. Soc.* **113**, 9320–9329 (1991).
143. Crampton, M. R. & Robotham, I. A. Acidities of some substituted ammonium ions in dimethyl sulfoxide. *J. Chem. Res.* 22–23 (1997) doi:10.1039/a606020j.
144. Yang, Q. *et al.* Holistic prediction of the pK_a in diverse solvents based on a machine-learning approach. *Angew. Chem. Int. Ed.* **59**, 19282–19291 (2020).
145. Guthrie, J. P. Hydrolysis of esters of oxy acids: pK_a values for strong acids; Brønsted relationship for attack of water at methyl; free energies of hydrolysis of esters of oxy acids; and a linear relationship between free energy of hydrolysis and pK_a holding over a range of 20 pK units. *Can. J. Chem.* **56**, 2342–2354 (1978).

146. Appel, A. M. & Helm, M. L. Determining the overpotential for a molecular electrocatalyst. *ACS Catal.* **4**, 630–633 (2014).
147. Hall, H. K. Jr. Correlation of the base strengths of amines 1. *J. Am. Chem. Soc.* **79**, 5441–5444 (1957).
148. Sung, K. & Chen, C.-C. Kinetics and mechanism of acid-catalyzed hydrolysis of cyclohexyl isocyanide and pKa determination of *N*-cyclohexylnitrilium ion. *Tetrahedron Lett.* **42**, 4845–4848 (2001).
149. van Leusen, A. M., van Leusen, D. & Czakó, B. *p*-Tolylsulfonylmethyl Isocyanide. in *Encyclopedia of Reagents for Organic Synthesis* (John Wiley & Sons, Ltd, 2008).
150. Bordwell, F. G., Van der Puy, M. & Vanier, N. R. Carbon acids. 8. The trimethylammonio group as a model for assessing the polar effects of electron-withdrawing groups. *J. Org. Chem.* **41**, 1883–1885 (1976).
151. Porter, G. R., Rydon, H. N. & Schofield, J. A. Mechanism of esterase action: nature of the reactive serine residue in enzymes inhibited by organo-phosphorus compounds. *Nature* **182**, 927–927 (1958).
152. Rossini, E., Bochevarov, A. D. & Knapp, E. W. Empirical conversion of pKa values between different solvents and interpretation of the parameters: application to water, acetonitrile, dimethyl sulfoxide, and methanol. *ACS Omega* **3**, 1653–1662 (2018).
153. Clayden, J., Greeves, N. & Warren, S. *Organic Chemistry*. (Oxford University Press, London, 2012).
154. Gilbert, A. K., Zhao, Y., Otteson, C. E. & Pluth, M. D. Development of acid-mediated H₂S/COS donors that respond to a specific pH window. *J. Org. Chem.* **84**, 14469–14475 (2019).
155. Edelsbrunner, Letscher, & Zomorodian. Topological persistence and simplification. *Discrete Comput. Geom.* **28**, 511–533 (2002).
156. Zomorodian, A. & Carlsson, G. Computing persistent homology. in *Proceedings of the twentieth annual symposium on Computational geometry* 347–356 (ACM, Brooklyn New York USA, 2004). doi:10.1145/997817.997870.
157. Joharinad, P. & Jost, J. *Mathematical Principles of Topological and Geometric Data Analysis*. (Springer, Cham, 2023). doi:10.1007/978-3-031-33440-5.
158. Santra, G., Sylvetsky, N. & Martin, J. M. L. Minimally empirical double-hybrid functionals trained against the GMTKN55 database: revDSD-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. *J. Phys. Chem. A* **123**, 5129–5143 (2019).

159. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).
160. Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **32**, 1456–1465 (2011).
161. Goerigk, L. *et al.* A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.* **19**, 32184–32215 (2017).
162. Aoto, Y. A., De Lima Batista, A. P., Köhn, A. & De Oliveira-Filho, A. G. S. How to arrive at accurate benchmark values for transition metal compounds: computation or experiment? *J. Chem. Theory Comput.* **13**, 5291–5316 (2017).
163. Jacquemin, D., Duchemin, I. & Blase, X. 0–0 energies using hybrid schemes: benchmarks of TD-DFT, CIS(D), ADC(2), CC2, and BSE/ *GW* formalisms for 80 real-life compounds. *J. Chem. Theory Comput.* **11**, 5340–5359 (2015).
164. Bannwarth, C. *et al.* Extended tight-binding quantum chemistry methods. *WIREs Comput. Mol. Sci.* **11**, e1493 (2021).
165. Bannwarth, C., Ehlert, S. & Grimme, S. GFN2-xTB—an accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory Comput.* **15**, 1652–1671 (2019).
166. Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **8**, 1057 (2006).
167. Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **7**, 3297 (2005).
168. Chai, J.-D. & Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **10**, 6615 (2008).
169. Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **90**, 1007–1023 (1989).
170. Tomasi, J., Mennucci, B. & Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **105**, 2999–3094 (2005).
171. Fehér, P. P., Madarász, Á. & Stirling, A. A practice-oriented benchmark strategy to predict the UV-Vis spectra of organic photocatalysts. *Chemistry-Methods* **3**, e202200069 (2023).

172. Gaussian Inc. Creating UV/Visible plots from the results of excited states calculations. <https://gaussian.com/uvvisplot/> (2017).
173. Ali, A. *et al.* TD-DFT benchmark for UV-visible spectra of fused-ring electron acceptors using global and range-separated hybrids. *Phys. Chem. Chem. Phys.* **22**, 7864–7874 (2020).
174. Deno, N. C., Friedman, Norman., Hodge, J. D., MacKay, F. Patrick. & Saines, George. The hydride transfer nature of the reduction of carbonium ions by HBr, HI and a Pt and an Ir hydride. *J. Am. Chem. Soc.* **84**, 4713–4715 (1962).
175. Turro, N. J., Chow, M.-F. & Rigaudy, J. Thermolysis of anthracene endoperoxides. Concerted vs. diradical mechanisms. Microscopic reversibility in endothermic chemiluminescent reactions. *J. Am. Chem. Soc.* **101**, 1300–1302 (1979).
176. Fidler, H., Lauer, A., Freyer, W., Koeppe, B. & Heyne, K. Photochemistry of anthracene-9,10-endoperoxide. *J. Phys. Chem. A* **113**, 6289–6296 (2009).
177. Lauer, A., Dobryakov, A. L., Kovalenko, S. A., Fidler, H. & Heyne, K. Dual photochemistry of anthracene-9,10-endoperoxide studied by femtosecond spectroscopy. *Phys. Chem. Chem. Phys.* **13**, 8723–8732 (2011).
178. Borrelli, R. L. & Coleman, C. S. *Differential Equations: A Modeling Perspective*. (Wiley, New York, 2004).
179. Gorban, A. N., Radulescu, O. & Zinovyev, A. Y. Asymptotology of chemical reaction networks. *Chem. Eng. Sci.* **65**, 2310–2324 (2010).
180. Radulescu, O., Gorban, A. N., Zinovyev, A. & Lilienbaum, A. Robust simplifications of multiscale biochemical networks. *BMC Syst. Biol.* **2**, 86 (2008).
181. Yablonsky, G. S., Olea, M. & Marin, G. B. Temporal analysis of products: basic principles, applications, and theory. *J. Catal.* **216**, 120–134 (2003).
182. Horn, F. General first order kinetics. *Berichte Bunsenges. Für Phys. Chem.* **75**, 1191–1201 (1971).
183. Yablonskii, G., Bykov, V. & Gorban, A. *Kinetic Models of Catalytic Reactions (in Russian)*. (Nauka, Novosibirsk, 1983).
184. Kvasnička, V. Formal first-order chemical kinetics. *Chem. Pap.* **41**, 145–169 (1987).
185. Gorban, A., Bykov, V. & Yablonskii, G. *Essays on Chemixal Relaxation*. (Nauka, Novosibirsk, 1986).
186. Rugh, W. *Linear System Theory*. (Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1996).

187. Alligood, K. T., Sauer, T. & Yorke, J. A. *Chaos: An Introduction to Dynamical Systems*. (Springer, New York, 1996).
188. Dormand, J. R. & Prince, P. J. A family of embedded Runge-Kutta formulae. *J. Comput. Appl. Math.* **6**, 19–26 (1980).
189. Sukharev, A. Optimal method of constructing best uniform approximations for functions of a certain class. *USSR Comput. Math. Math. Phys.* **18**, 21–31 (1978).
190. Cooper, D. A. Learning lipschitz functions. *Int. J. Comput. Math.* **59**, 15–26 (1995).
191. Zabinsky, Z. B., Smith, R. L. & Kristinsdottir, B. P. Optimal estimation of univariate black-box Lipschitz functions with upper and lower error bounds. *Comput. Oper. Res.* **30**, 1539–1553 (2003).
192. Cooper, D. A. An improved bound for learning Lipschitz functions. *Commun. Appl. Anal.* **10**, 19–28 (2006).
193. Bambah, R. P. & Gupta, H. On lattice coverings by spheres. in *Proceedings of the National Institute of Sciences of India* vol. 20 25–52 (Indian National Science Academy, 1954).
194. Few, L. Covering space by spheres. *Mathematika* **3**, 136–139 (1956).
195. Bell, E. R., Raley, J. H., Rust, F. F., Seubold, F. H. & Vaughan, W. E. Reactions of free radicals associated with low temperature oxidation of paraffins. *Discuss. Faraday Soc.* **10**, 242 (1951).
196. Benson, S. W. Kinetics of pyrolysis of alkyl hydroperoxides and their O–O bond dissociation energies. *J. Chem. Phys.* **40**, 1007–1013 (1964).
197. Jones, P. J., Riser, B. & Zhang, J. Flash pyrolysis of *t*-butyl hydroperoxide and di-*t*-butyl peroxide: evidence of roaming in the decomposition of organic hydroperoxides. *J. Phys. Chem. A* **121**, 7846–7853 (2017).
198. Bhat, S. & Chandrasekaran, S. Oxygenation of alkenes with *t*-BuOOH catalysed by β -cyclodextrin borate. *Tetrahedron Lett.* **37**, 3581–3584 (1996).
199. Bloodworth, A. J. & Courtneidge, J. L. Oxymetallation. Part 17. *t*-Butyl peroxymercuration and subsequent demercuration of phenylcyclopropane. *J. Chem. Soc. Perkin I* 1807–1809 (1982).
200. Terent'ev, A. O., Sharipov, M. Yu., Krylov, I. B., Gaidarenko, D. V. & Nikishin, G. I. Manganese triacetate as an efficient catalyst for bisperoxidation of styrenes. *Org. Biomol. Chem.* **13**, 1439–1445 (2015).

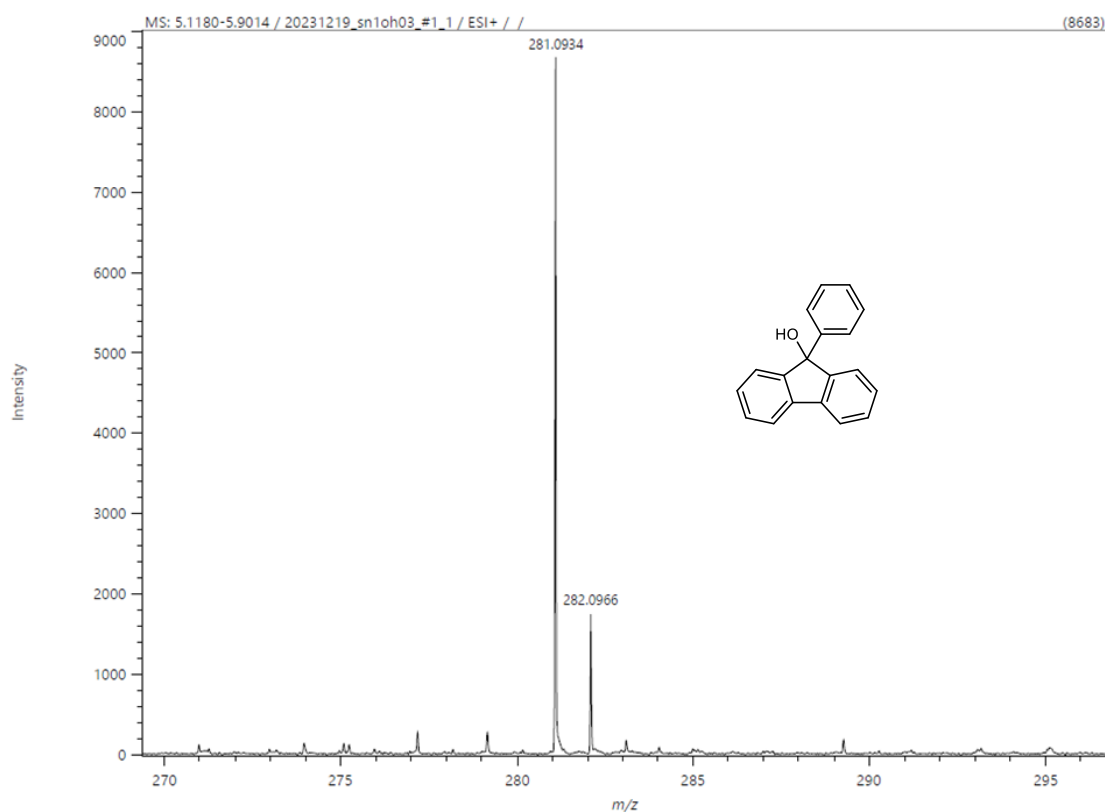
201. Minisci, F. *et al.* Kharasch and metalloporphyrin catalysis in the functionalization of alkanes, alkenes, and alkylbenzenes by t-BuOOH. Free radical mechanisms, solvent effect, and relationship with the Gif Reaction. *J. Am. Chem. Soc.* **117**, 226–232 (1995).

11. Mass spectra

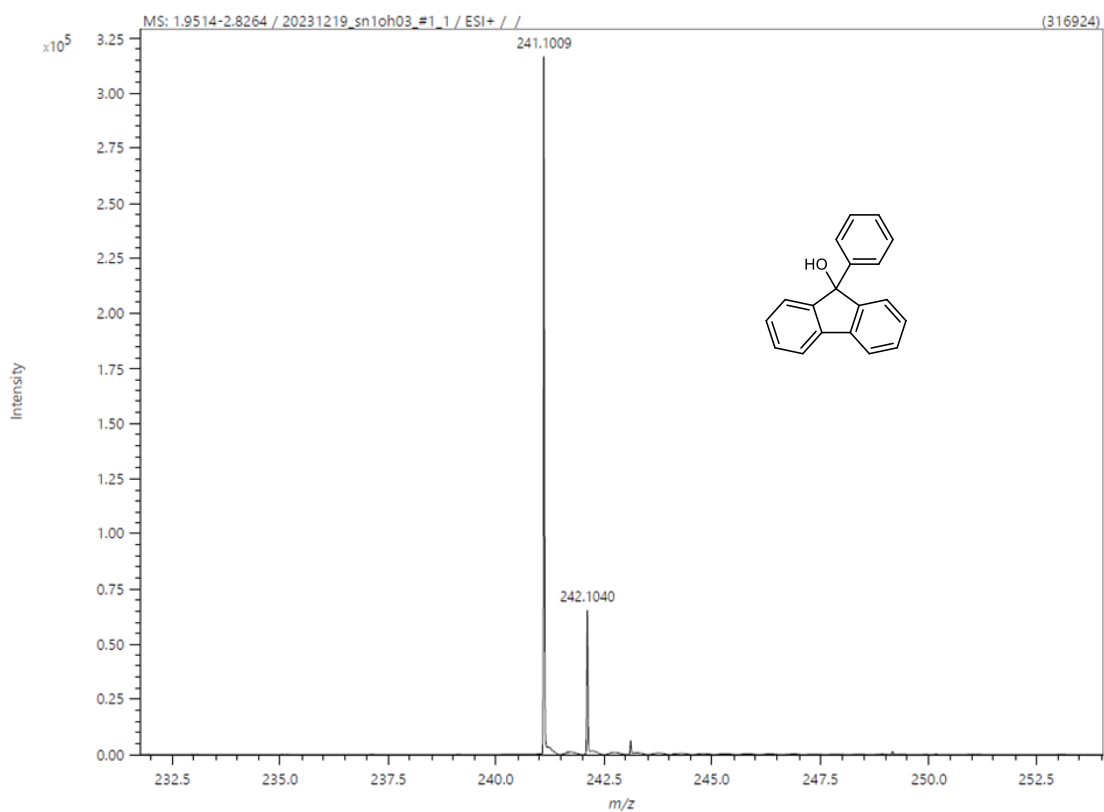
Content list

10. Mass spectra	258
High resolution mass spectrum of <u>14a</u> (ESI–TOF).....	259
High resolution mass spectrum of <u>14a</u> (ESI–TOF).....	259
High resolution mass spectrum of <u>14b</u> (ESI–TOF).....	260
High resolution mass spectrum of <u>13a</u> (ESI–TOF).....	260
High resolution mass spectrum of <u>13a</u> (ESI–TOF).....	261
High resolution mass spectrum of <u>13b</u> (ESI–TOF).....	261
High resolution mass spectrum of <u>15a</u> (ESI–TOF).....	262
High resolution mass spectrum of <u>15a</u> (ESI–TOF).....	262
High resolution mass spectrum of <u>19f</u> and <u>19g</u> (ESI–TOF).....	263
High resolution mass spectrum of <u>19h</u> (ESI–TOF).....	263
High resolution mass spectrum of EAB (ESI–TOF).....	264
High resolution mass spectrum of <u>15b</u> (ESI–TOF).....	264
High resolution mass spectrum of <u>15b</u> (ESI–TOF).....	265
High resolution mass spectrum of <u>16a</u> (ESI–TOF).....	265
High resolution mass spectrum of <u>19k</u> (ESI–TOF).....	266
High resolution mass spectrum of <u>19m</u>	266
High resolution mass spectrum of <u>19l</u> (ESI–TOF).....	267
High resolution mass spectrum of <u>19n</u> (ESI–TOF).....	267
High resolution mass spectrum of <u>19o</u> (ESI–TOF).....	268
High resolution mass spectrum of <u>19r</u> (ESI–TOF).....	268
High resolution mass spectrum of <u>19p</u>	269
High resolution mass spectrum of <u>19j</u> (ESI–TOF).....	269
High resolution mass spectrum of <u>19e</u> (ESI–TOF).....	270
High resolution mass spectrum of <u>19h</u> (ESI–TOF).....	270
High-resolution mass spectrum of 2-(butylamino)-2-(4-nitrophenyl)- <i>N</i> -(tosylmethyl)acetamide (ESI–TOF).....	271
High-resolution mass spectrum of <u>17g</u> (ESI–TOF).....	271

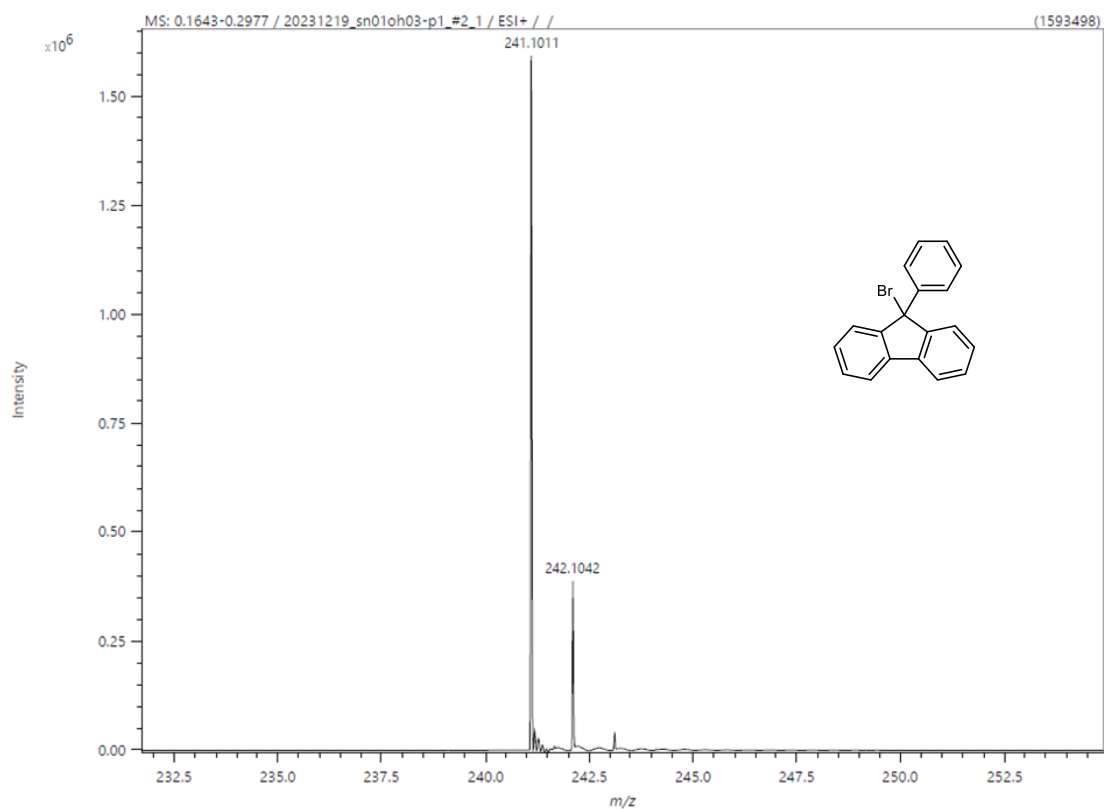
High resolution mass spectrum of 14a (ESI-TOF).



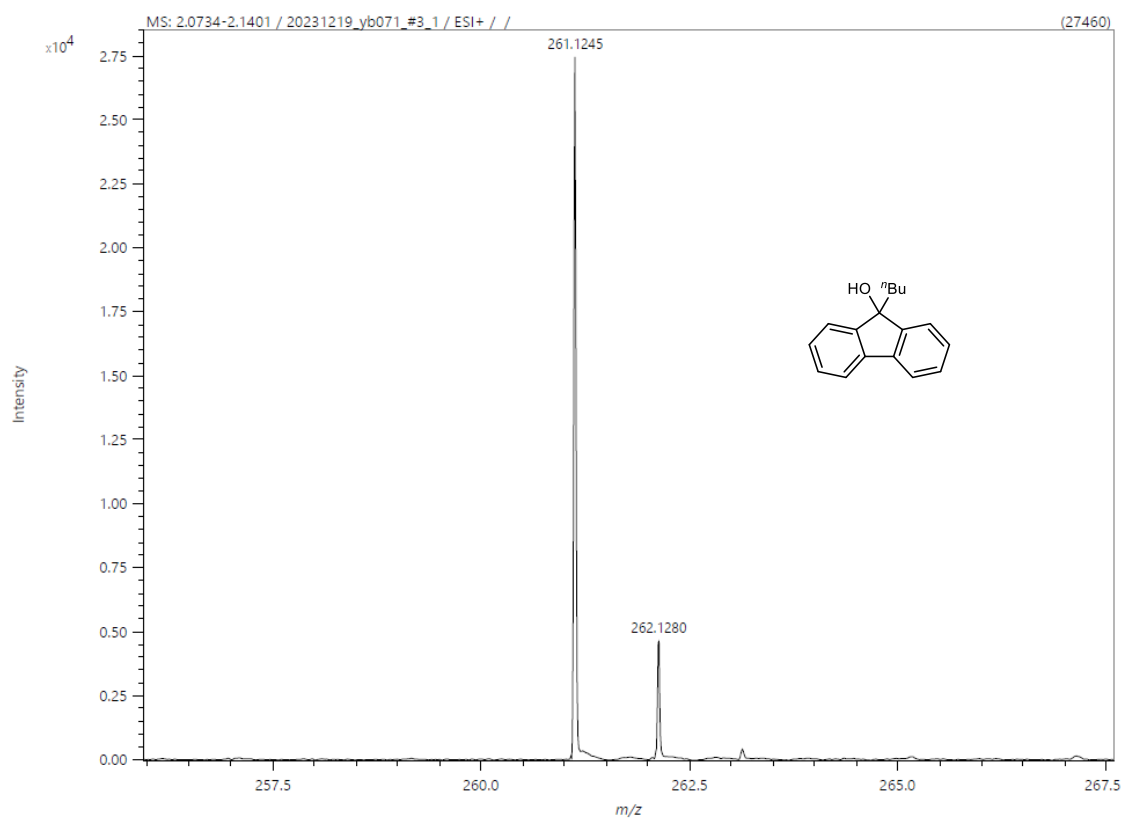
High resolution mass spectrum of 14a (ESI-TOF)



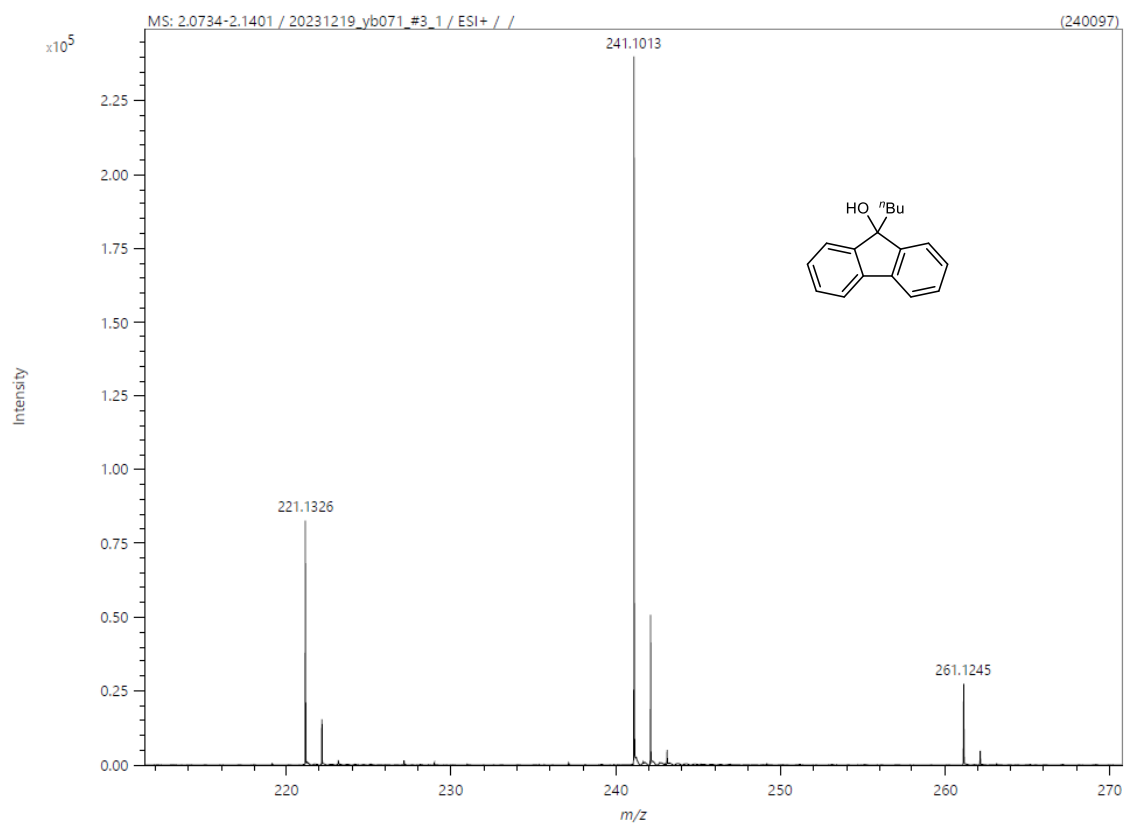
High resolution mass spectrum of 14b (ESI-TOF).



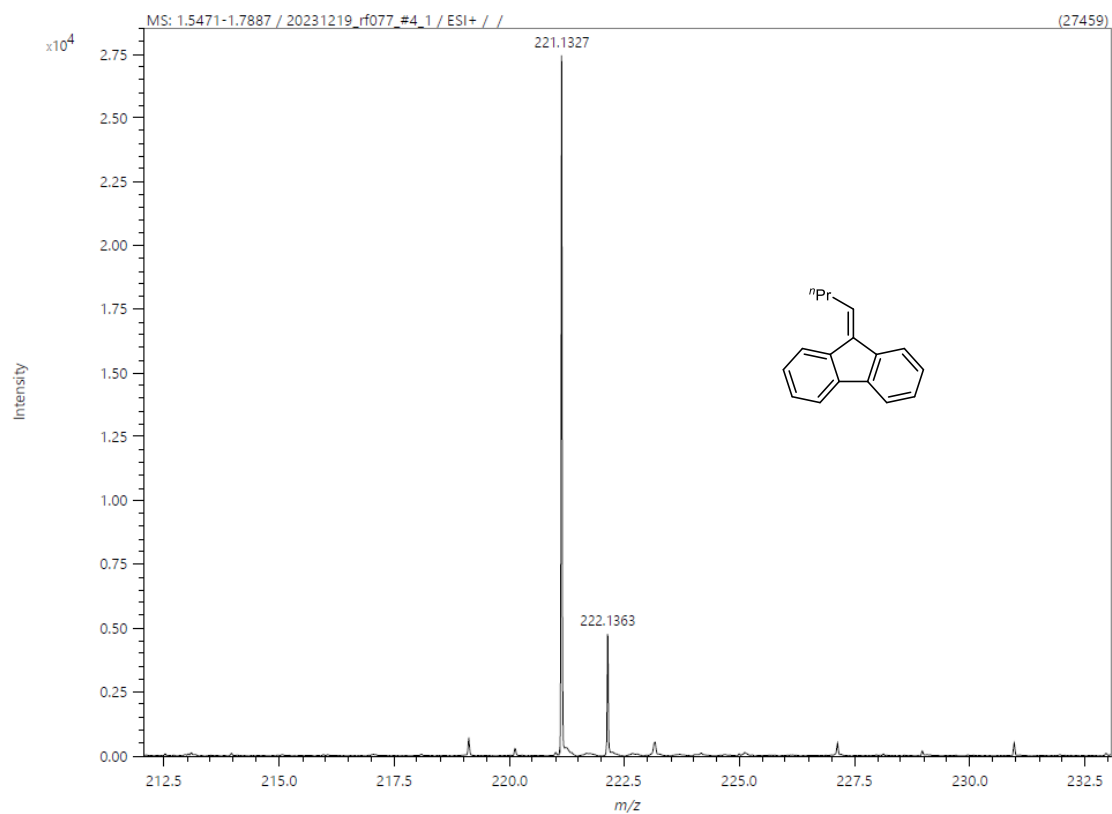
High resolution mass spectrum of 13a (ESI-TOF)



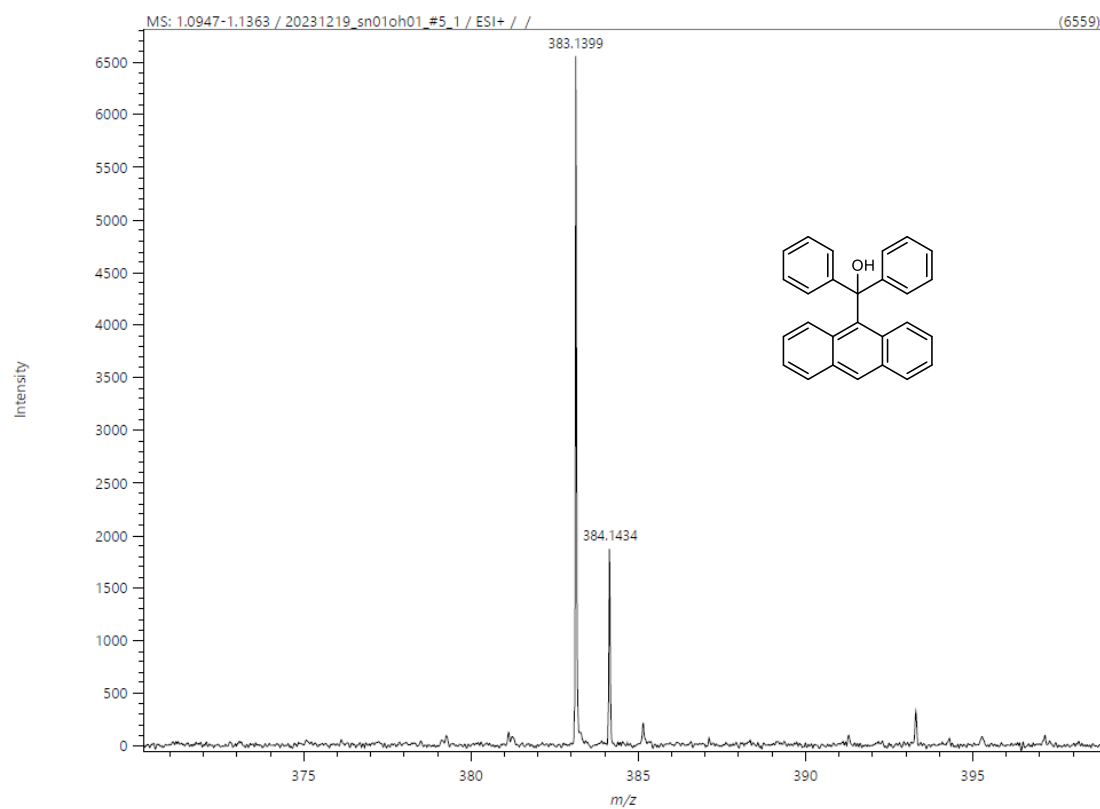
High resolution mass spectrum of **13a** (ESI-TOF)



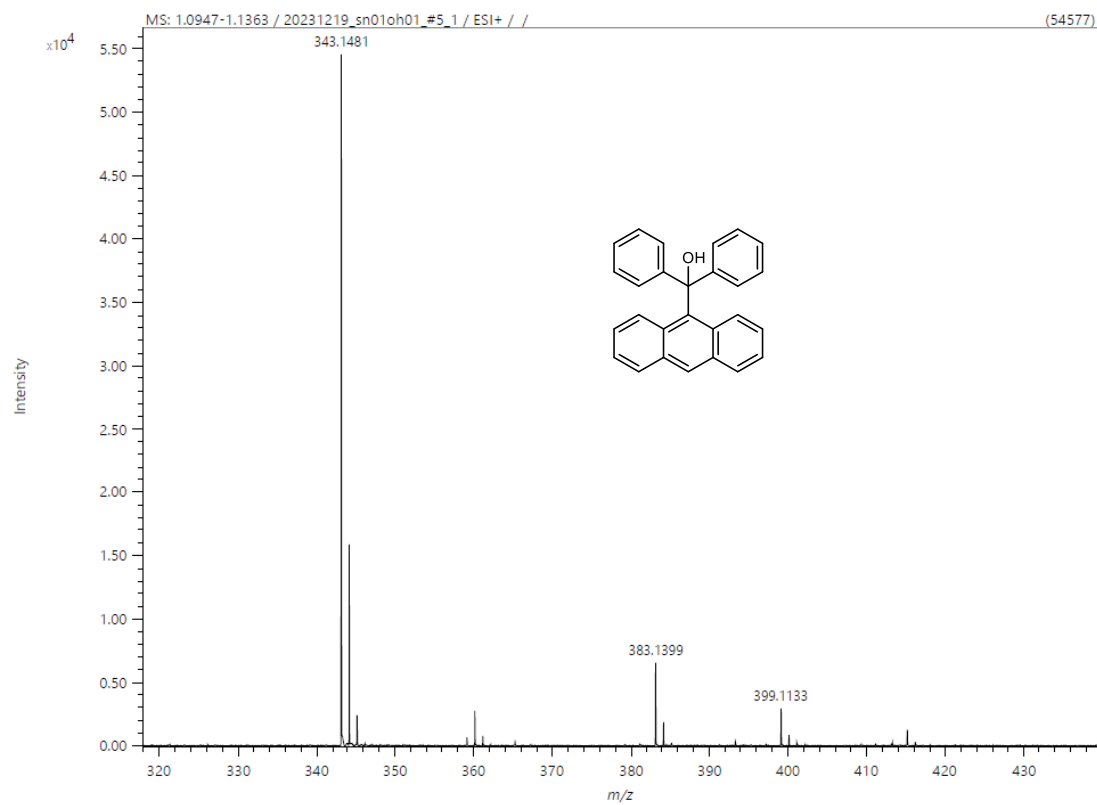
High resolution mass spectrum of **13b** (ESI-TOF).



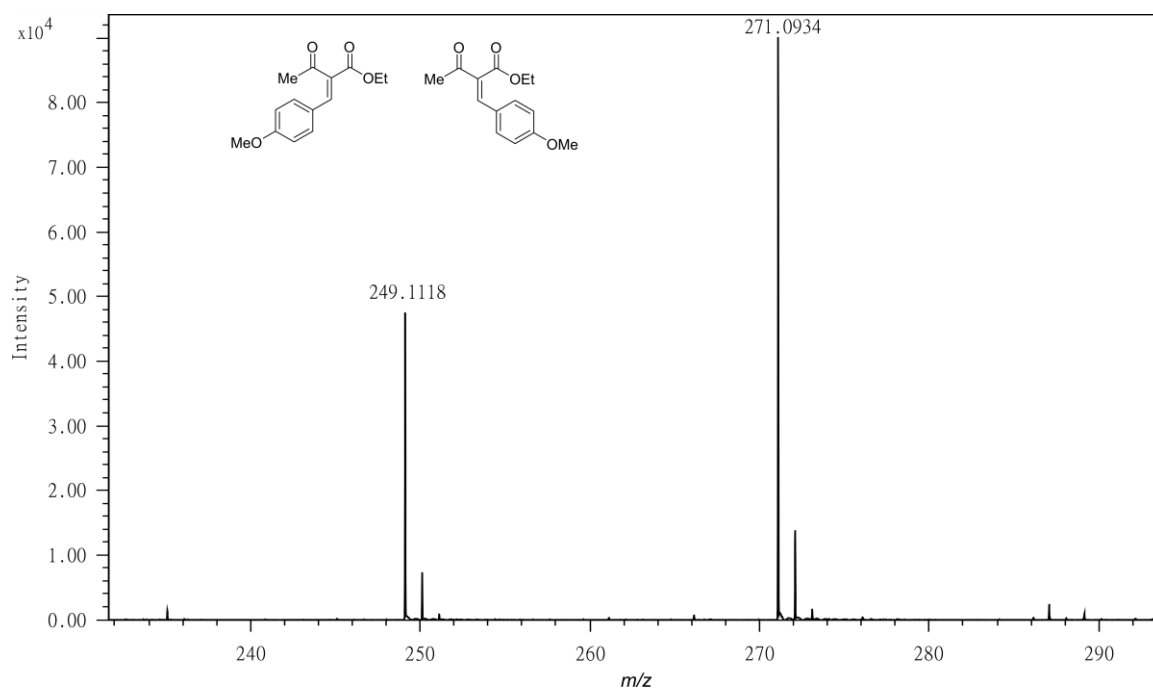
High resolution mass spectrum of 15a (ESI-TOF).



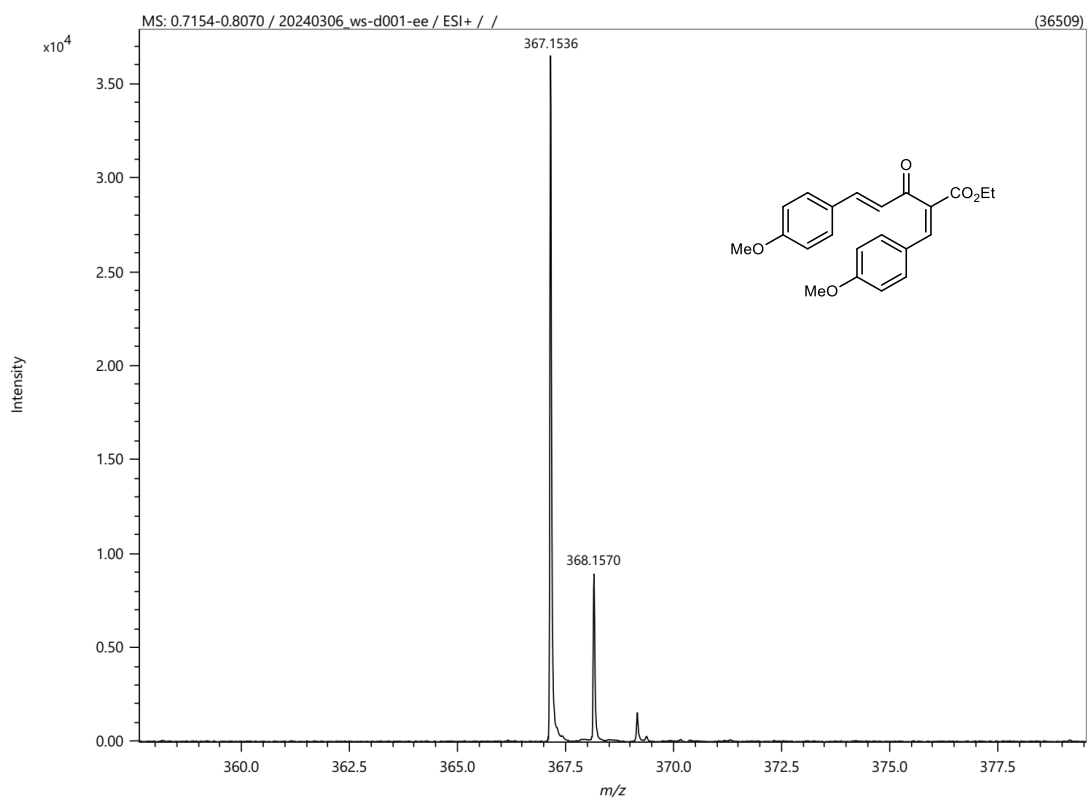
High resolution mass spectrum of 15a (ESI-TOF).



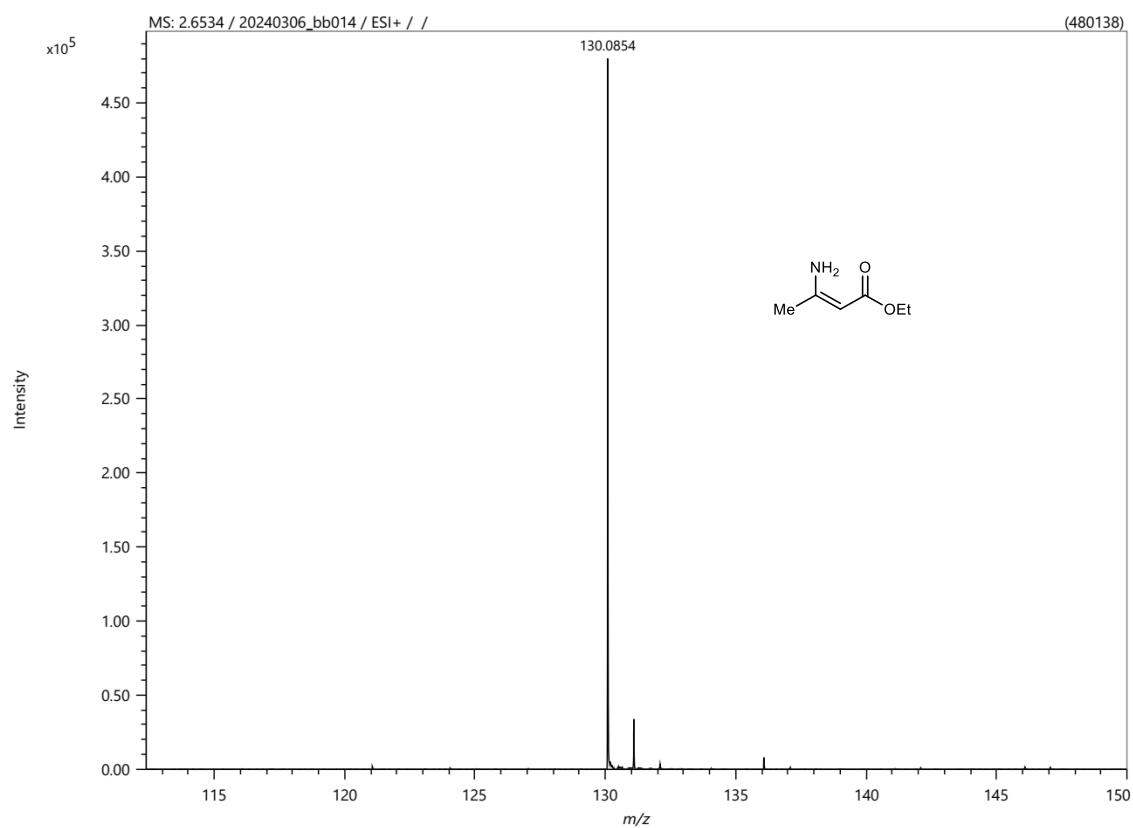
High resolution mass spectrum of 19f and 19g (ESI-TOF).



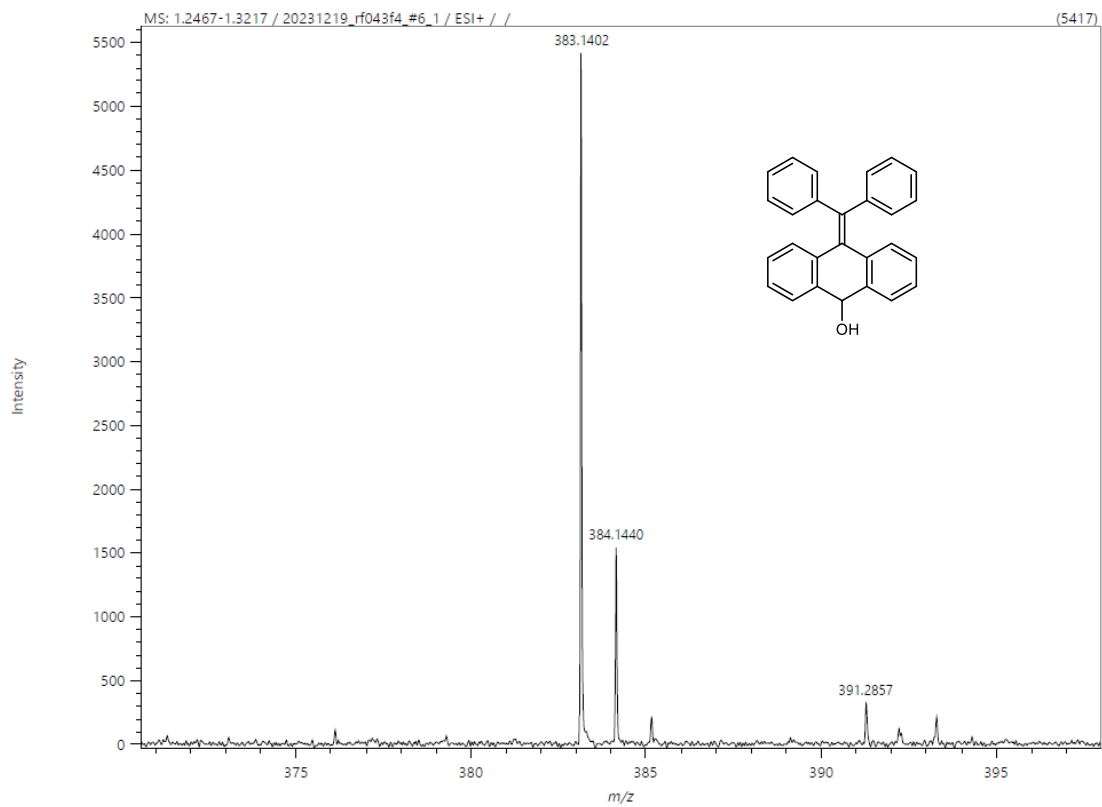
High resolution mass spectrum of 19h (ESI-TOF).



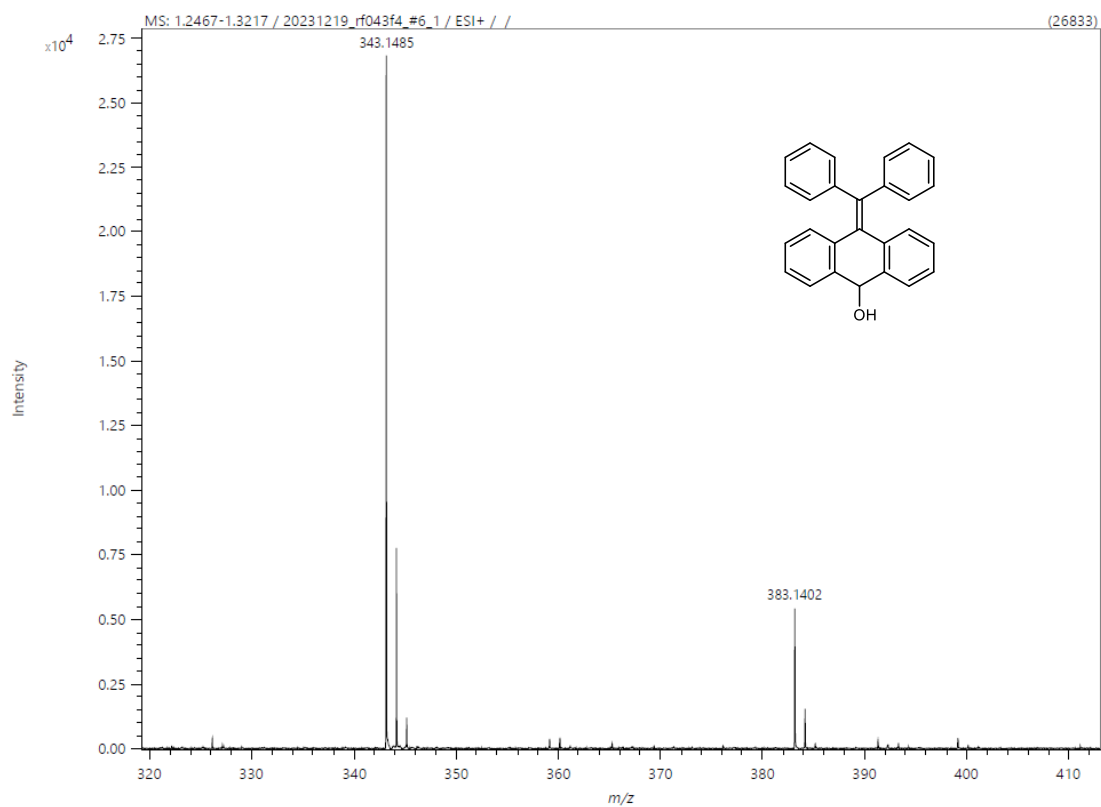
High resolution mass spectrum of EAB (ESI-TOF)



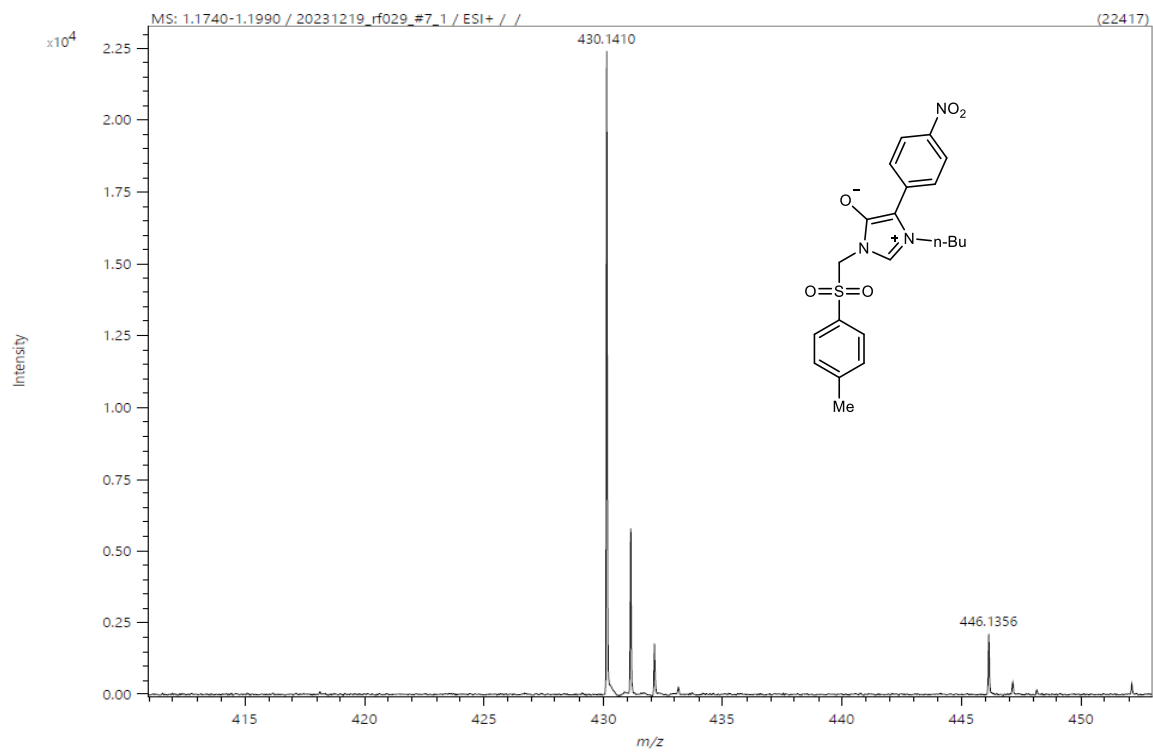
High resolution mass spectrum of 15b (ESI-TOF).



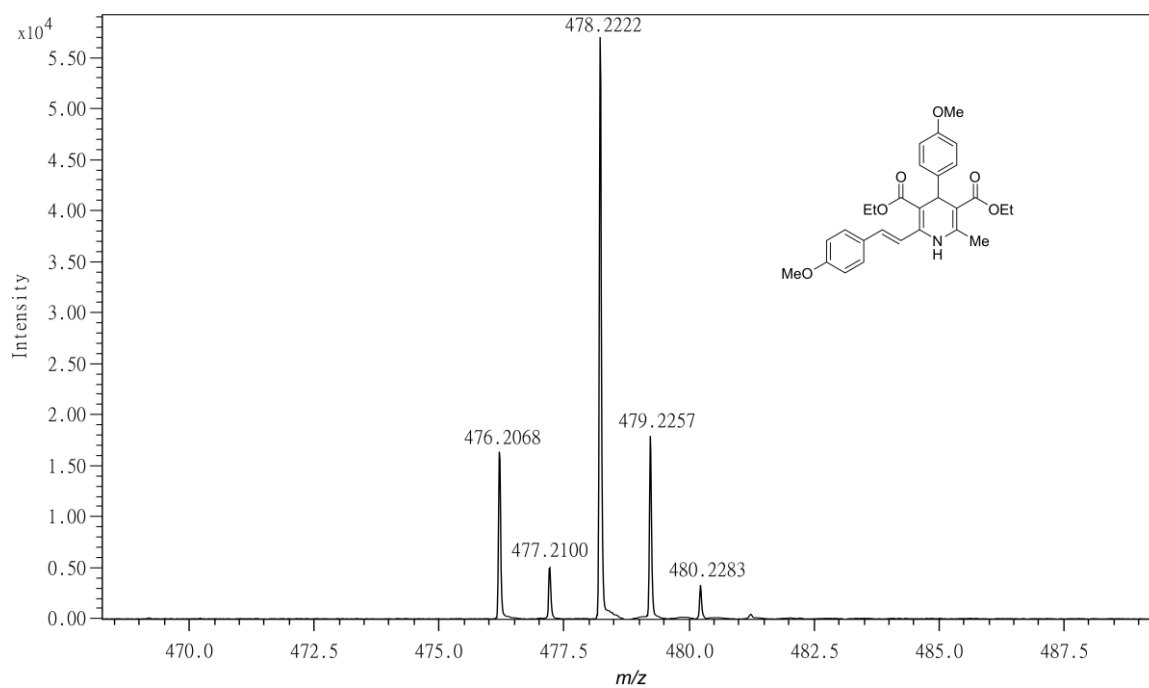
High resolution mass spectrum of 15b (ESI-TOF).



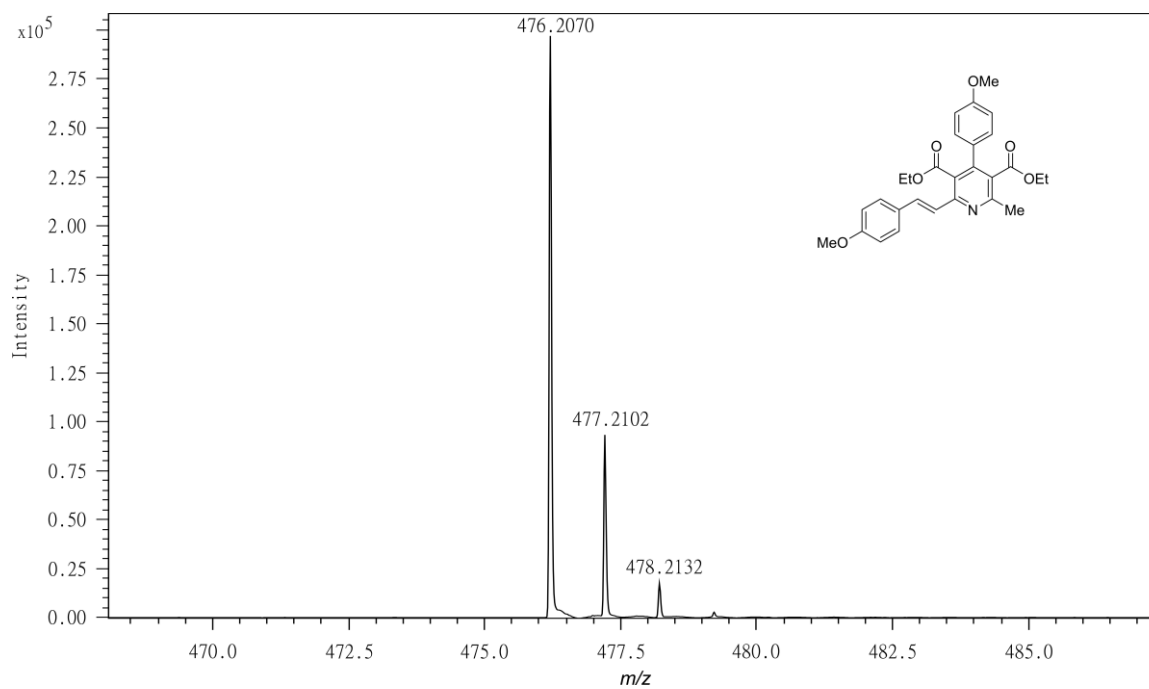
High resolution mass spectrum of 16a (ESI-TOF).



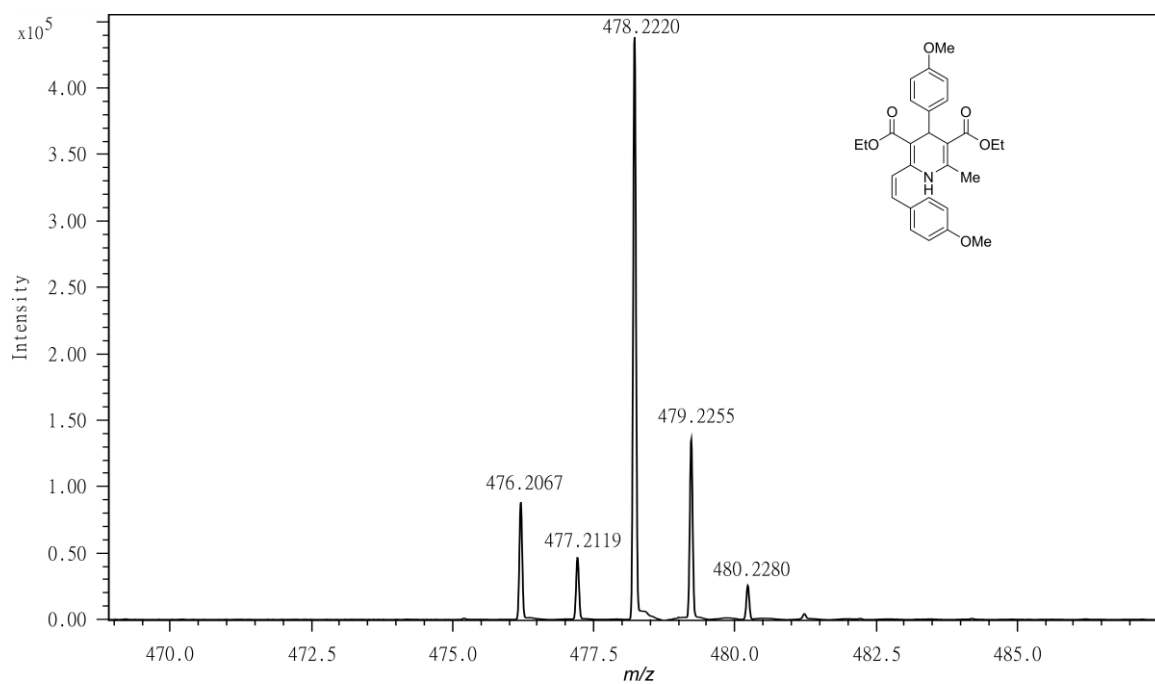
High resolution mass spectrum of 19k (ESI-TOF).



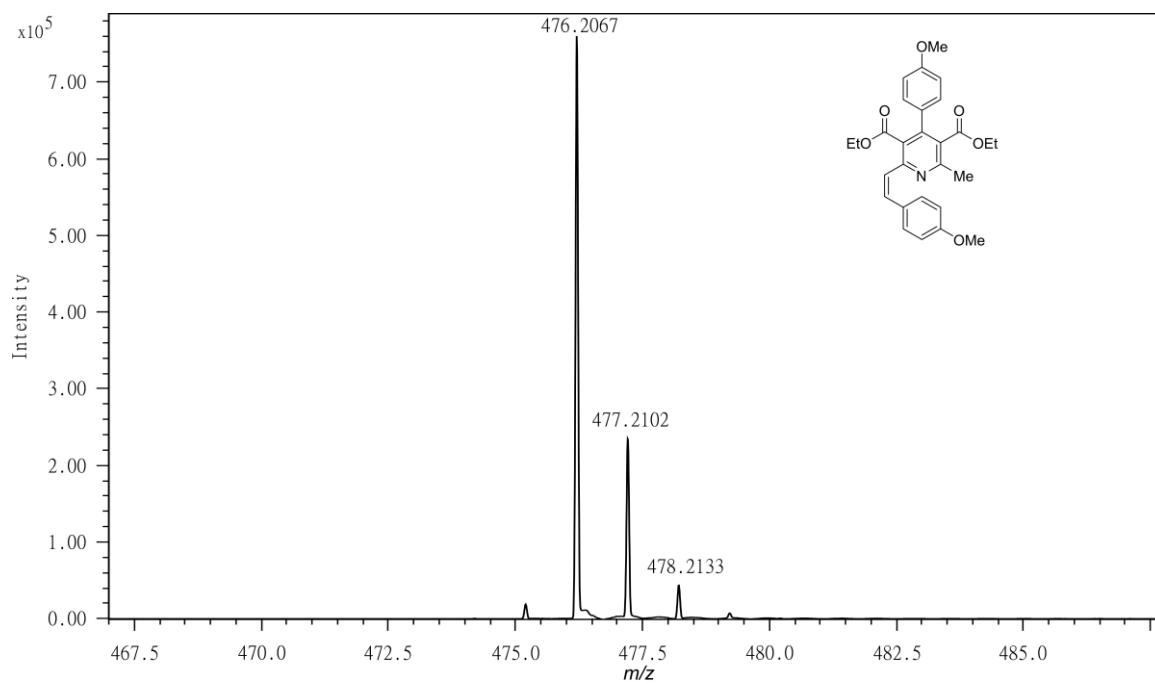
High resolution mass spectrum of 19m.



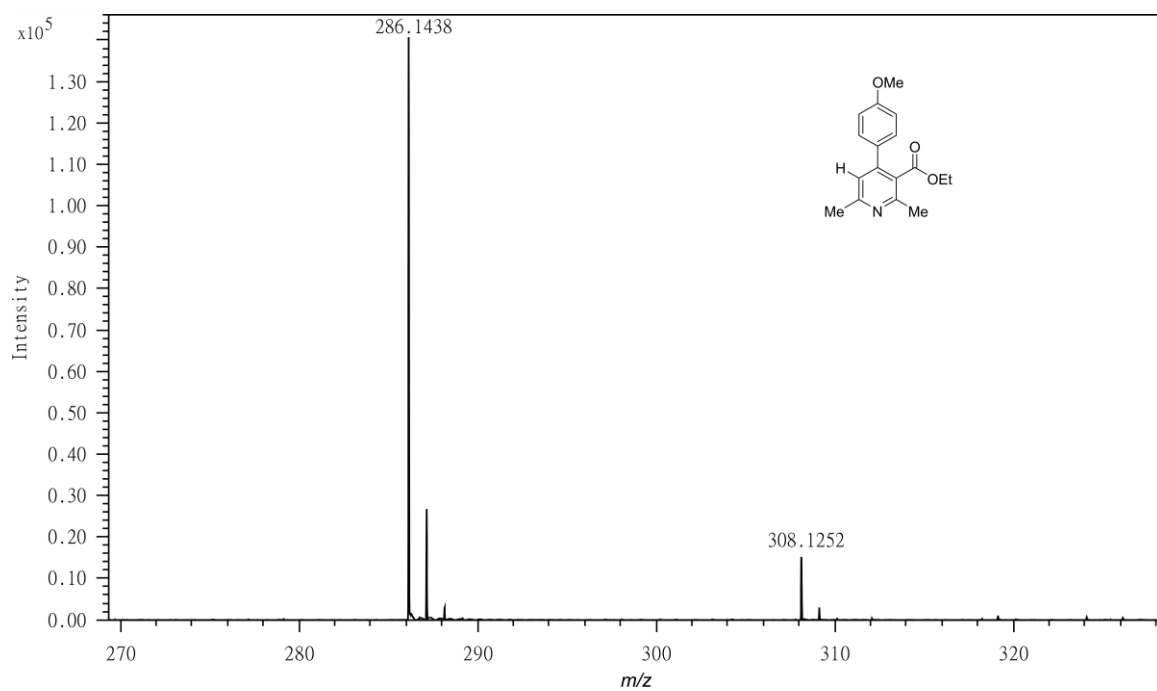
High resolution mass spectrum of 19l (ESI-TOF).



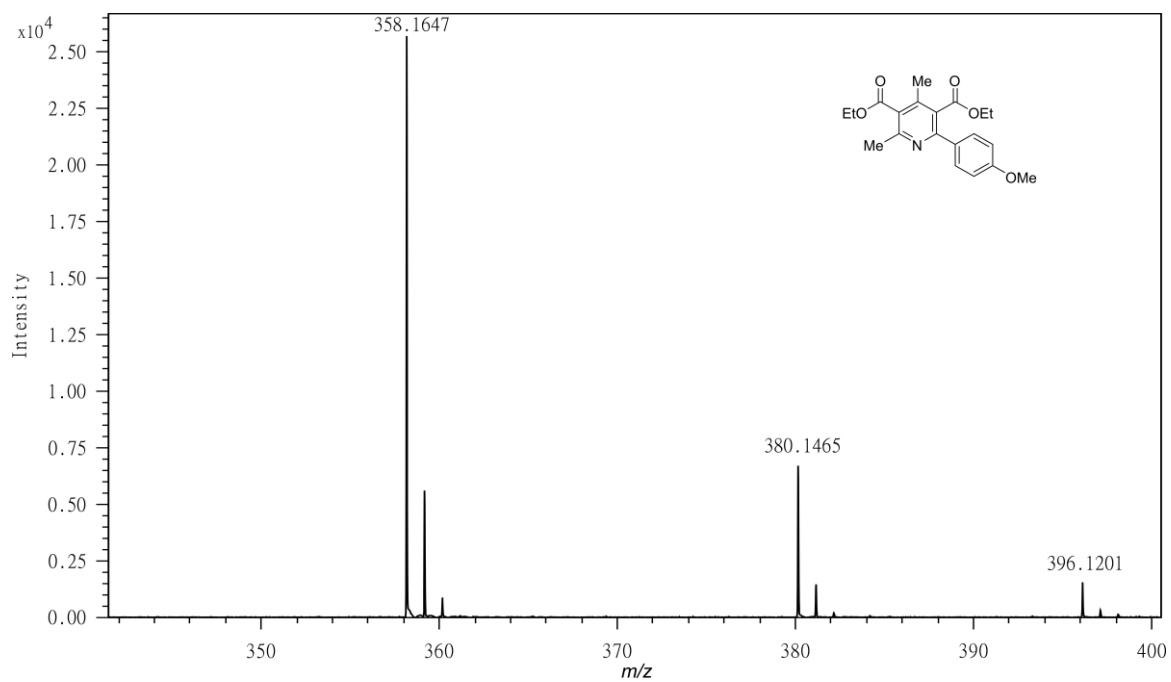
High resolution mass spectrum of 19n (ESI-TOF)



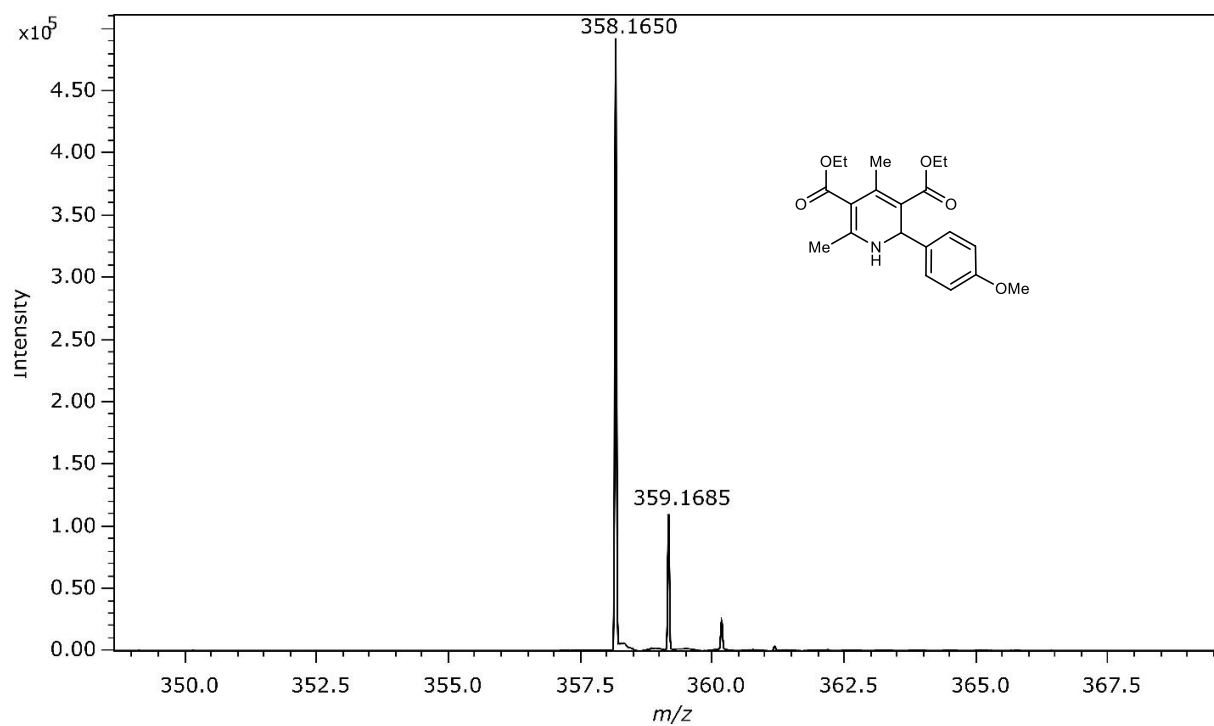
High resolution mass spectrum of 19o (ESI-TOF).



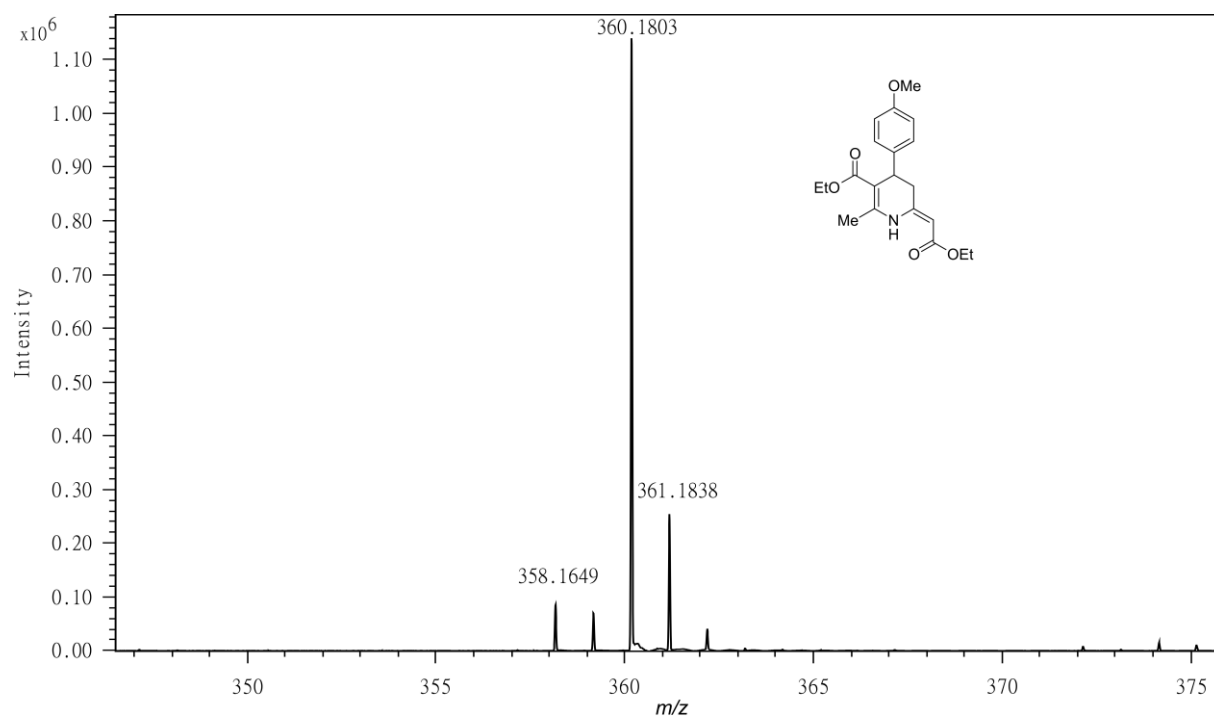
High resolution mass spectrum of 19r (ESI-TOF)



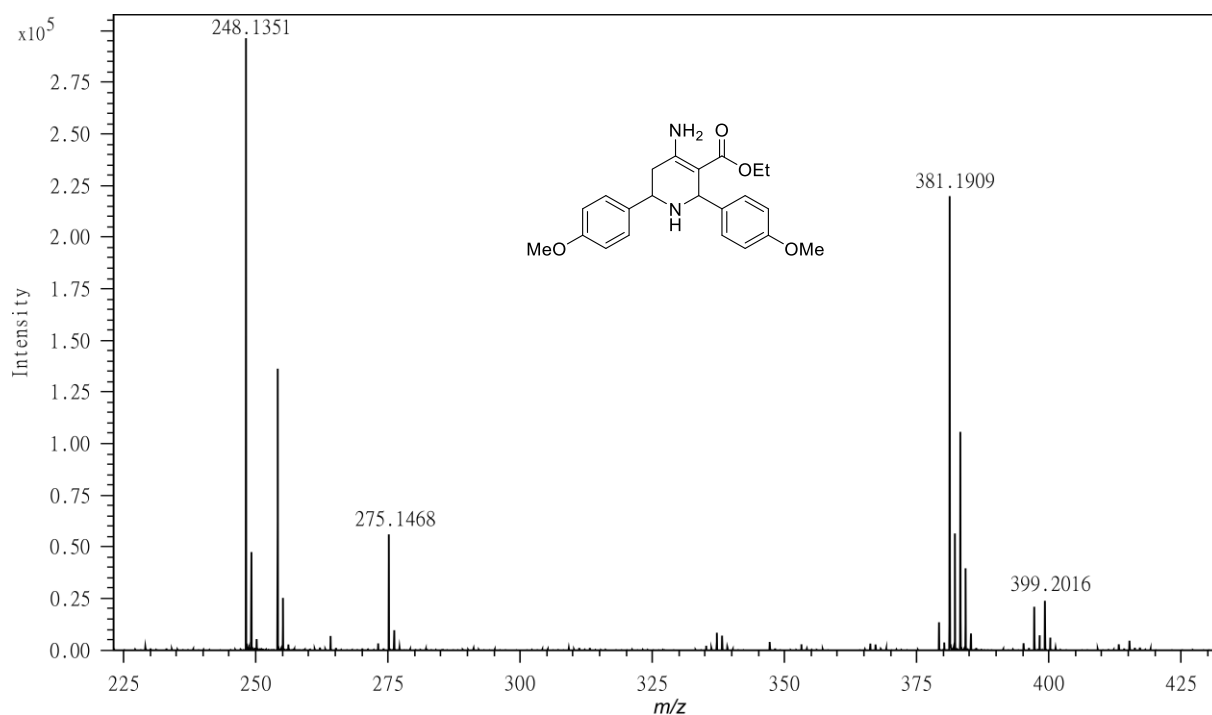
High resolution mass spectrum of **19p**



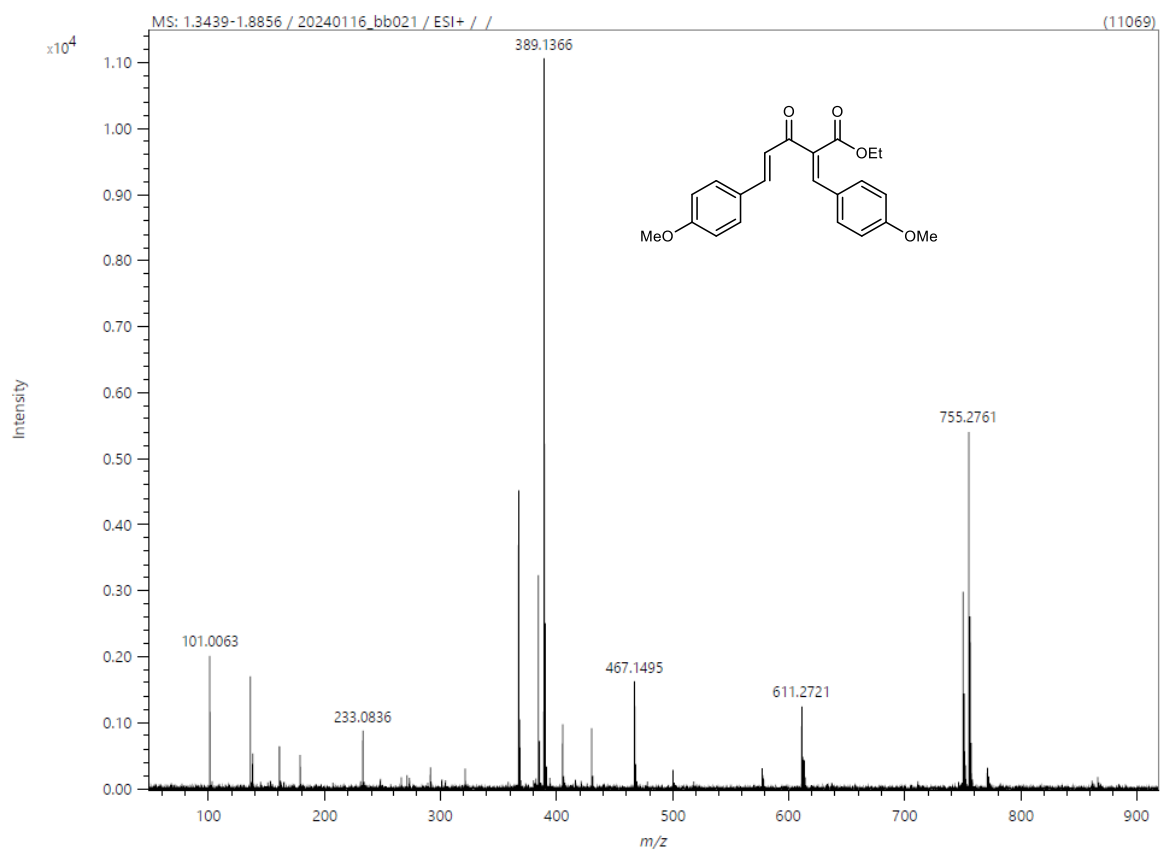
High resolution mass spectrum of **19j** (ESI-TOF)



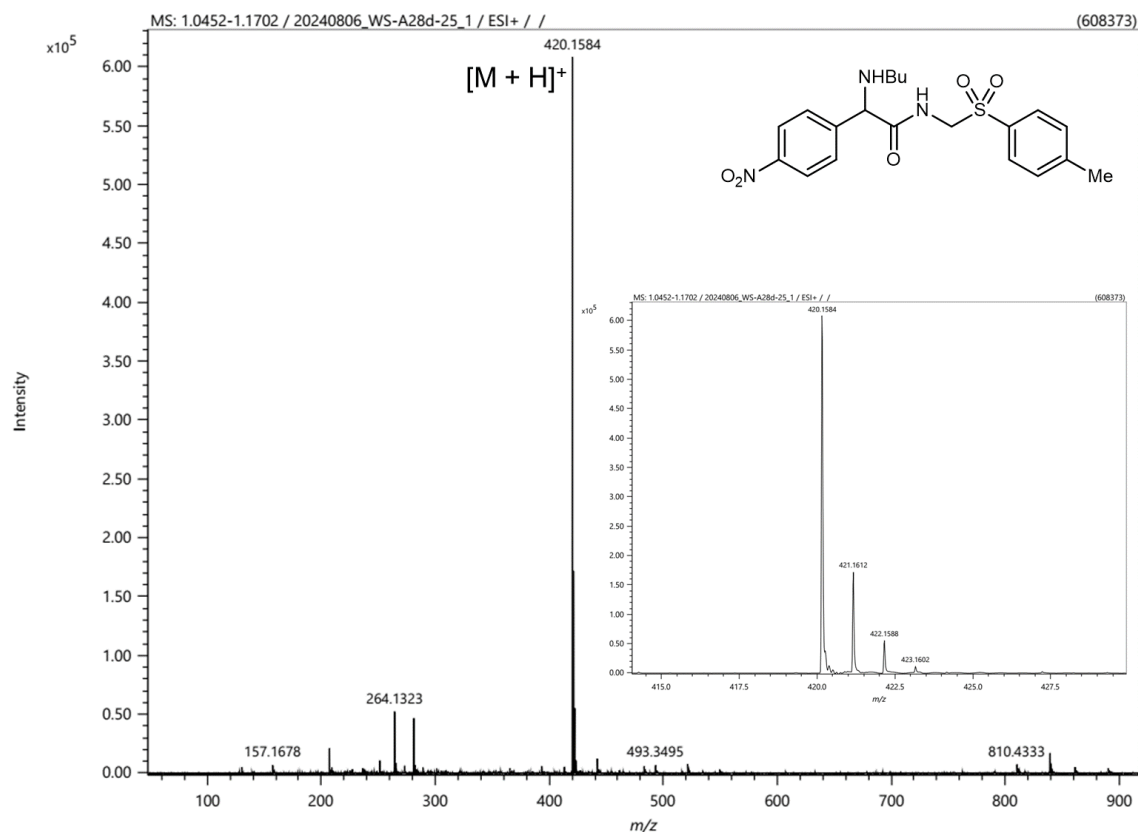
High resolution mass spectrum of **19e** (ESI-TOF)



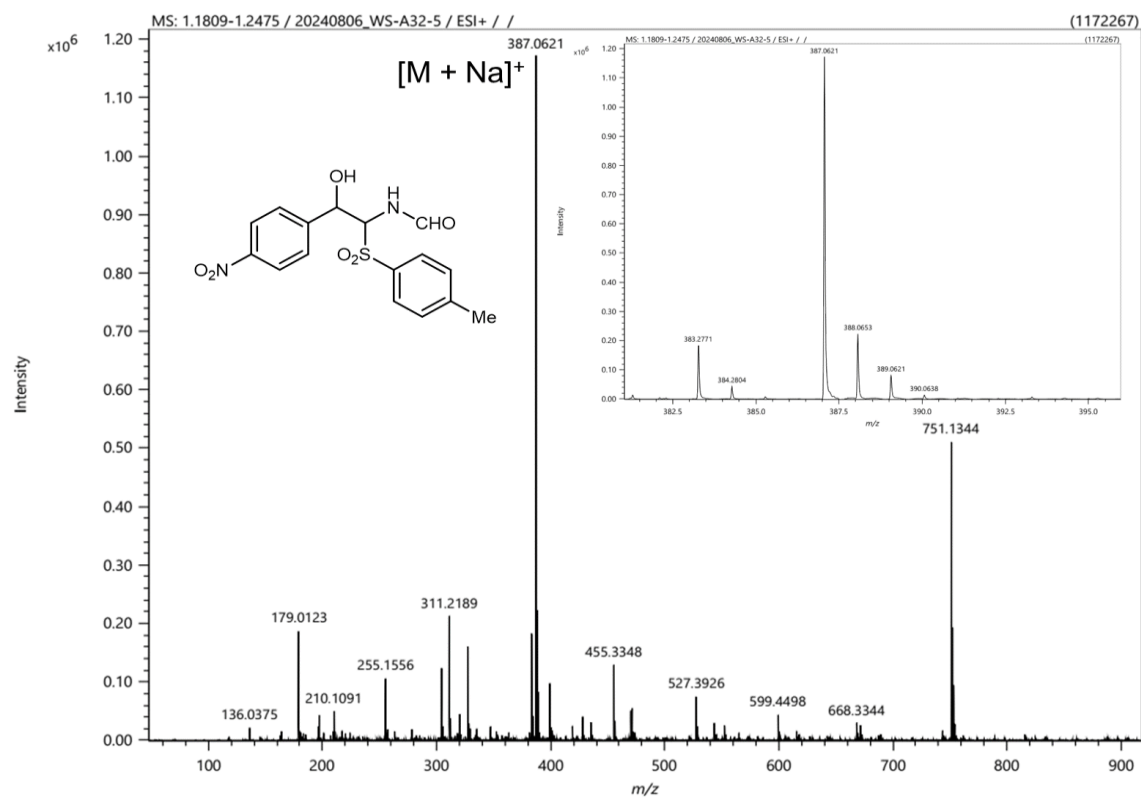
High resolution mass spectrum of **19h** (ESI-TOF).



High-resolution mass spectrum of 2-(butylamino)-2-(4-nitrophenyl)-*N*-(tosylmethyl)acetamide (ESI-TOF)



High-resolution mass spectrum of **17g** (ESI-TOF)



12. NMR spectra

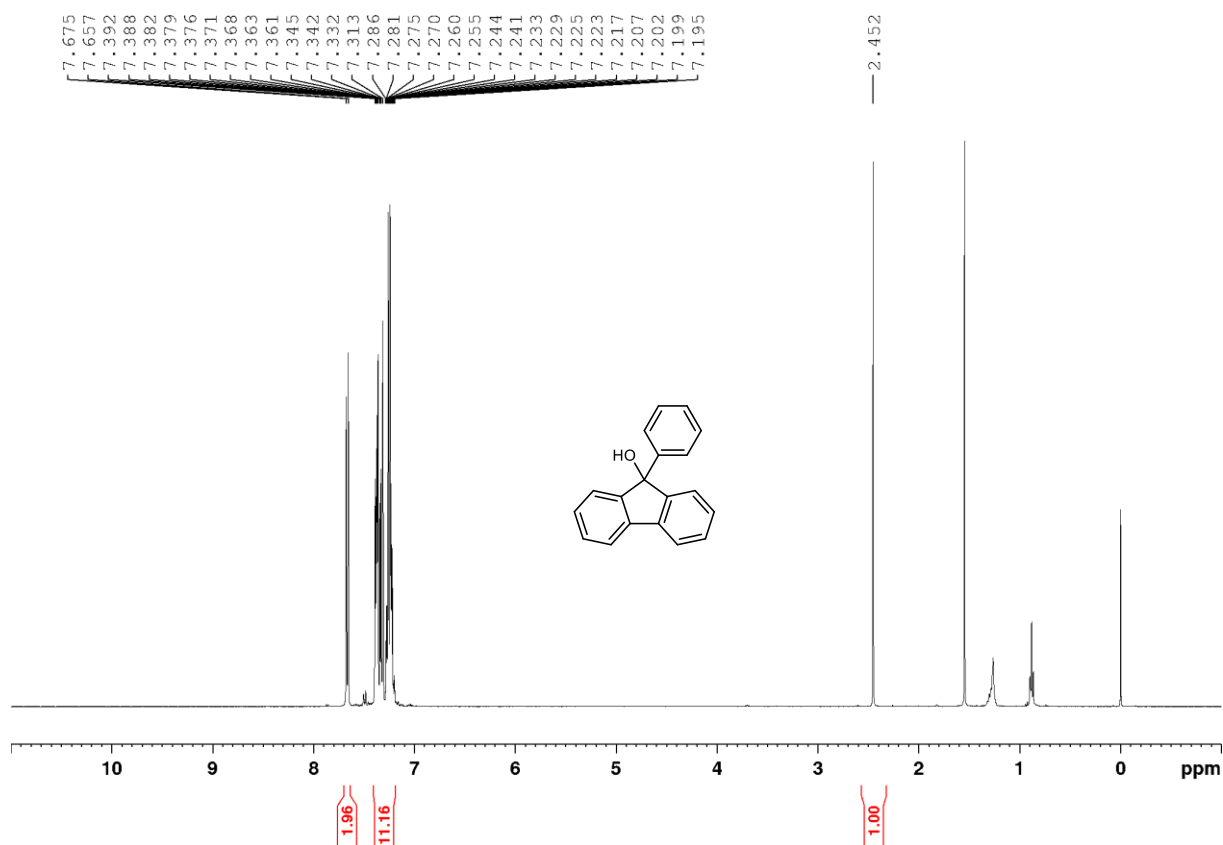
Content list

¹ H NMR spectrum of <u>14a</u> (400 MHz, chloroform- <i>d</i>).....	275
¹³ C NMR spectrum of <u>14a</u> (100 MHz, chloroform- <i>d</i>).....	275
¹ H NMR spectrum of <u>14b</u> (400 MHz, chloroform- <i>d</i>).	276
¹³ C NMR spectrum of <u>14b</u> (100 MHz, chloroform- <i>d</i>).	276
¹ H NMR spectrum of <u>13a</u> (400 MHz, chloroform- <i>d</i>).....	277
¹³ C NMR spectrum of <u>13a</u> (100 MHz, chloroform- <i>d</i>).....	277
¹ H NMR spectrum of <u>13b</u> (400 MHz, chloroform- <i>d</i>).	278
¹³ C NMR spectrum of <u>13b</u> (100 MHz, chloroform- <i>d</i>).	278
¹ H NMR spectrum of <u>15a</u> (400 MHz, chloroform- <i>d</i>).....	279
¹³ C NMR spectrum of <u>15a</u> (100 MHz, chloroform- <i>d</i>).....	279
¹ H NMR spectrum of <u>19d</u> (400 MHz, chloroform- <i>d</i>)	280
¹³ C NMR spectrum of <u>19d</u> (100 MHz, chloroform- <i>d</i>)	280
¹ H NMR spectrum of <u>19f</u> (400 MHz, chloroform- <i>d</i>).	281
¹³ C NMR spectrum of <u>19f</u> (100 MHz, chloroform- <i>d</i>).....	281
¹ H NMR spectrum of <u>19g</u> (400 MHz, chloroform- <i>d</i>).	282
¹³ C NMR spectrum of <u>19g</u> (100 MHz, chloroform- <i>d</i>).	282
¹ H NMR spectrum of <u>19h</u> (400 MHz, chloroform- <i>d</i>).	283
¹³ C NMR spectrum of <u>19h</u> (100 MHz, chloroform- <i>d</i>).	283
NOESY spectrum of <u>19h</u> (chloroform- <i>d</i>).....	284
HSQC–TOCSY spectrum of <u>19h</u>	285
¹ H NMR spectrum of EAB (400 MHz, DMSO- <i>d</i> ₆).	286
¹³ C NMR spectrum of EAB (100 MHz, DMSO- <i>d</i> ₆).	286
¹ H NMR spectrum of <u>15b</u> (400 MHz, dichloromethene- <i>d</i> ₂).	287
¹³ C NMR spectrum of <u>15b</u> (100 MHz, dichloromethene- <i>d</i> ₂).	287
¹ H NMR spectrum of <u>15g</u> (400 MHz, DMSO- <i>d</i> ₆).	288
¹ H NMR spectrum of <u>16a</u> (400 MHz, DMSO- <i>d</i> ₆).....	288
¹³ C NMR spectrum of <u>16a</u> (100 MHz, DMSO- <i>d</i> ₆).....	289
¹ H NMR spectrum of <u>19k</u> (400 MHz, chloroform- <i>d</i>)	289
¹³ C NMR spectrum of <u>19k</u> (100 MHz, chloroform- <i>d</i>)	290
COSY spectrum of <u>19k</u> (chloroform- <i>d</i>)	291
NOESY spectrum of <u>19k</u> (chloroform- <i>d</i>).....	292
HSQC–TOCSY spectrum of <u>19k</u> (chloroform- <i>d</i>).	293

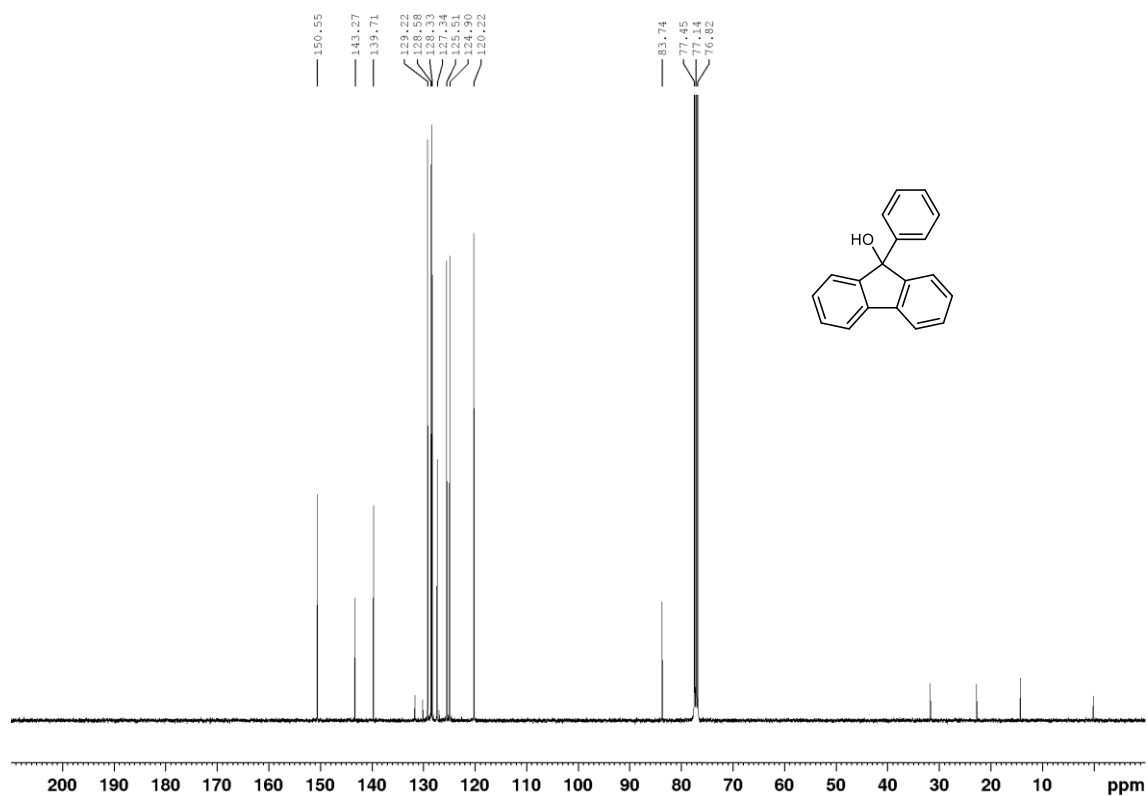
HMBC spectrum of <u>19k</u> (chloroform- <i>d</i>)	294
¹ H NMR spectrum of <u>19m</u> (400 MHz, chloroform- <i>d</i>)	295
¹³ C NMR spectrum of <u>19m</u> (100 MHz, chloroform- <i>d</i>)	295
¹ H NMR spectrum of <u>19l</u> (400 MHz, chloroform- <i>d</i>)	296
¹³ C NMR spectrum of <u>19l</u> (100 MHz, chloroform- <i>d</i>)	296
NOESY spectrum of <u>19l</u> (chloroform- <i>d</i>).	297
HSQC–TOCSY spectrum of <u>19l</u> (chloroform- <i>d</i>).	298
HMBC spectrum of <u>19l</u> (chloroform- <i>d</i>)	299
¹ H NMR spectrum of <u>19n</u> (400 MHz, chloroform- <i>d</i>).	300
¹³ C NMR spectrum of <u>19n</u> (100 MHz, chloroform- <i>d</i>)	300
NOESY spectrum of <u>19n</u> (chloroform- <i>d</i>).....	301
HSQC–TOCSY spectrum of <u>19n</u> (chloroform- <i>d</i>)	302
HMBC spectrum of <u>19n</u> (chloroform- <i>d</i>)	303
¹ H NMR spectrum of <u>19o</u> (400 MHz, chloroform- <i>d</i>)	304
¹³ C NMR spectrum of <u>19o</u> (100 MHz, chloroform- <i>d</i>)	304
NOESY spectrum of <u>19o</u> (chloroform- <i>d</i>).....	305
HSQC–TOCSY spectrum of <u>19o</u> (chloroform- <i>d</i>)	306
HMBC spectrum of <u>19o</u> (chloroform- <i>d</i>).	307
¹ H NMR spectrum of <u>19r</u> (400 MHz, chloroform- <i>d</i>)	308
¹³ C NMR spectrum of <u>19r</u> (100 MHz, chloroform- <i>d</i>).	308
NOESY spectrum of <u>19r</u> (chloroform- <i>d</i>)	309
HSQC–TOCSY spectrum of <u>19r</u> (chloroform- <i>d</i>).	310
HMBC spectrum of <u>19r</u> (chloroform- <i>d</i>).	311
¹ H NMR spectrum of <u>19p</u> (400 MHz, chloroform- <i>d</i>).	312
¹³ C NMR spectrum of <u>19p</u> (100 MHz, chloroform- <i>d</i>).	312
¹ H NMR spectrum of <u>19j</u> (400 MHz, chloroform- <i>d</i>)	313
¹³ C NMR spectrum of <u>19j</u> (100 MHz, chloroform- <i>d</i>)	313
COSY spectrum of <u>19j</u> (chloroform- <i>d</i>)	314
NOESY spectrum of <u>19j</u> (chloroform- <i>d</i>)	315
HSQC–TOCSY spectrum of <u>19j</u> (chloroform- <i>d</i>).	316
HMBC spectrum of <u>19j</u> (chloroform- <i>d</i>).	317
¹ H NMR spectrum of <u>19e</u> (400 MHz, chloroform- <i>d</i>).....	318
¹³ C NMR spectrum of <u>19e</u> (100 MHz, chloroform- <i>d</i>).....	318
NOESY spectrum of <u>19e</u> (chloroform- <i>d</i>).....	319
HSQC–TOCSY spectrum of <u>19e</u> (chloroform- <i>d</i>).	320

HMBC spectrum of <u>19e</u> (chloroform- <i>d</i>).	321
¹ H NMR spectrum of <u>19h</u> (400 MHz, chloroform- <i>d</i>).	322
¹³ C NMR spectrum of <u>19h</u> (100 MHz, chloroform- <i>d</i>).	322
¹ H NMR (400 MHz, CDCl ₃ , 300 K) spectrum of sample 1 in Figure S66	323
¹ H NMR (400 MHz, CDCl ₃ , 300 K) spectrum of sample 2 in Figure S66	323
¹ H NMR (400 MHz, CDCl ₃ , 300 K) spectrum of sample 3 in Figure S66	324
¹ H NMR (400 MHz, CDCl ₃ , 300 K) spectrum of sample 4 in Figure S66	324
¹ H NMR (400 MHz, CDCl ₃ , 300 K) spectrum of sample 5 in Figure S66	325
¹ H NMR (400 MHz, CDCl ₃ , 300 K) spectrum of sample 6 in Figure S66	325
¹ H NMR spectrum of <u>19e</u> -BOC (400 MHz, DMSO- <i>d</i> ₆).	326
¹ H NMR spectrum of <u>19e</u> -BOC with addition of D ₂ O (400 MHz, DMSO- <i>d</i> ₆).	326
¹³ C NMR spectrum of <u>19e</u> -BOC (100 MHz, DMSO- <i>d</i> ₆).	327
ROESY spectrum of <u>19e</u> -BOC (DMSO- <i>d</i> ₆).	328
HSQC–TOCSY spectrum of <u>19e</u> -BOC (DMSO- <i>d</i> ₆).	329
HMBC spectrum of <u>19e</u> -BOC (DMSO- <i>d</i> ₆).	330
¹ H NMR spectrum of <u>19i</u> (400 MHz, DCM- <i>d</i> ₂).	331
¹³ C NMR spectrum of <u>19i</u> (100 MHz, DCM- <i>d</i> ₂).	331
NOESY spectrum of <u>19i</u> (DCM- <i>d</i> ₂).	332
HSQC–TOCSY spectrum of <u>19i</u> (DCM- <i>d</i> ₂).	333
¹ H NMR spectrum of <u>17e</u> (400 MHz, CDCl ₃)	334
¹³ C NMR spectrum of <u>17e</u> (100 MHz, CDCl ₃)	334
¹ H NMR spectrum of 2-(Butylamino)-2-(4-nitrophenyl)- <i>N</i> -(tosylmethyl)acetamide (400 MHz, DMSO- <i>d</i> ₆)	335
¹³ C NMR spectrum of 2-(Butylamino)-2-(4-nitrophenyl)- <i>N</i> -(tosylmethyl)acetamide (100 MHz, DMSO- <i>d</i> ₆).	335
¹ H NMR spectrum of <u>17g</u> (600 MHz, DMSO- <i>d</i> ₆)	336
¹³ C NMR spectrum of <u>17g</u> (150 MHz, DMSO- <i>d</i> ₆).	336

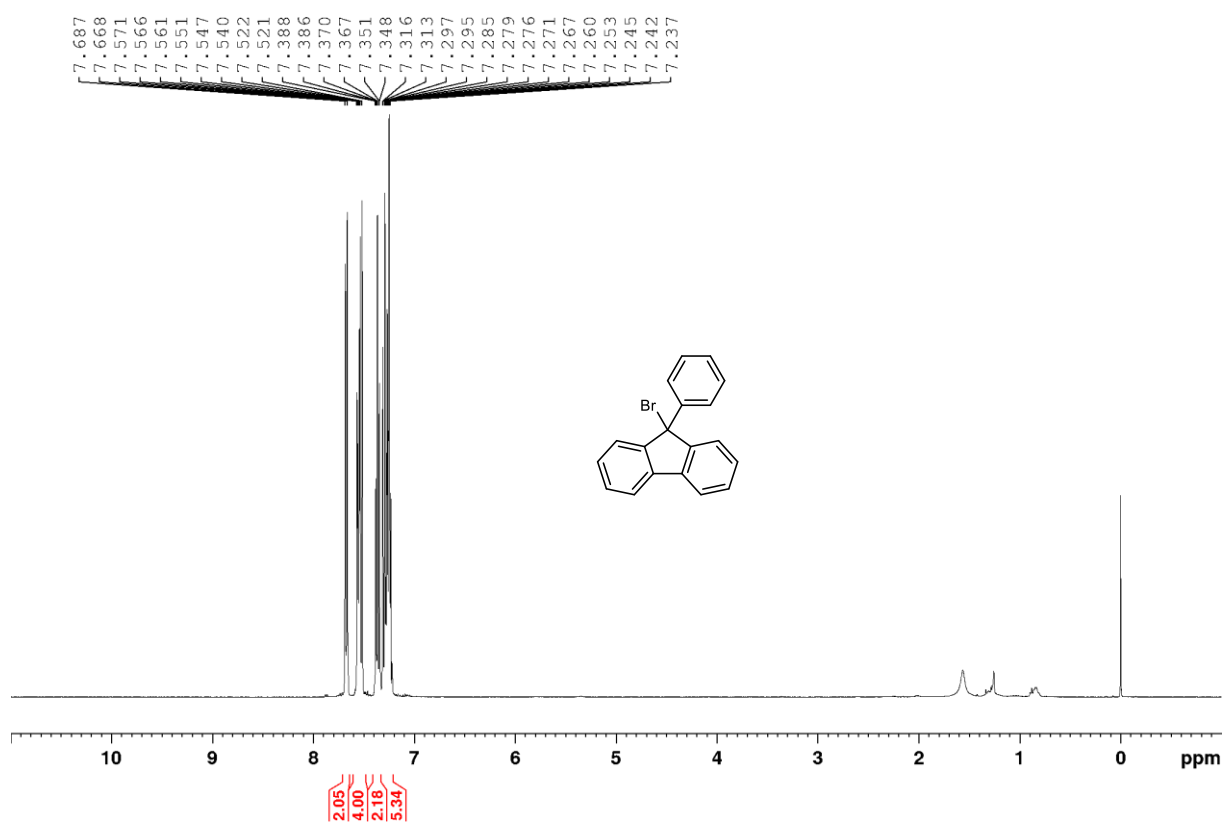
^1H NMR spectrum of 14a (400 MHz, chloroform-*d*)



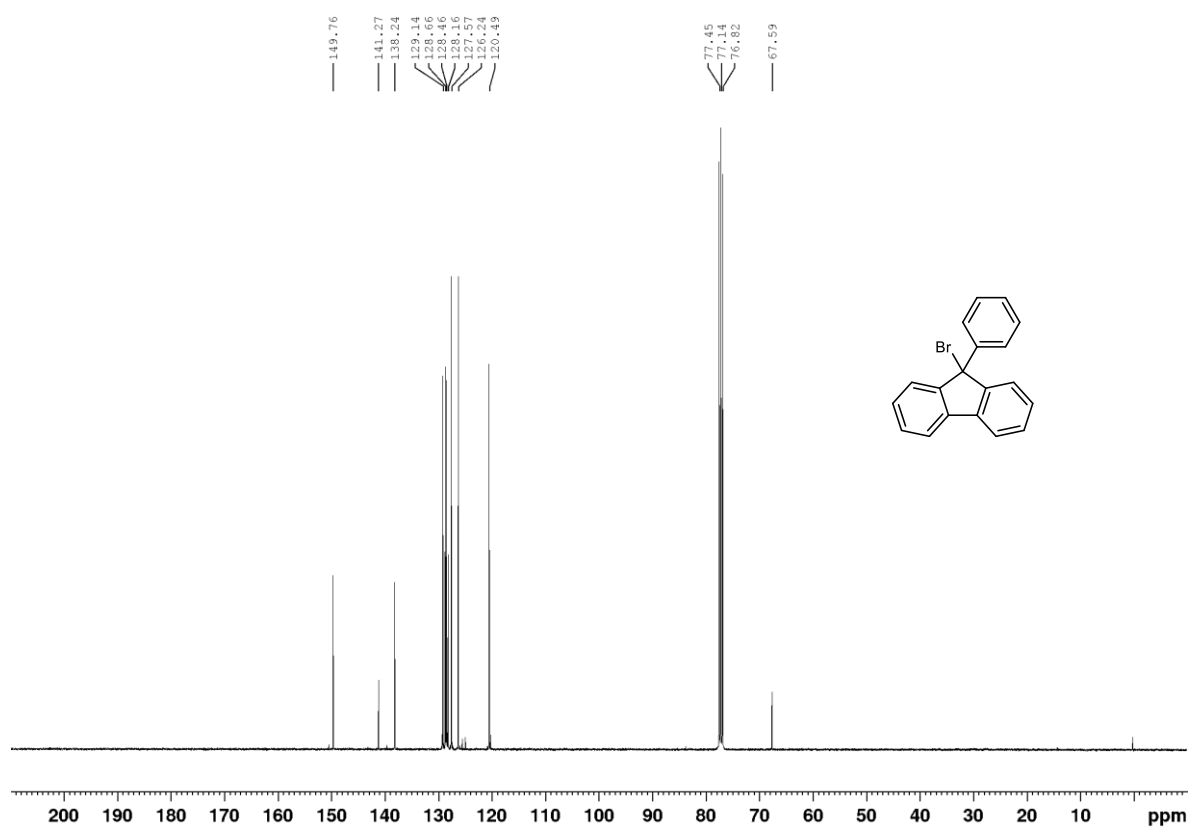
^{13}C NMR spectrum of 14a (100 MHz, chloroform-*d*).



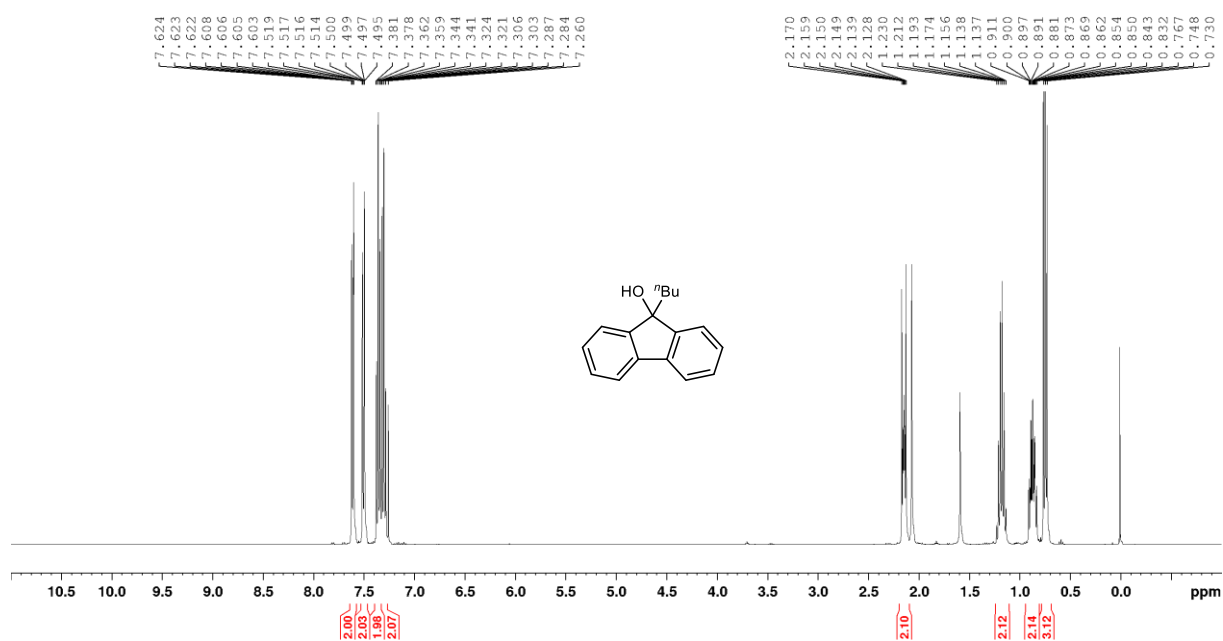
¹H NMR spectrum of 14b (400 MHz, chloroform-*d*).



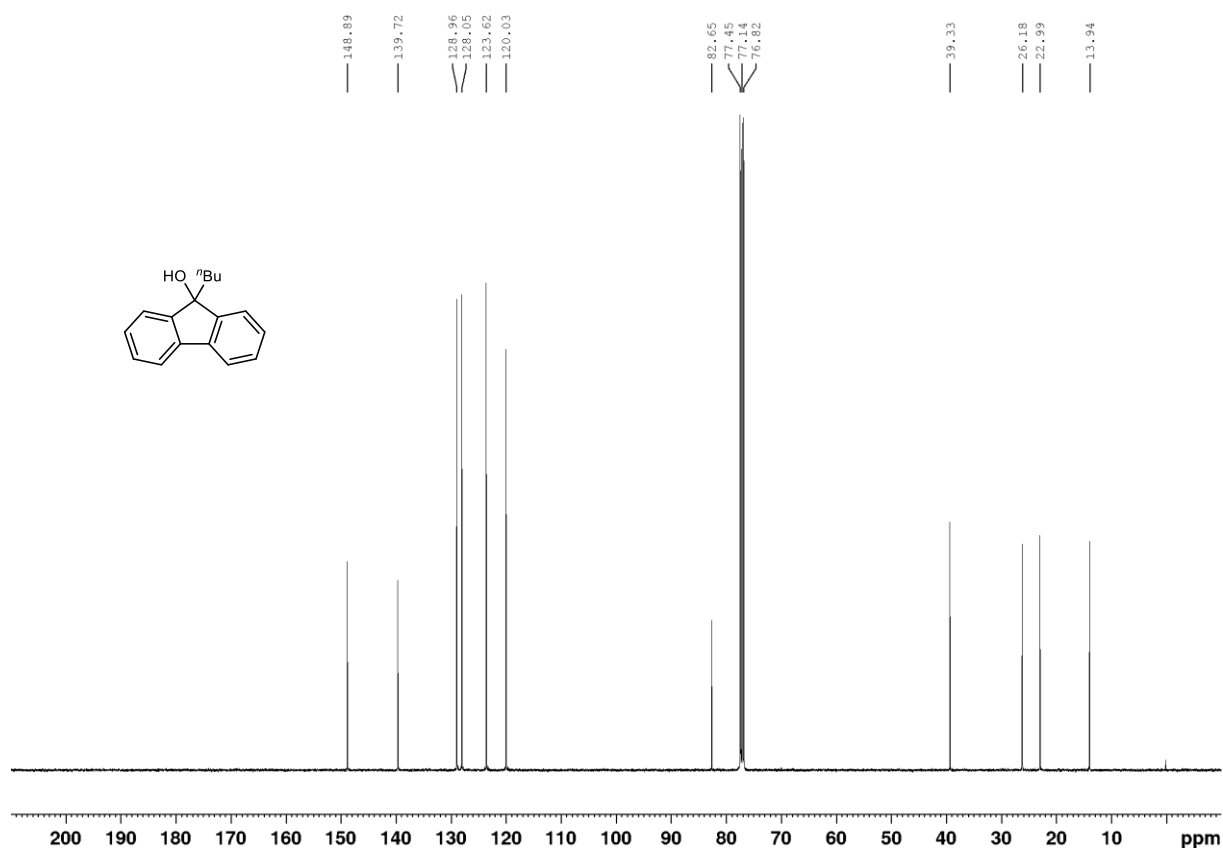
¹³C NMR spectrum of 14b (100 MHz, chloroform-*d*).



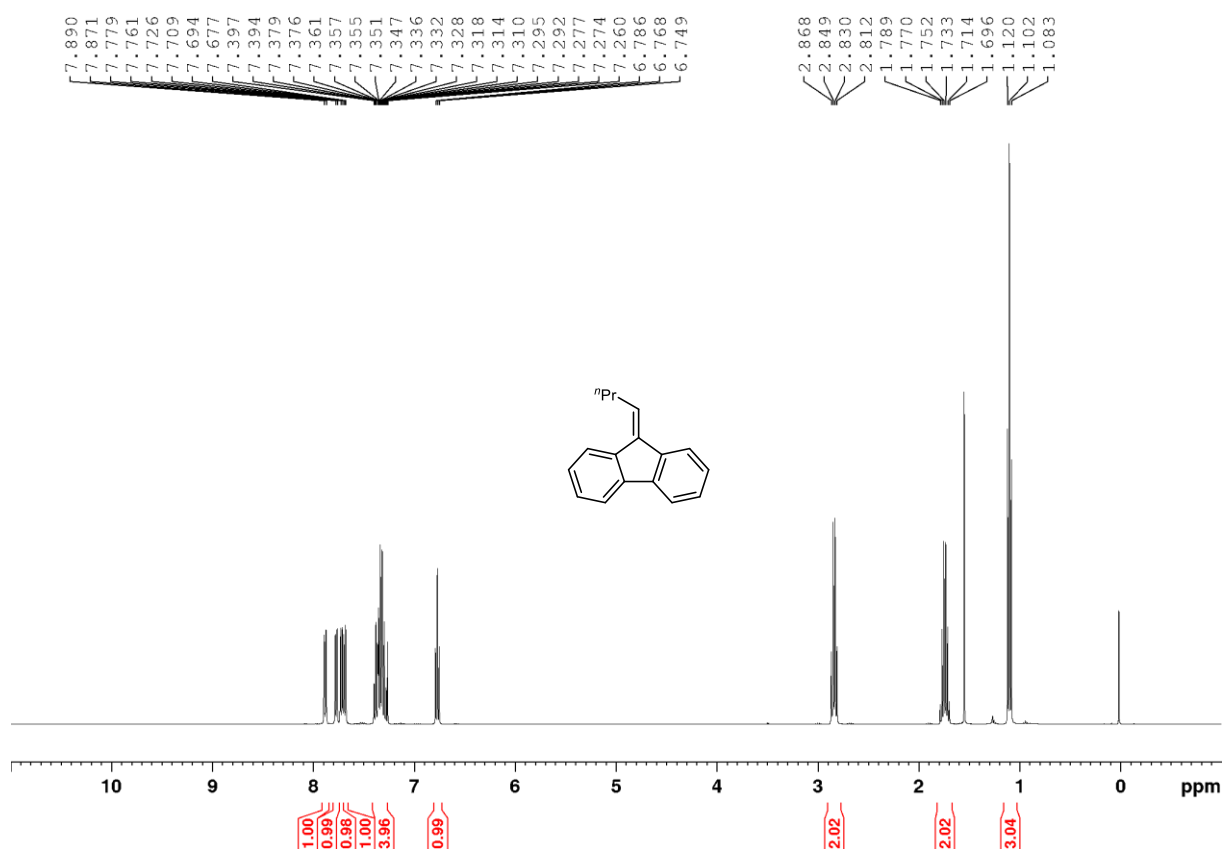
^1H NMR spectrum of 13a (400 MHz, chloroform-*d*).



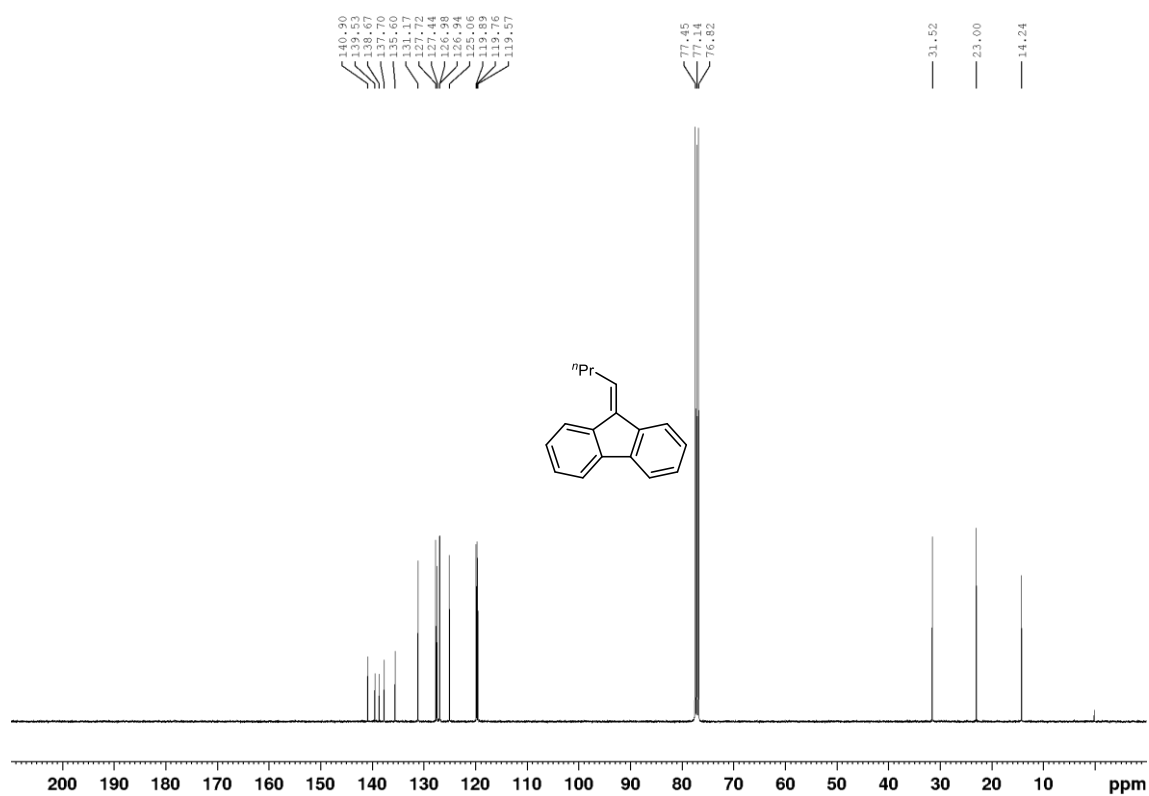
^{13}C NMR spectrum of 13a (100 MHz, chloroform-*d*).



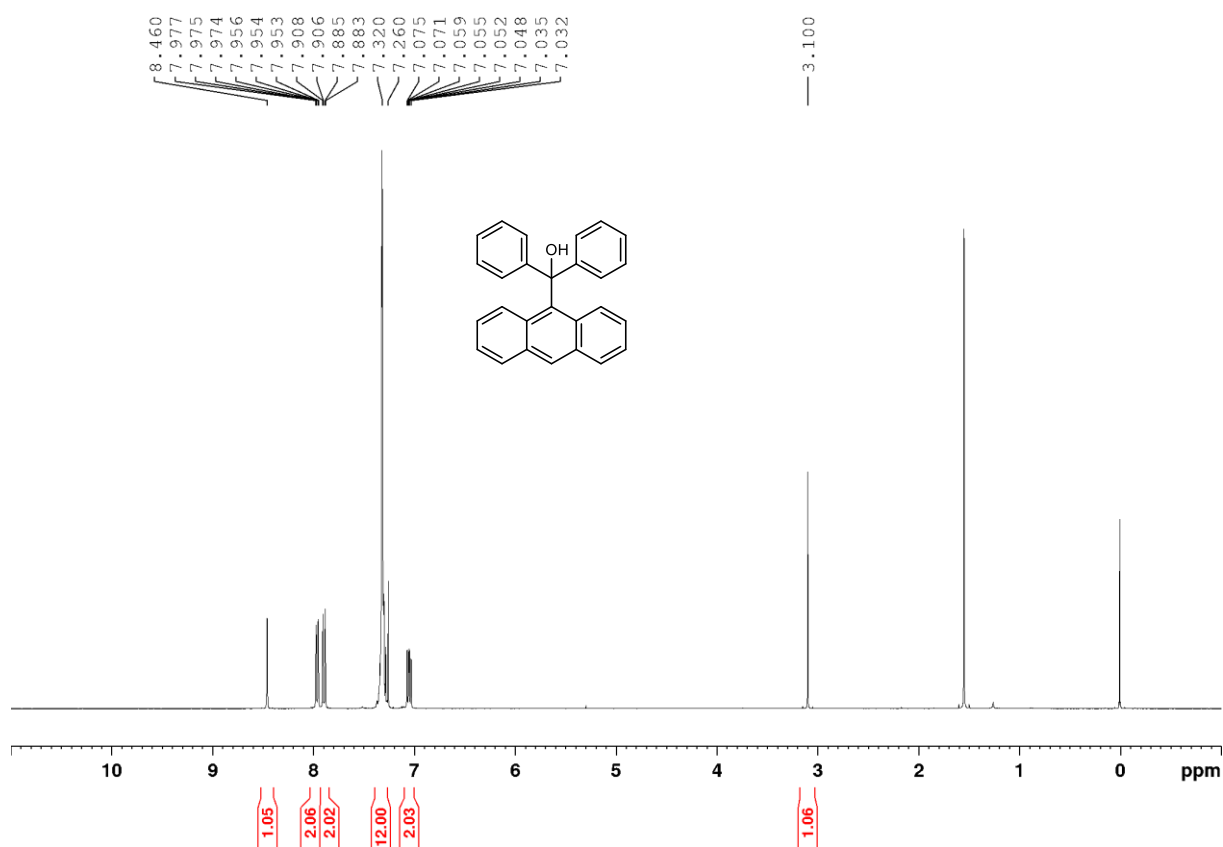
^1H NMR spectrum of **13b (400 MHz, chloroform-*d*).**



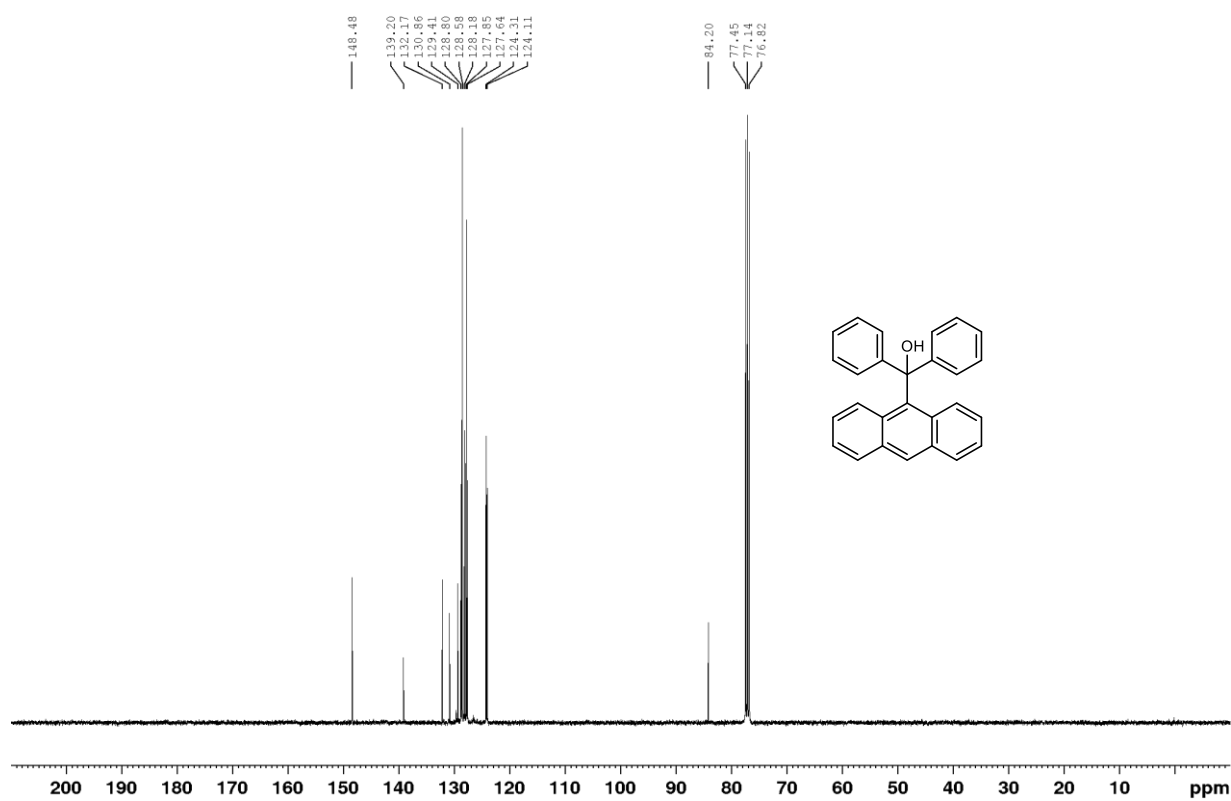
^{13}C NMR spectrum of **13b (100 MHz, chloroform-*d*).**



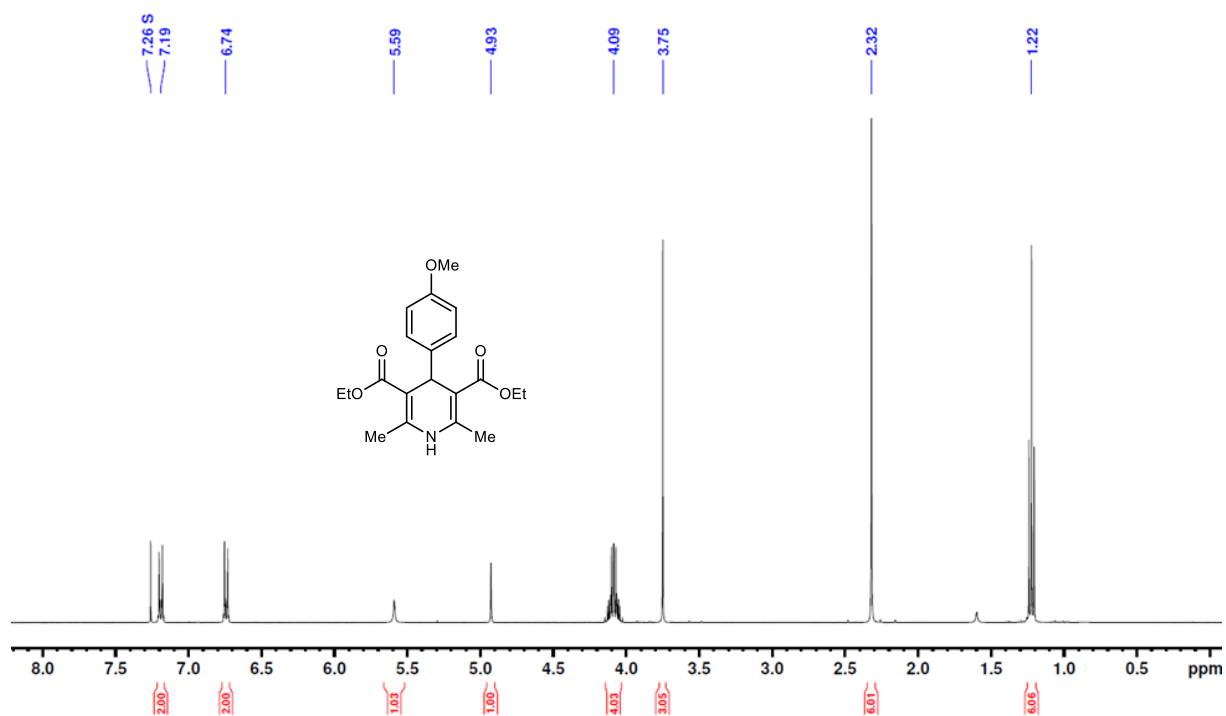
¹H NMR spectrum of 15a (400 MHz, chloroform-*d*).



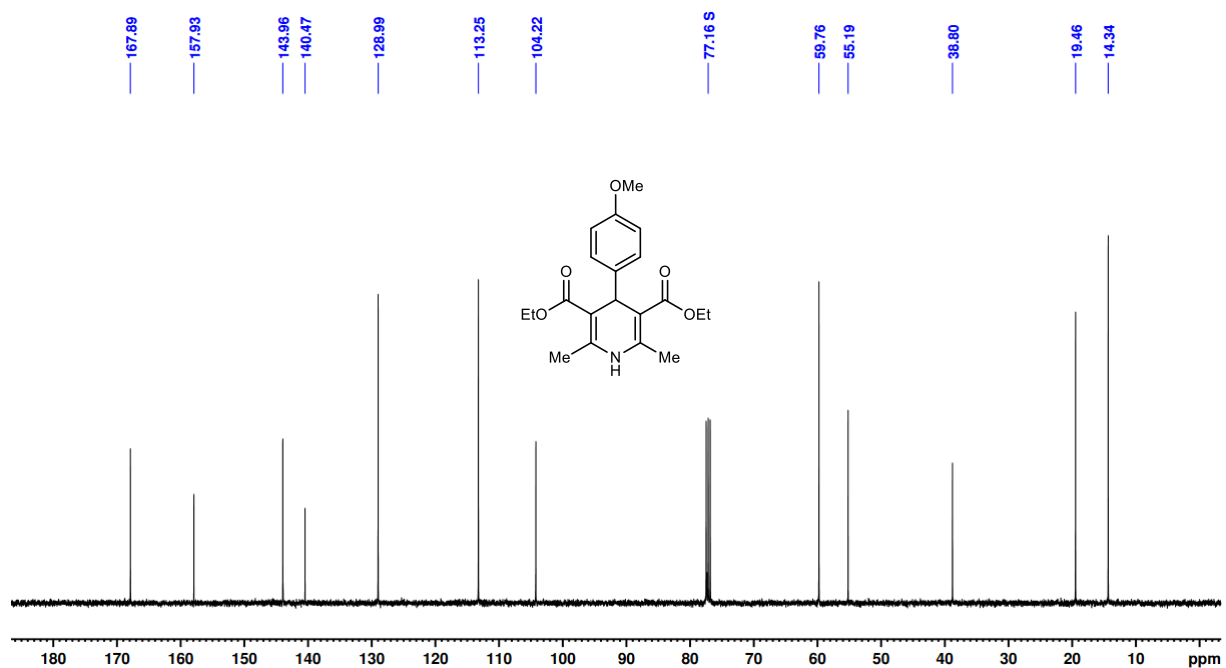
¹³C NMR spectrum of 15a (100 MHz, chloroform-*d*).



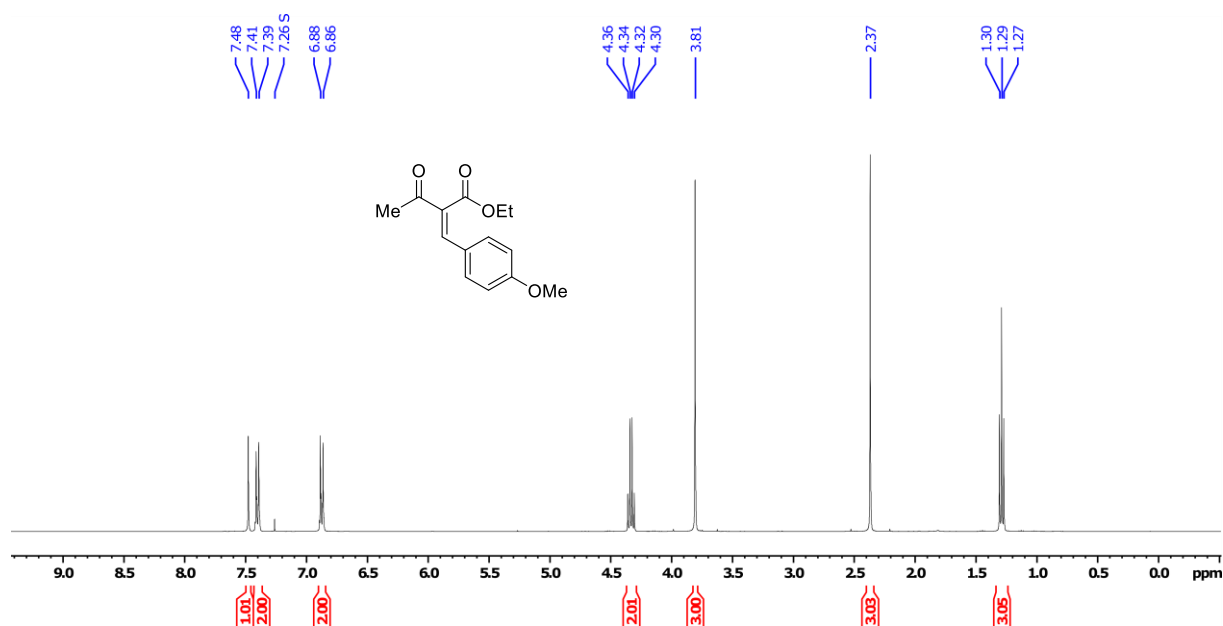
¹H NMR spectrum of 19d (400 MHz, chloroform-*d*)



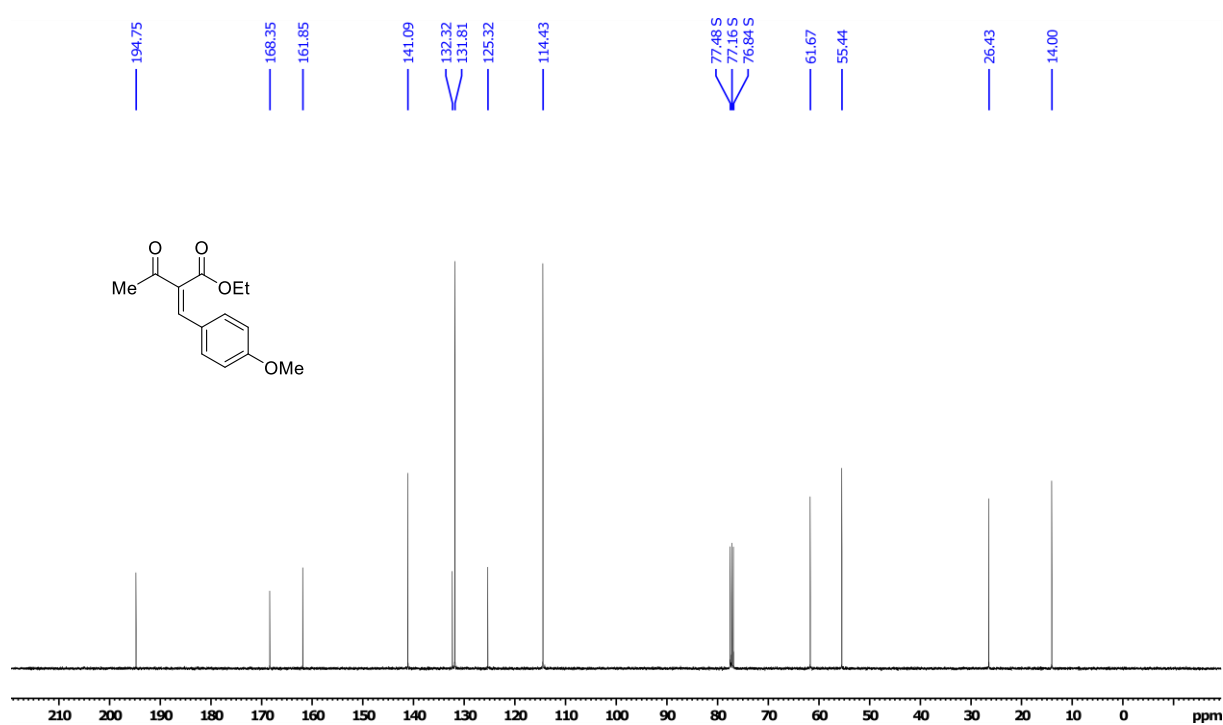
¹³C NMR spectrum of 19d (100 MHz, chloroform-*d*)



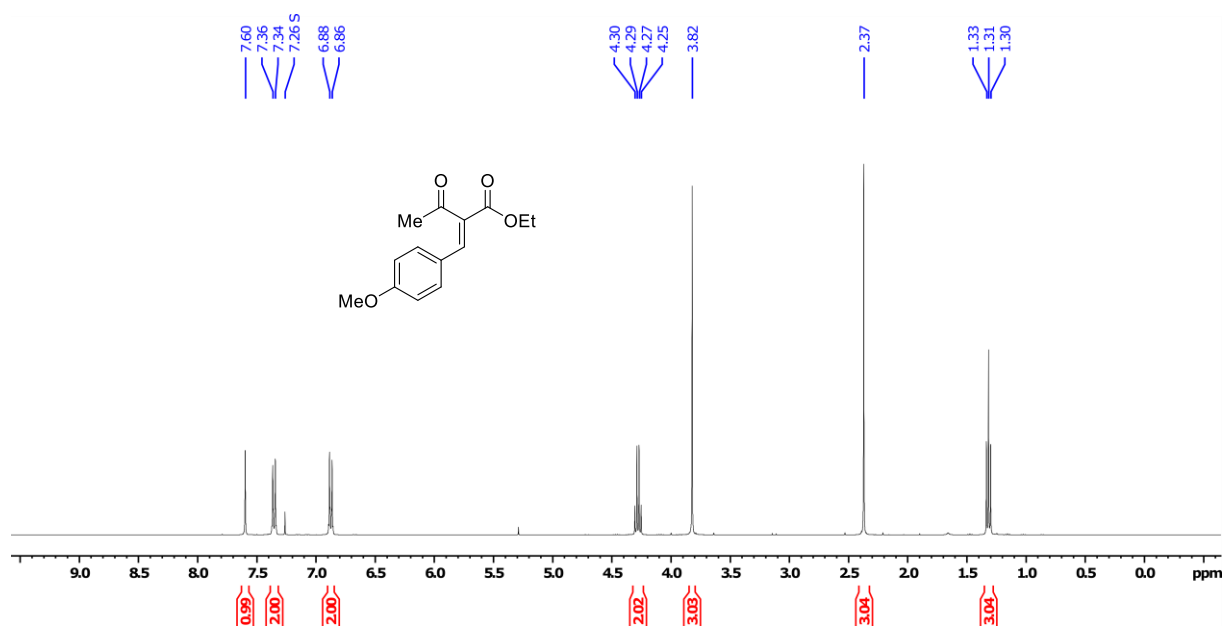
^1H NMR spectrum of 19f (400 MHz, chloroform-*d*).



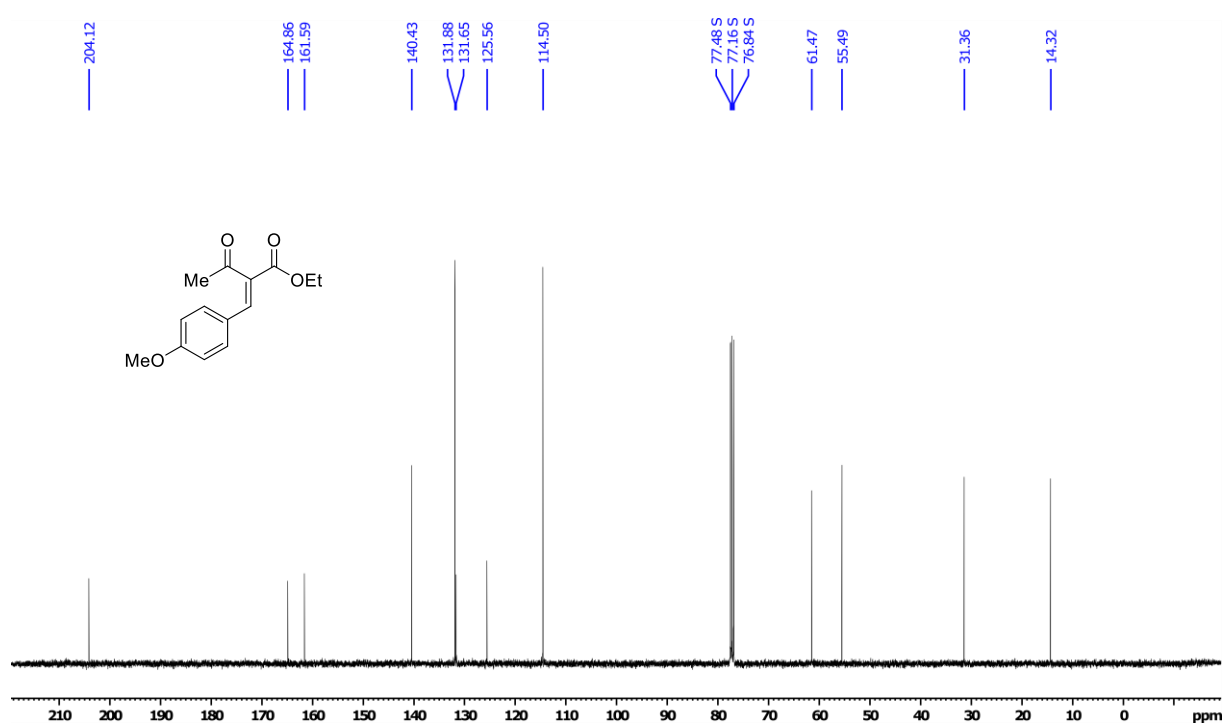
^{13}C NMR spectrum of 19f (100 MHz, chloroform-*d*).



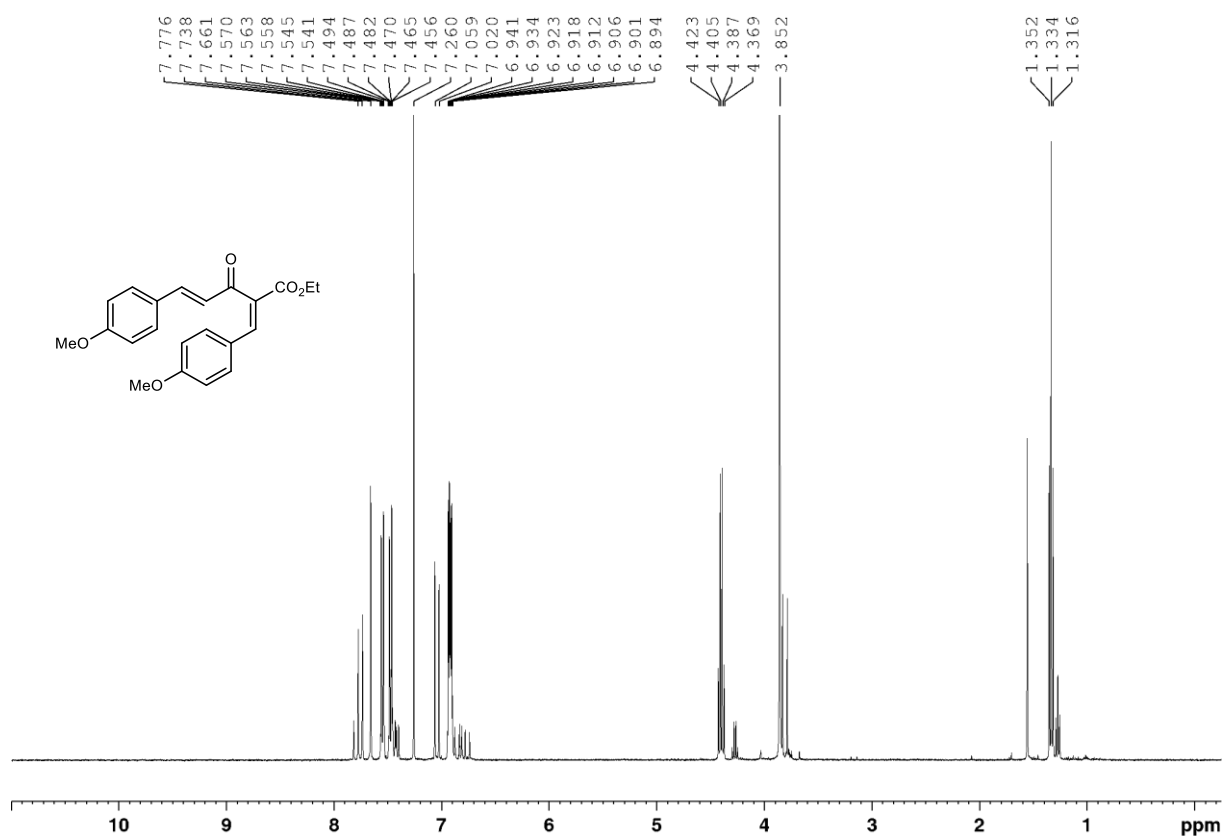
¹H NMR spectrum of 19g (400 MHz, chloroform-*d*).



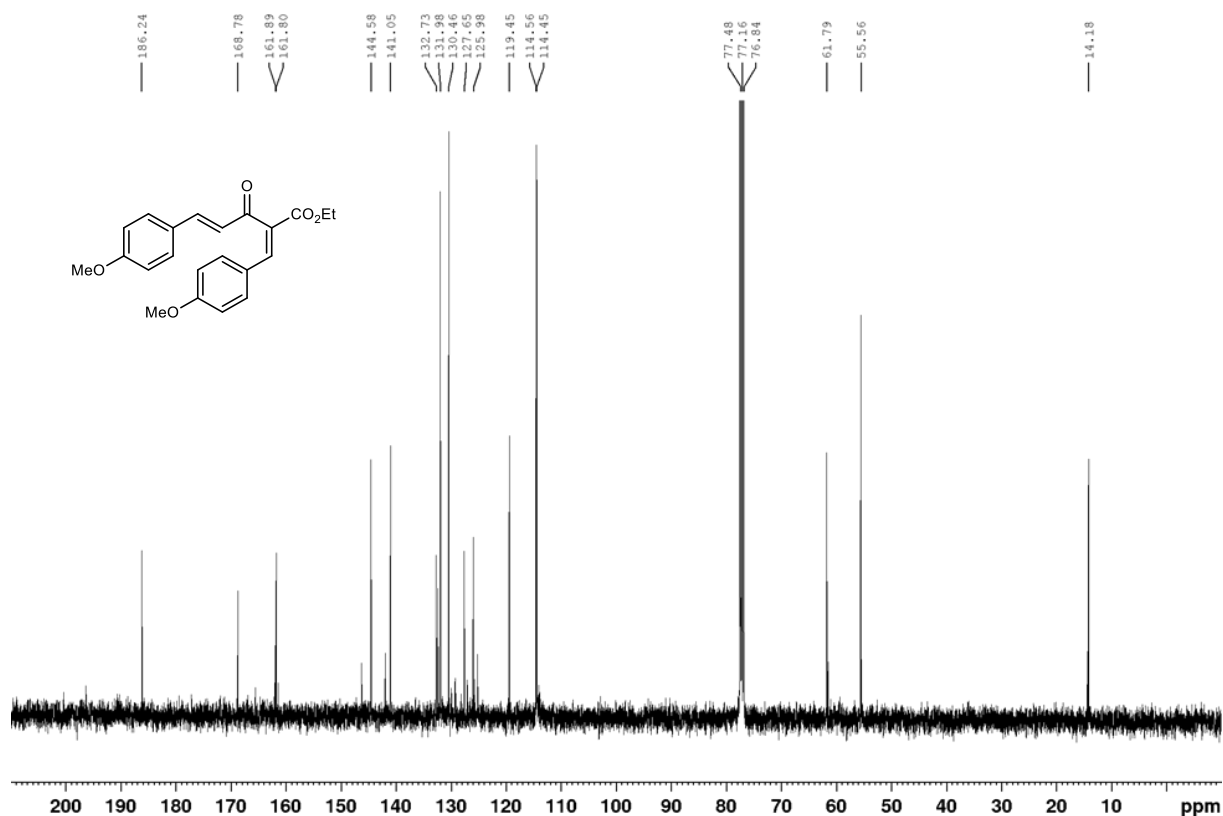
¹³C NMR spectrum of 19g (100 MHz, chloroform-*d*).



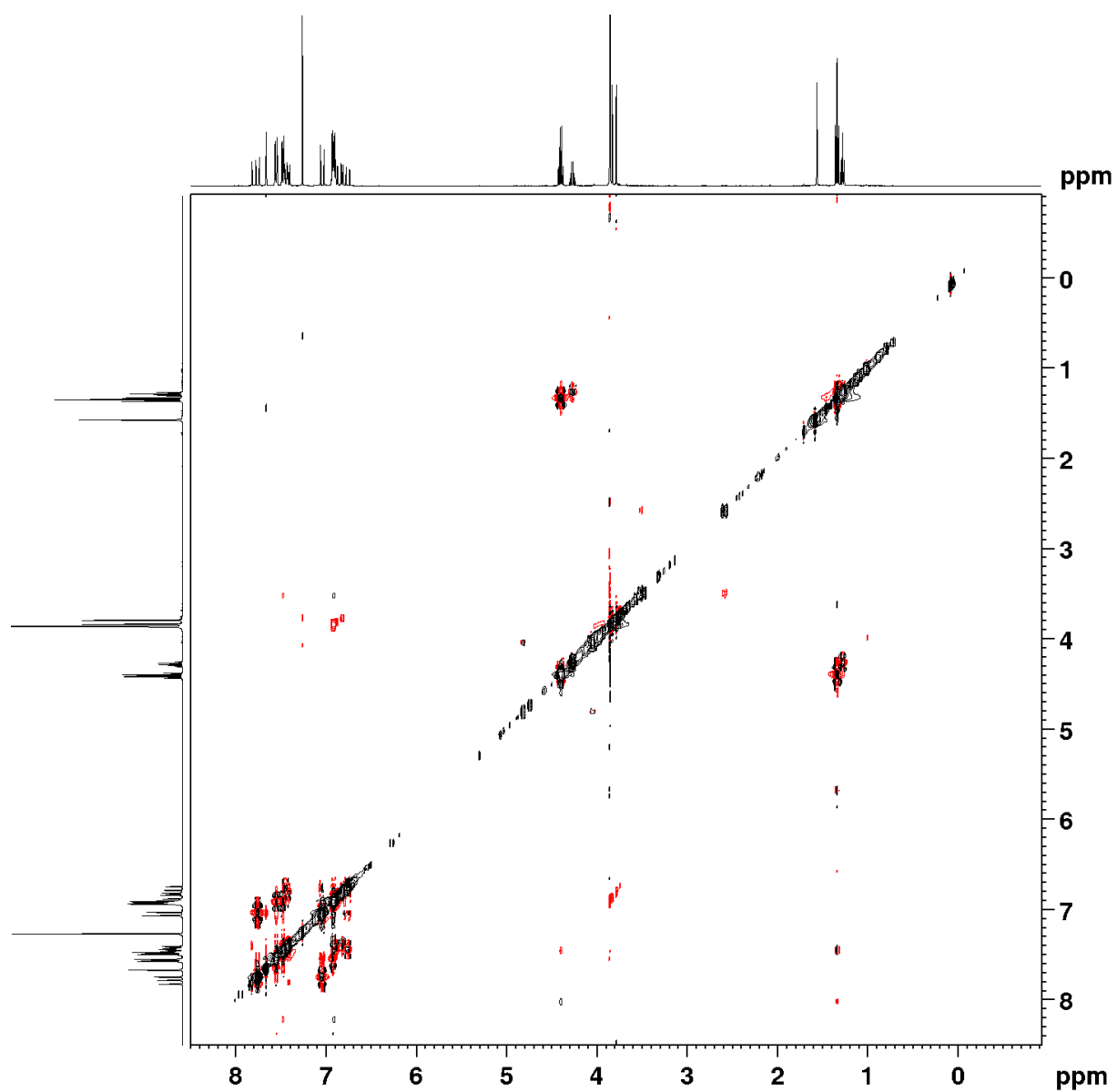
¹H NMR spectrum of 19h (400 MHz, chloroform-*d*).



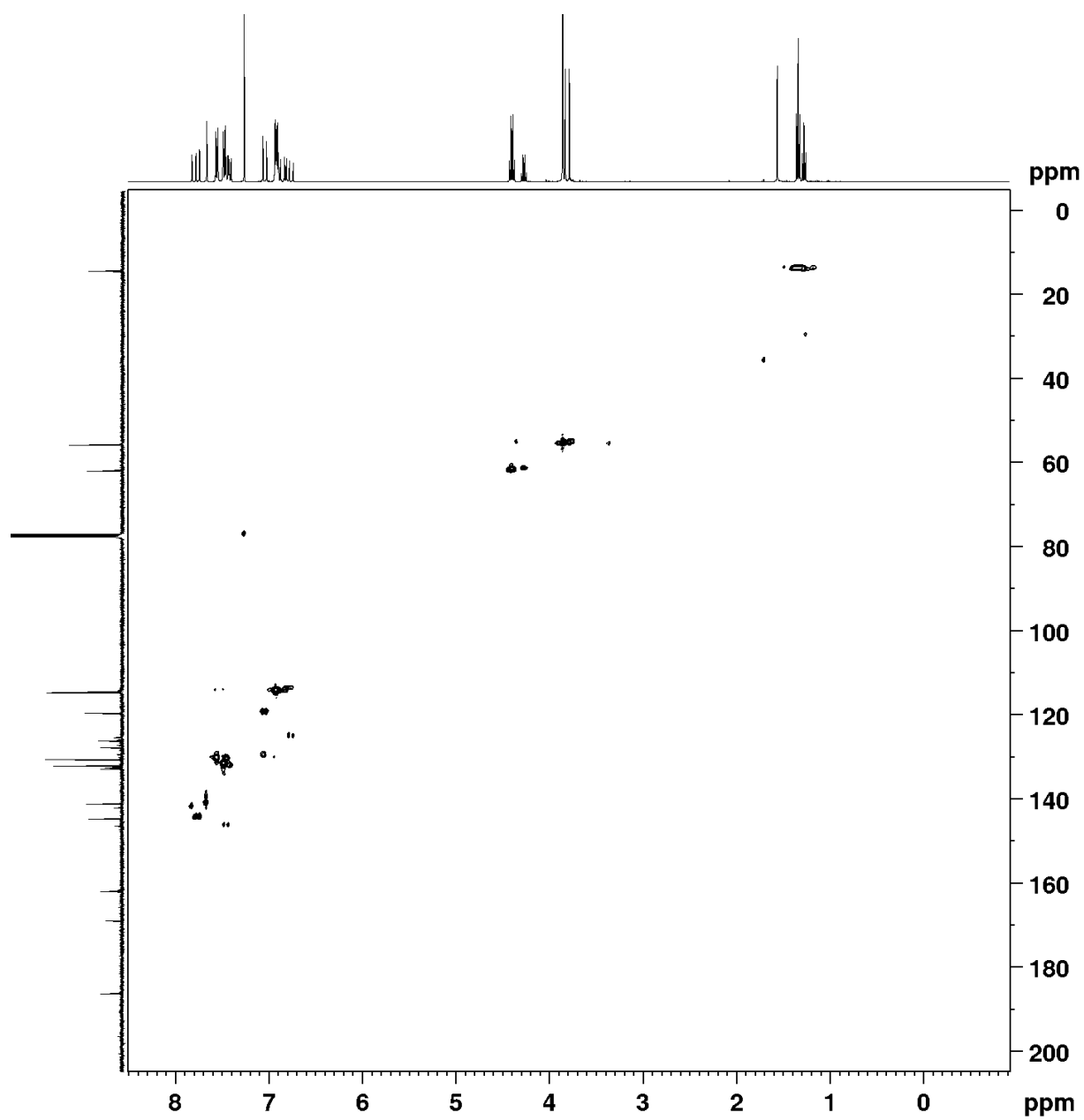
¹³C NMR spectrum of 19h (100 MHz, chloroform-*d*).



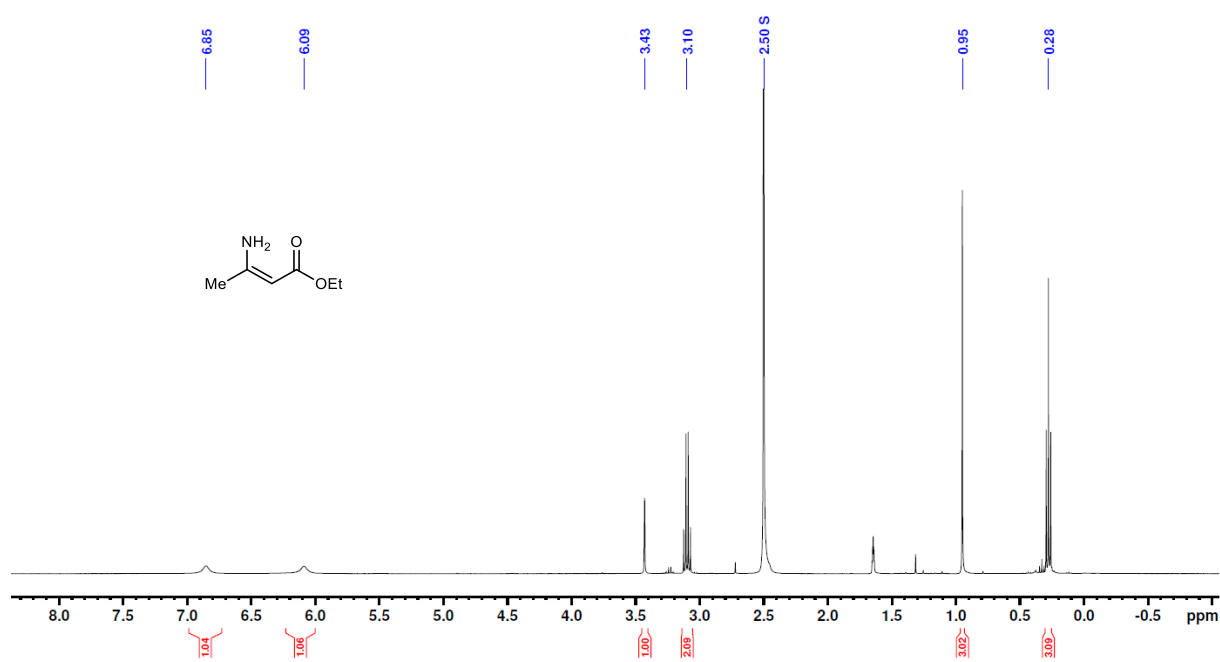
NOESY spectrum of **19h** (chloroform-*d*).



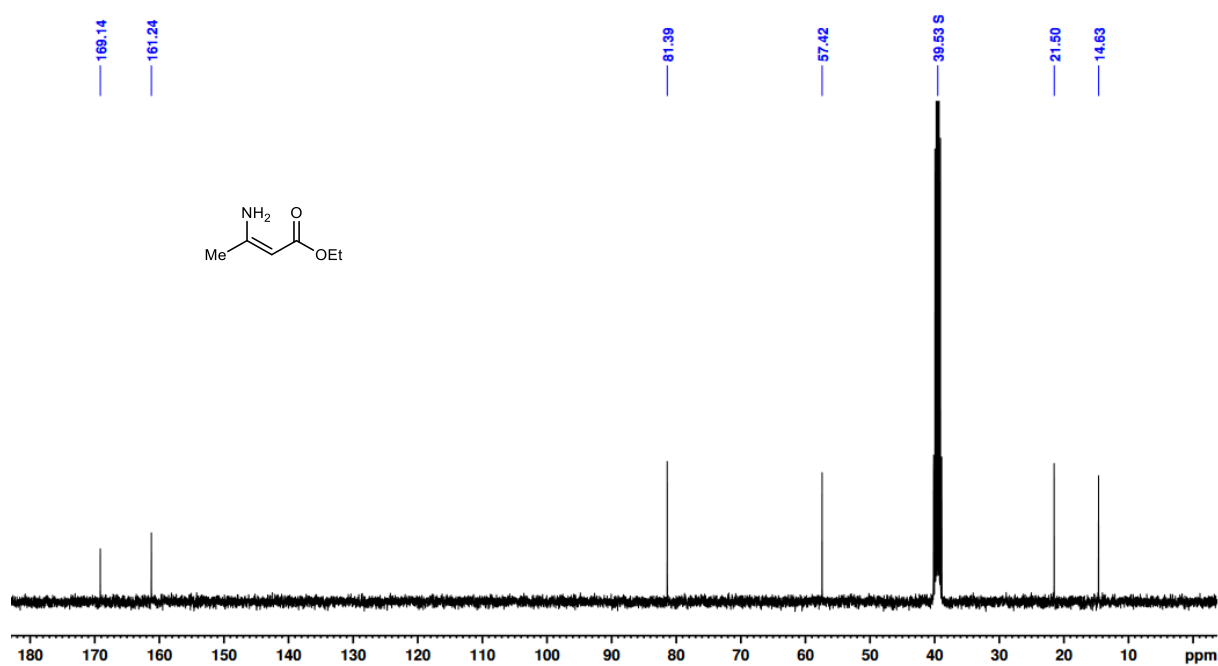
HSQC-TOCSY spectrum of 19h.



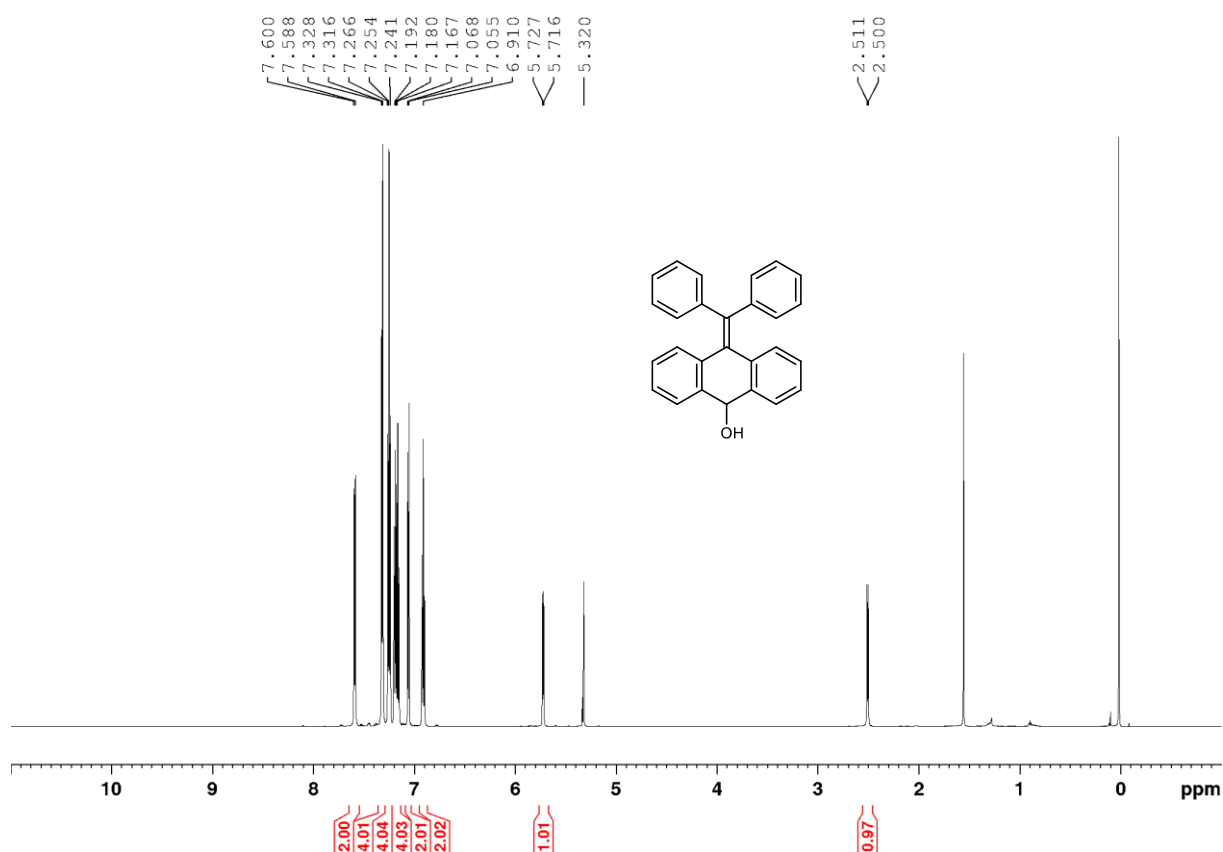
^1H NMR spectrum of EAB (400 MHz, $\text{DMSO}-d_6$).



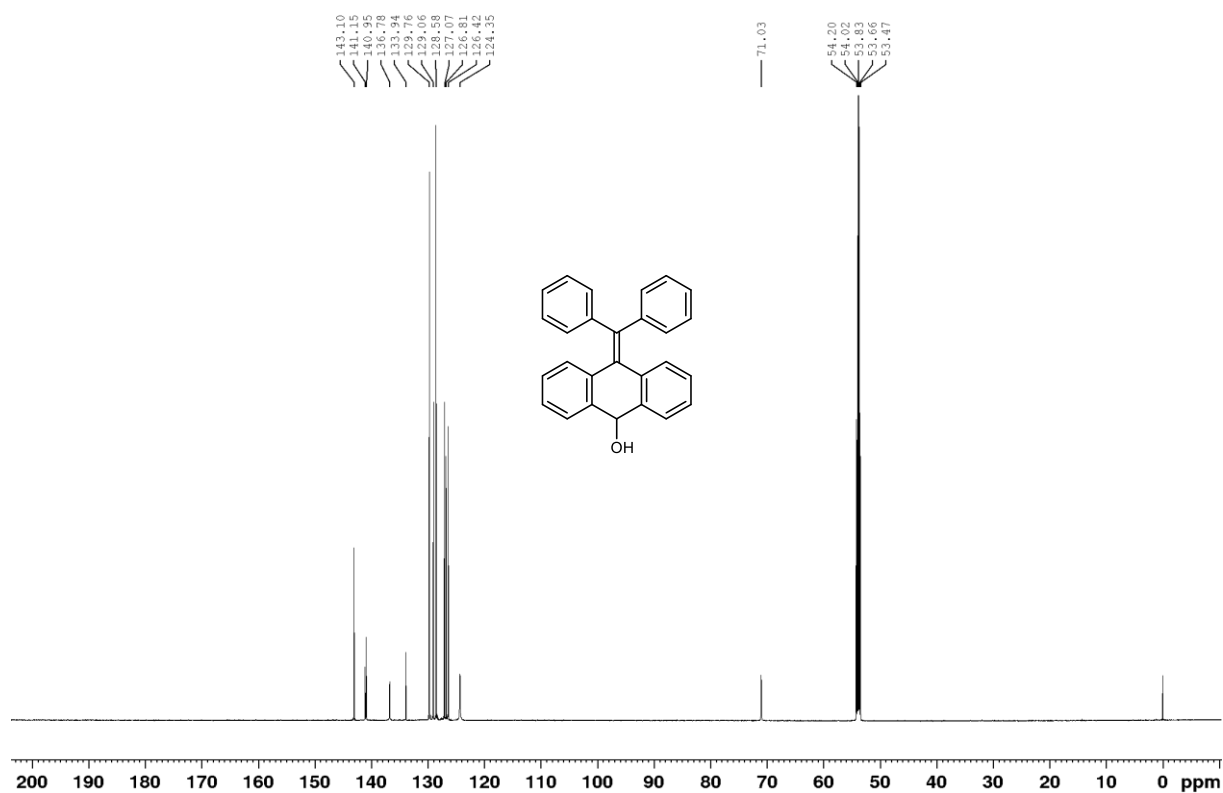
^{13}C NMR spectrum of EAB (100 MHz, $\text{DMSO}-d_6$).



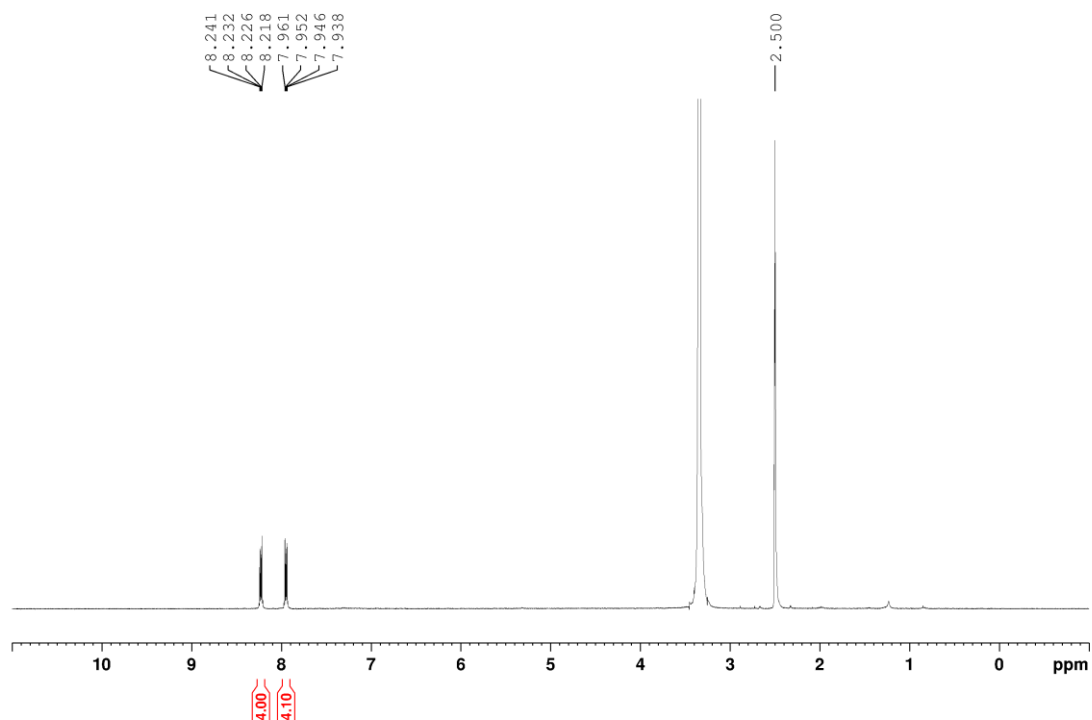
¹H NMR spectrum of 15b (400 MHz, dichloromethene-*d*₂).



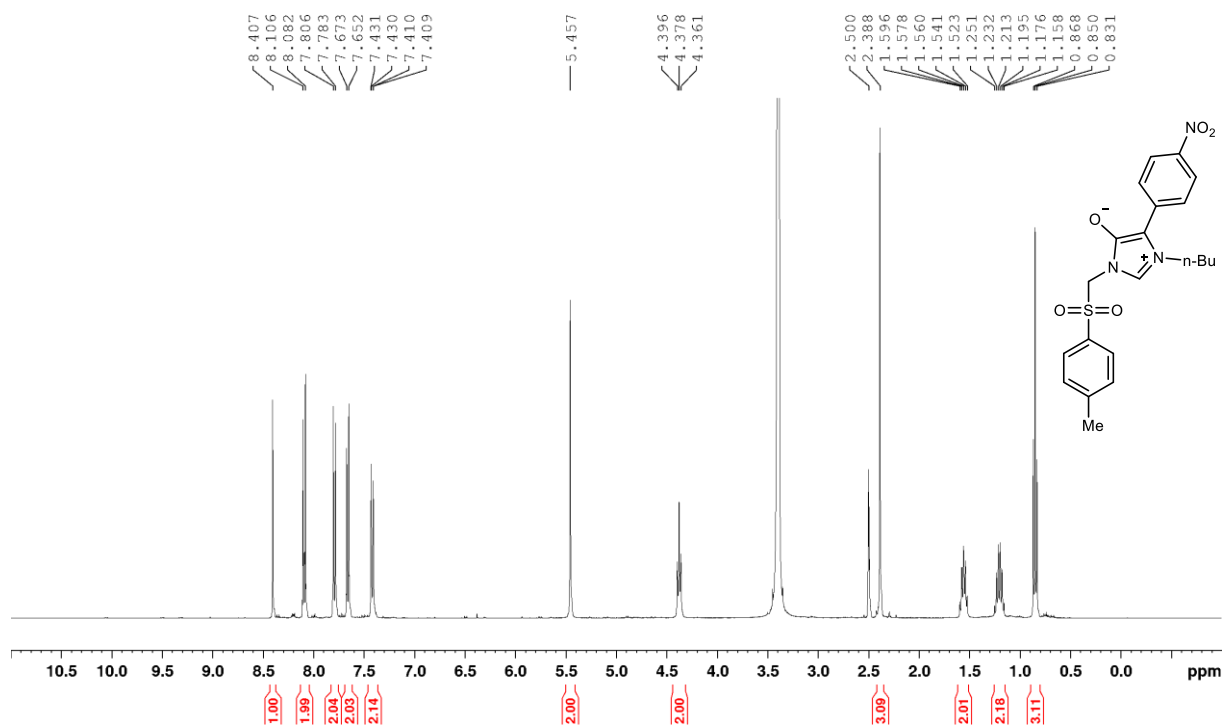
¹³C NMR spectrum of 15b (100 MHz, dichloromethene-*d*₂).



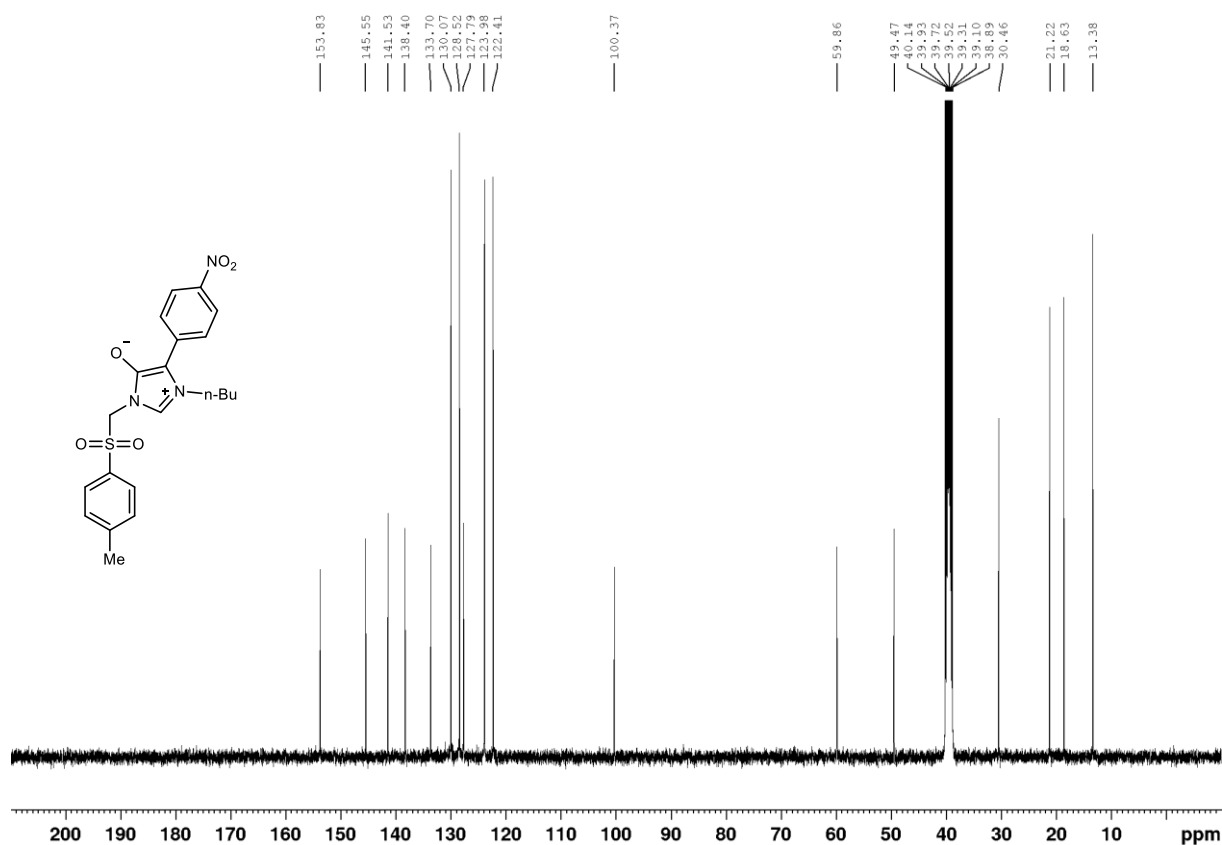
¹H NMR spectrum of 15g (400 MHz, DMSO-*d*₆).



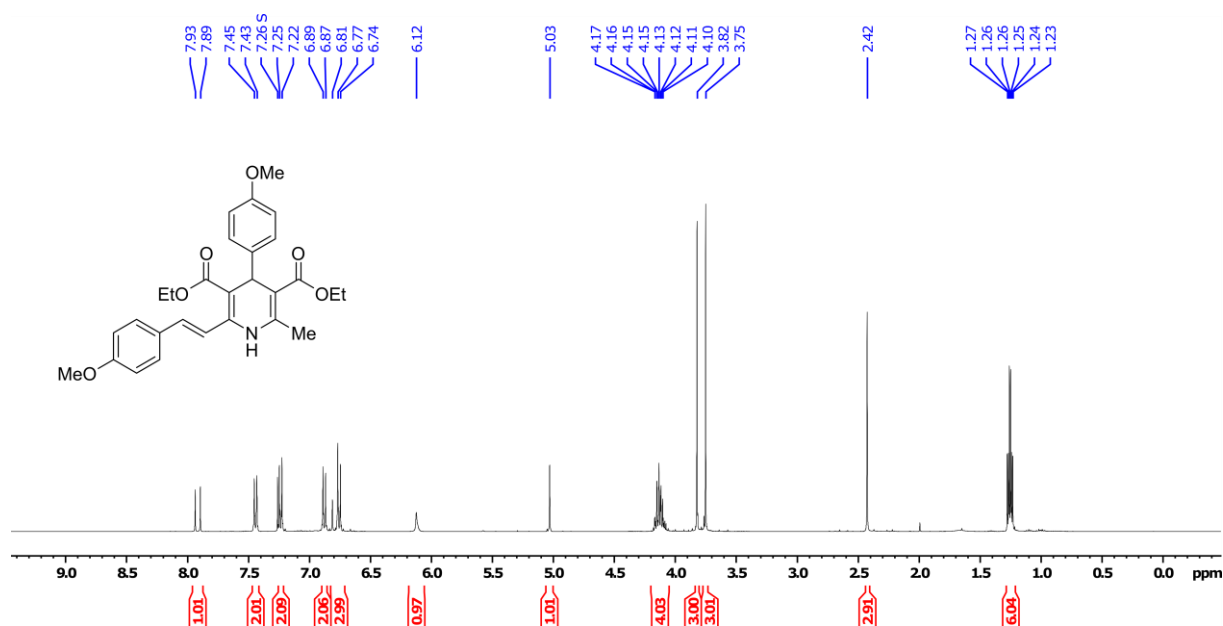
¹H NMR spectrum of 16a (400 MHz, DMSO-*d*₆).



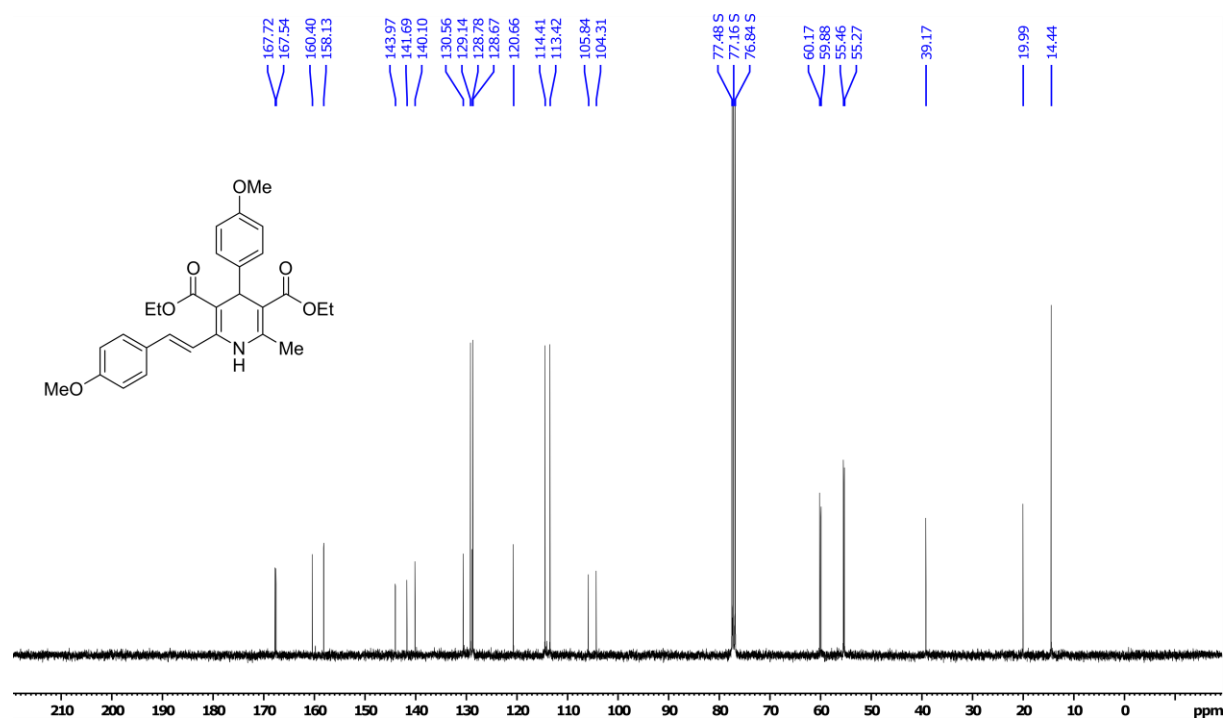
¹³C NMR spectrum of 16a (100 MHz, DMSO-*d*₆)



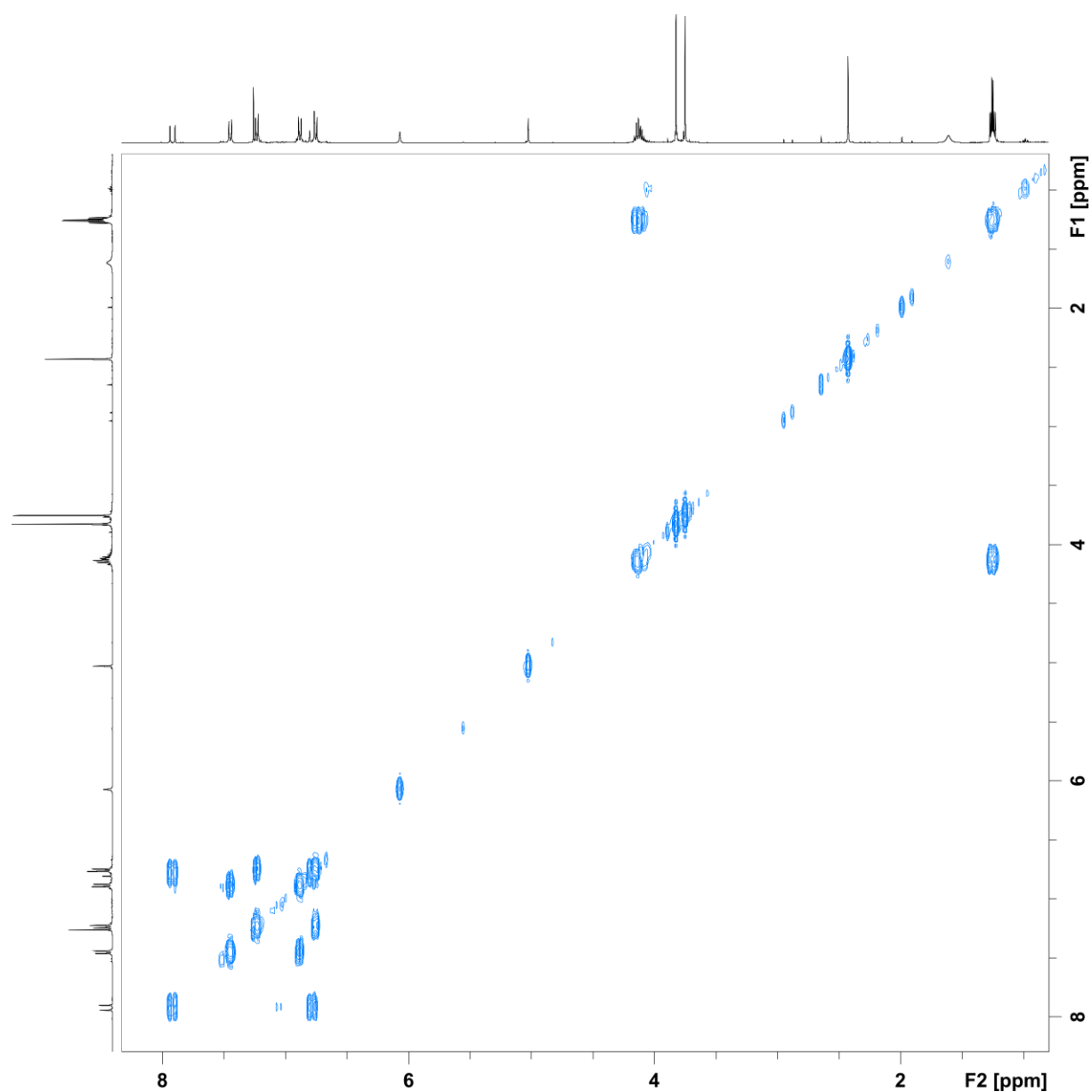
¹H NMR spectrum of 19k (400 MHz, chloroform-*d*)



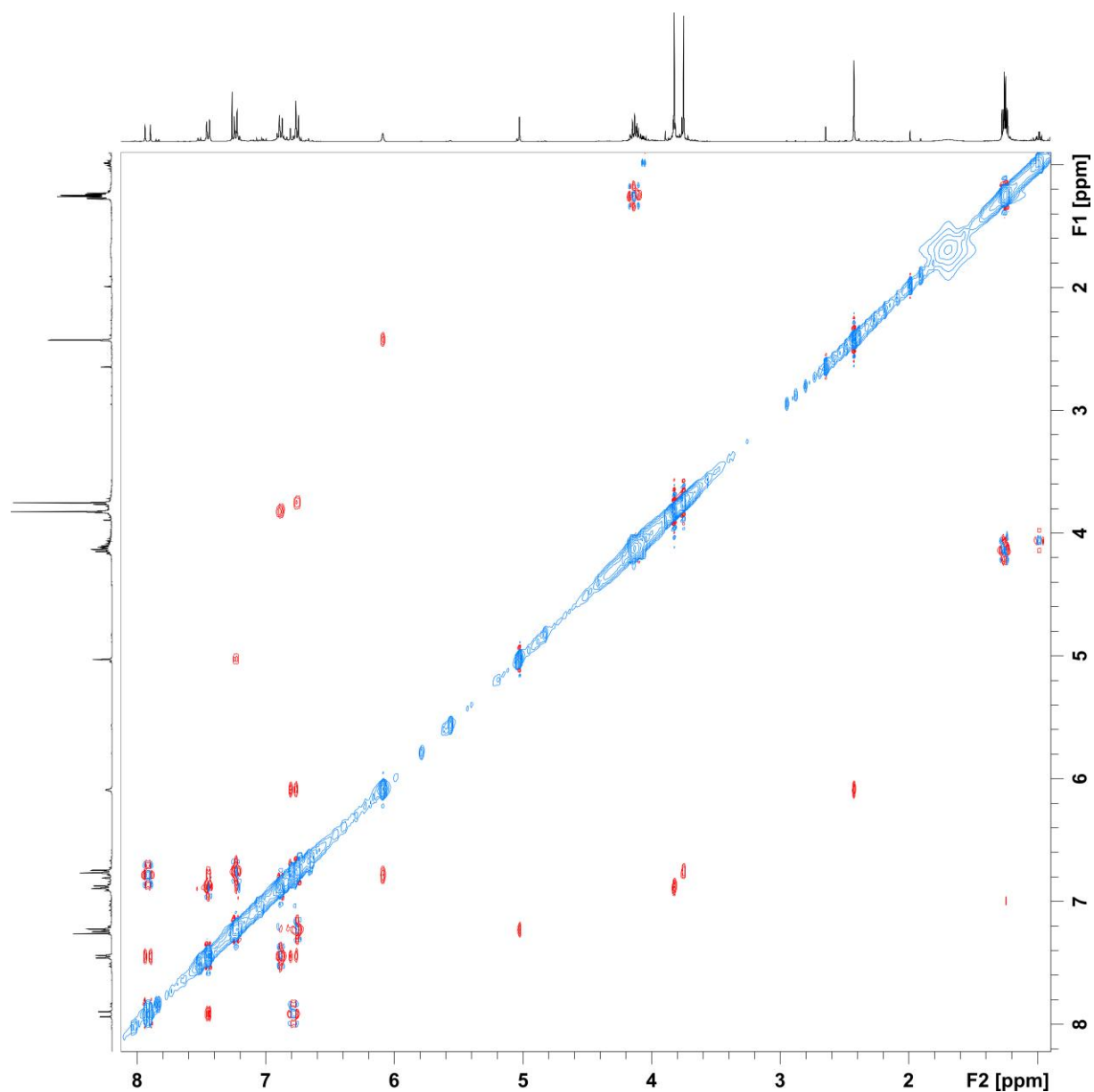
¹³C NMR spectrum of 19k (100 MHz, chloroform-*d*)



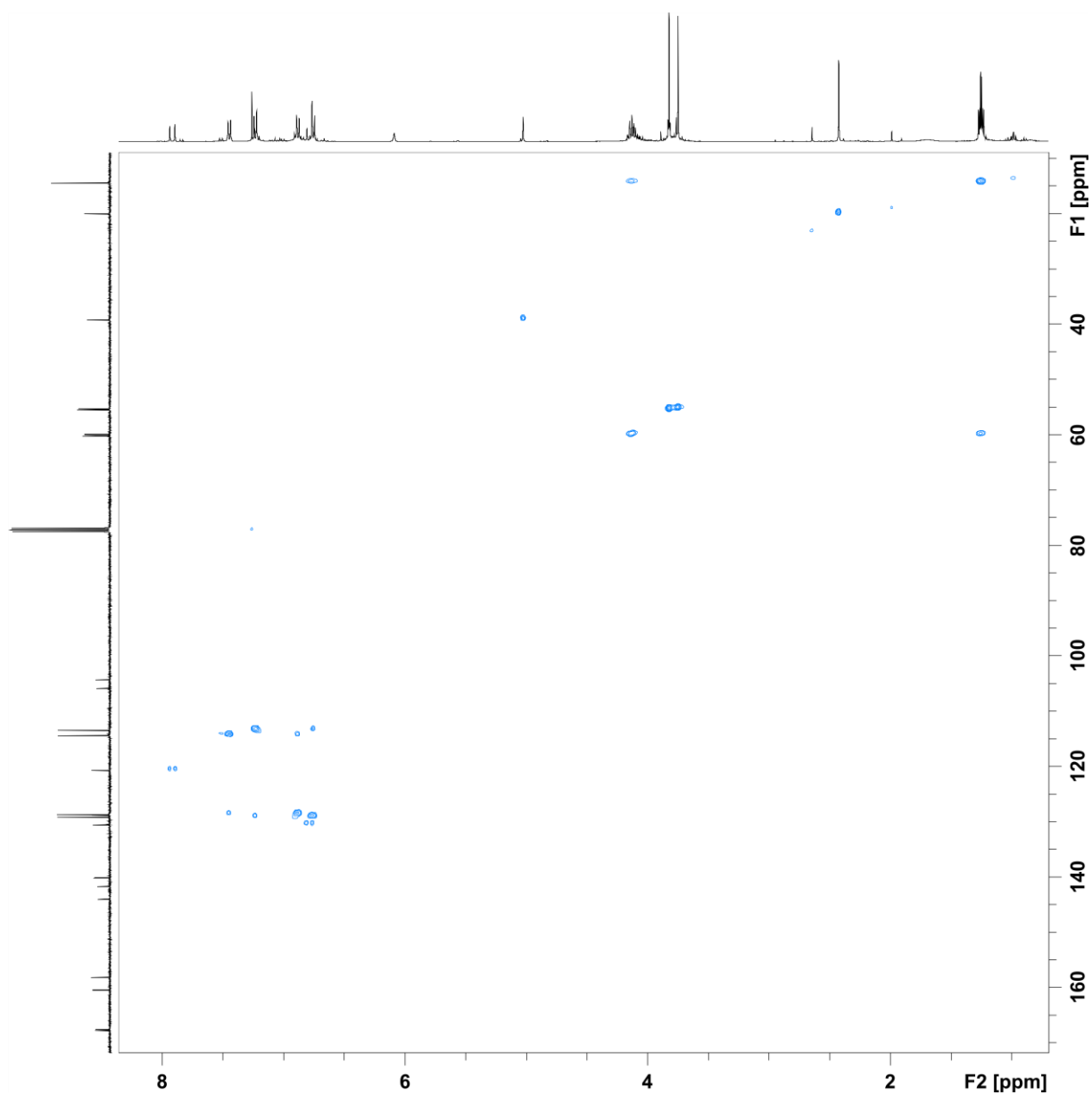
COSY spectrum of **19k** (chloroform-*d*)



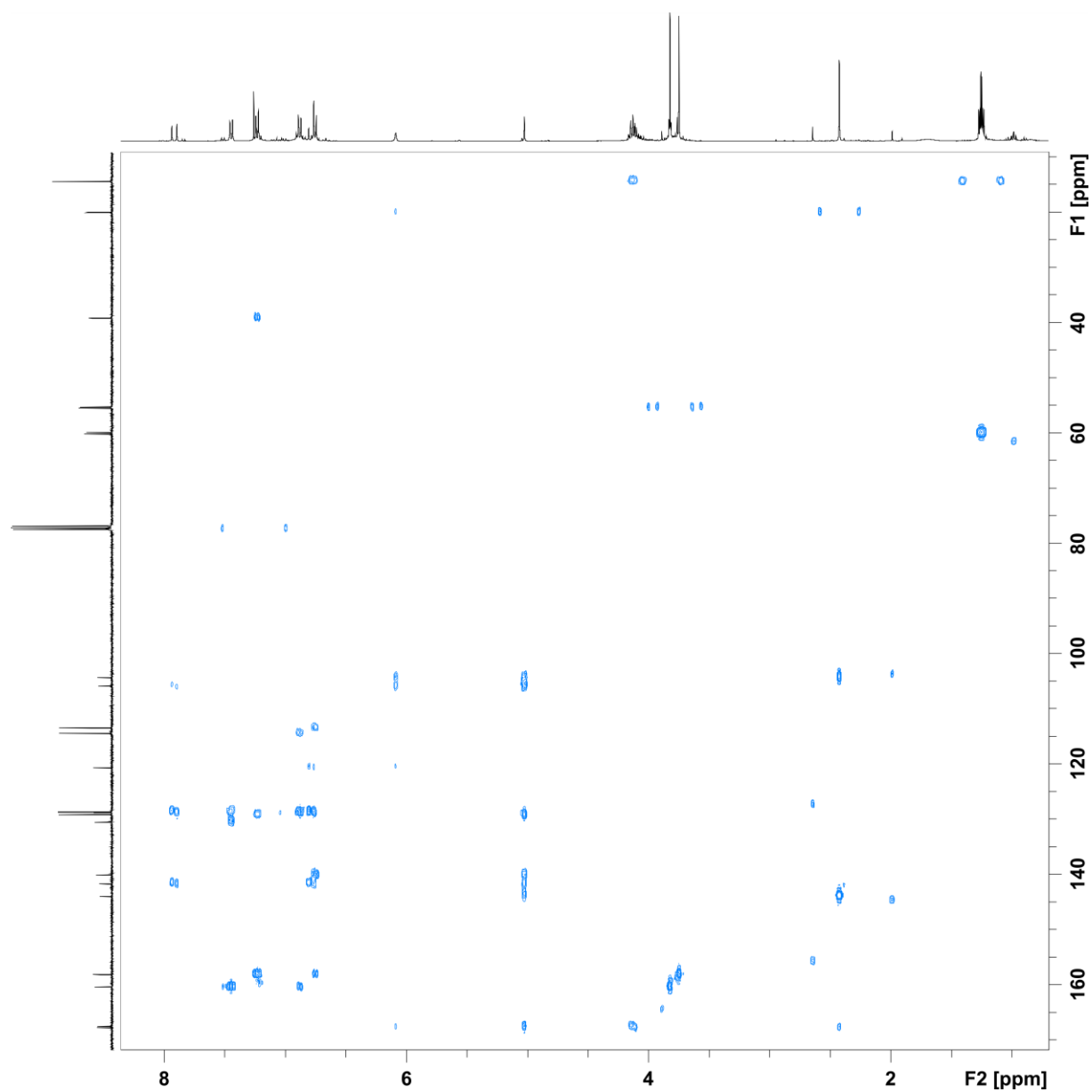
NOESY spectrum of **19k** (chloroform-*d*)



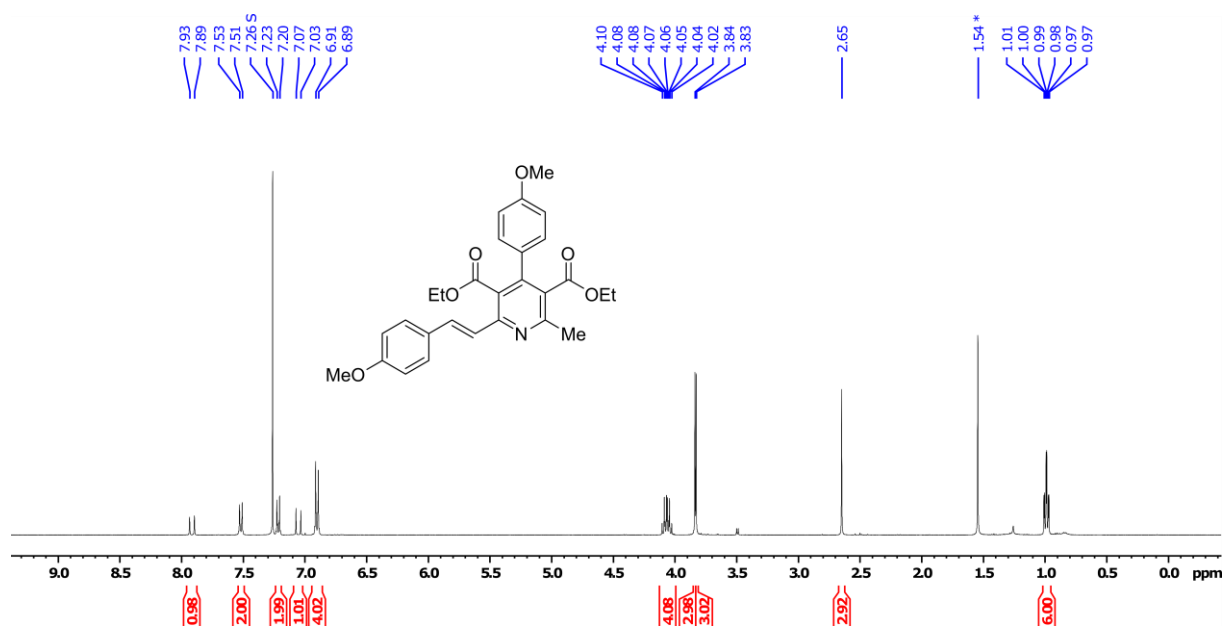
HSQC-TOCSY spectrum of 19k (chloroform-*d*).



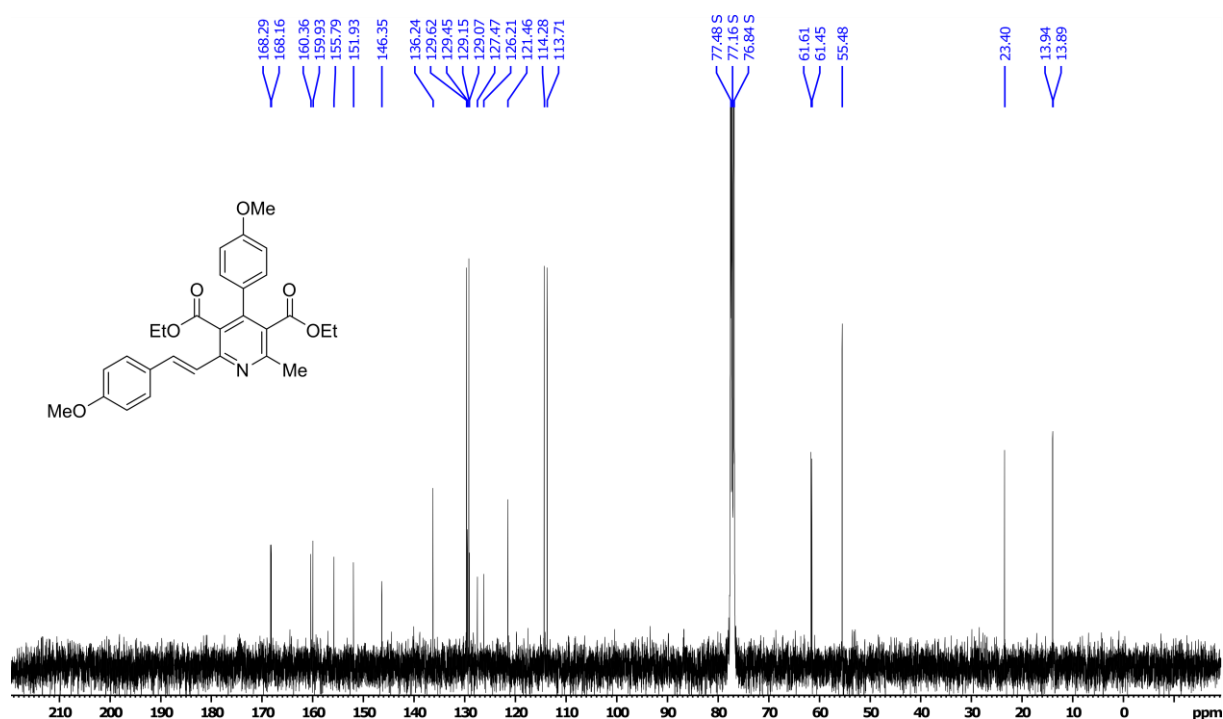
HMBC spectrum of **19k** (chloroform-*d*)



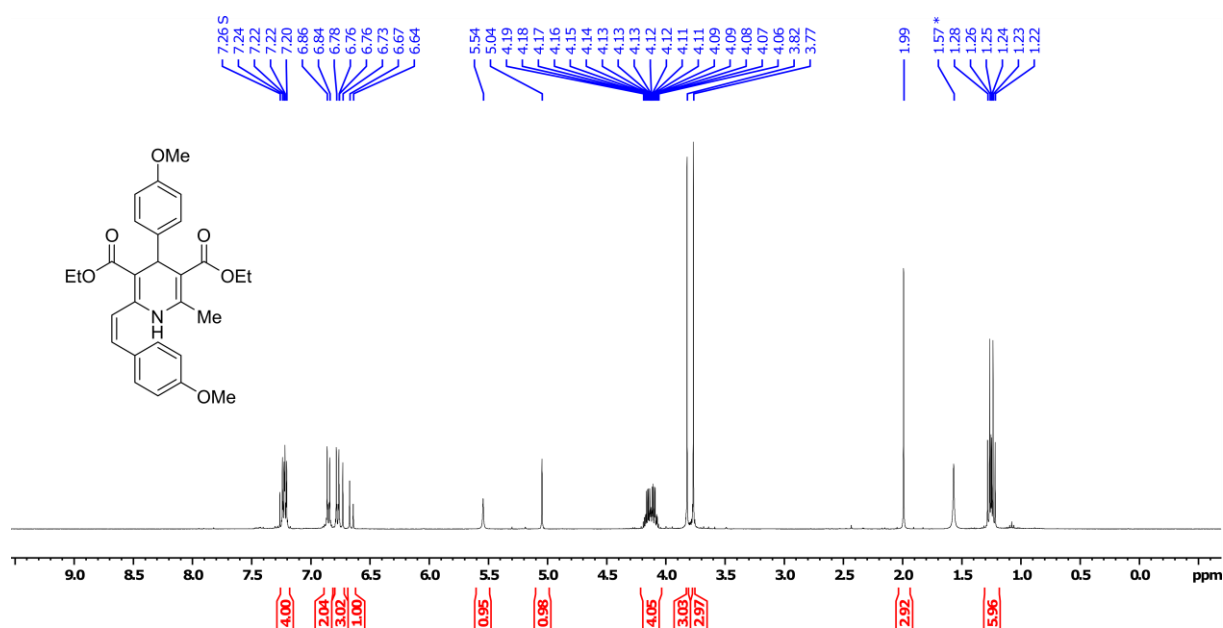
¹H NMR spectrum of 19m (400 MHz, chloroform-*d*)



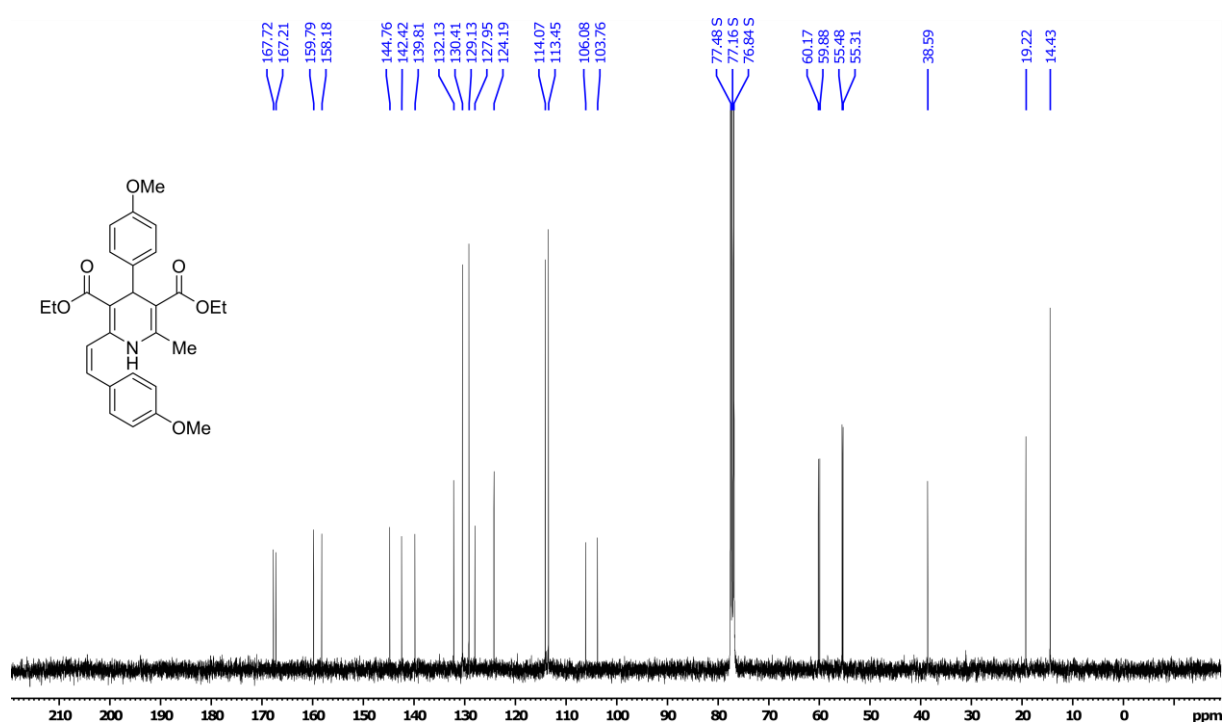
¹³C NMR spectrum of 19m (100 MHz, chloroform-*d*)



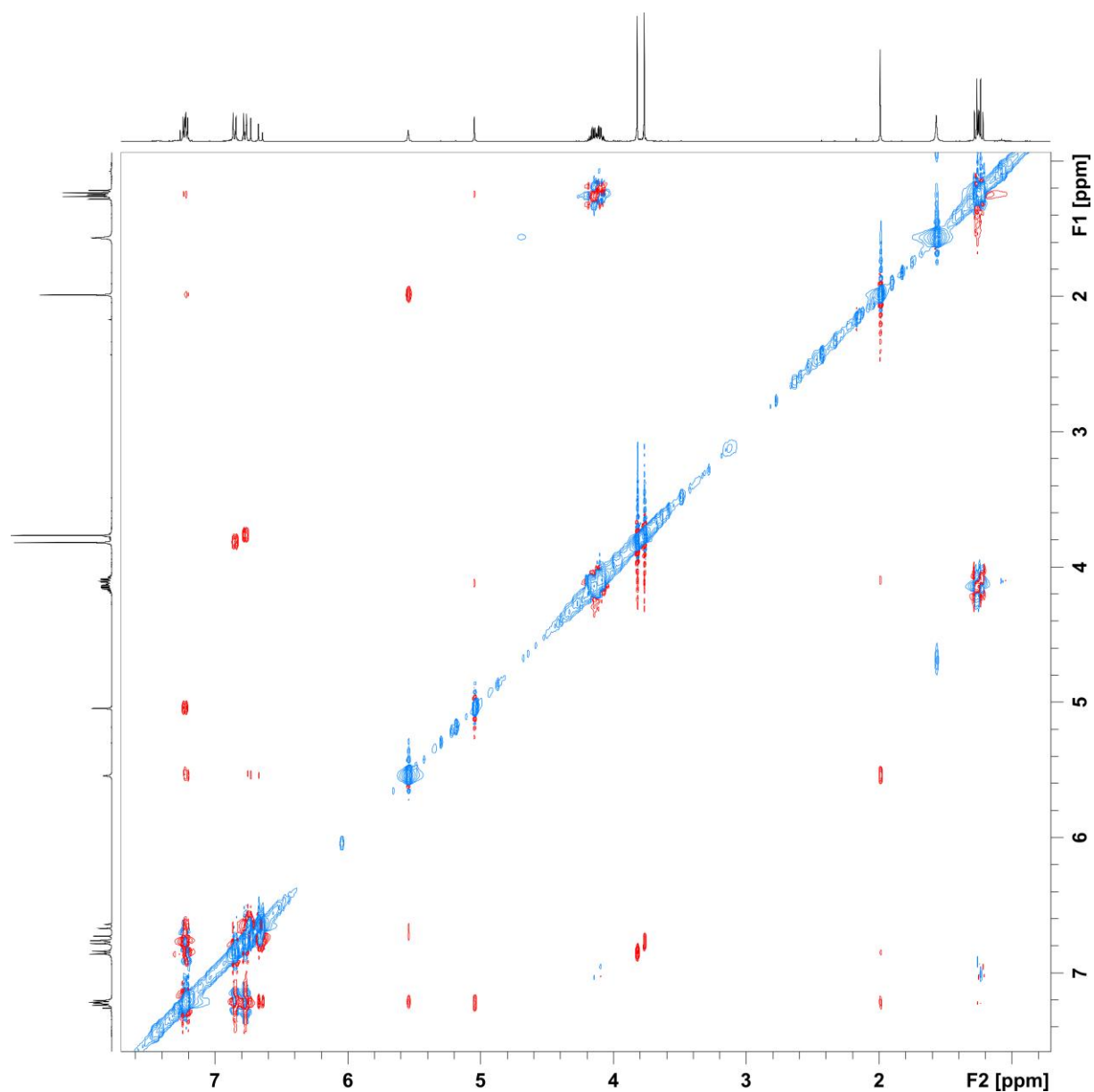
^1H NMR spectrum of 19l (400 MHz, chloroform-*d*)



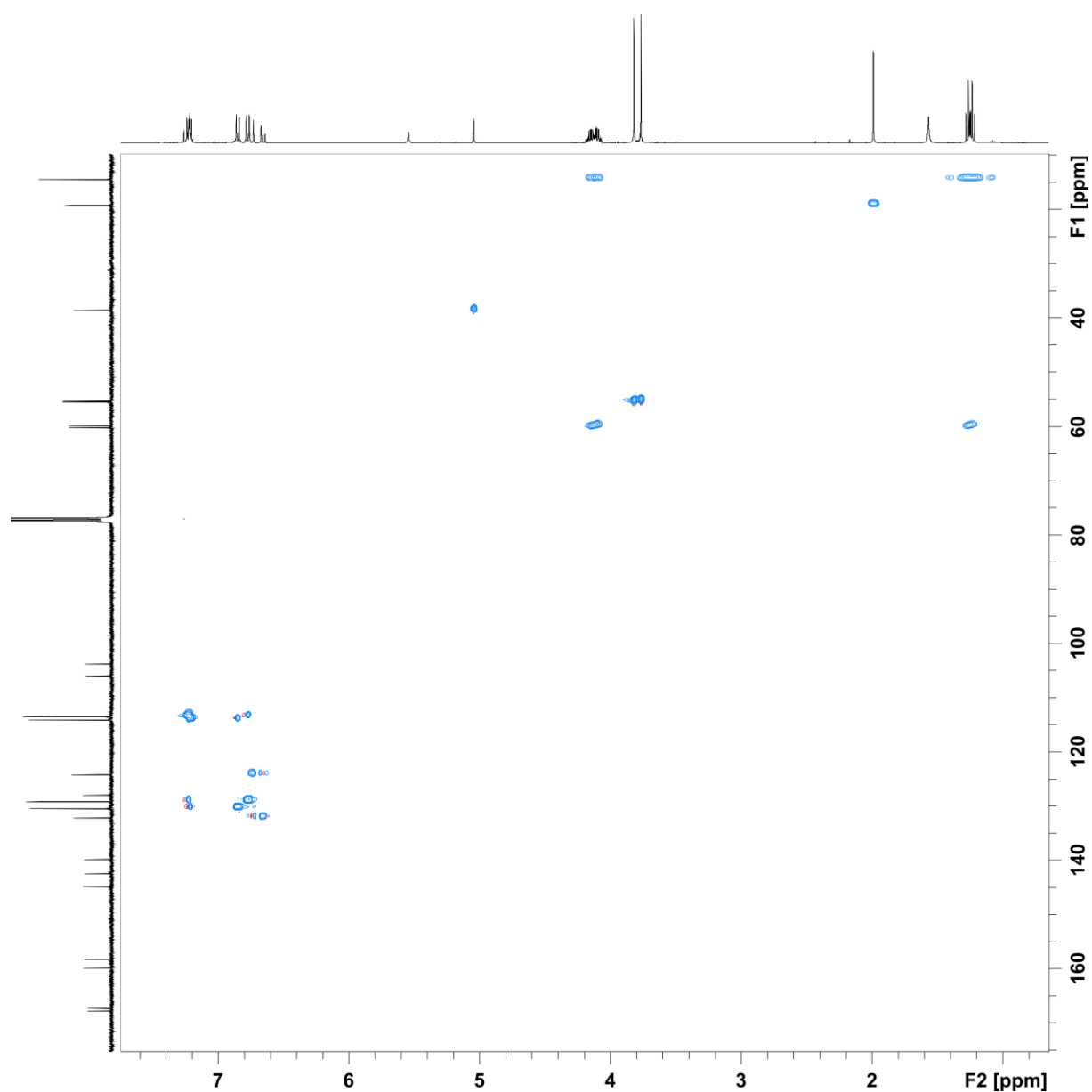
^{13}C NMR spectrum of 19l (100 MHz, chloroform-*d*)



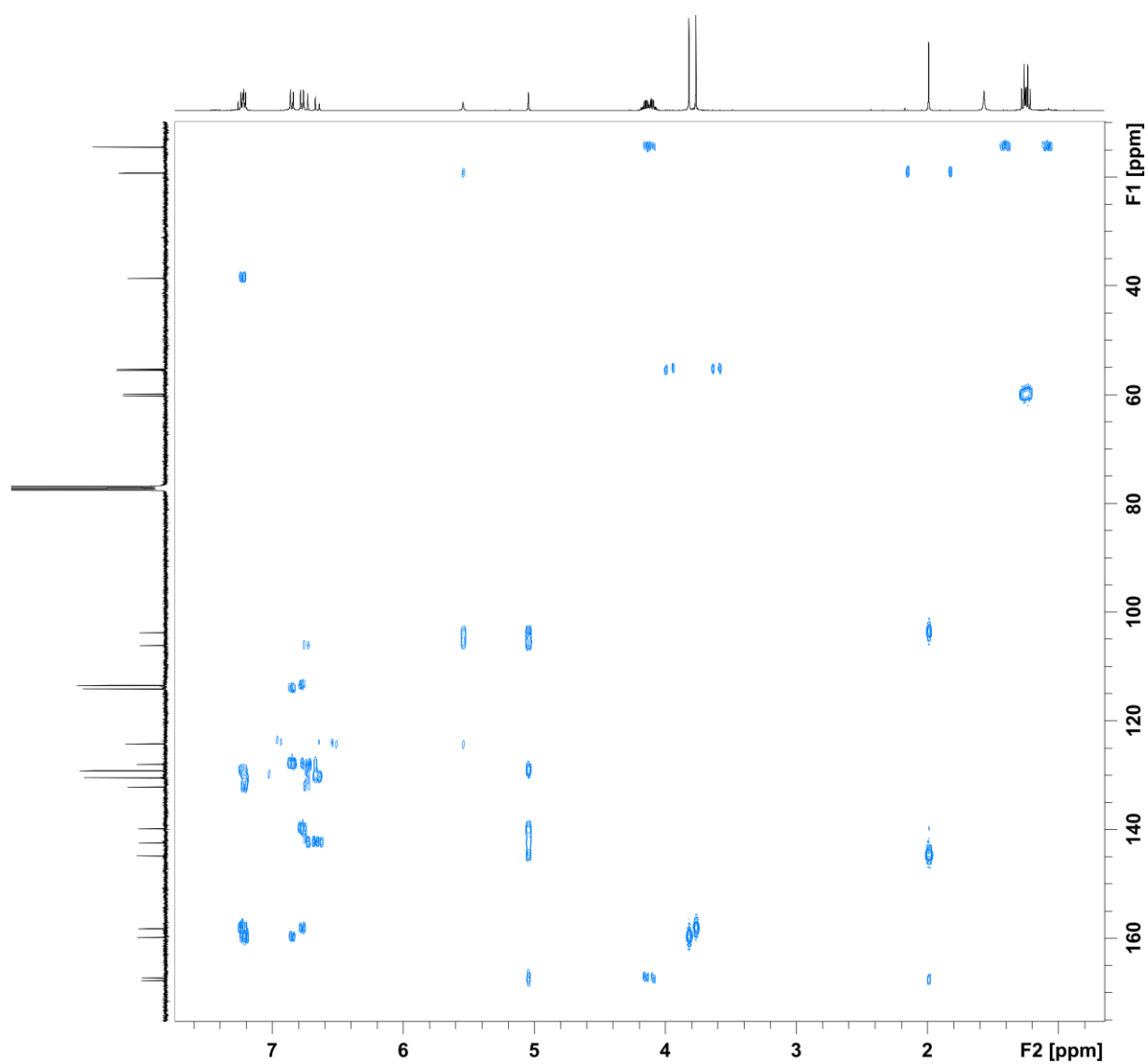
NOESY spectrum of **19l** (chloroform-*d*).



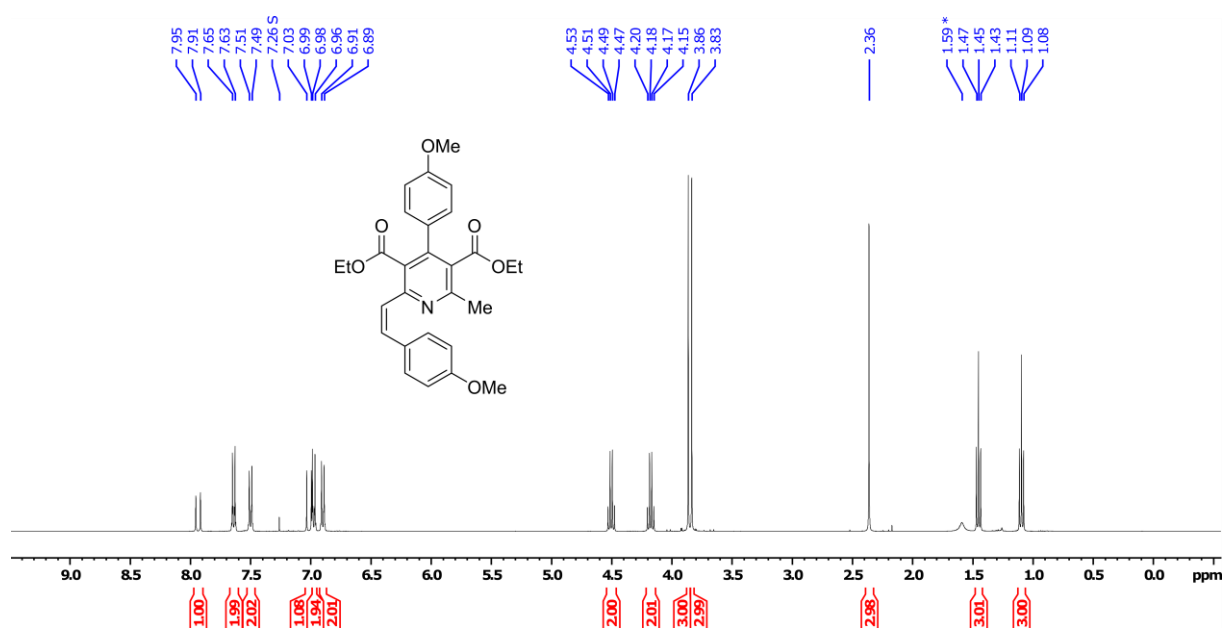
HSQC-TOCSY spectrum of 19l (chloroform-*d*).



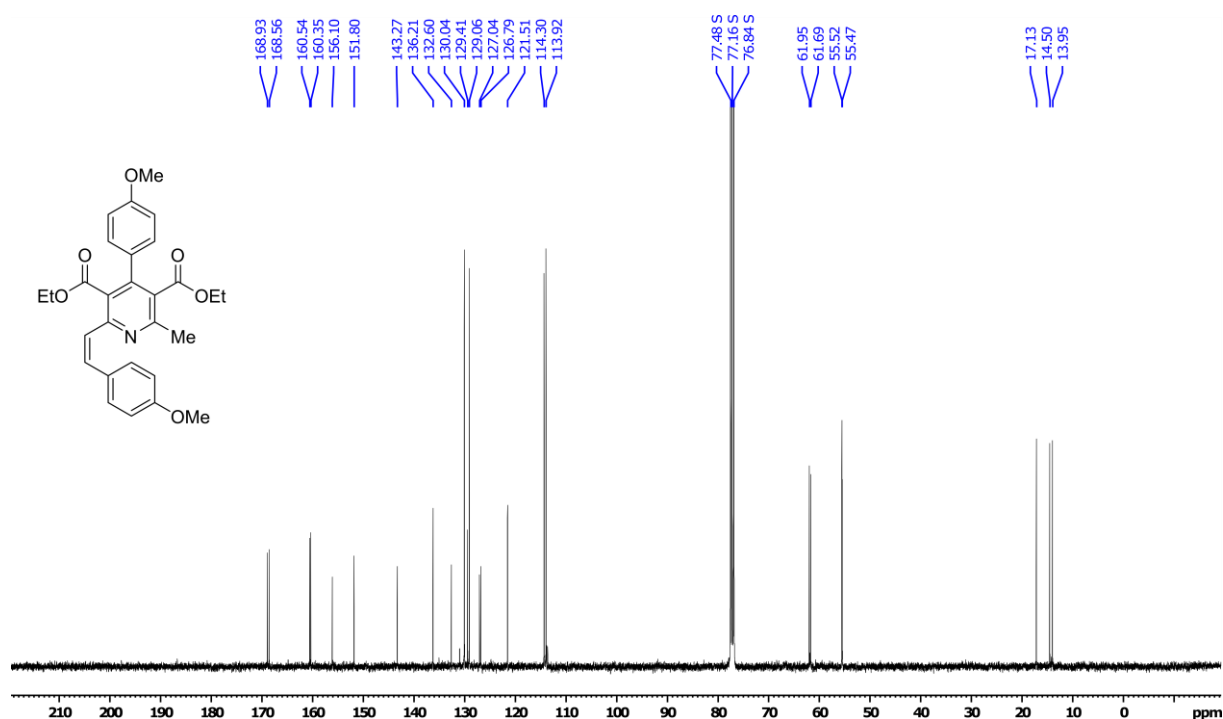
HMBC spectrum of 19l (chloroform-*d*)



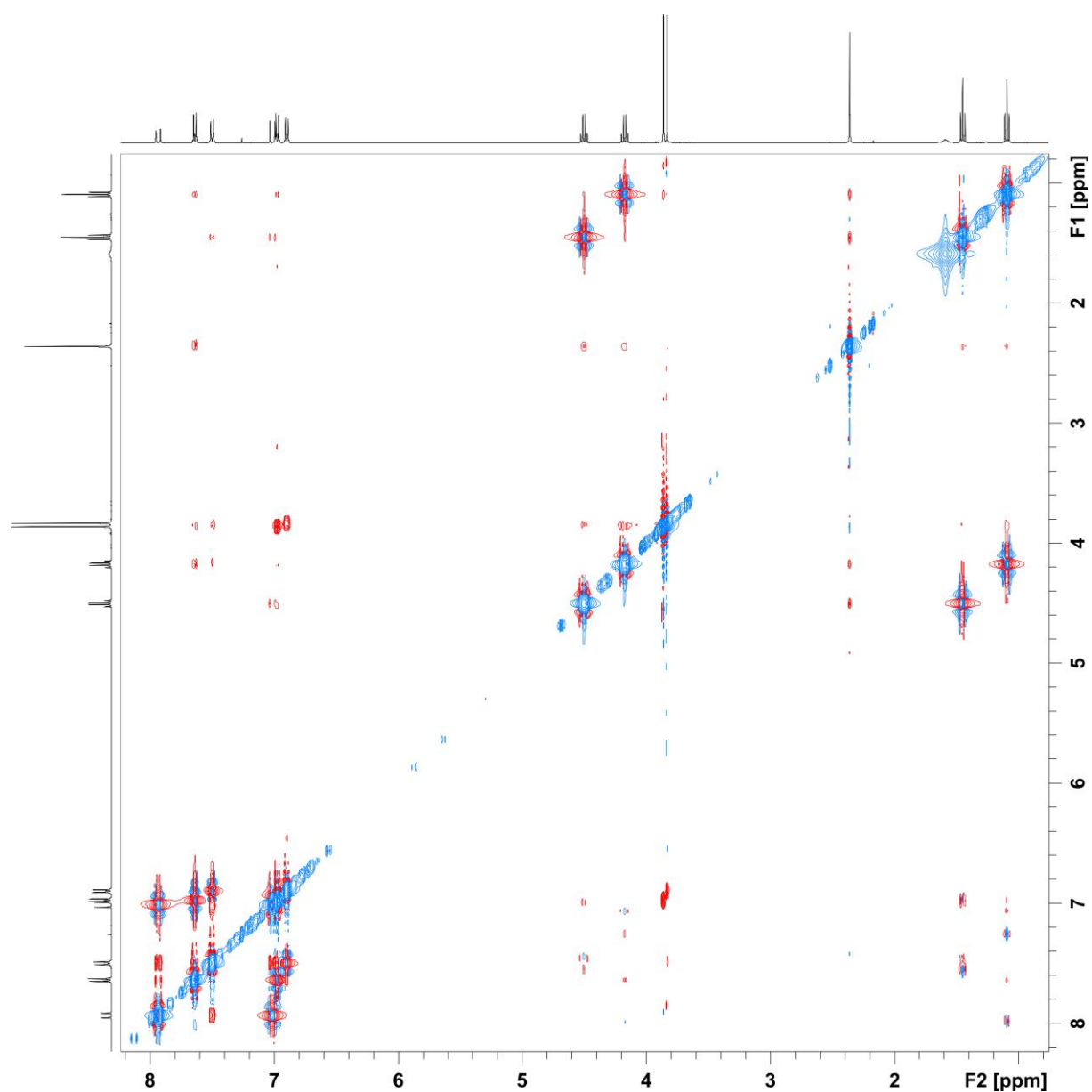
^1H NMR spectrum of 19n (400 MHz, chloroform-*d*).



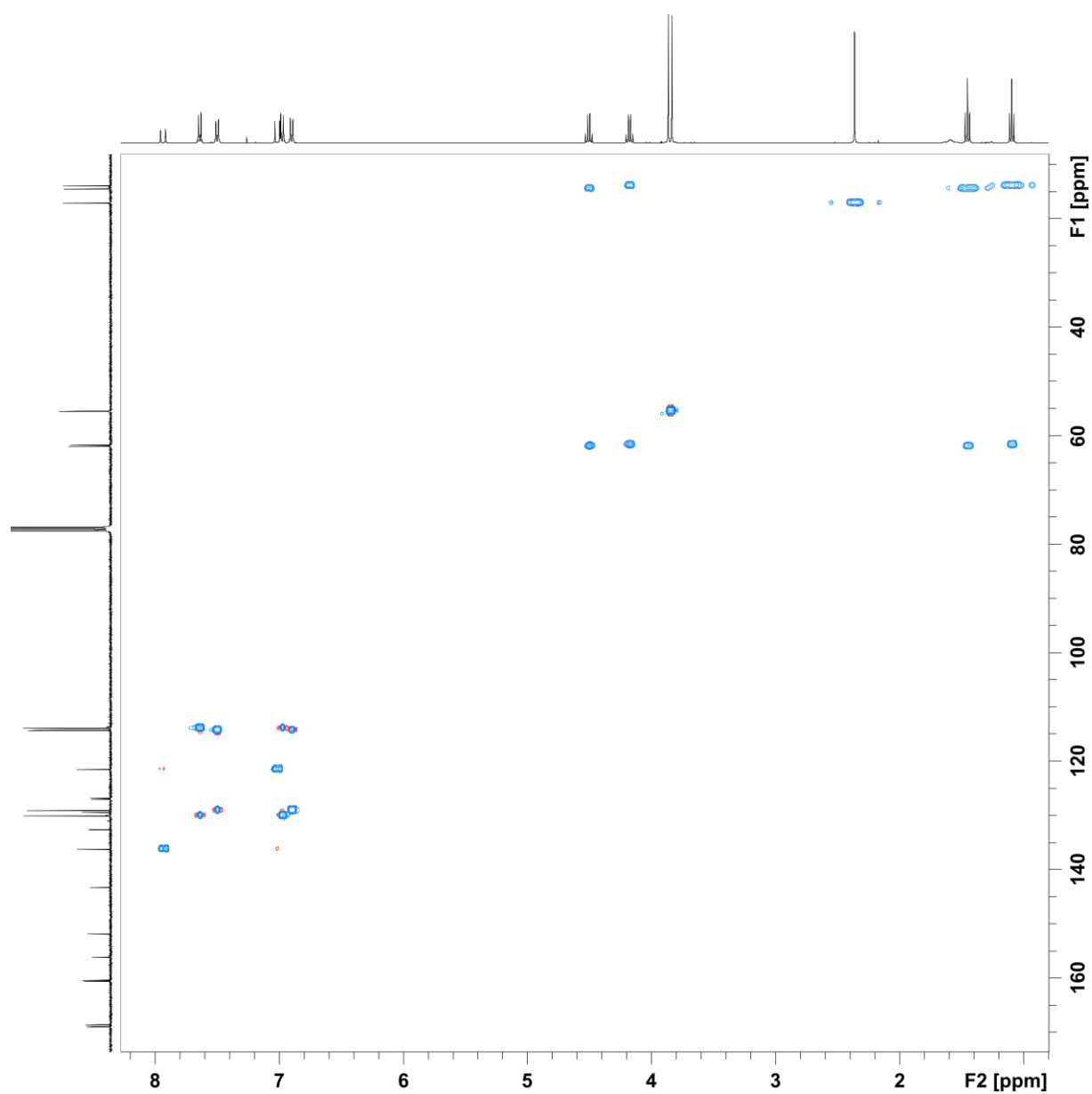
^{13}C NMR spectrum of 19n (100 MHz, chloroform-*d*)



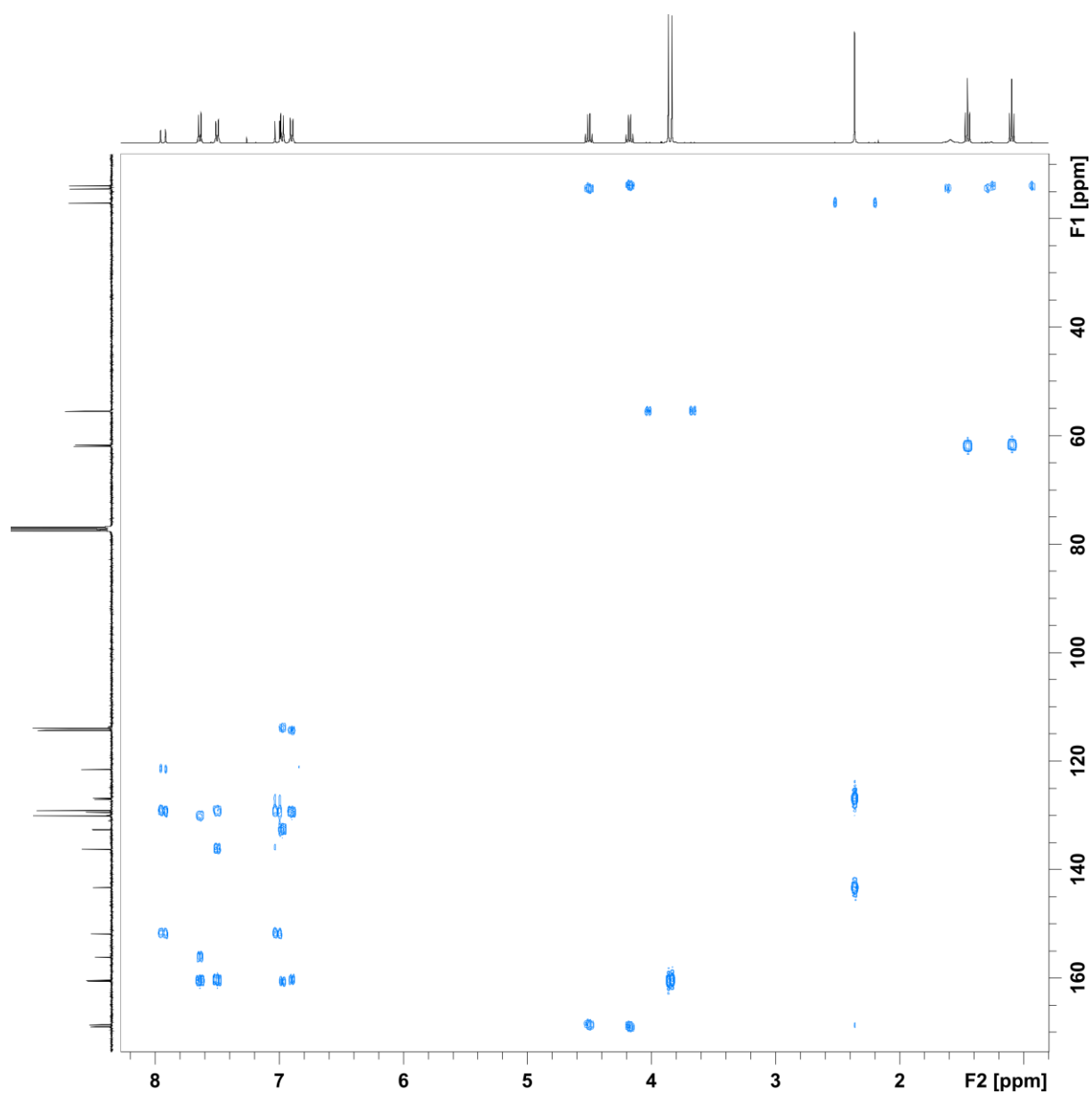
NOESY spectrum of **19n** (chloroform-*d*)



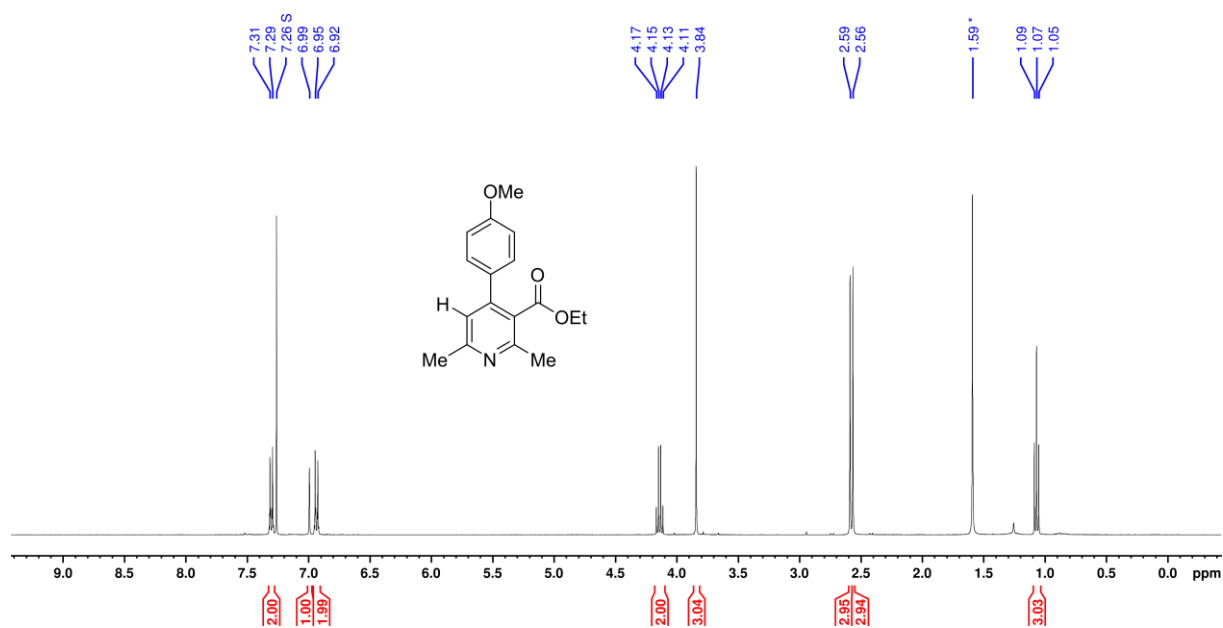
HSQC-TOCSY spectrum of 19n (chloroform-*d*)



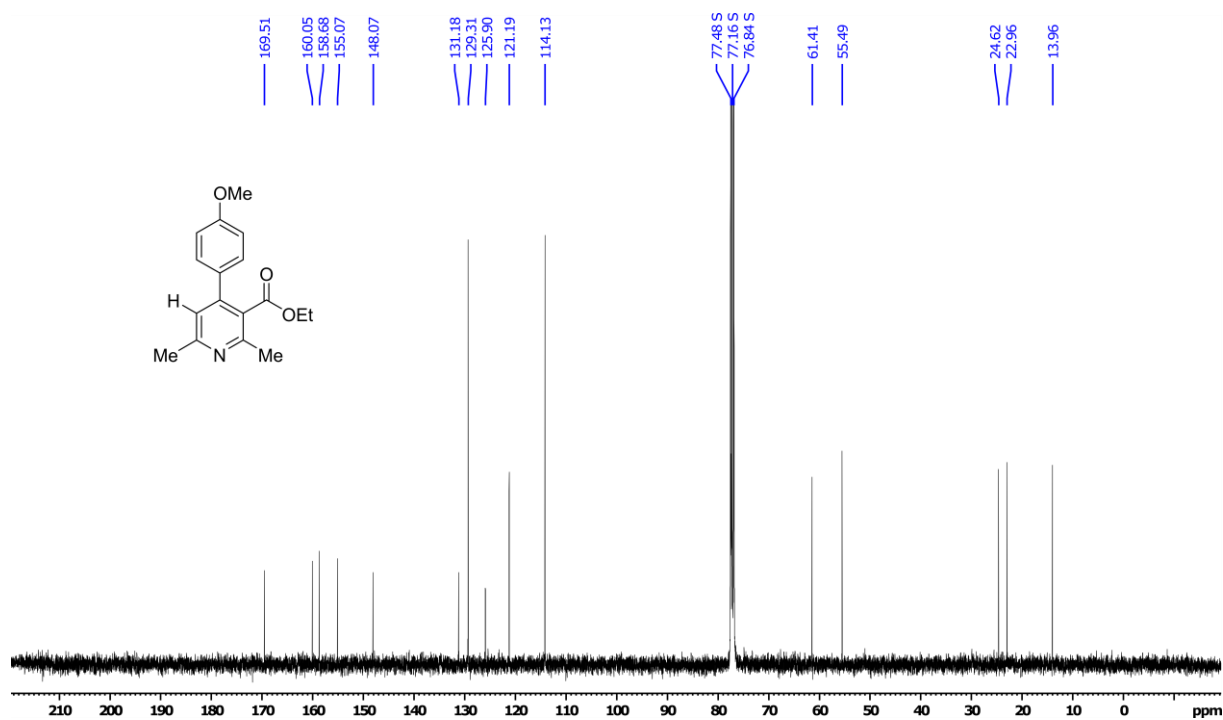
HMBC spectrum of **19n** (chloroform-*d*)



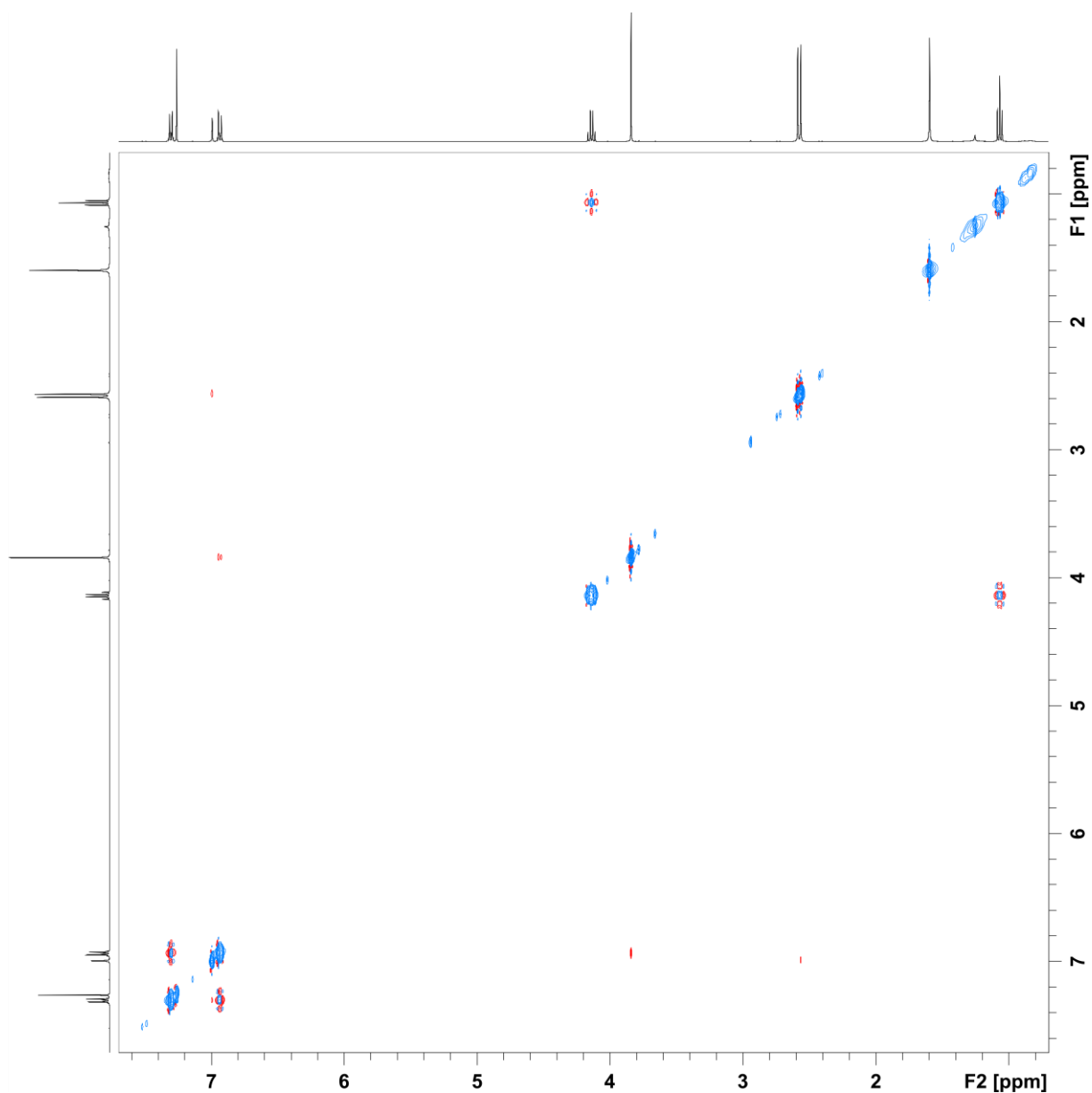
¹H NMR spectrum of 19o (400 MHz, chloroform-*d*)



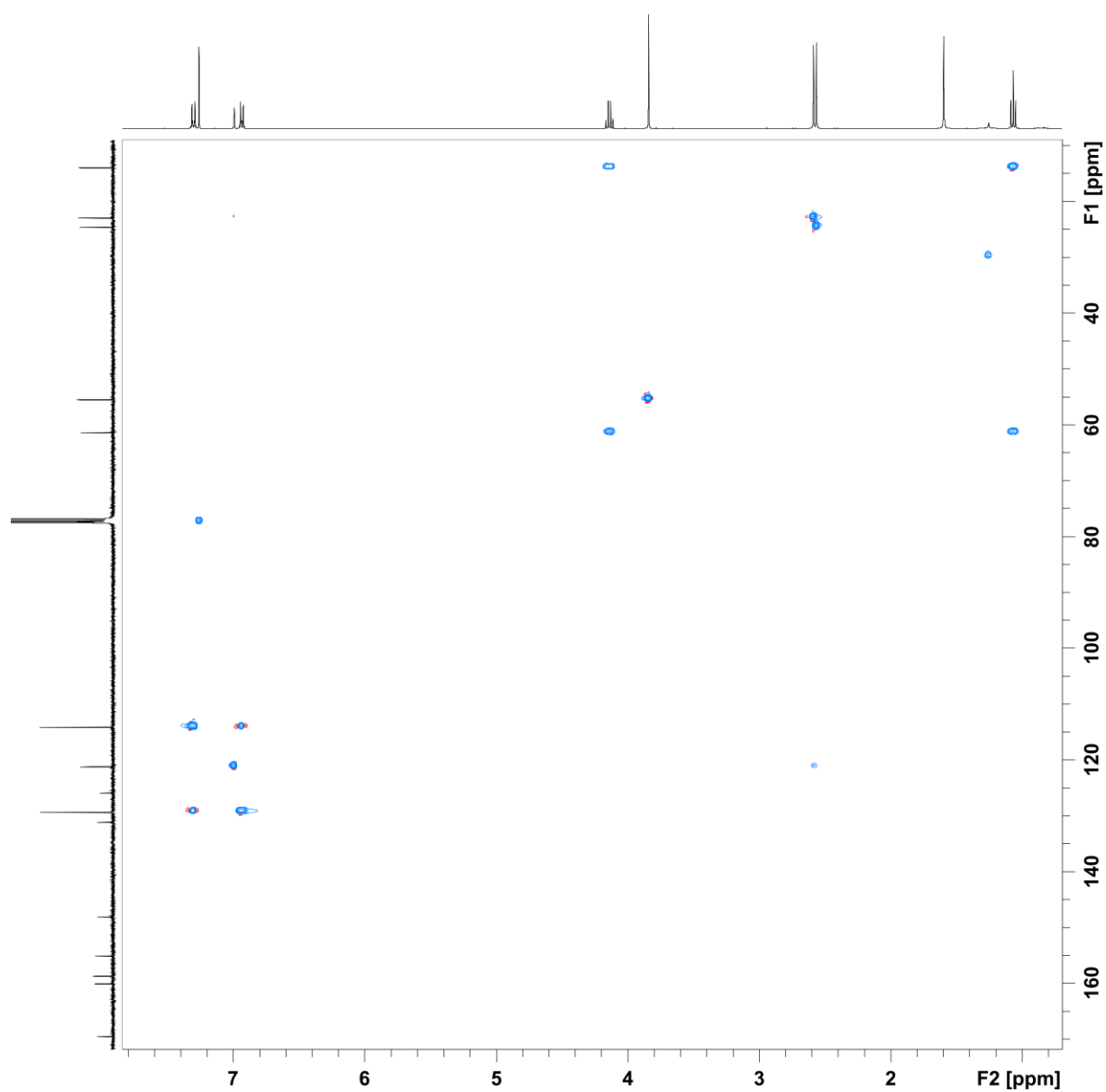
¹³C NMR spectrum of 19o (100 MHz, chloroform-*d*)



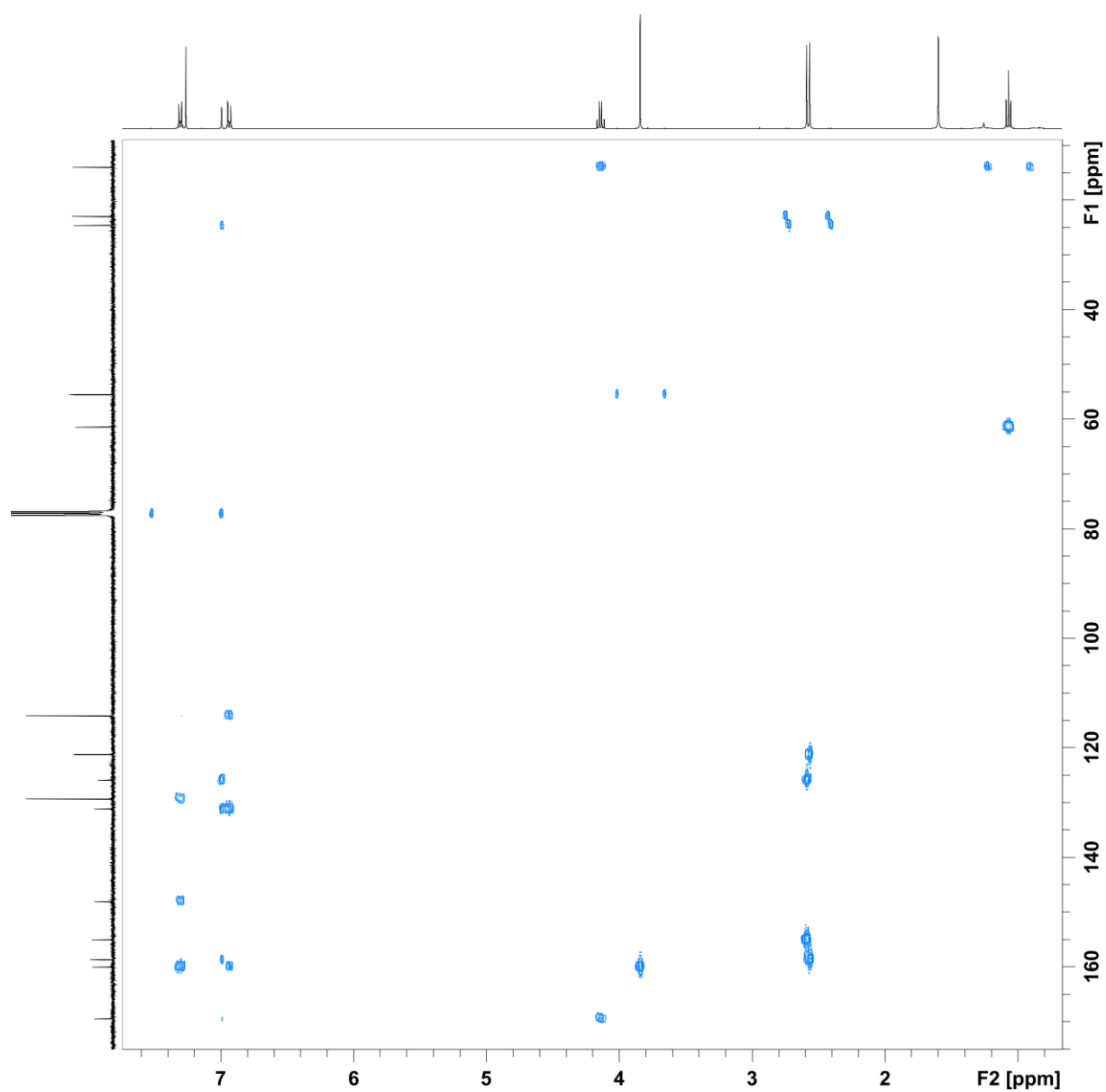
NOESY spectrum of 19o (chloroform-*d*)



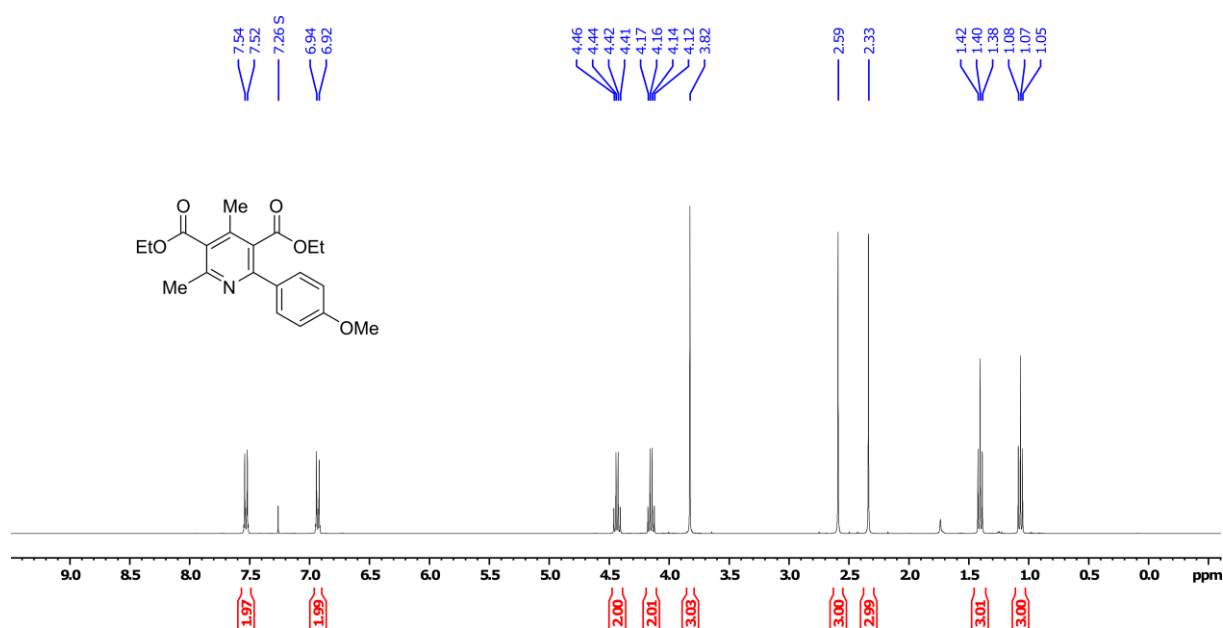
HSQC-TOCSY spectrum of 19o (chloroform-*d*)



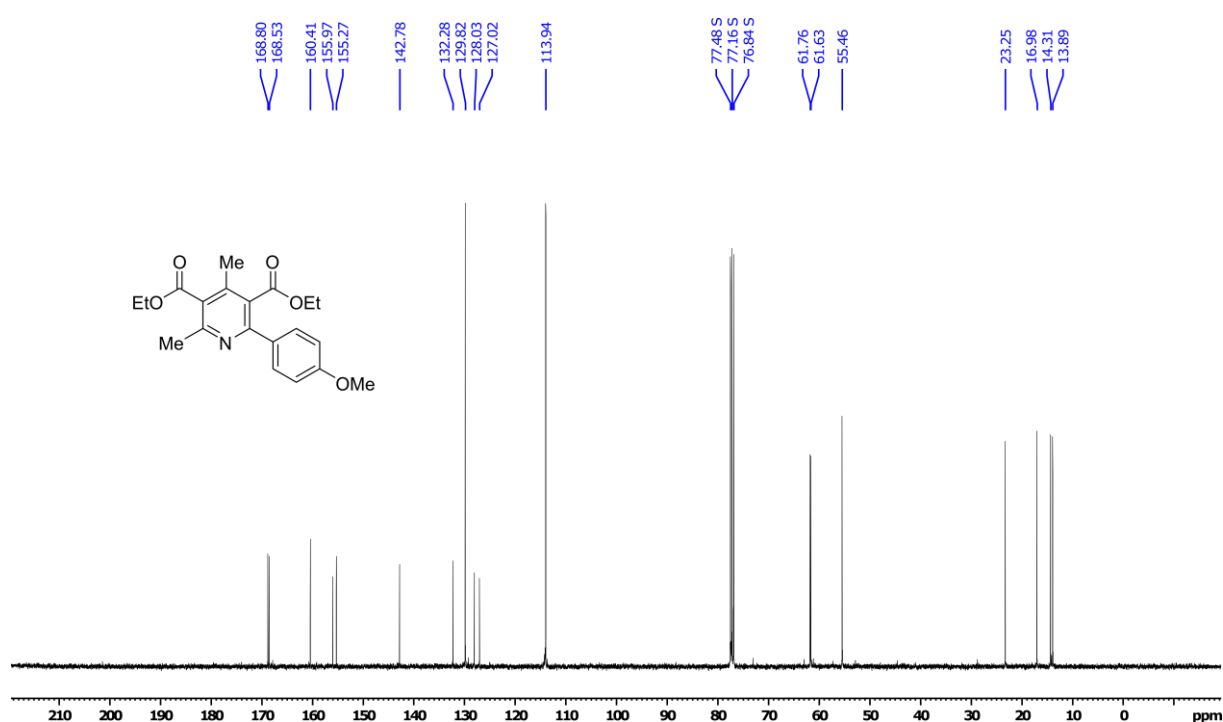
HMBC spectrum of 19o (chloroform-*d*).



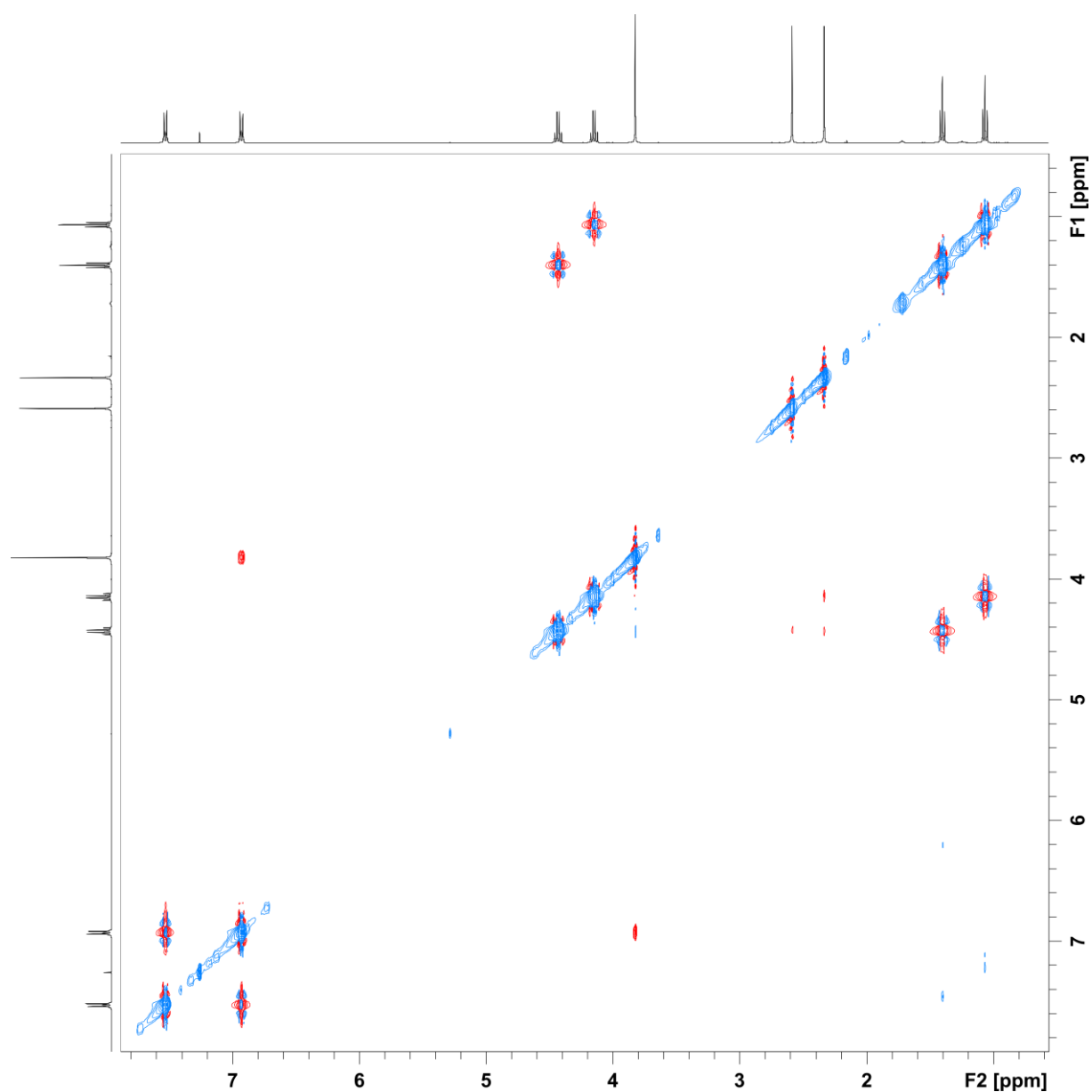
^1H NMR spectrum of 19r (400 MHz, chloroform-*d*)



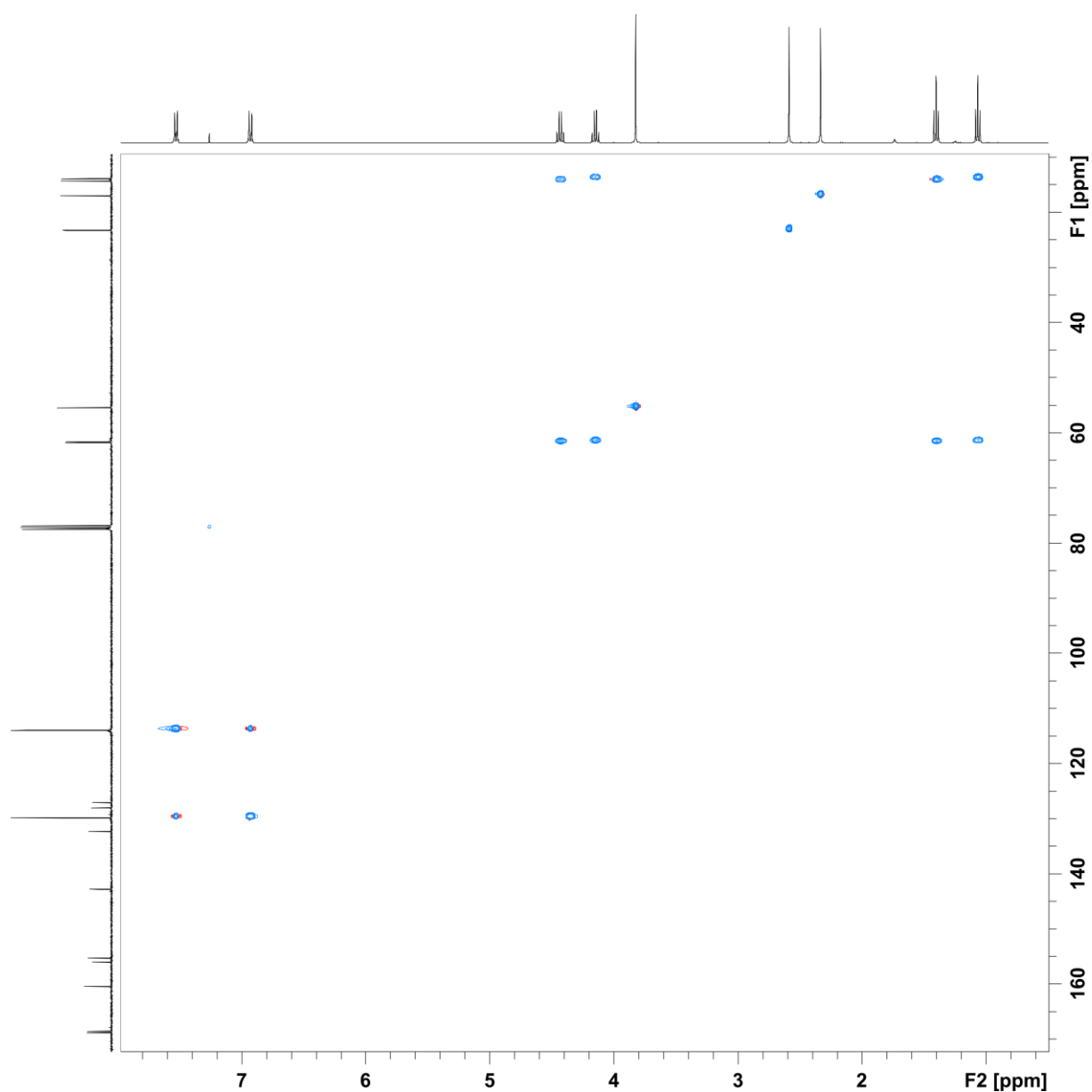
^{13}C NMR spectrum of 19r (100 MHz, chloroform-*d*).



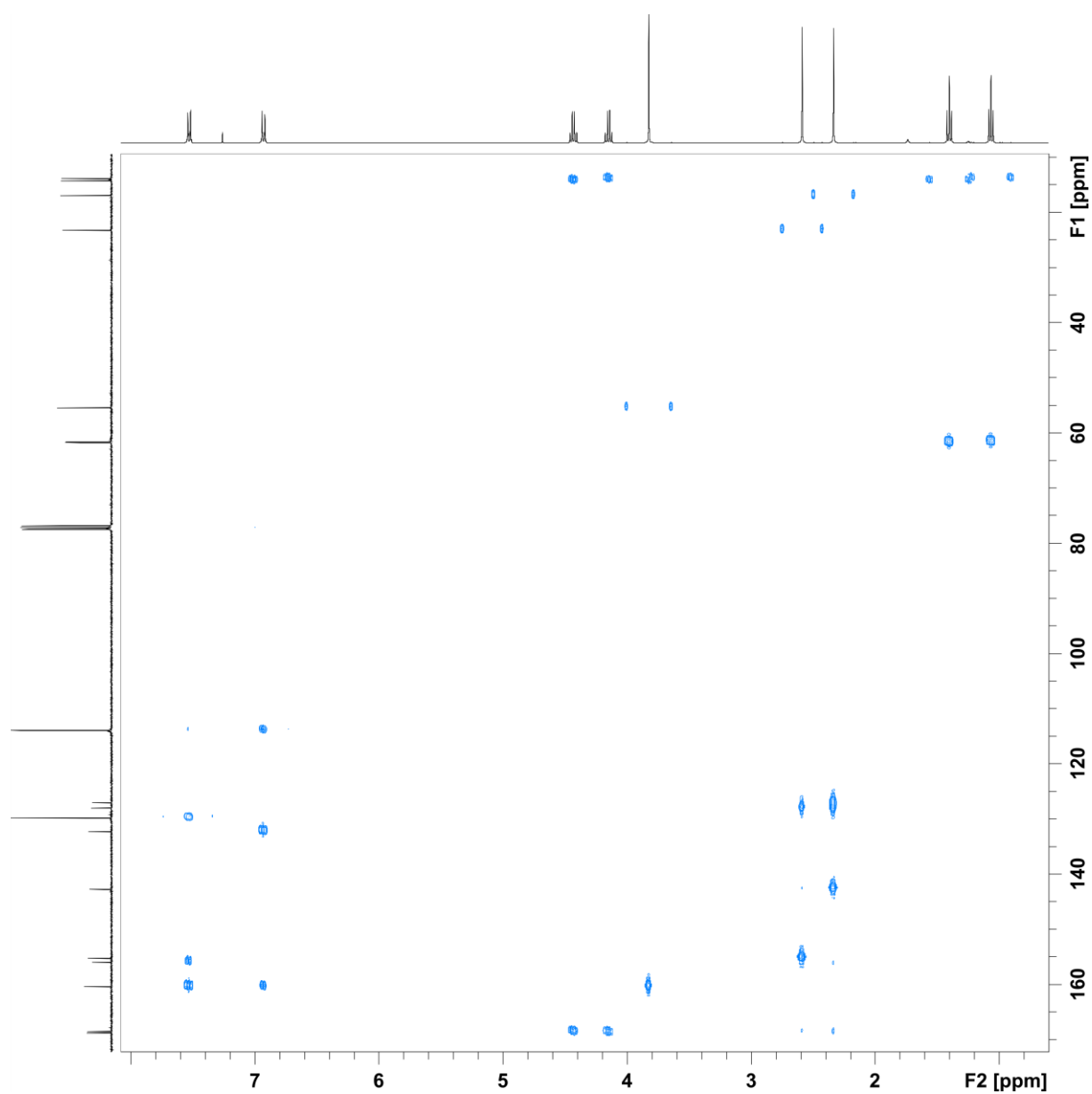
NOESY spectrum of **19r** (chloroform-*d*)



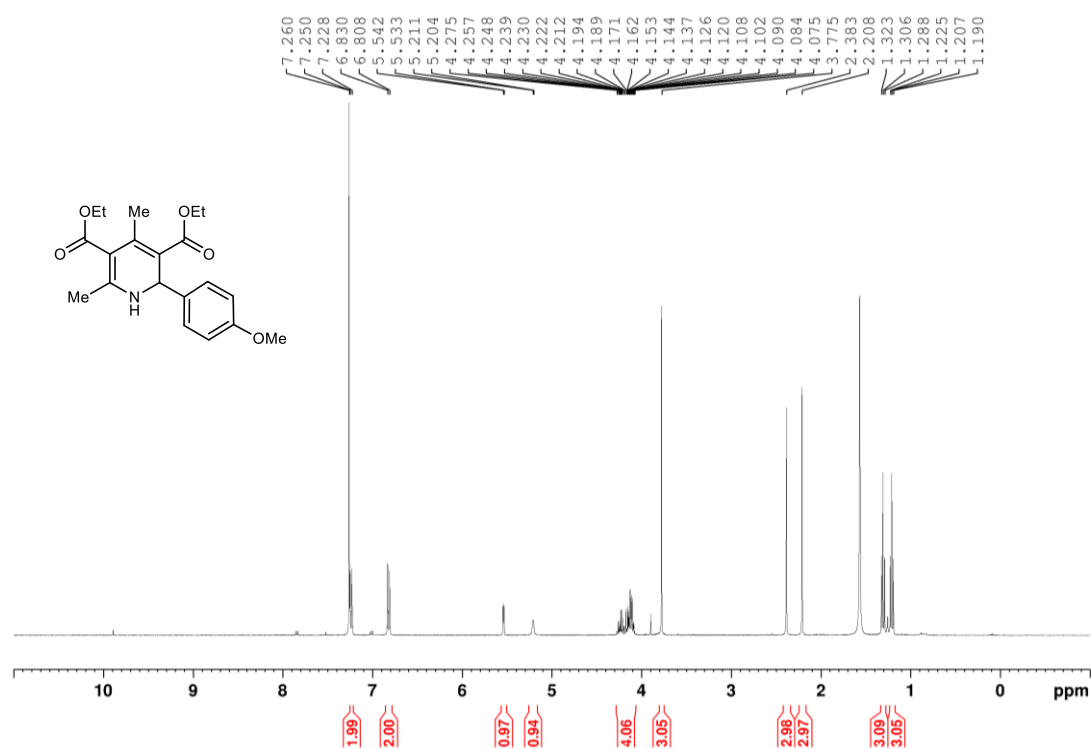
HSQC-TOCSY spectrum of **19r** (chloroform-*d*).



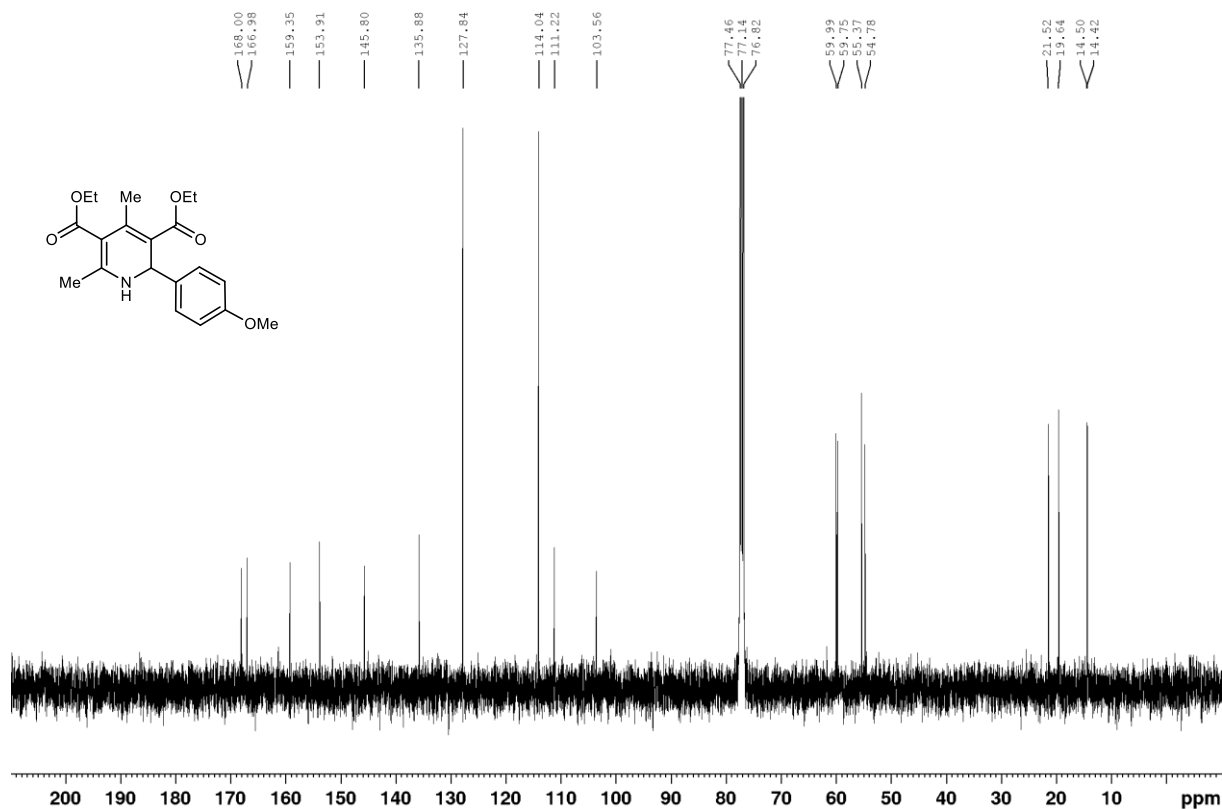
HMBC spectrum of **19r** (chloroform-*d*).



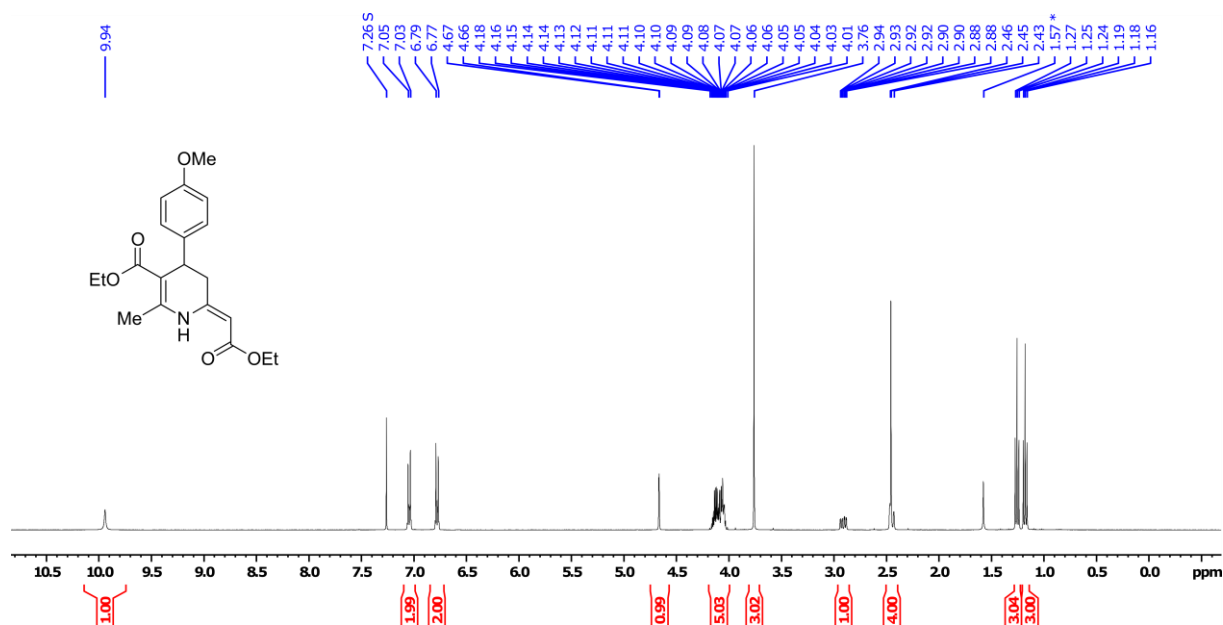
¹H NMR spectrum of 19p (400 MHz, chloroform-*d*).



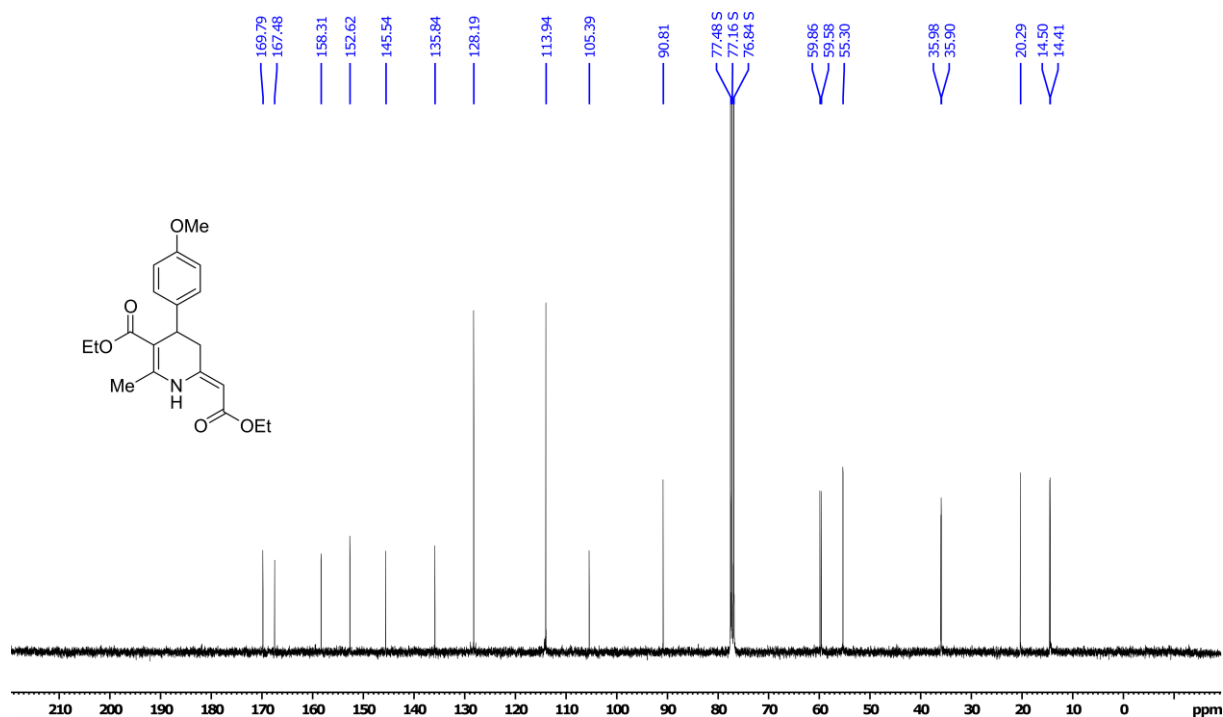
¹³C NMR spectrum of 19p (100 MHz, chloroform-*d*).



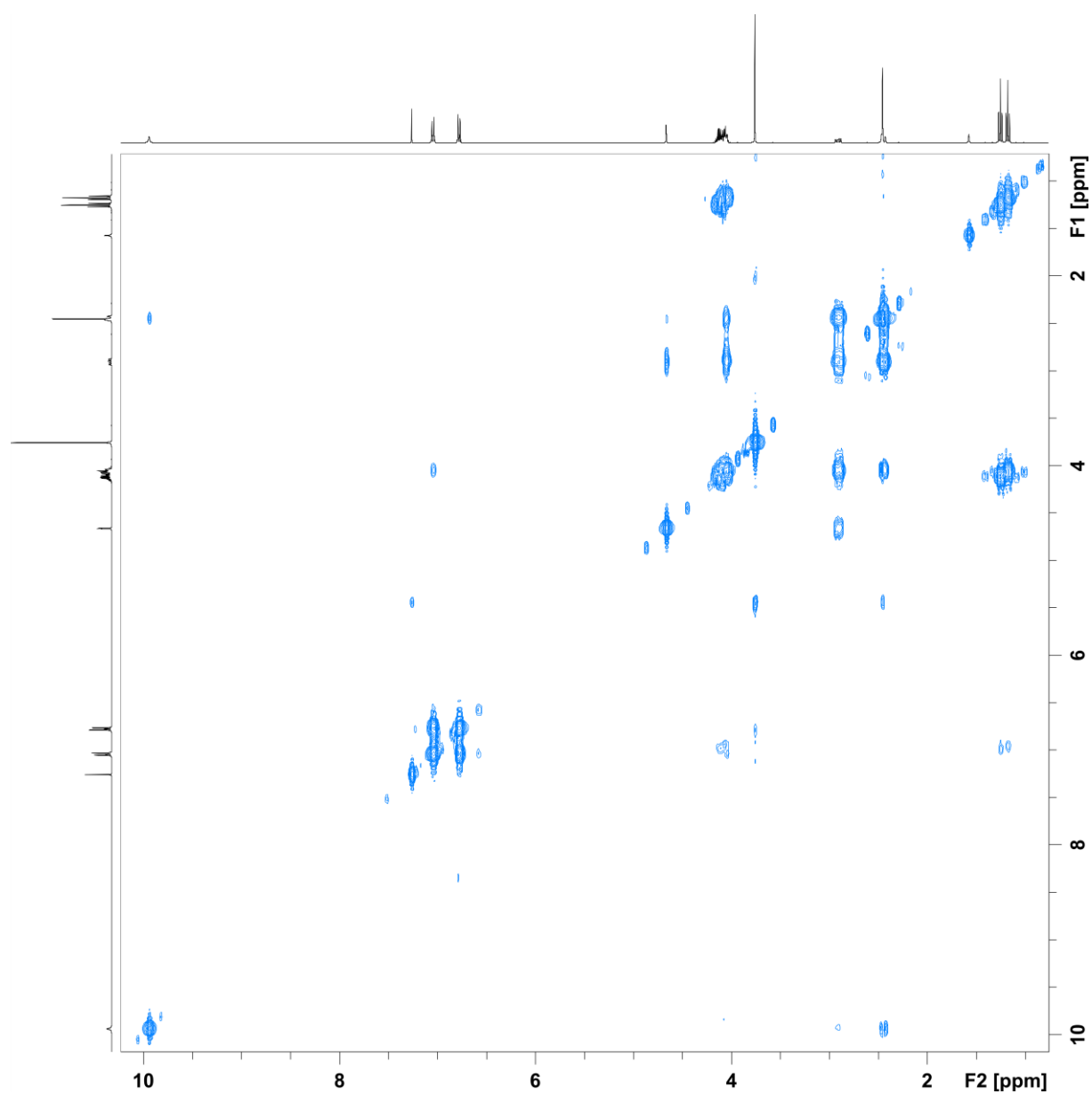
^1H NMR spectrum of 19j (400 MHz, chloroform-*d*)



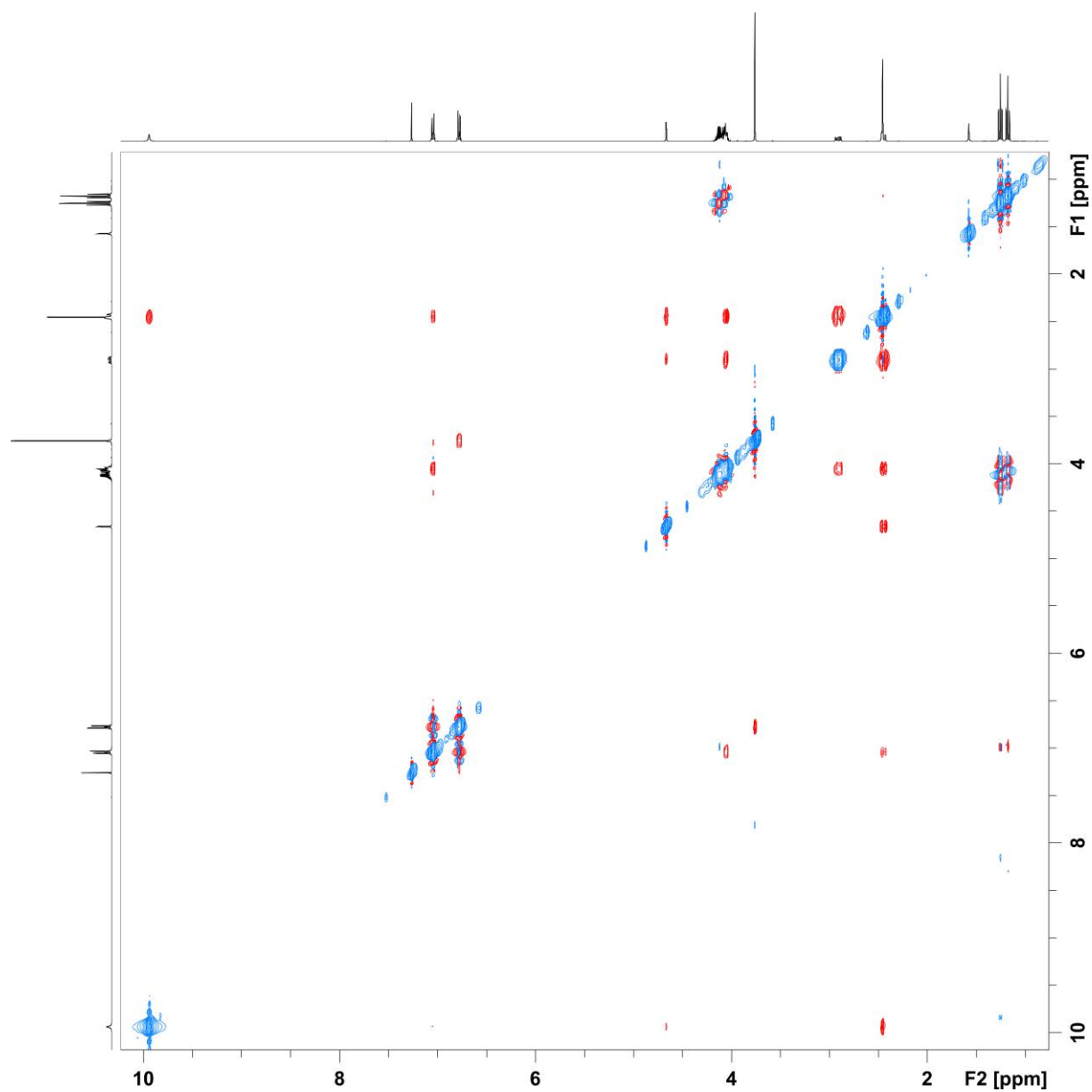
^{13}C NMR spectrum of 19j (100 MHz, chloroform-*d*)



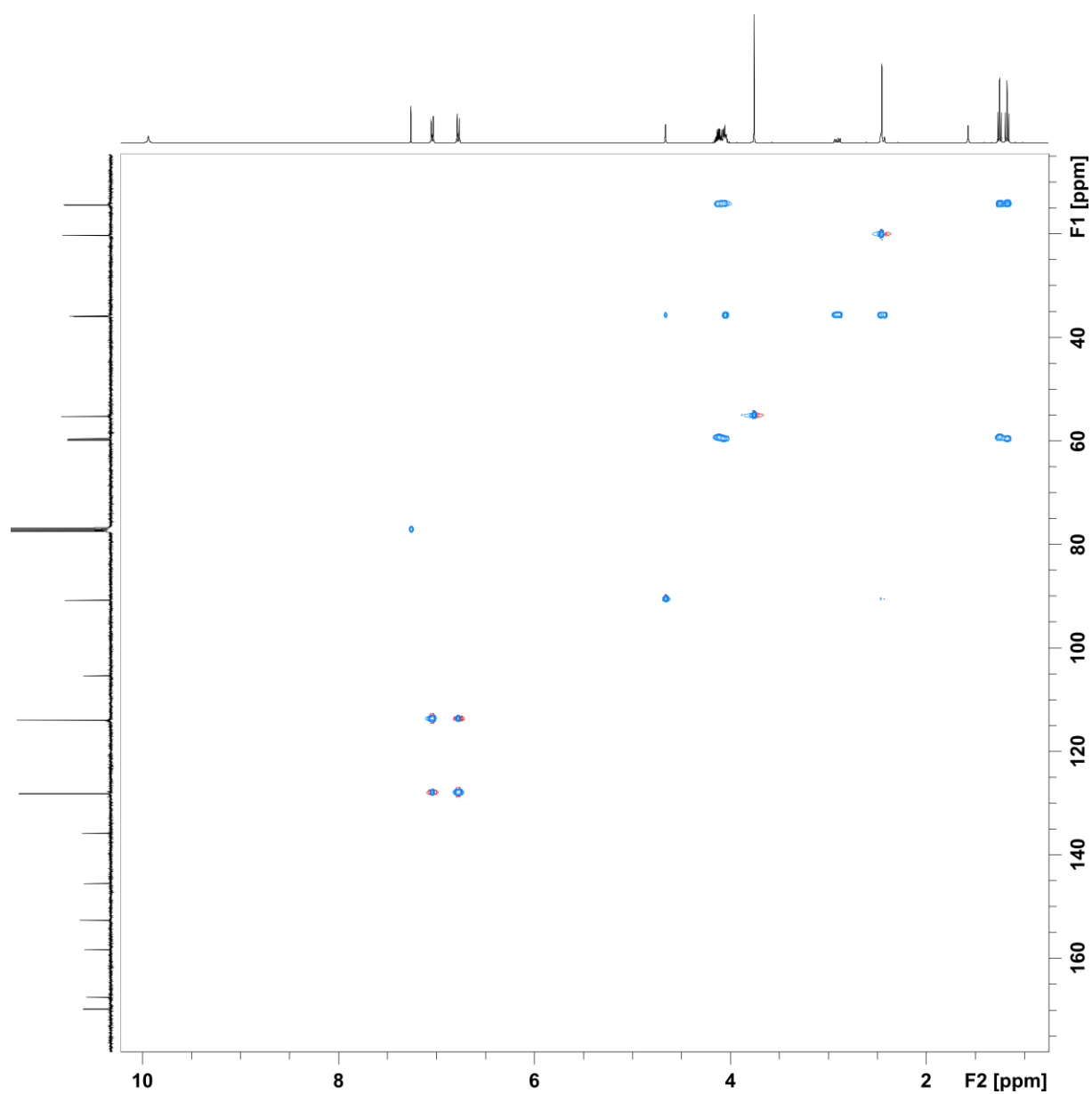
COSY spectrum of 19i (chloroform-*d*)



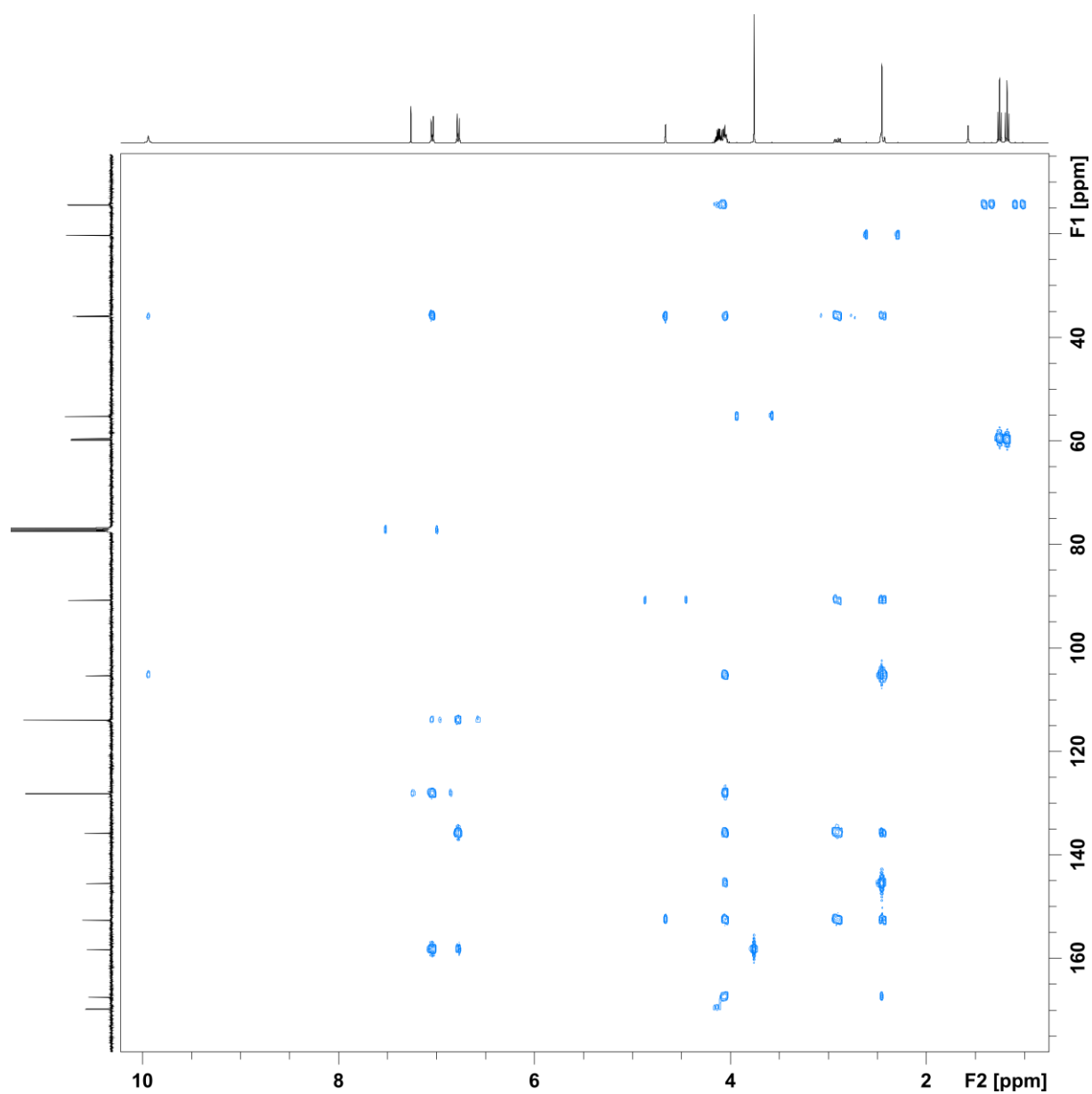
NOESY spectrum of **19j** (chloroform-*d*)



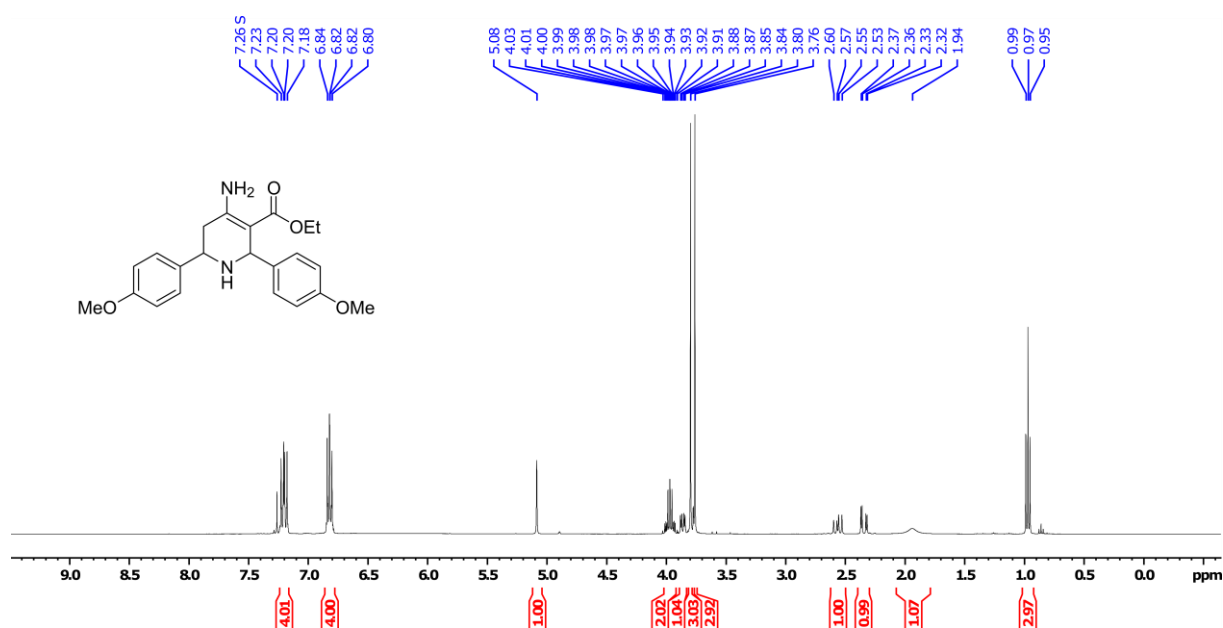
HSQC-TOCSY spectrum of **19j** (chloroform-*d*).



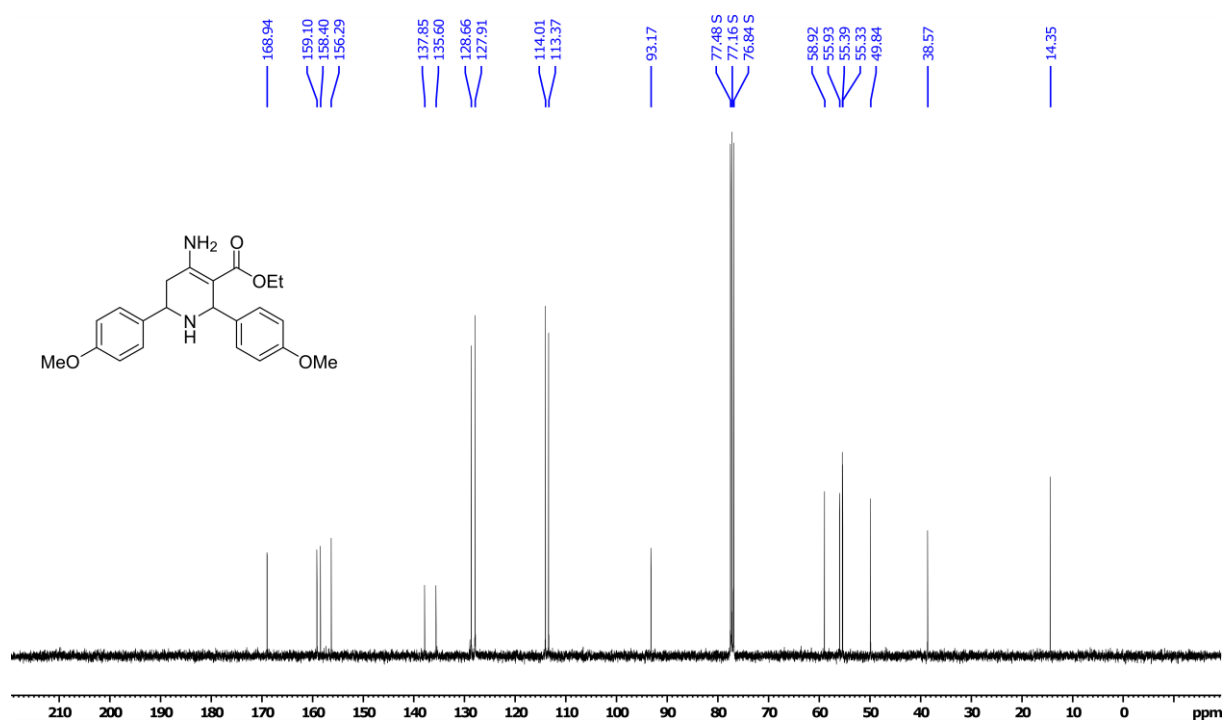
HMBC spectrum of **19j** (chloroform-*d*).



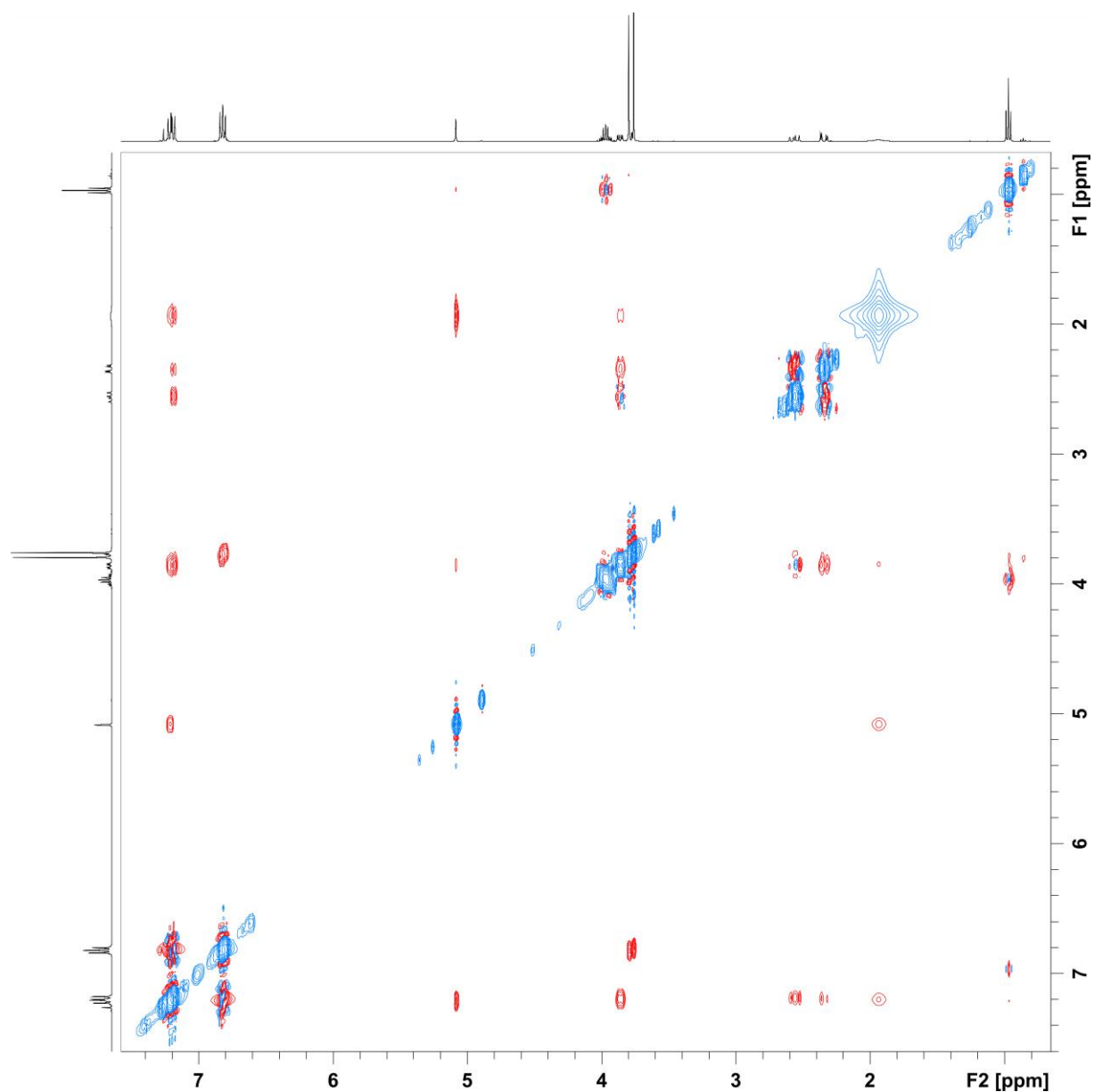
¹H NMR spectrum of 19e (400 MHz, chloroform-*d*).



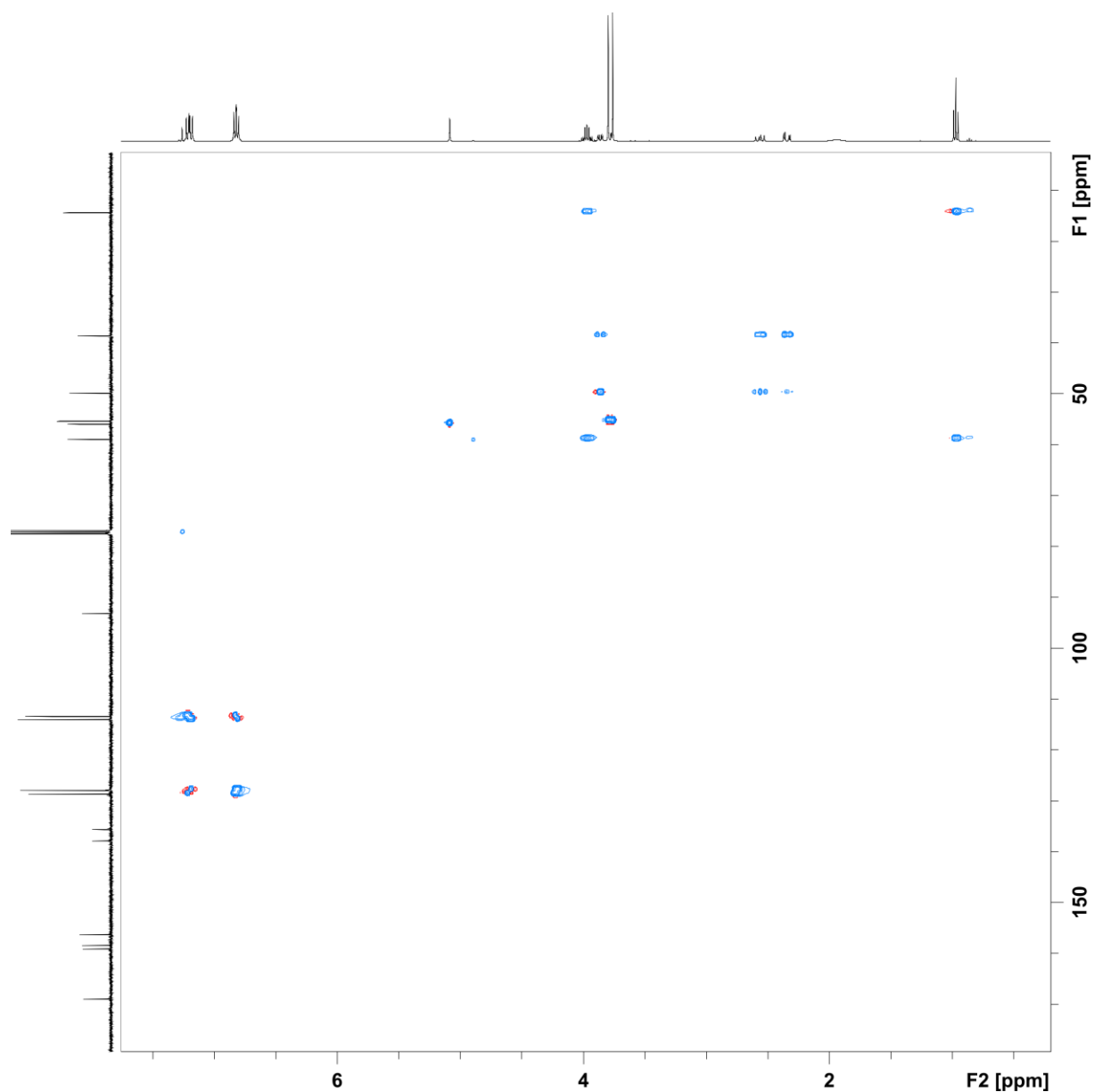
¹³C NMR spectrum of 19e (100 MHz, chloroform-*d*).



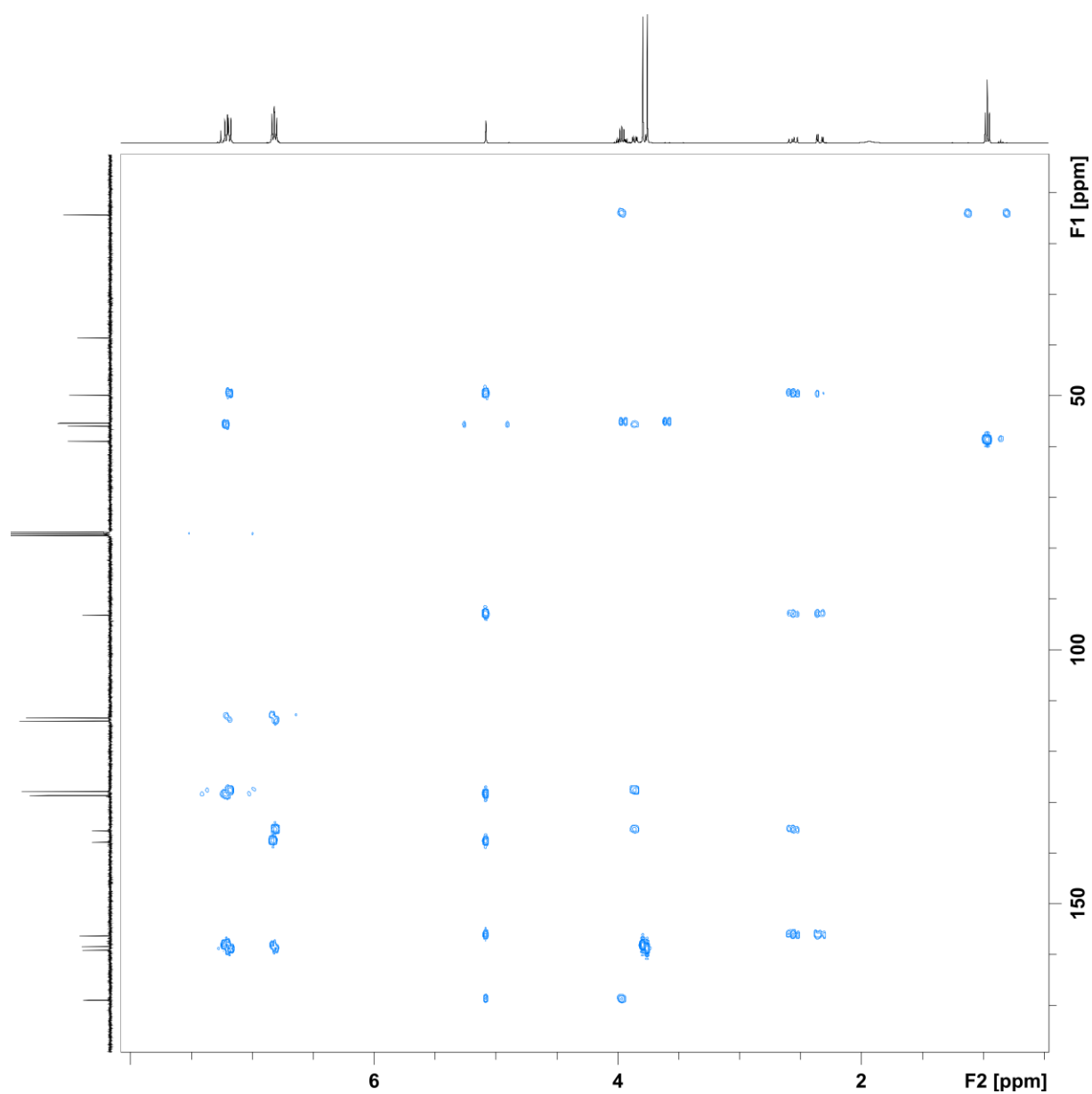
NOESY spectrum of **19e** (chloroform-*d*)



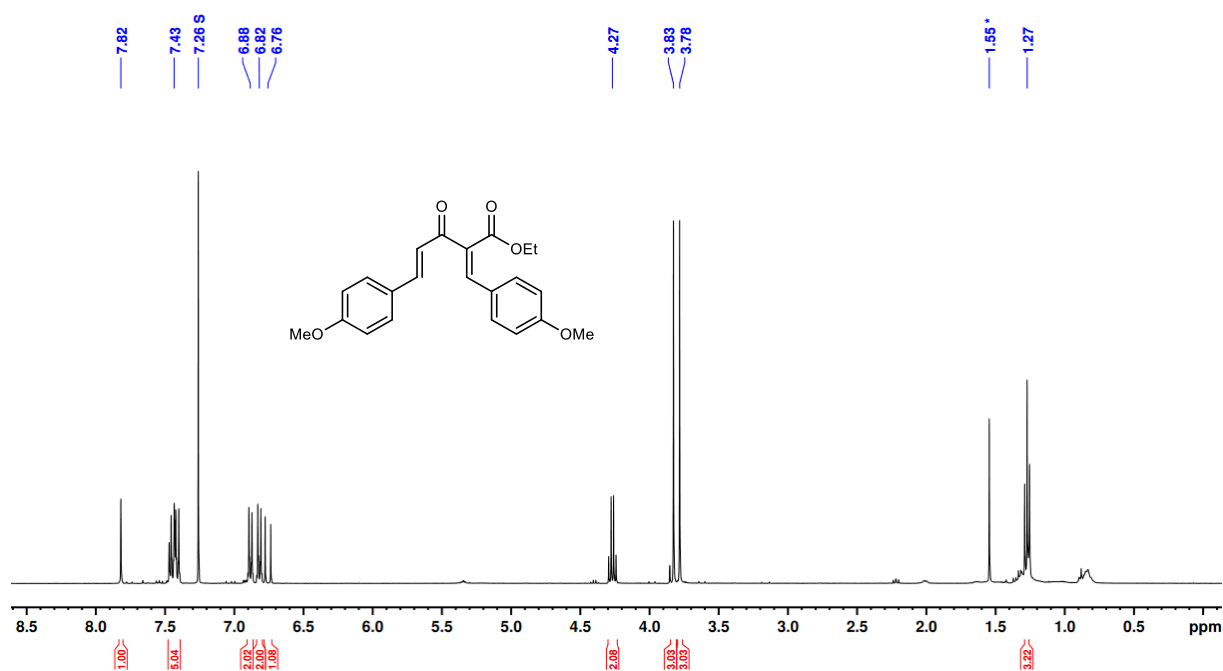
HSQC-TOCSY spectrum of 19e (chloroform-*d*).



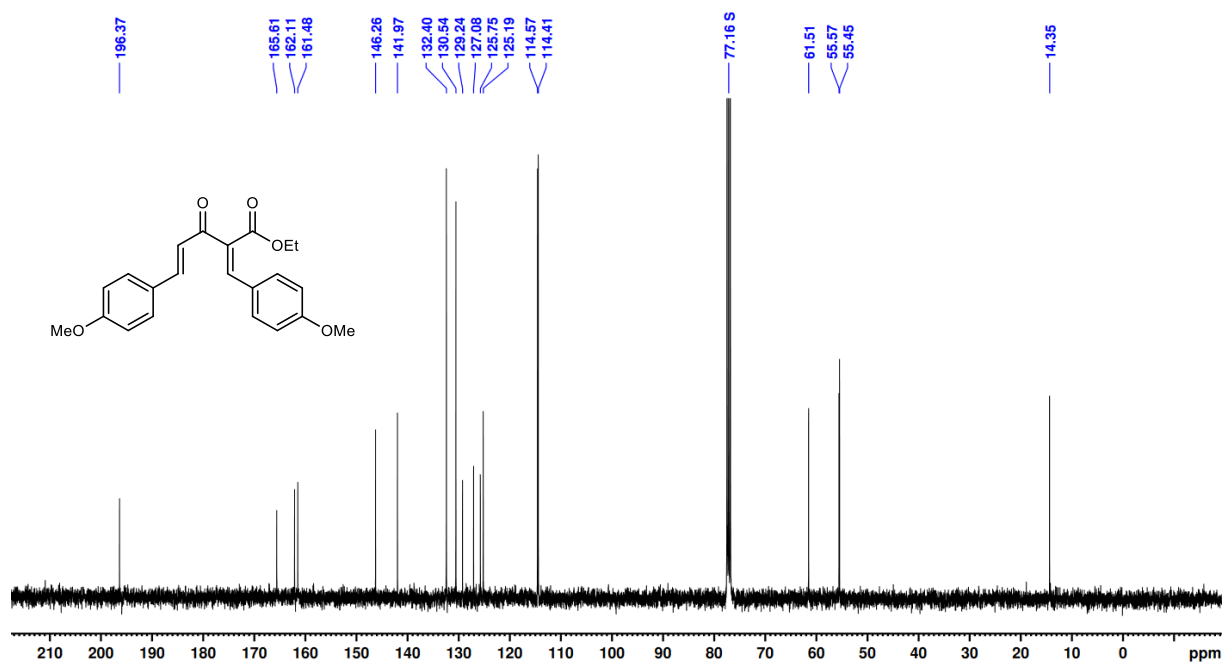
HMBC spectrum of **19e** (chloroform-*d*).



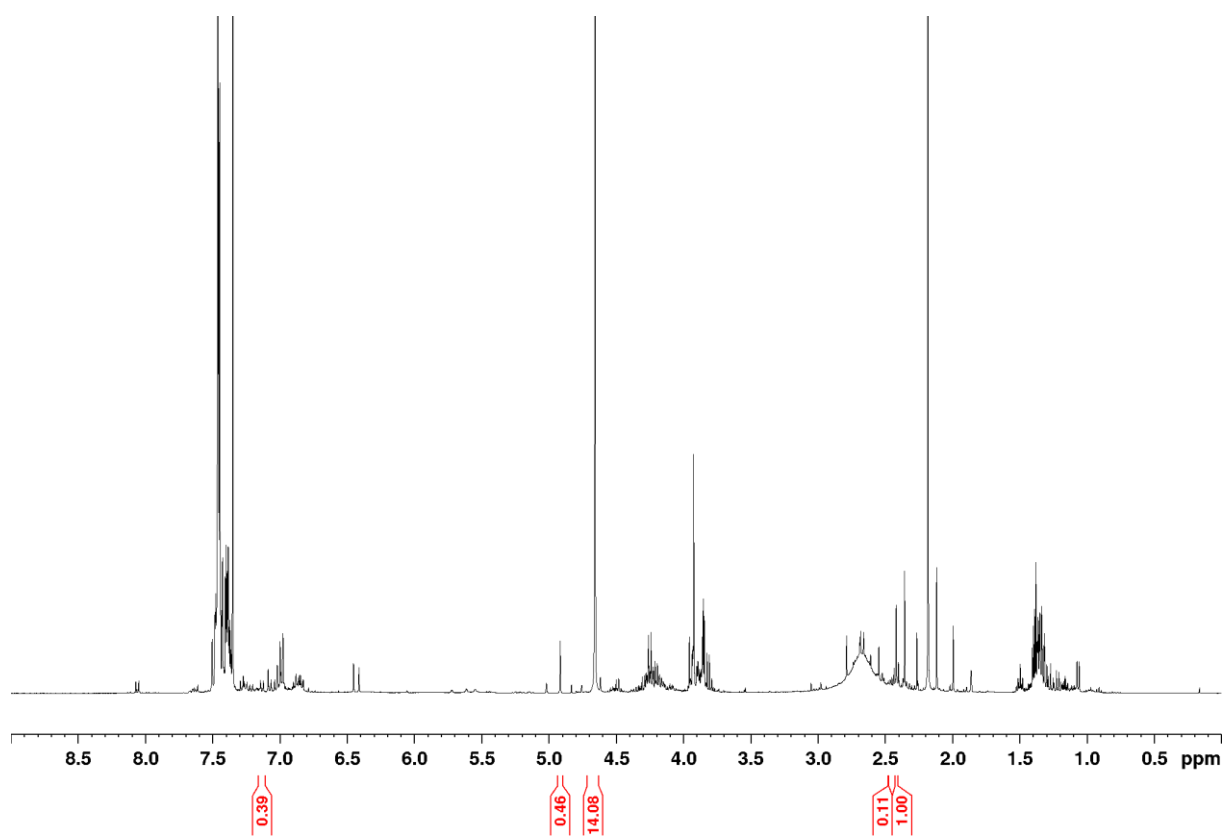
¹H NMR spectrum of 19h (400 MHz, chloroform-*d*).



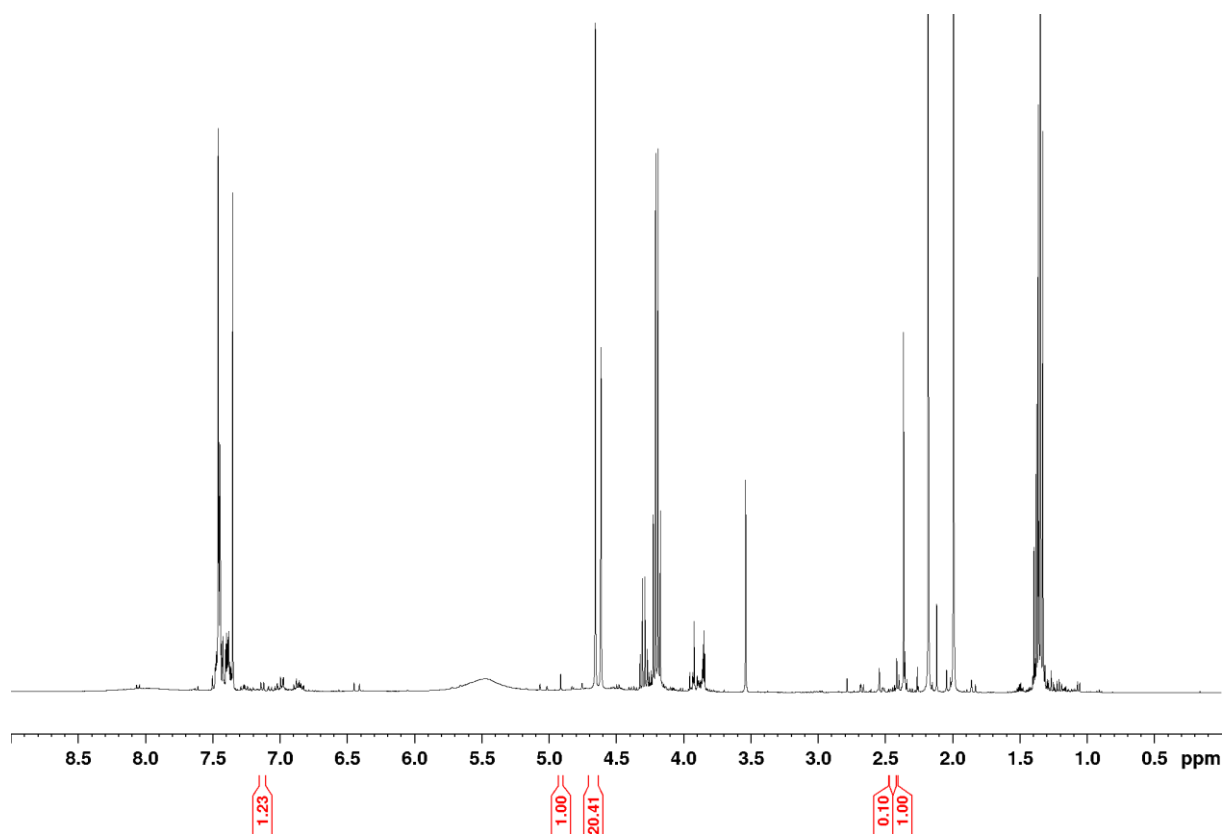
¹³C NMR spectrum of 19h (100 MHz, chloroform-*d*).



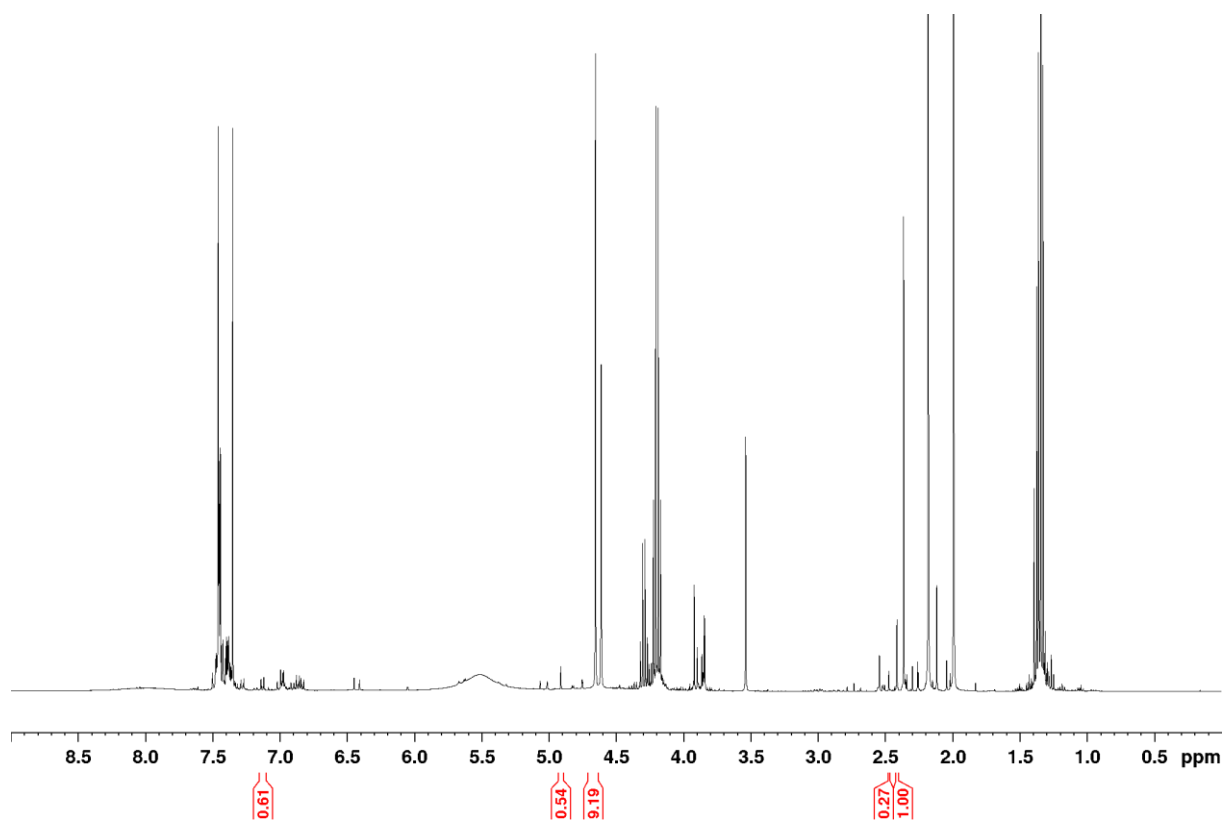
^1H NMR (400 MHz, CDCl_3 , 300 K) spectrum of sample 1 in Figure S68



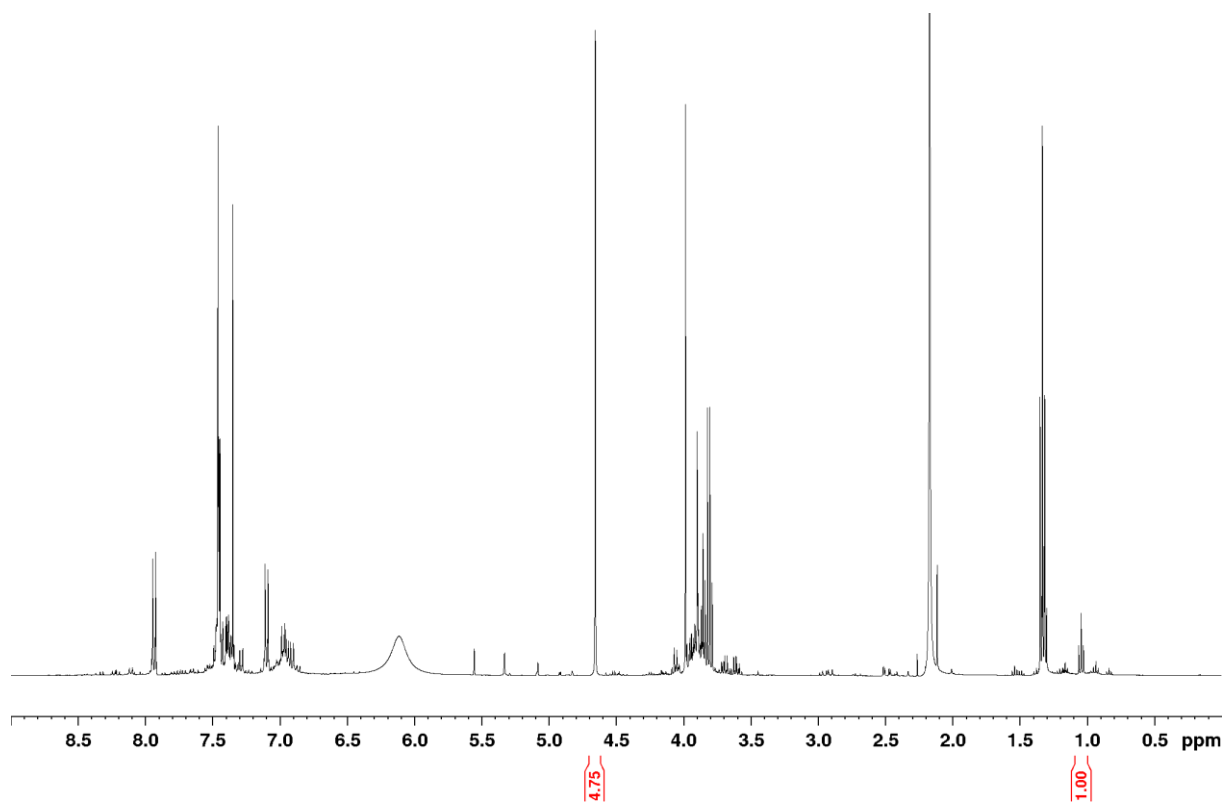
^1H NMR (400 MHz, CDCl_3 , 300 K) spectrum of sample 2 in Figure S68



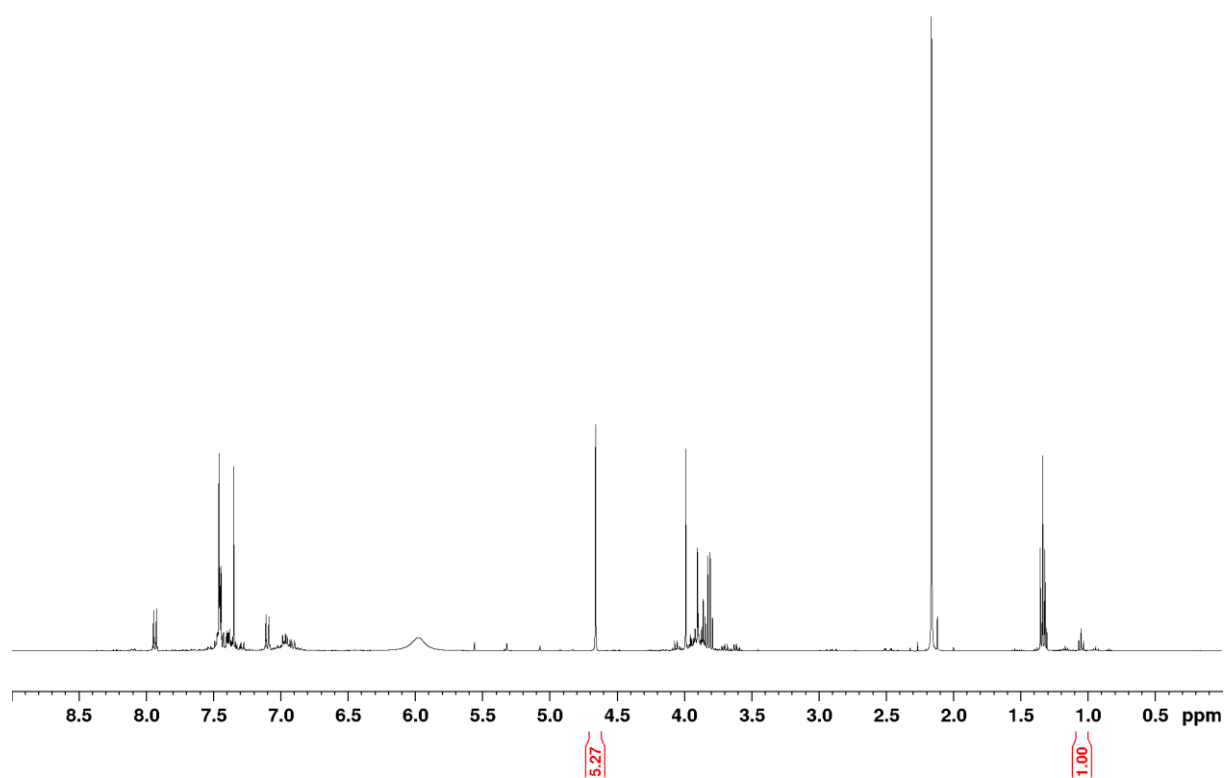
^1H NMR (400 MHz, CDCl_3 , 300 K) spectrum of sample 3 in Figure S68



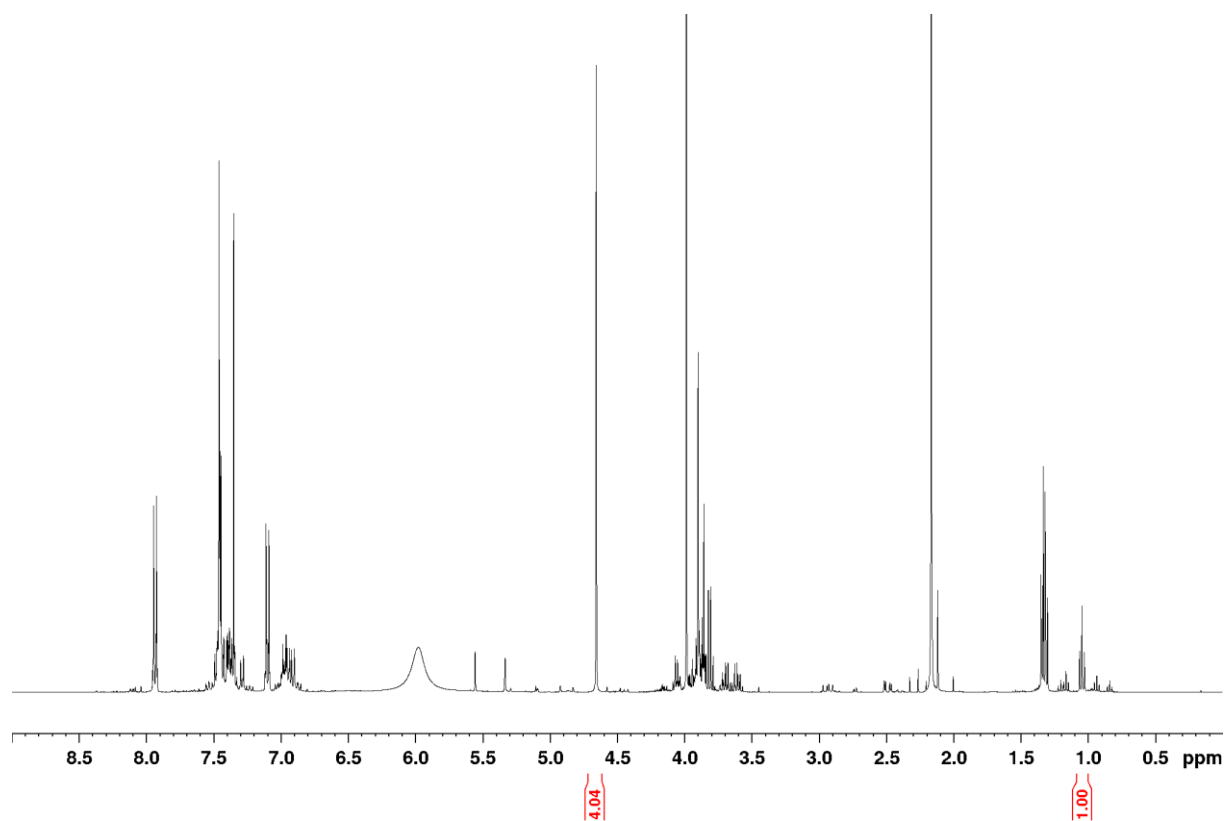
^1H NMR (400 MHz, CDCl_3 , 300 K) spectrum of sample 4 in Figure S68



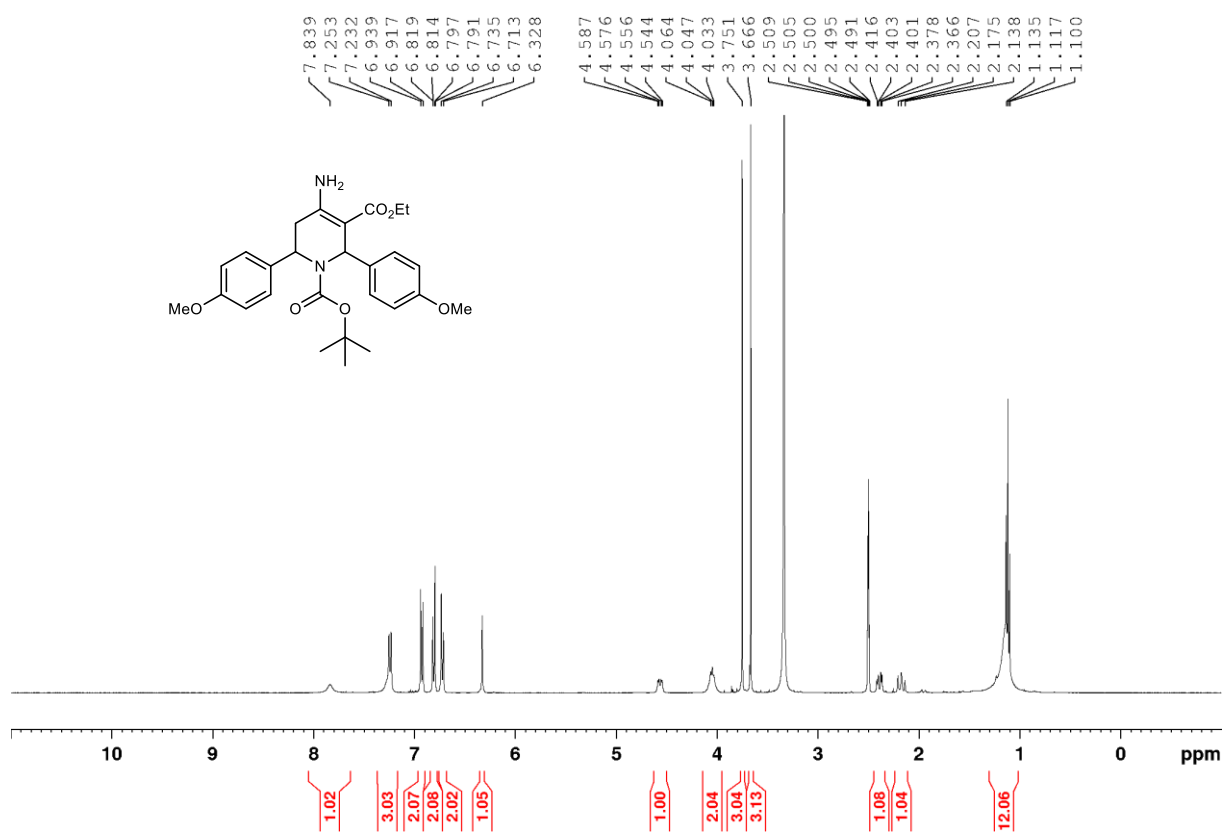
^1H NMR (400 MHz, CDCl_3 , 300 K) spectrum of sample 5 in Figure S68



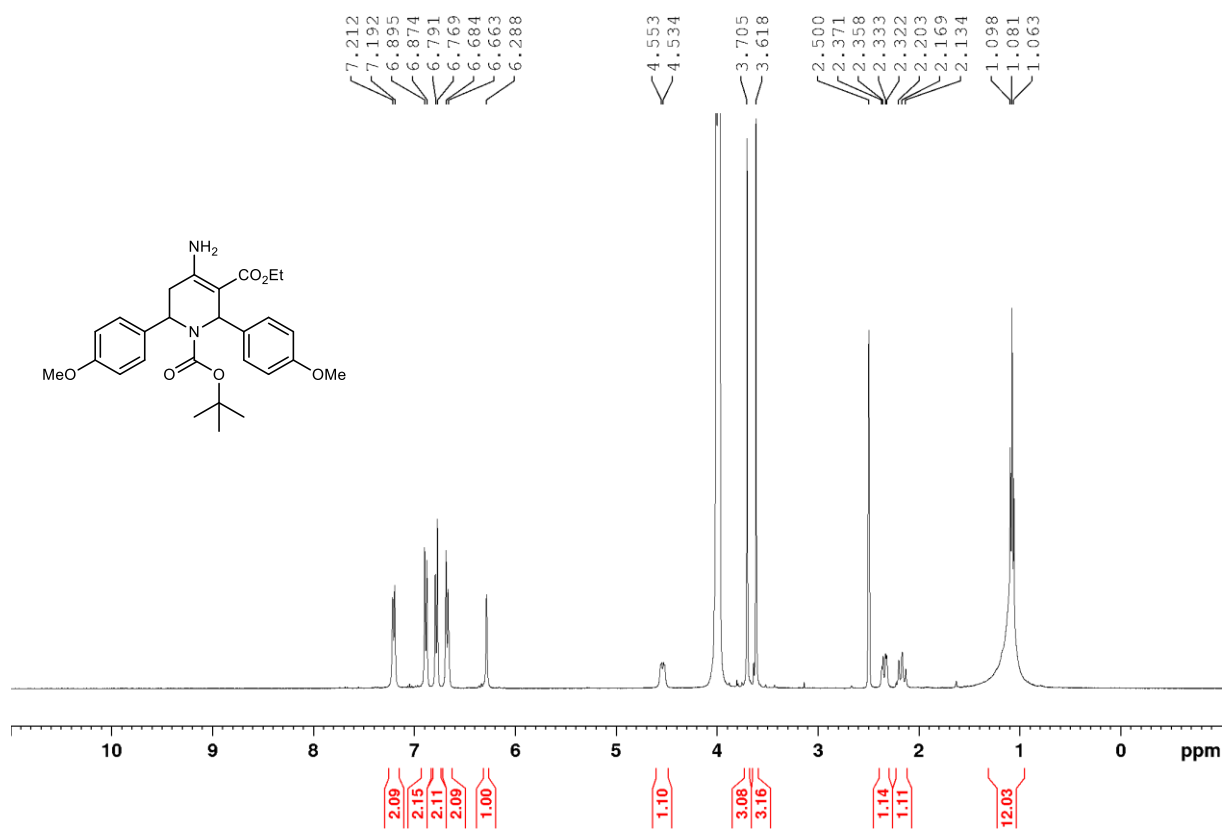
^1H NMR (400 MHz, CDCl_3 , 300 K) spectrum of sample 6 in Figure S68



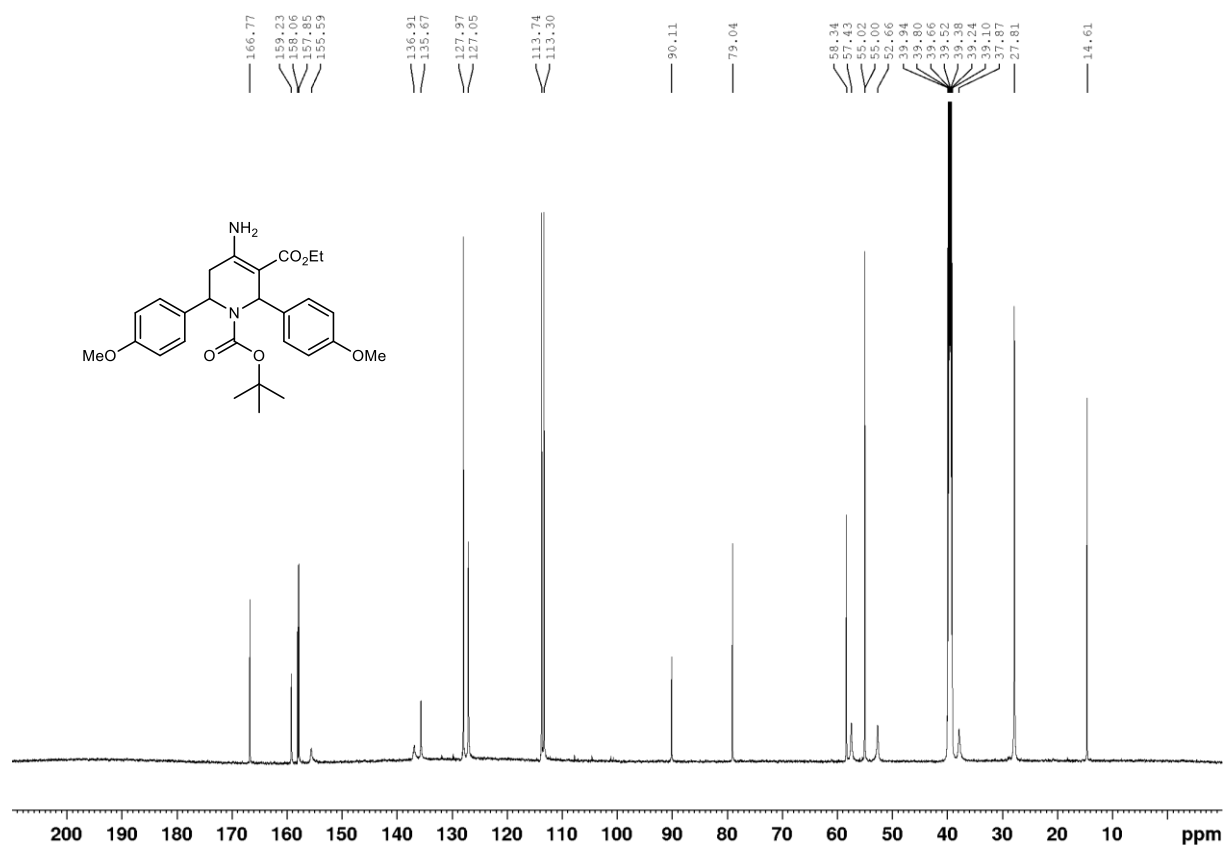
^1H NMR spectrum of 19e -BOC (400 MHz, $\text{DMSO}-d_6$).



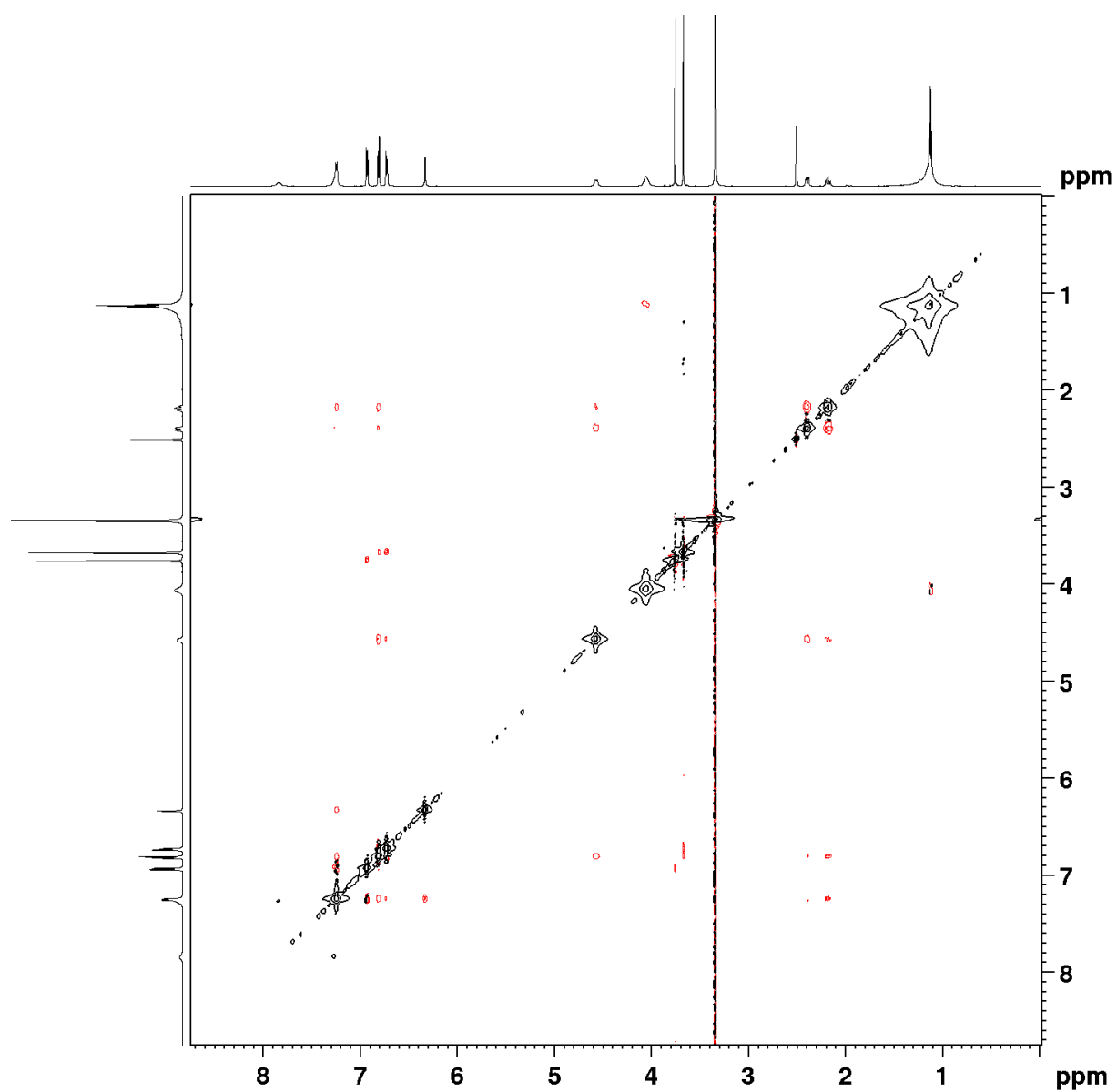
^1H NMR spectrum of 19e -BOC with addition of D_2O (400 MHz, $\text{DMSO}-d_6$).



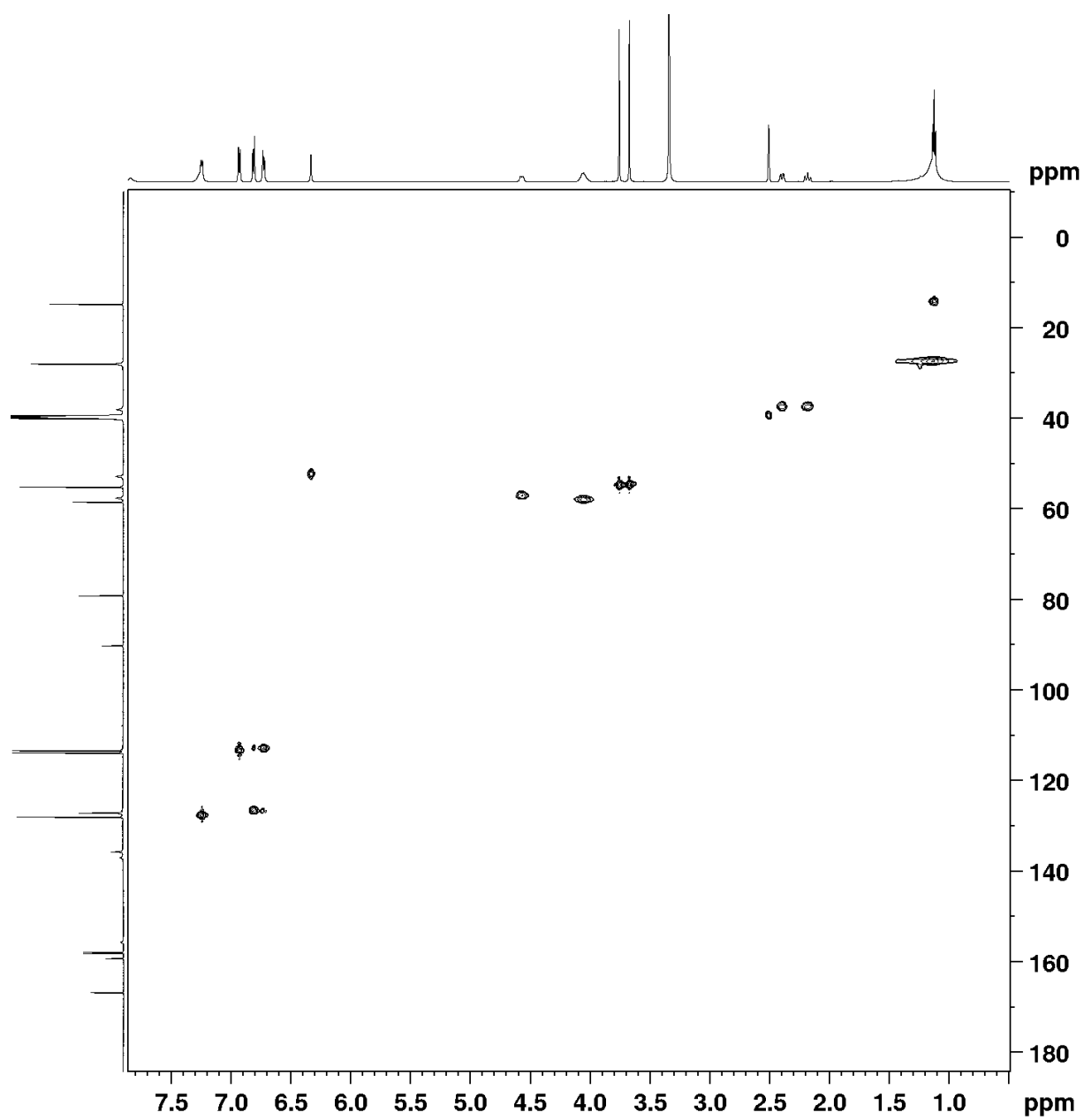
^{13}C NMR spectrum of 19e -BOC (100 MHz, $\text{DMSO}-d_6$).



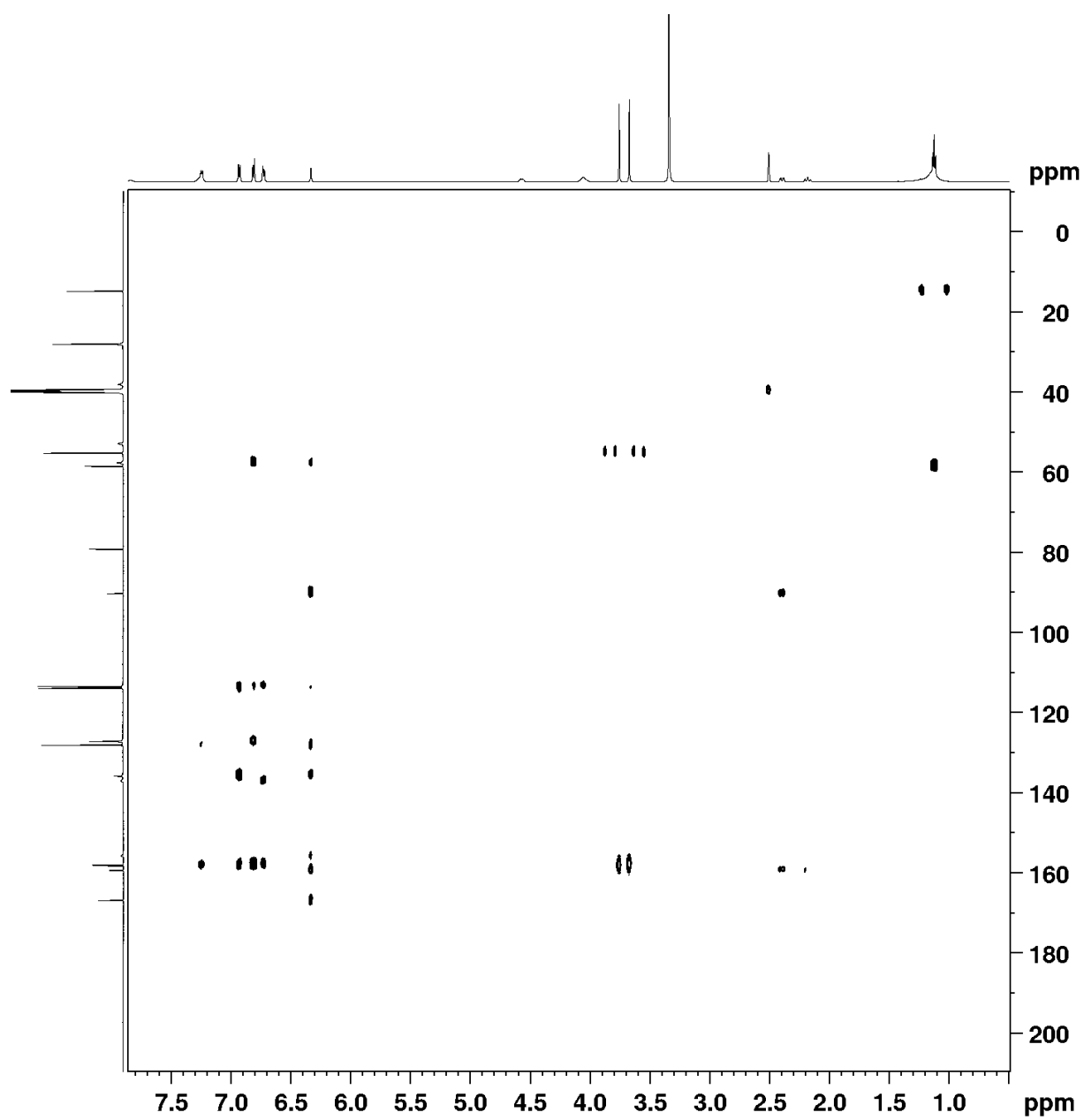
ROESY spectrum of **19e** -BOC (DMSO-*d*₆).



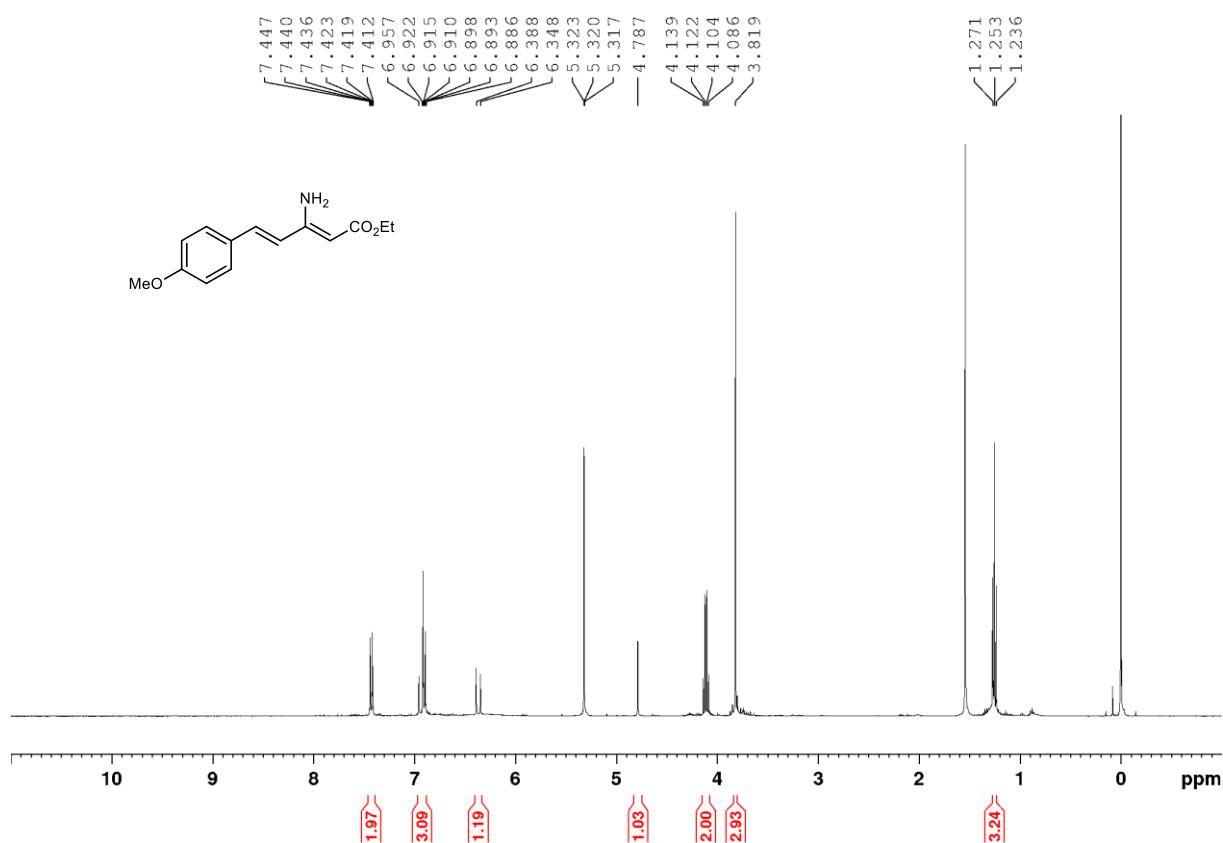
HSQC-TOCSY spectrum of 19e -BOC (DMSO-*d*₆).



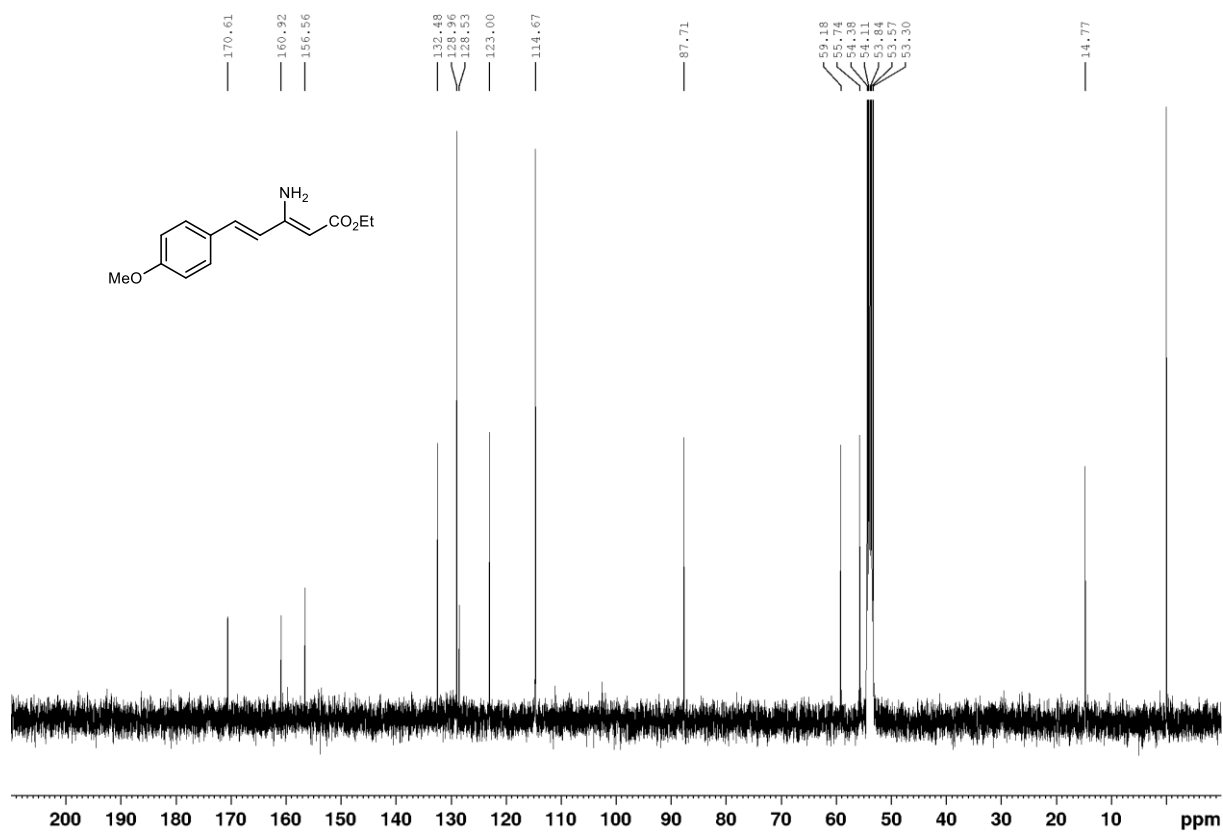
HMBC spectrum of 19e-BOC (DMSO-*d*₆).



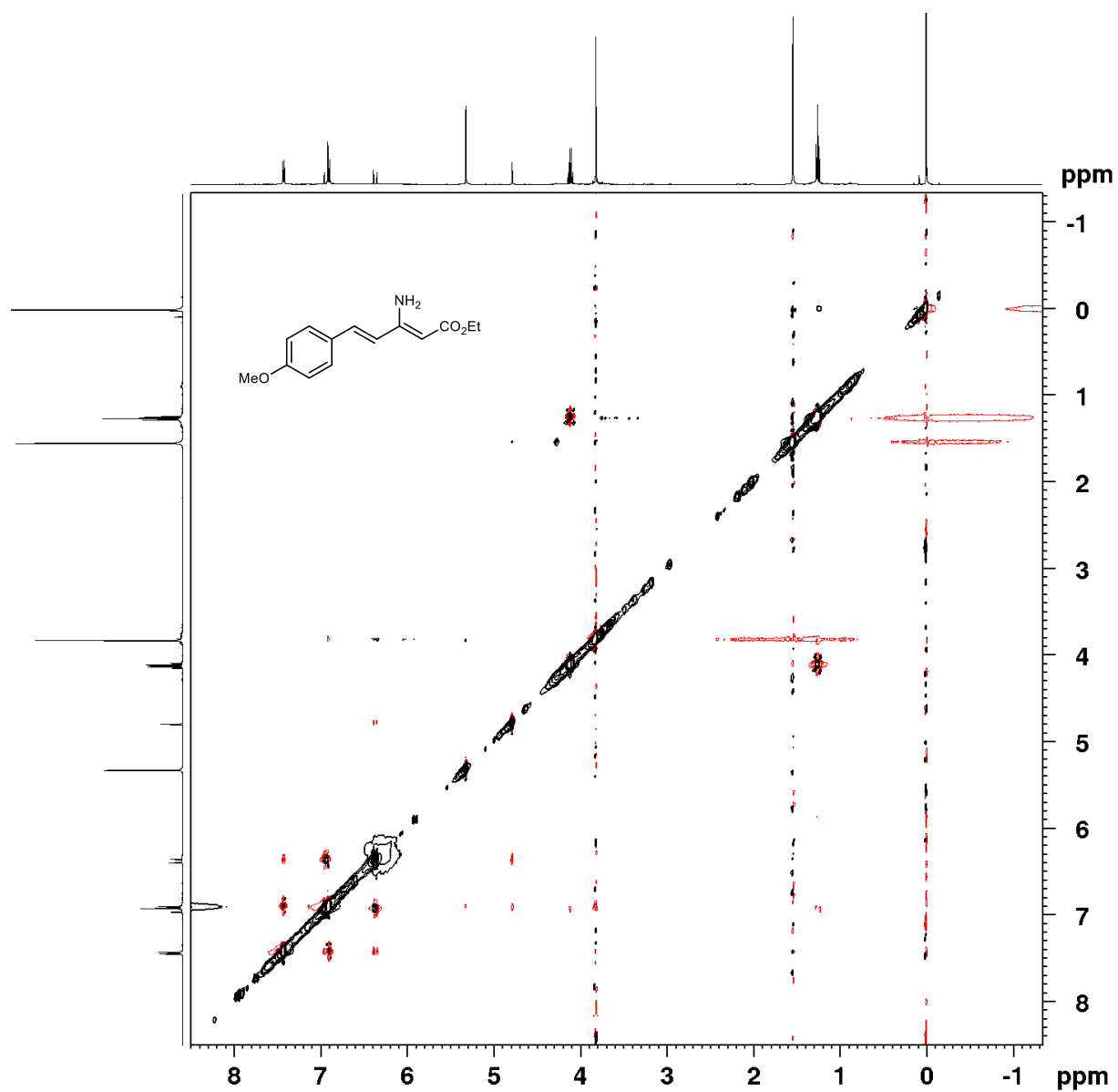
¹H NMR spectrum of 19i (400 MHz, DCM-*d*₂)



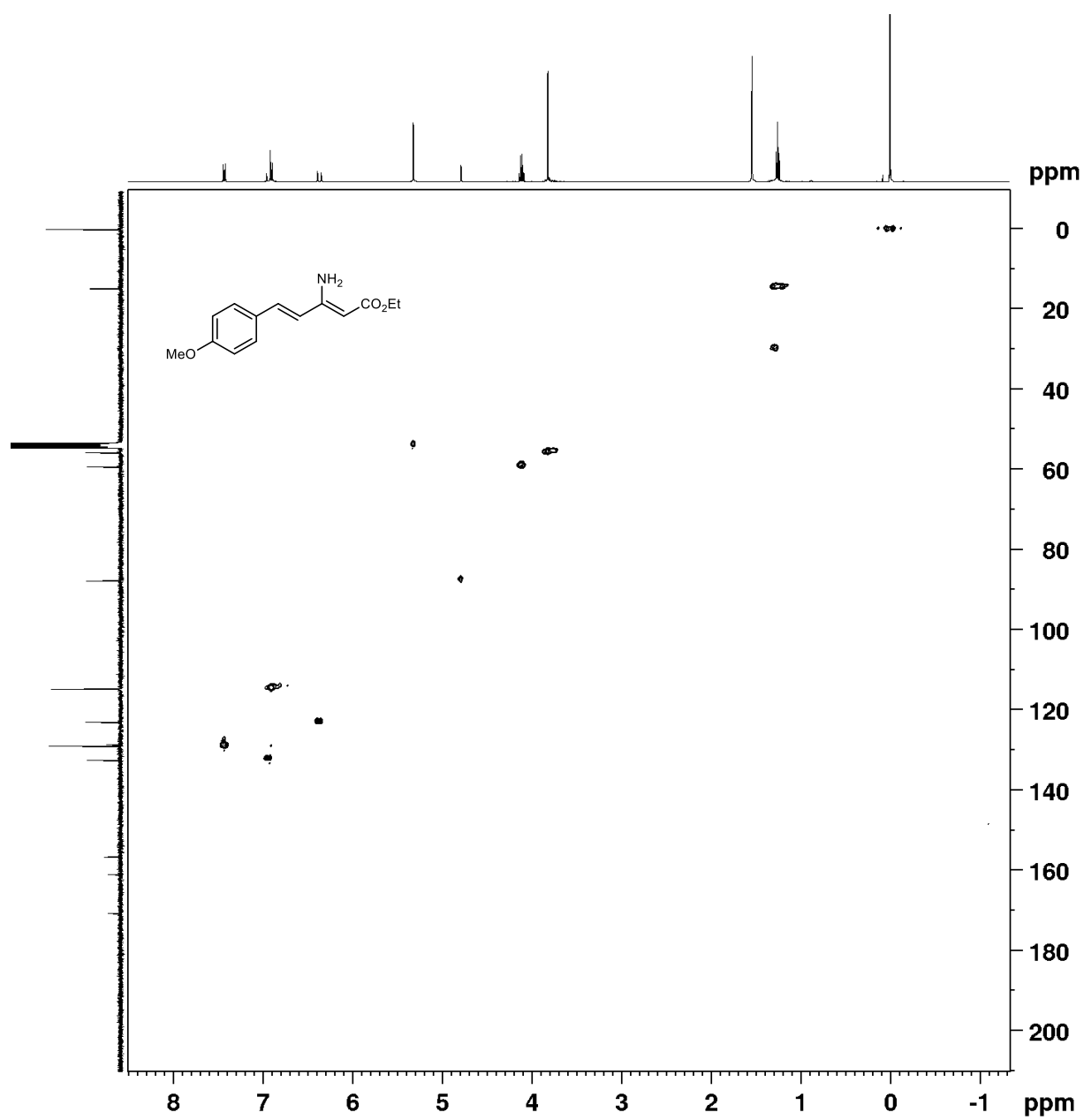
¹³C NMR spectrum of 19i (100 MHz, DCM-*d*₂)



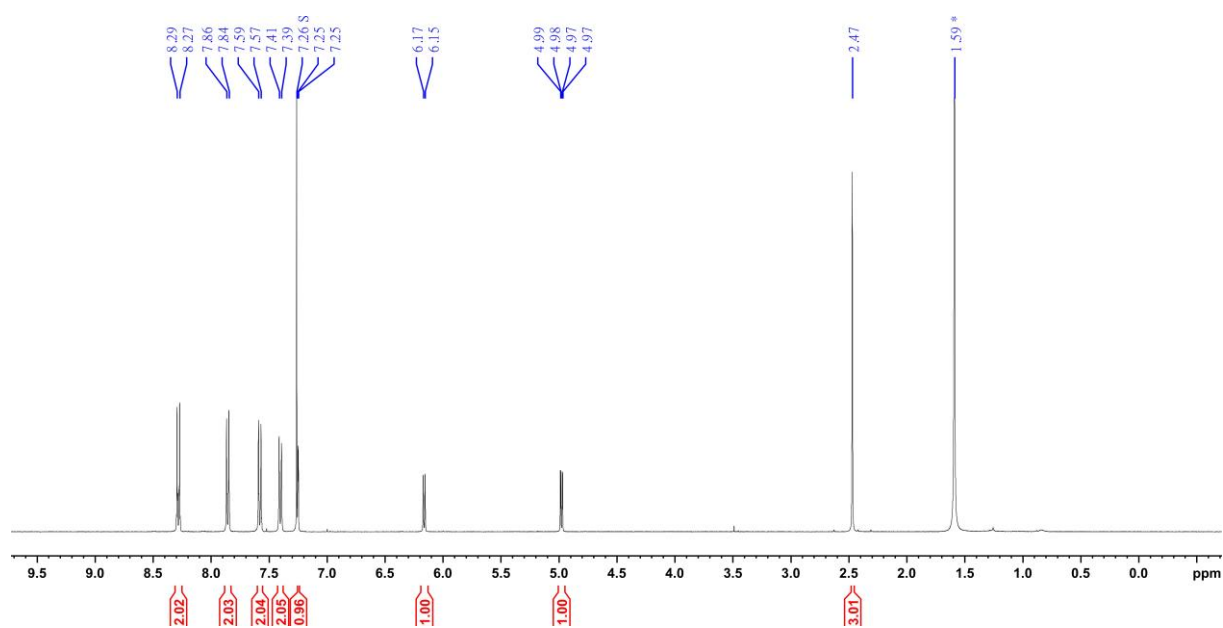
NOESY spectrum of **19i** (DCM-*d*₂)



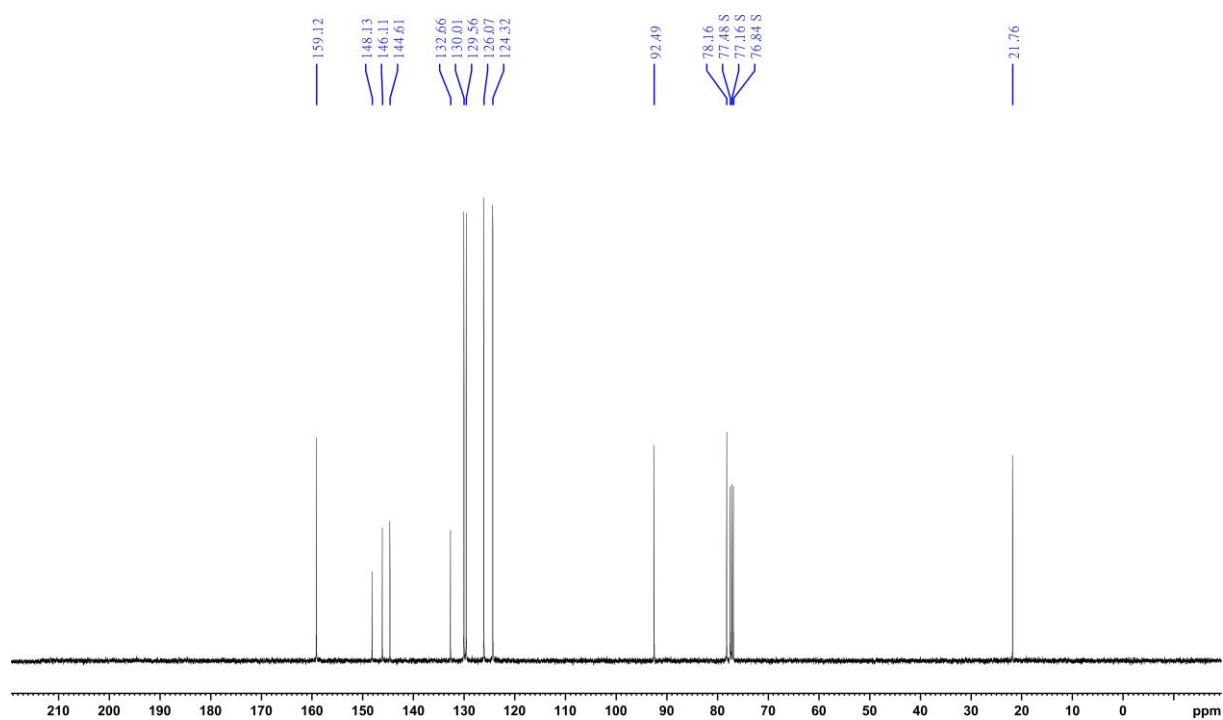
HSQC-TOCSY spectrum of 19i (DCM-*d*₂)



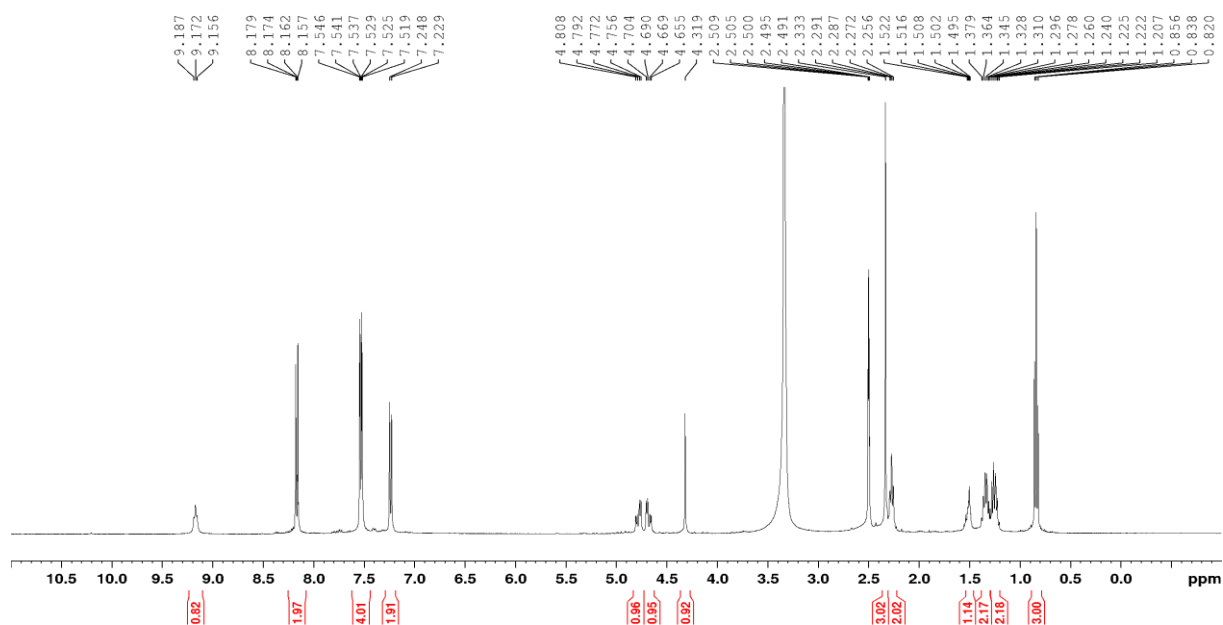
^1H NMR spectrum of 17e (400 MHz, CDCl_3)



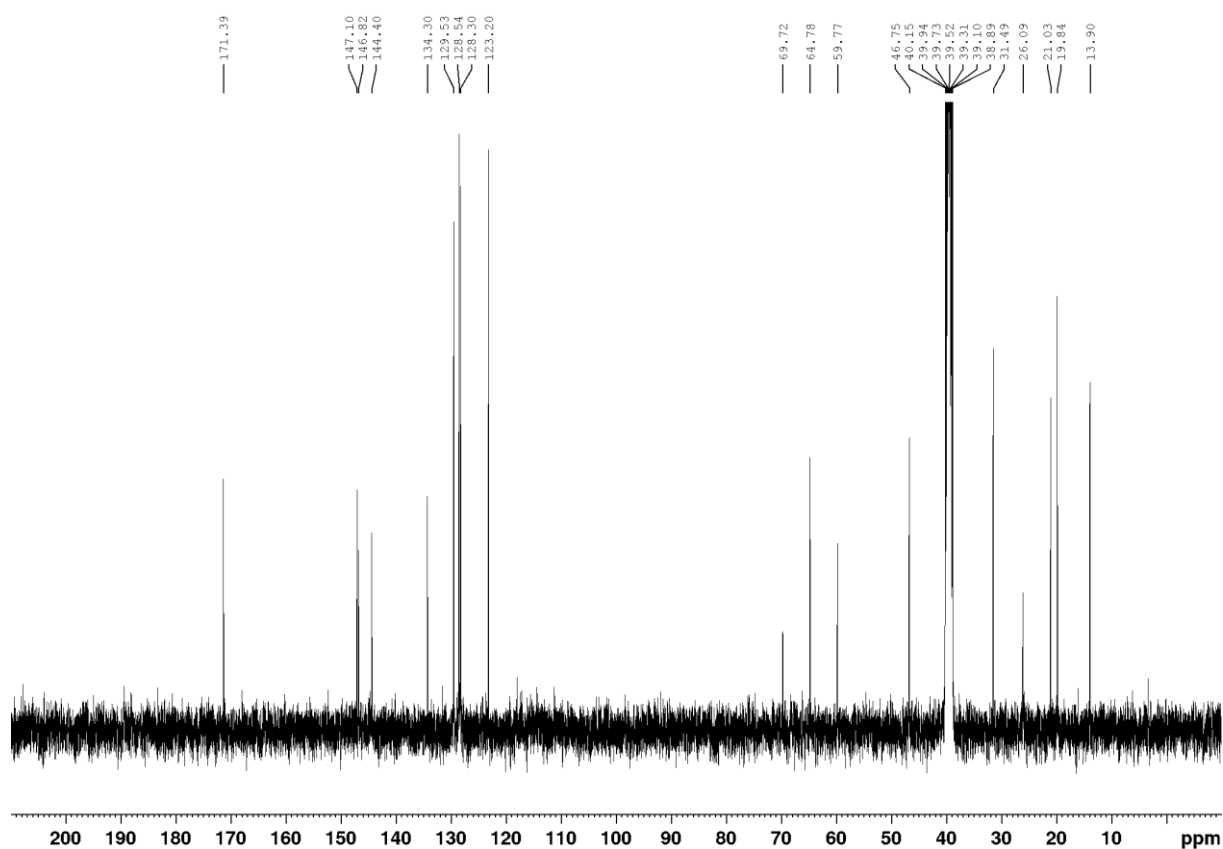
^{13}C NMR spectrum of 17e (100 MHz, CDCl_3)



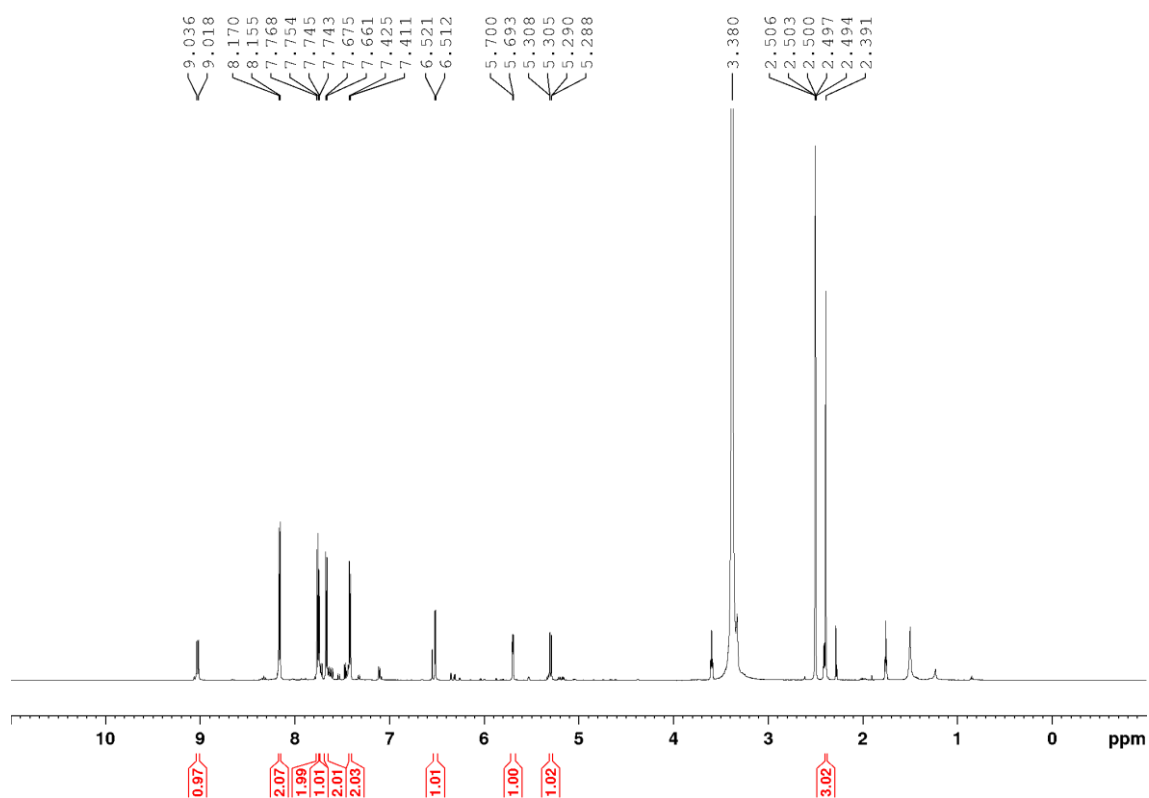
¹H NMR spectrum of 2-(Butylamino)-2-(4-nitrophenyl)-N-(tosylmethyl)acetamide (400 MHz, DMSO-*d*₆)



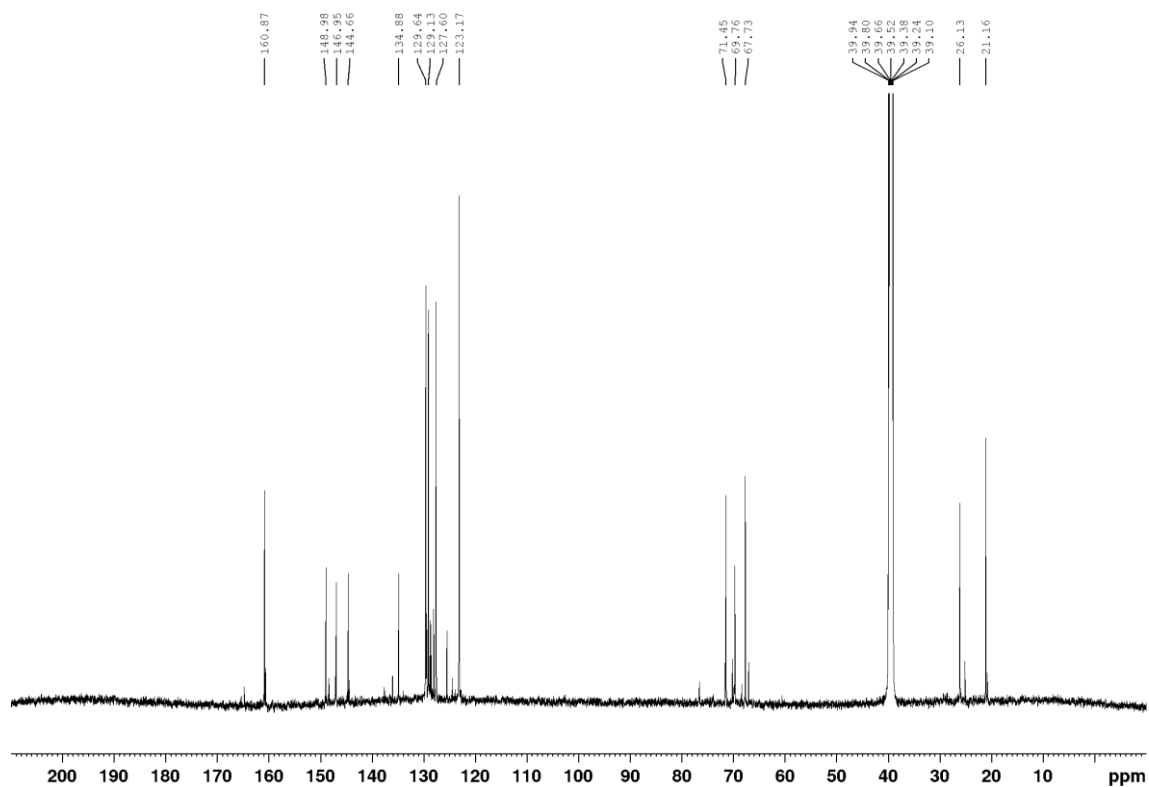
¹³C NMR spectrum of 2-(Butylamino)-2-(4-nitrophenyl)-N-(tosylmethyl)acetamide (100 MHz, DMSO-*d*₆).



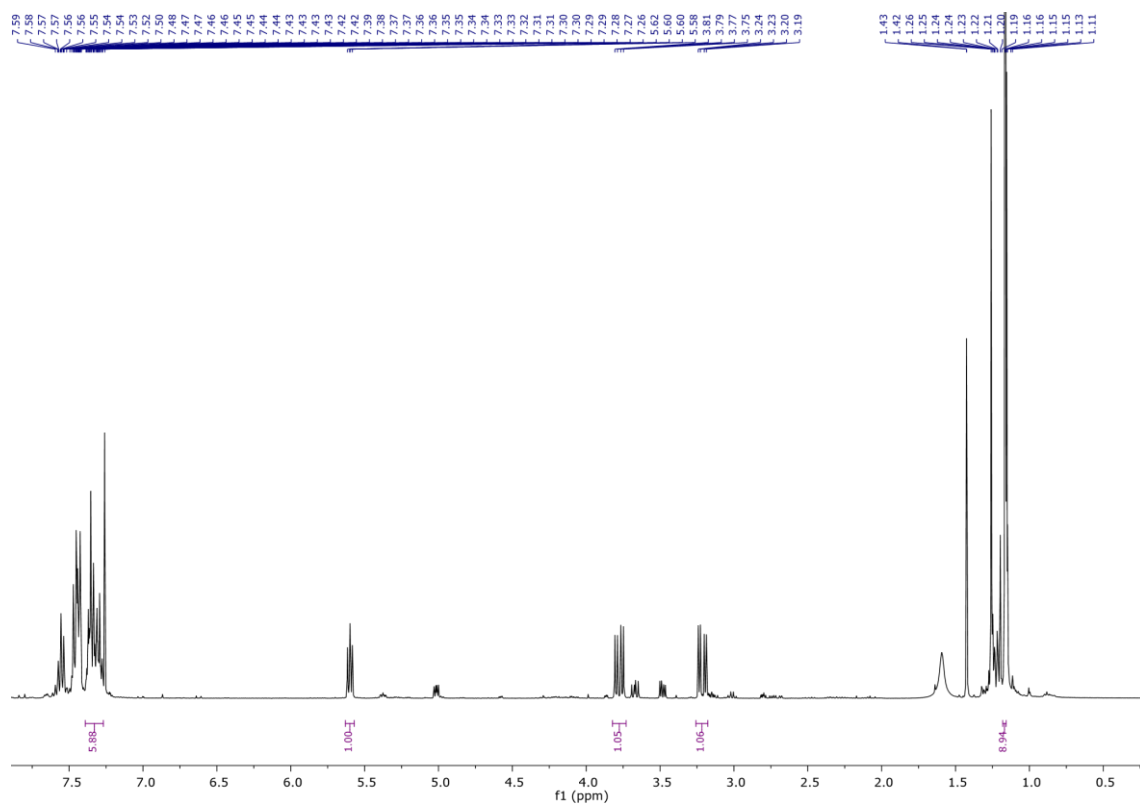
^1H NMR spectrum of 17g (600 MHz, $\text{DMSO-}d_6$)



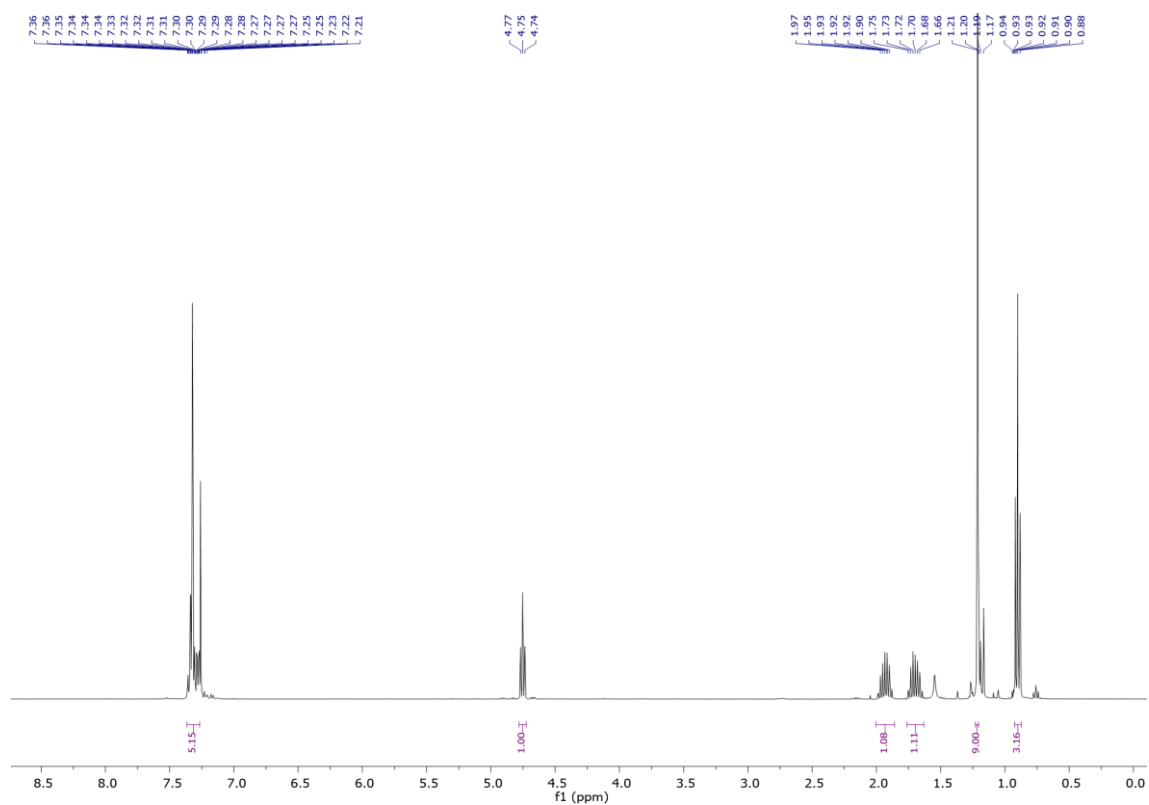
^{13}C NMR spectrum of 17g (150 MHz, $\text{DMSO-}d_6$).



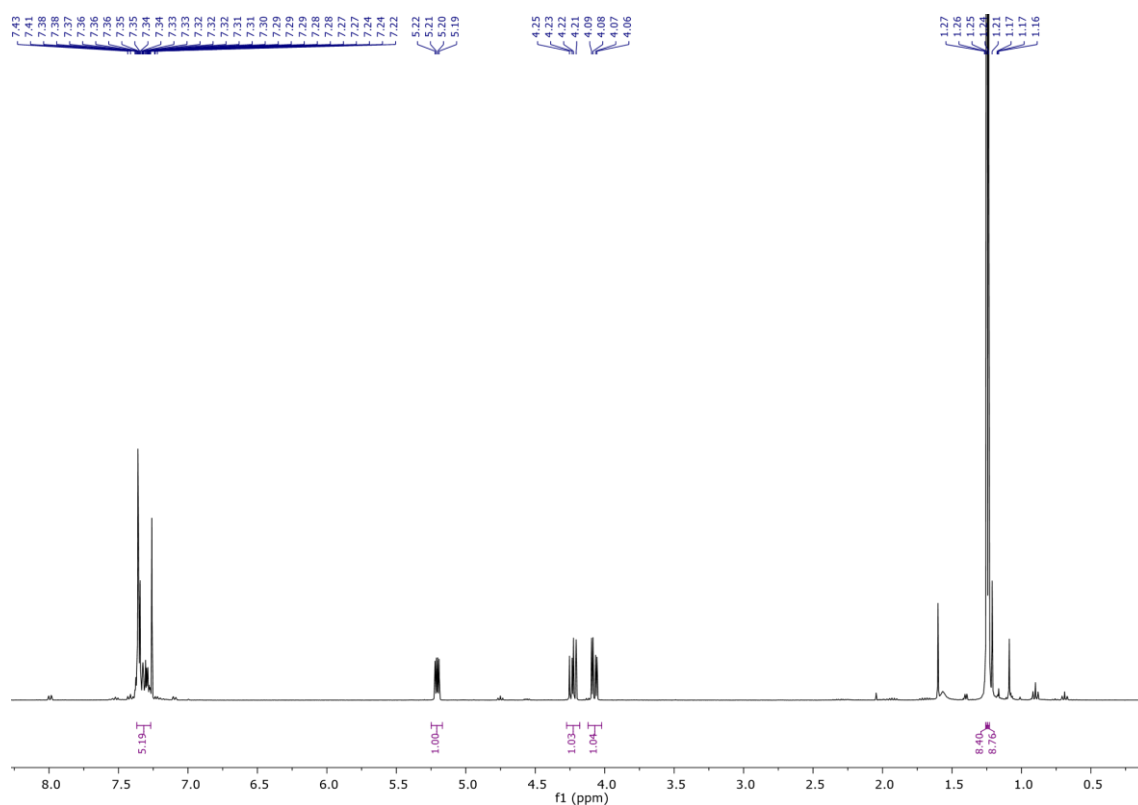
¹H NMR spectrum of 20d (400 MHz, CDCl₃)



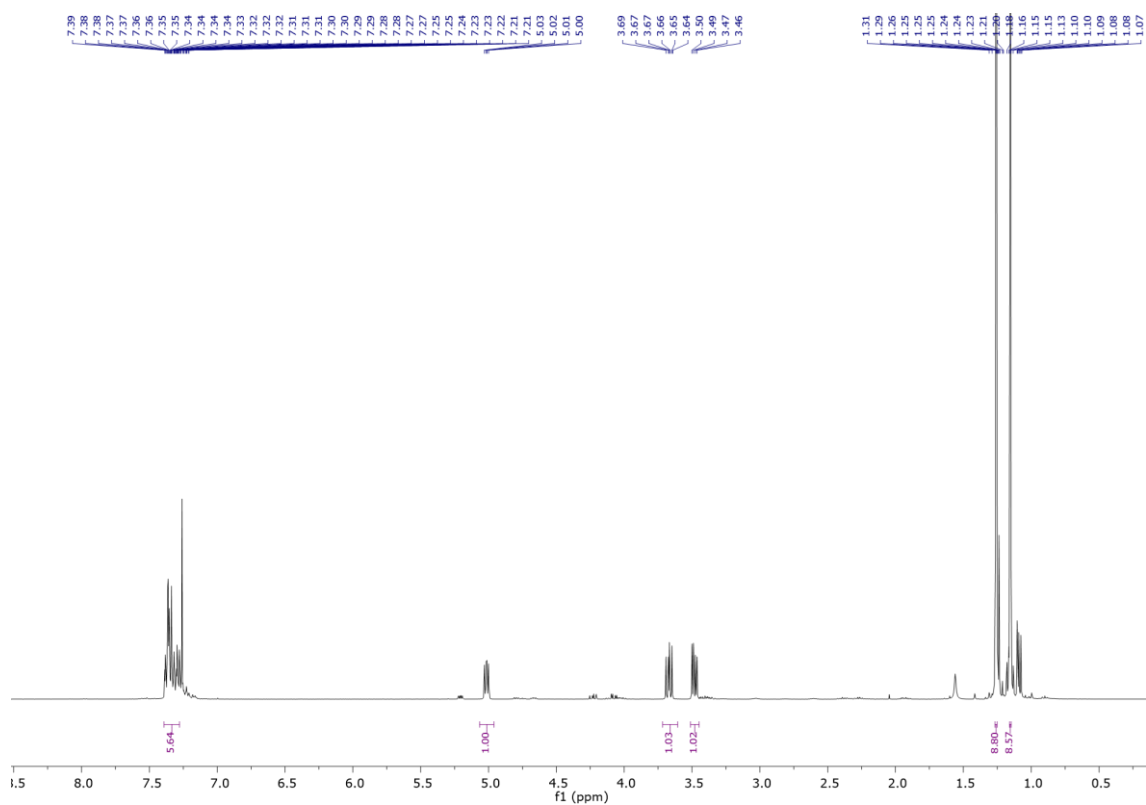
¹H NMR spectrum of 20e (400 MHz, CDCl₃)



¹H NMR spectrum of 20f (400 MHz, CDCl₃)



¹H NMR spectrum of 20g (400 MHz, CDCl₃)



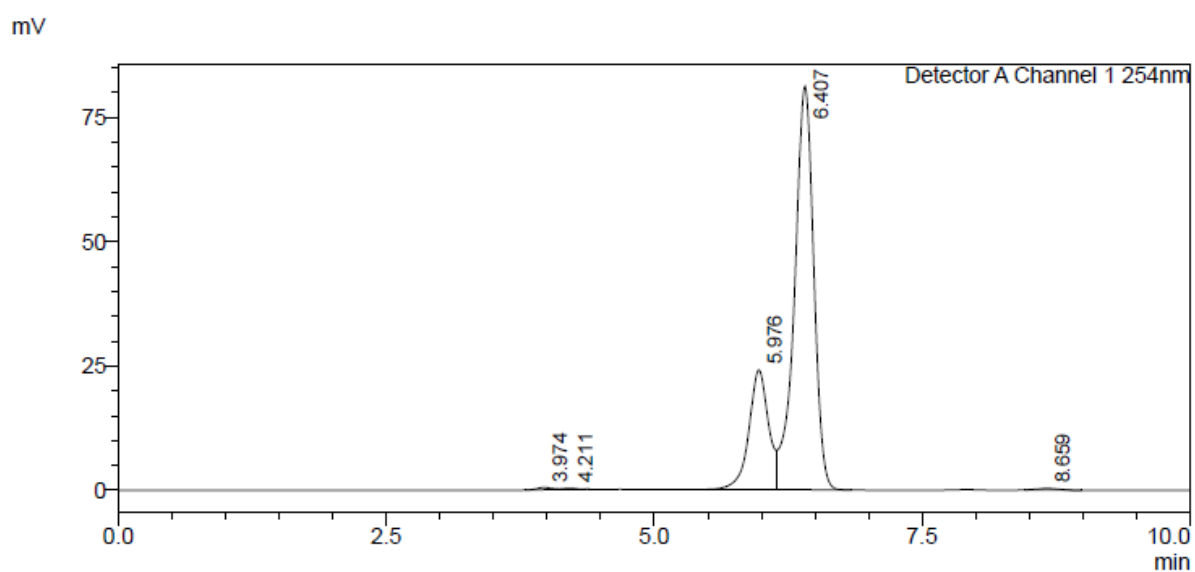
13. HPLC chromatograms

Content List

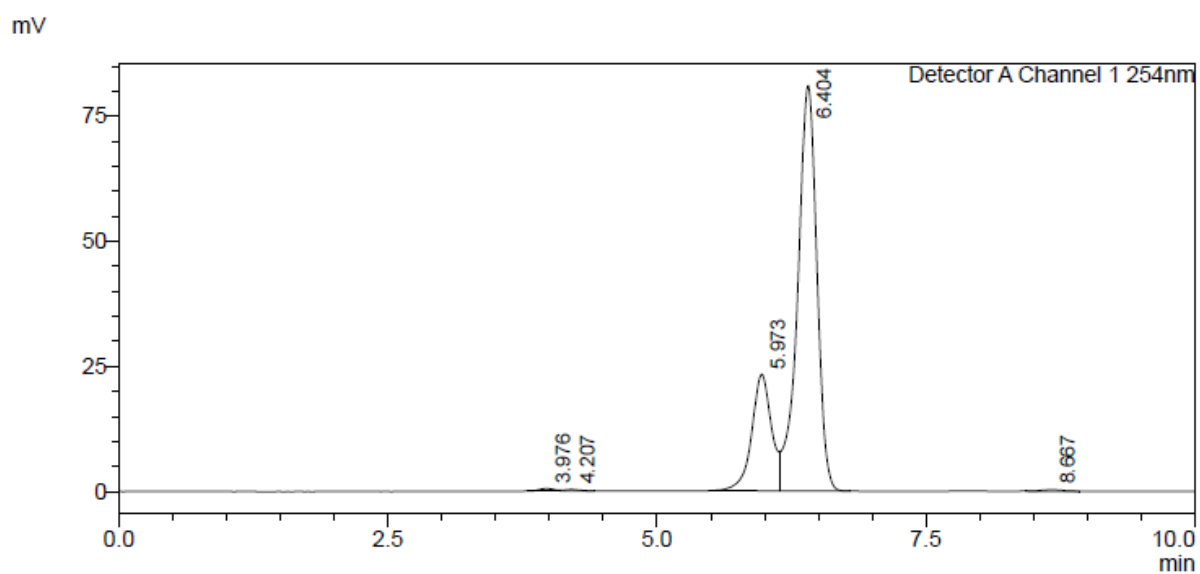
HPLC chromatogram of sample 0 from Table S6.....	341
HPLC chromatogram of sample 1 from Table S6.....	341
HPLC chromatogram of sample 2 from Table S6.....	342
HPLC chromatogram of sample 3 from Table S6.....	342
HPLC chromatogram of sample 4 from Table S6.....	343
HPLC chromatogram of sample 5 from Table S6.....	343
HPLC chromatogram of sample 6 from Table S6.....	344
HPLC chromatogram of sample 7 from Table S6.....	344
HPLC chromatogram of sample 8 from Table S6.....	345
HPLC chromatogram of sample 9 from Table S6.....	345
HPLC chromatogram of sample 10 from Table S6.....	346
HPLC chromatogram of sample 11 from Table S6.....	346
HPLC chromatogram of sample 12 from Table S6.....	347
HPLC chromatogram of sample 13 from Table S6.....	347
HPLC chromatogram of sample 14 from Table S6.....	348
HPLC chromatogram of sample 15 from Table S7.....	348
HPLC chromatogram of sample 16 from Table S7.....	349
HPLC chromatogram of sample 17 from Table S7.....	349
HPLC chromatogram of sample 18 from Table S7.....	350
HPLC chromatogram of sample 19 from Table S7.....	350
HPLC chromatogram of sample 20 from Table S7.....	351
HPLC chromatogram for condition 46, from Table S9.....	351
HPLC chromatogram for condition 49, from Table S9.....	352
HPLC chromatogram for condition 64, from Table S9.....	352
HPLC chromatogram for conditions 31, 91, 16, 4, from Table S9, top to bottom	353
HPLC chromatogram for conditions 34, 31, 91, 16, from Table S9	354
HPLC chromatogram for conditions 4, 46, 25, 10, from Table S9	355
HPLC chromatogram for conditions 1, 7, 55, 19, from Table S9	356
HPLC chromatogram for conditions 58, 28, 37, 43, from Table S9	357
HPLC chromatogram for conditions 16, 4, 46, 25, from Table S9	358
HPLC chromatogram for conditions 10, 64, 34, 31, from Table S9	359
HPLC chromatogram for condition 55, from Table S9.....	360

HPLC chromatogram for condition 61, from Table S9.....	360
HPLC chromatogram for condition 91, from Table S9.....	361
HPLC chromatogram for condition 85, from Table S9.....	361
HPLC chromatogram for conditions 52, 40, 85, 70, from Table S9	362
HPLC chromatogram for conditions 73, 76, 79, 82, from Table S9	363
HPLC chromatogram for conditions 19, 58, 28, 37, from Table S9	364
HPLC chromatogram for conditions 43, 49, 1, 7, from Table S9	365
HPLC chromatogram for conditions 70, 73, 76, 79, from Table S9	366
HPLC chromatogram for conditions 82, 61, 52, 40, from Table S9	367
HPLC chromatogram for conditions 22, 67, 88, 13, from Table S9	368
HPLC chromatogram for conditions 22, 67, 88, 13, from Table S9	369
HPLC chromatogram for conditions 31, 91, 16, 4, from Table S10.....	370
HPLC chromatogram for condition 46, from Table S10.....	371
HPLC chromatogram for condition 64, from Table S10.....	371
HPLC chromatogram for condition 91, from Table S10.....	371
HPLC chromatogram for condition 49, from Table S10.....	371
HPLC chromatogram for conditions 34, 31, 91, 16, from Table S10	372
HPLC chromatogram for conditions 4, 46, 25, 10, from Table S10	373
HPLC chromatogram for conditions 16, 4, 46, 25, from Table S10	374
HPLC chromatogram for conditions 10, 64, 34, 31, from Table S10	375
HPLC chromatogram for conditions 1, 7, 55, 19, from Table S10	376
HPLC chromatogram for conditions 58, 28, 37, 43, from Table S10	377
HPLC chromatogram for conditions 52, 40, 85, 70, from Table S10	378
HPLC chromatogram for conditions 73, 76, 79, 82, from Table S10	379
HPLC chromatogram for condition 61, from Table S10.....	380
HPLC chromatogram for conditions 55, from Table S10	380
HPLC chromatogram for condition 85, from Table S10.....	380
HPLC chromatograms for conditions 88, 13, 22, 67, from Table S10	381
HPLC chromatogram for conditions 19, 58, 28, 37, from Table S10	382
HPLC chromatogram for conditions 43, 49, 1, 7, from Table S10	383
HPLC chromatogram for conditions 70, 73, 76, 79, from Table S10	384
HPLC chromatogram for conditions 82, 61, 52, 40, from Table S10	385
HPLC chromatogram for conditions 22, 67, 88, 13, from Table S10	386

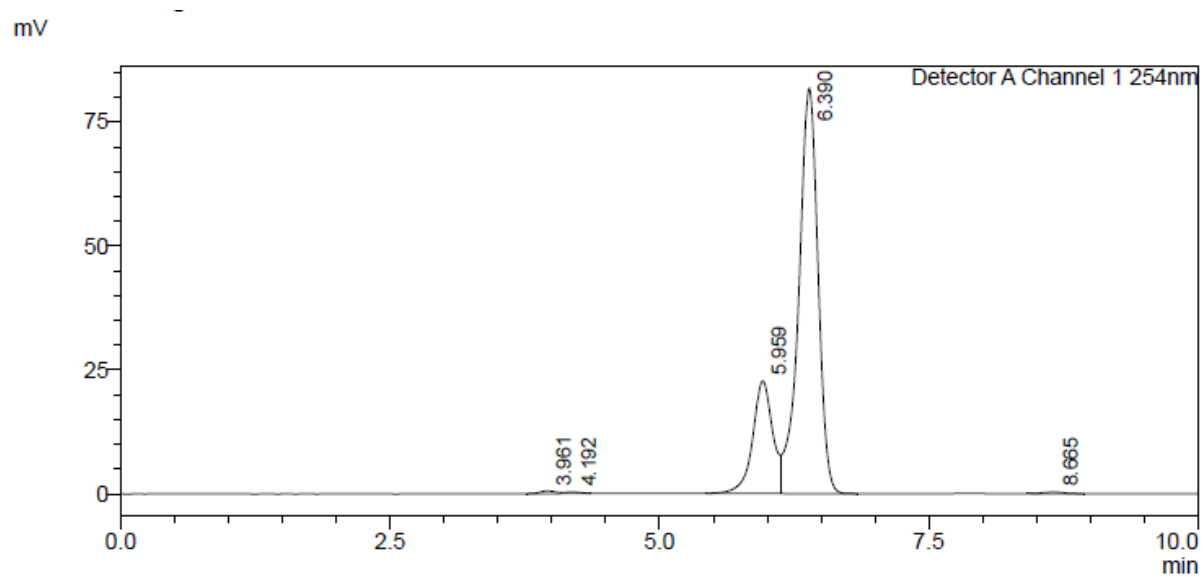
HPLC chromatogram of sample 0 from Table S9.



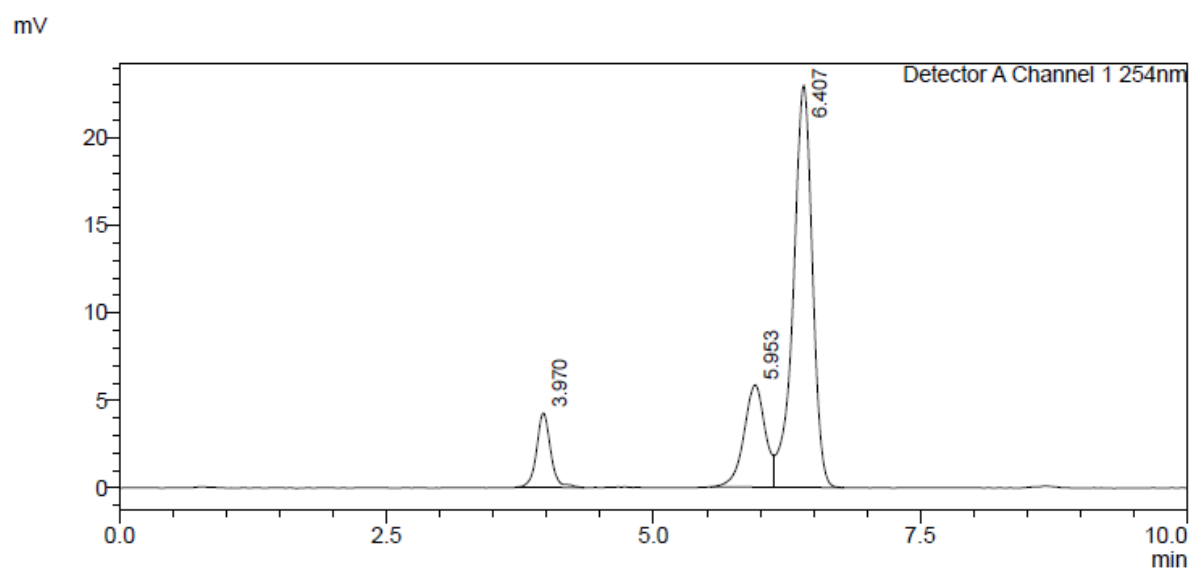
HPLC chromatogram of sample 1 from Table S6



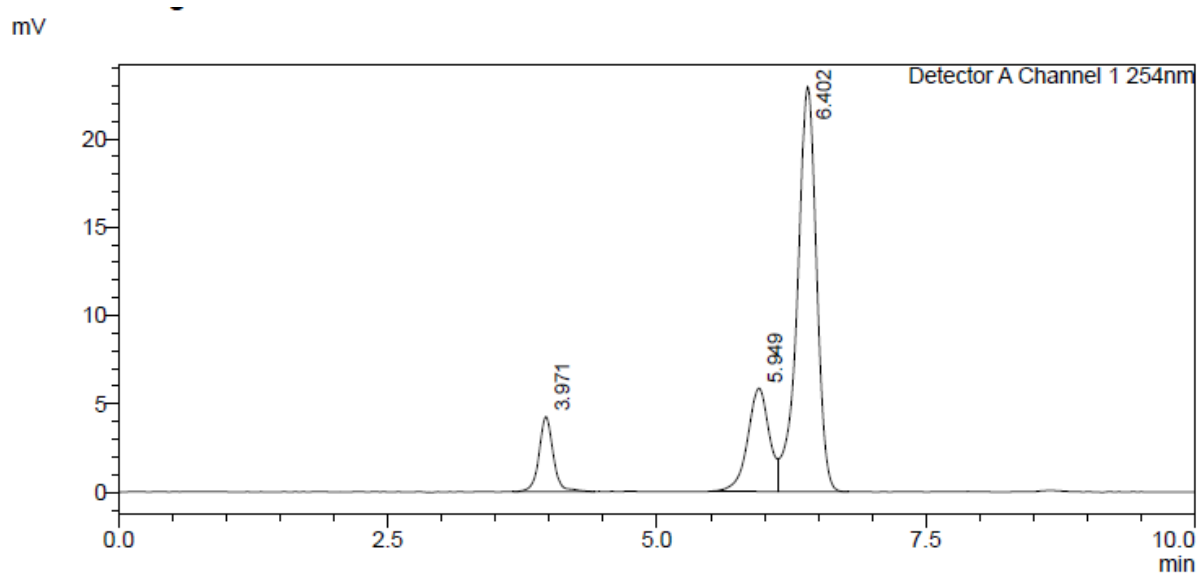
HPLC chromatogram of sample 2 from Table S9



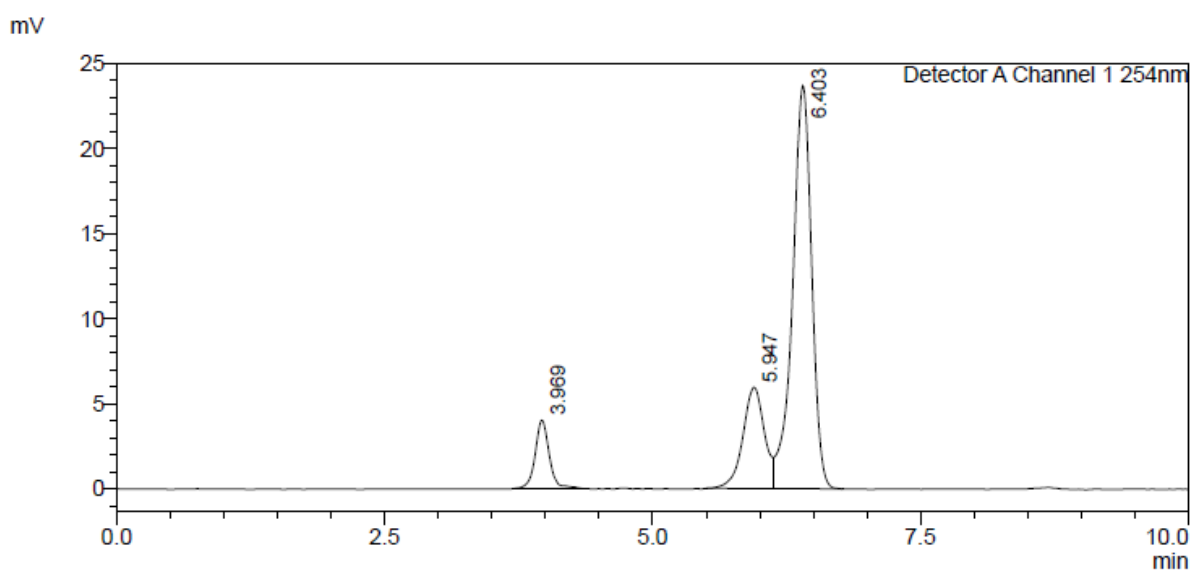
HPLC chromatogram of sample 3 from Table S9



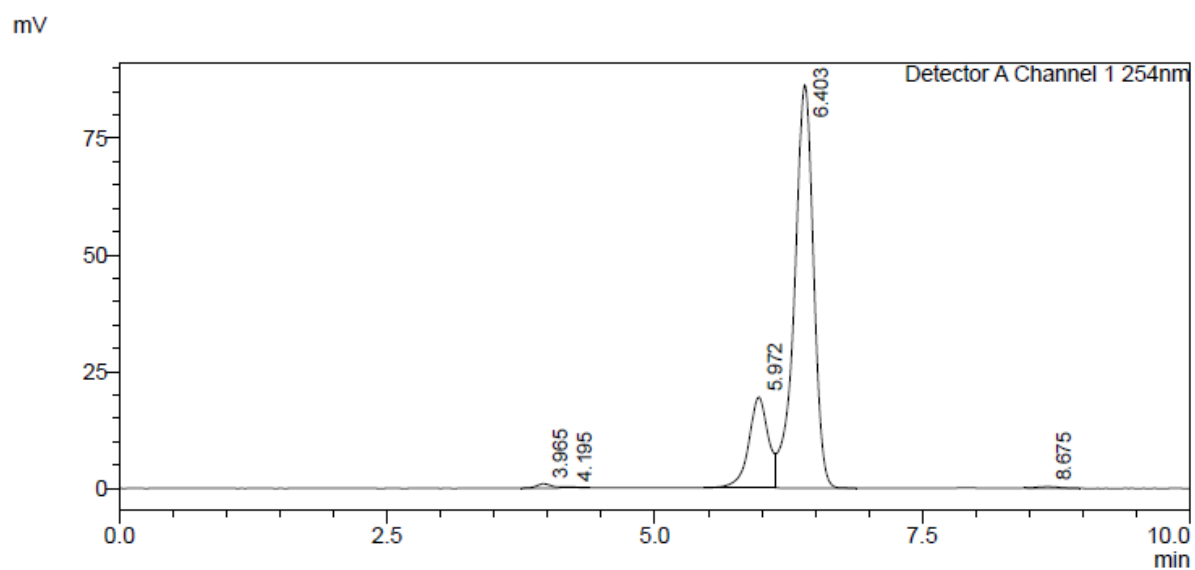
HPLC chromatogram of sample 4 from Table S9



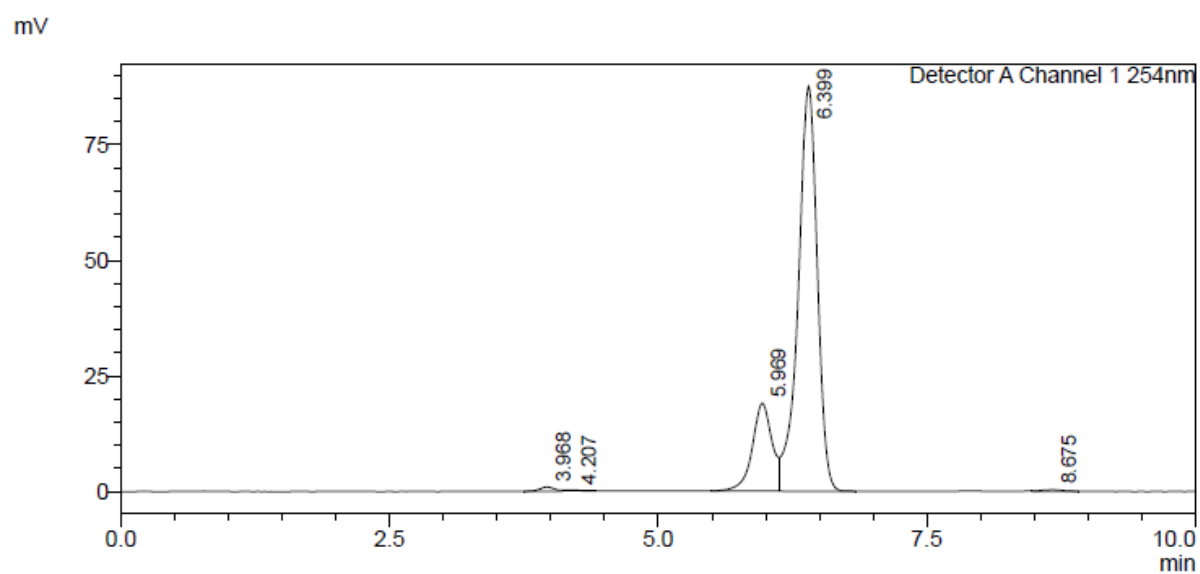
HPLC chromatogram of sample 5 from Table S9



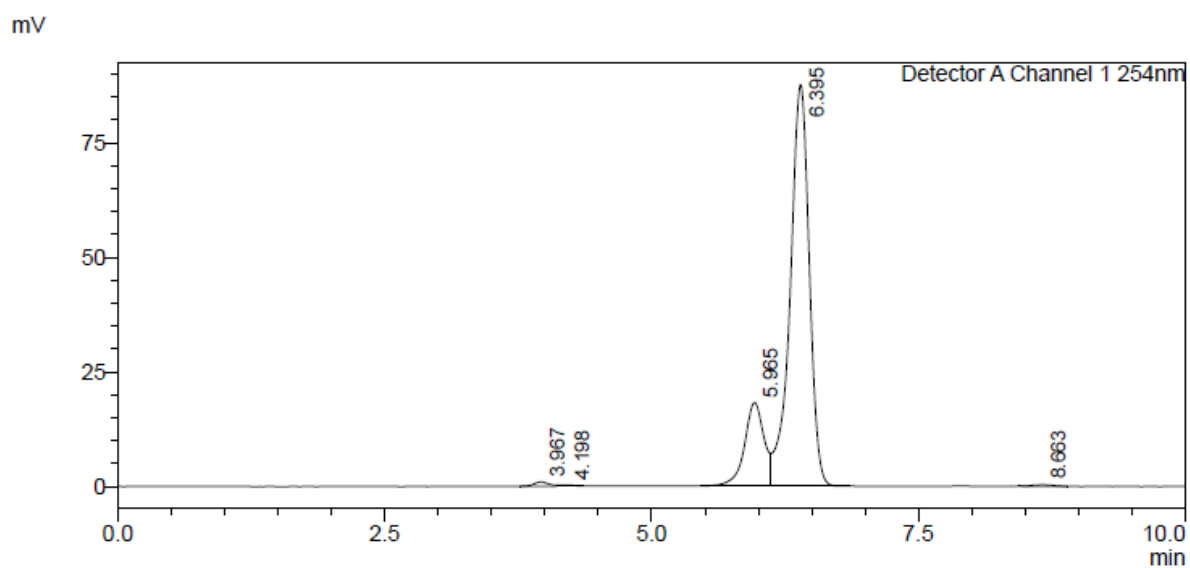
HPLC chromatogram of sample 6 from Table S9



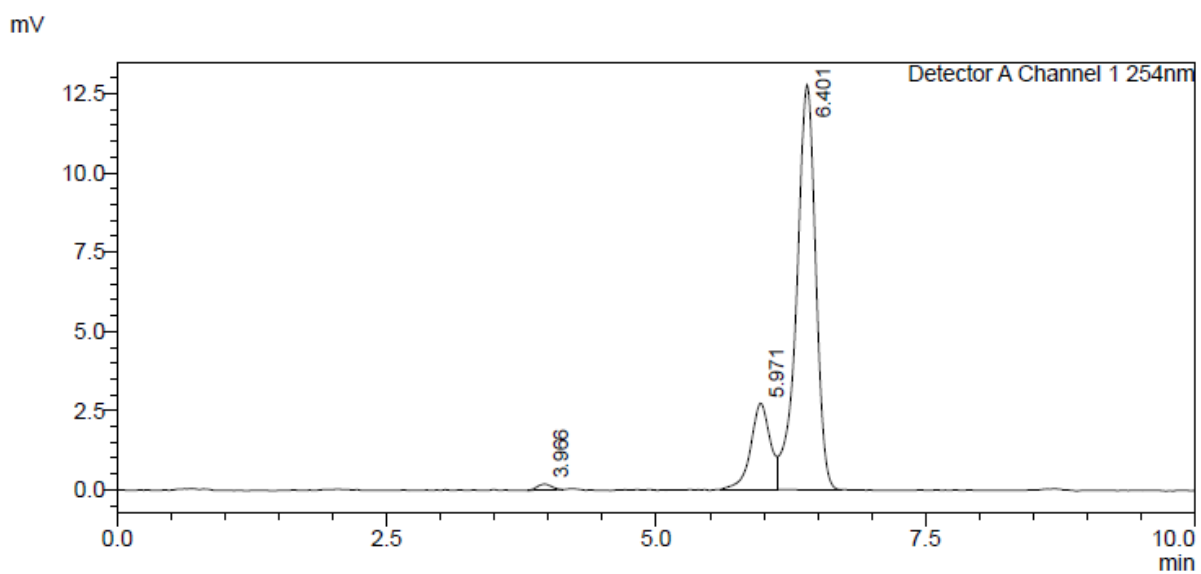
HPLC chromatogram of sample 7 from Table S9



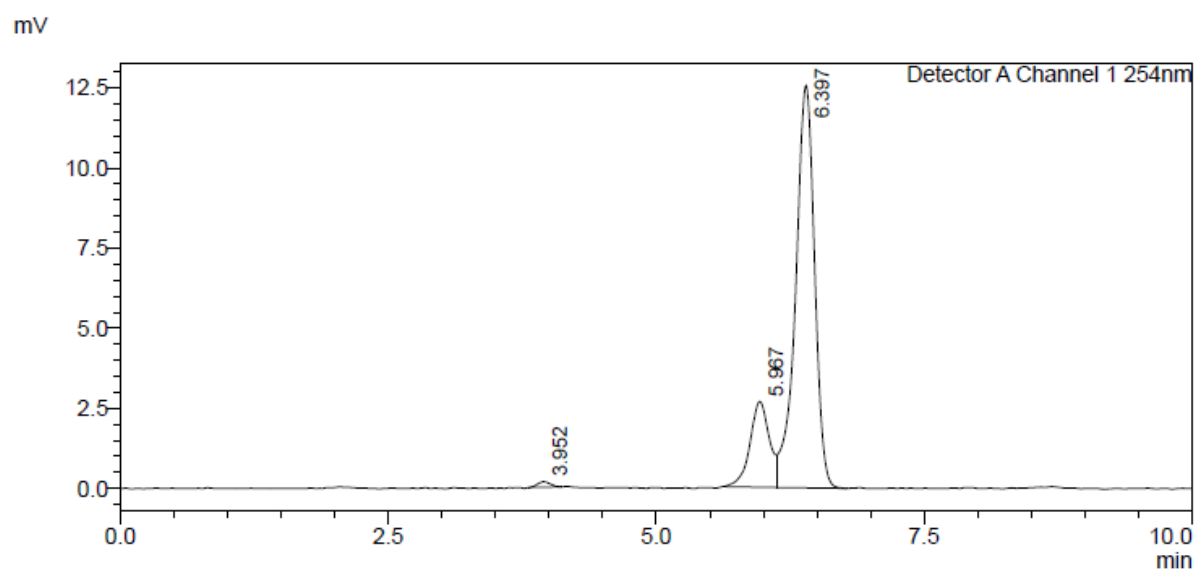
HPLC chromatogram of sample 8 from Table S9



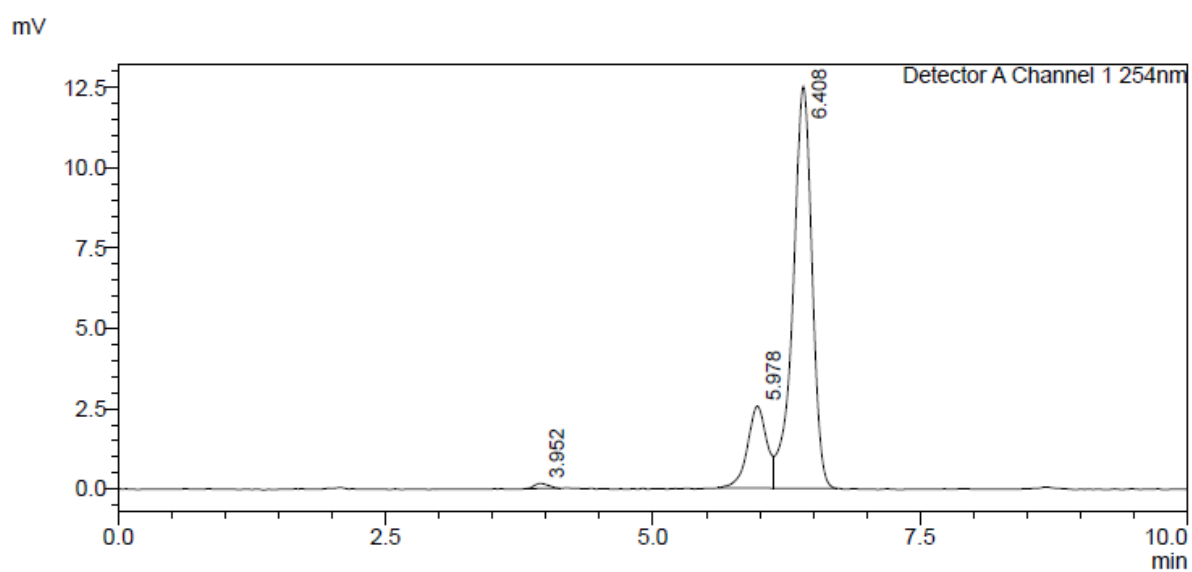
HPLC chromatogram of sample 9 from Table S9



HPLC chromatogram of sample 10 from Table S9

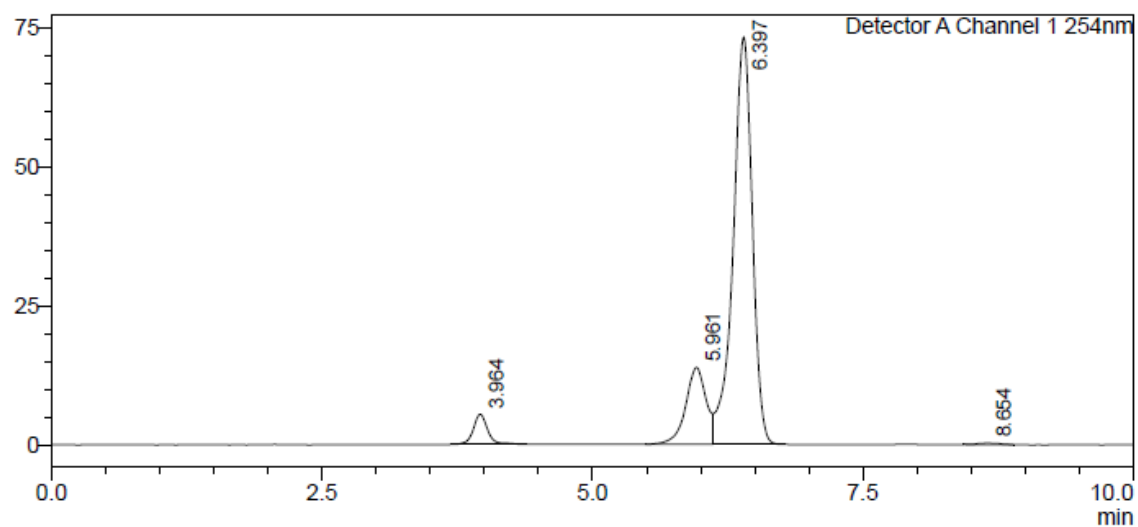


HPLC chromatogram of sample 11 from Table S9



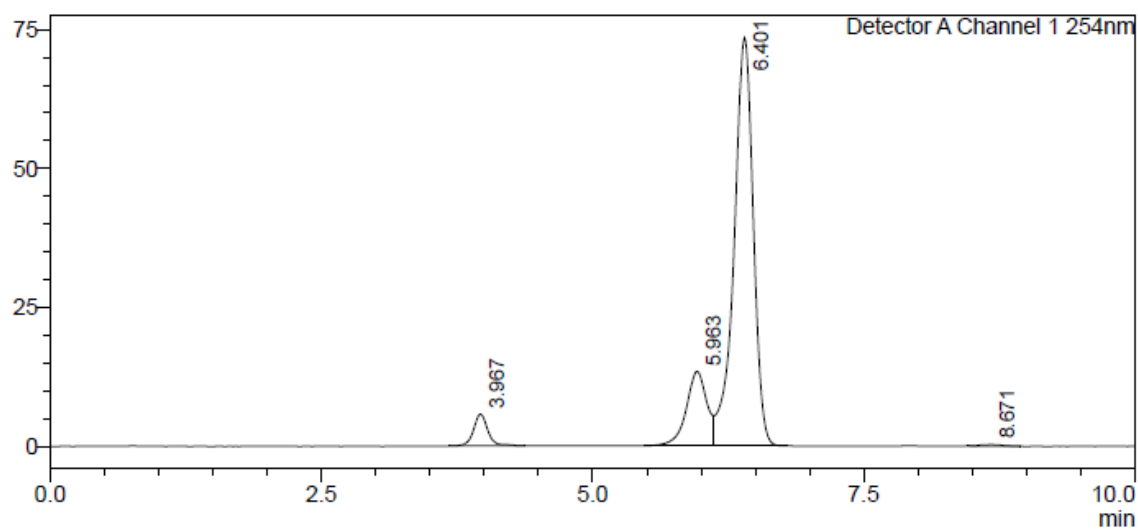
HPLC chromatogram of sample 12 from Table S9

mV

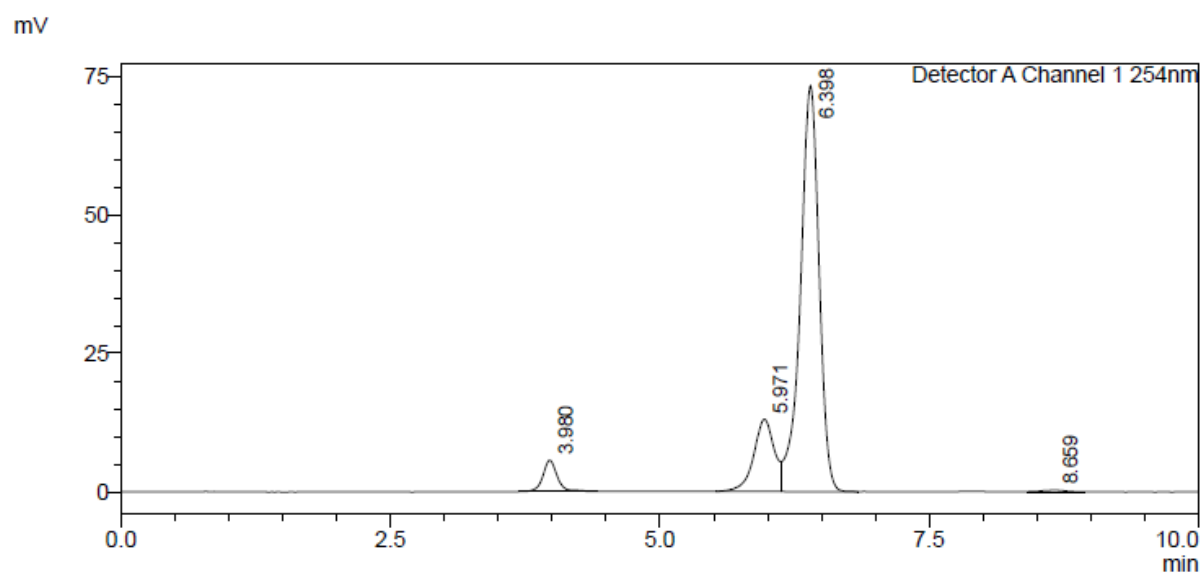


HPLC chromatogram of sample 13 from Table S9

mV

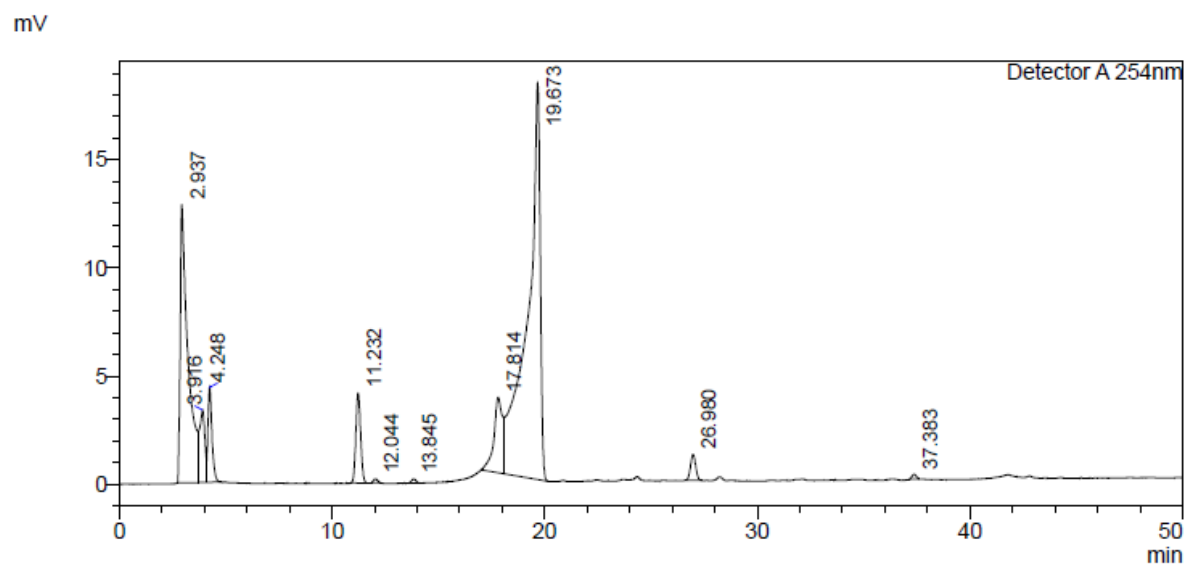


HPLC chromatogram of sample 14 from Table S9



HPLC chromatogram of sample 15 from

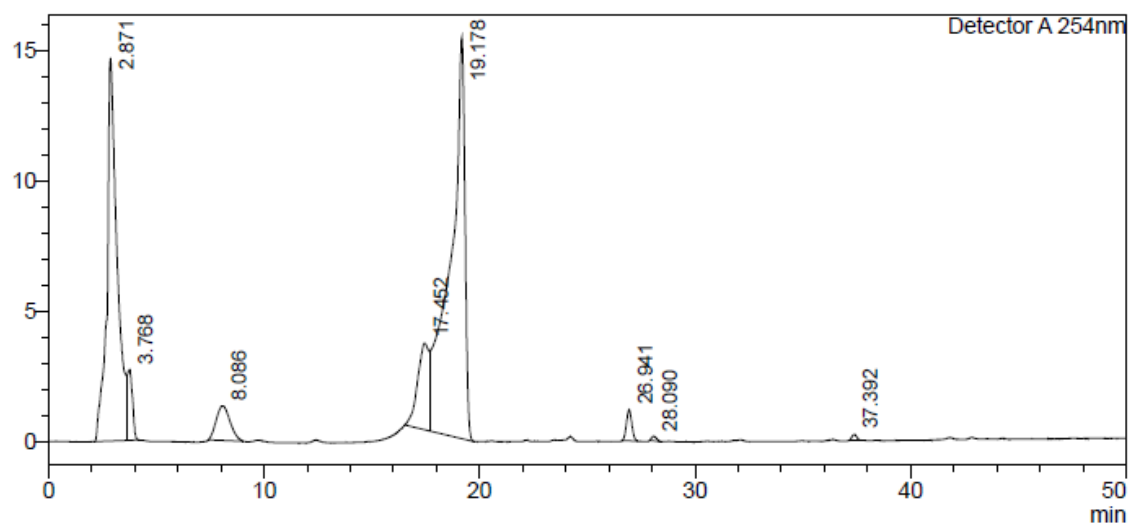
Table S10



HPLC chromatogram of sample 16 from

Table S10

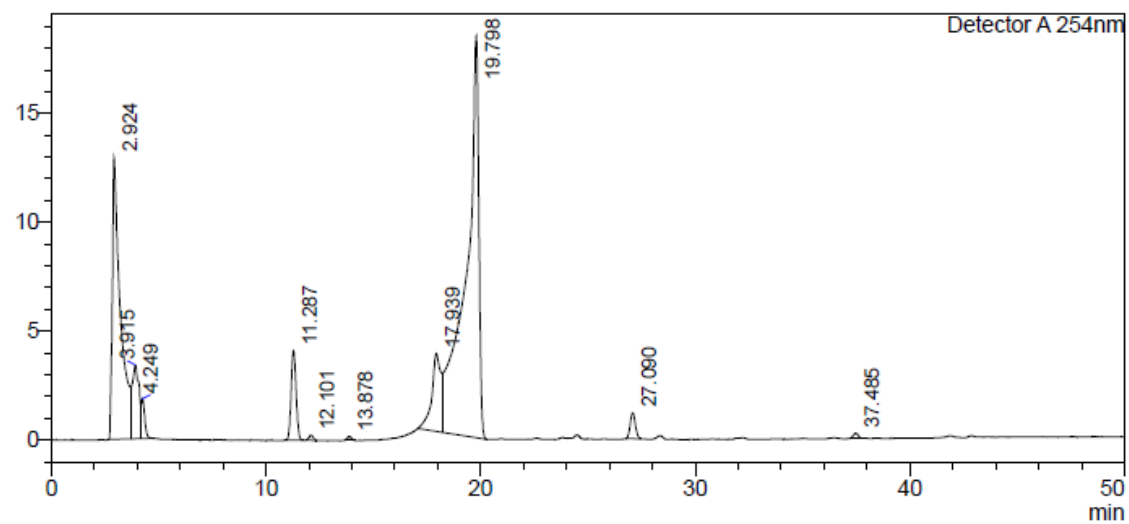
mV



HPLC chromatogram of sample 17 from

Table S10

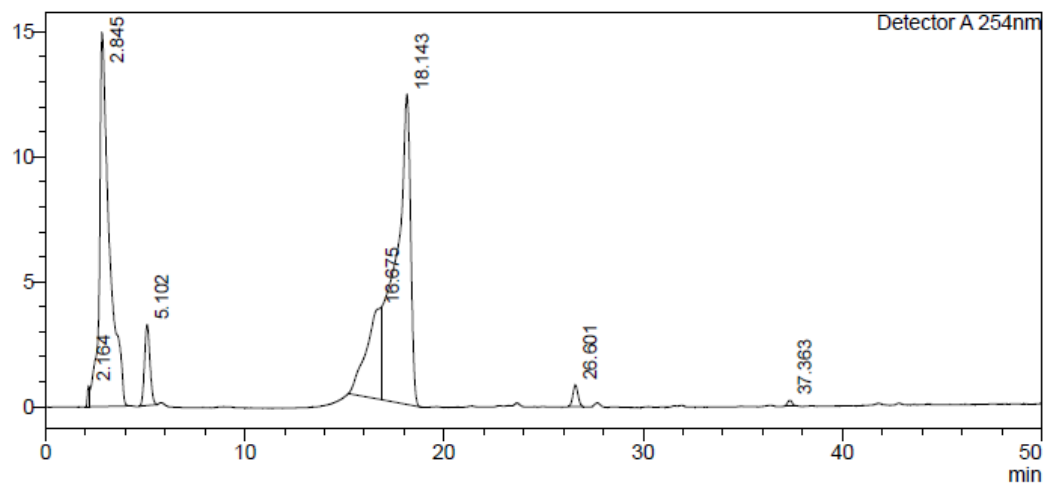
mV



HPLC chromatogram of sample 18 from

Table S10

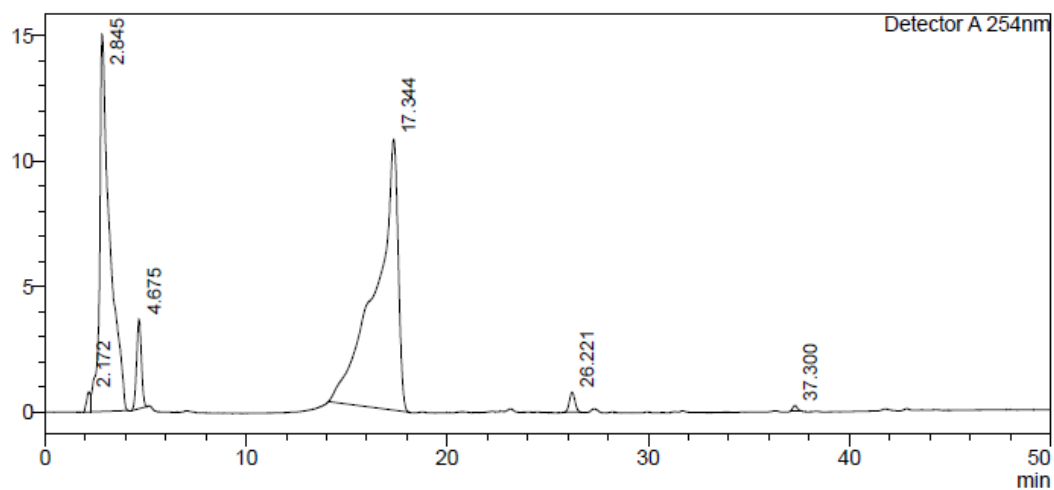
mV



HPLC chromatogram of sample 19 from

Table S10

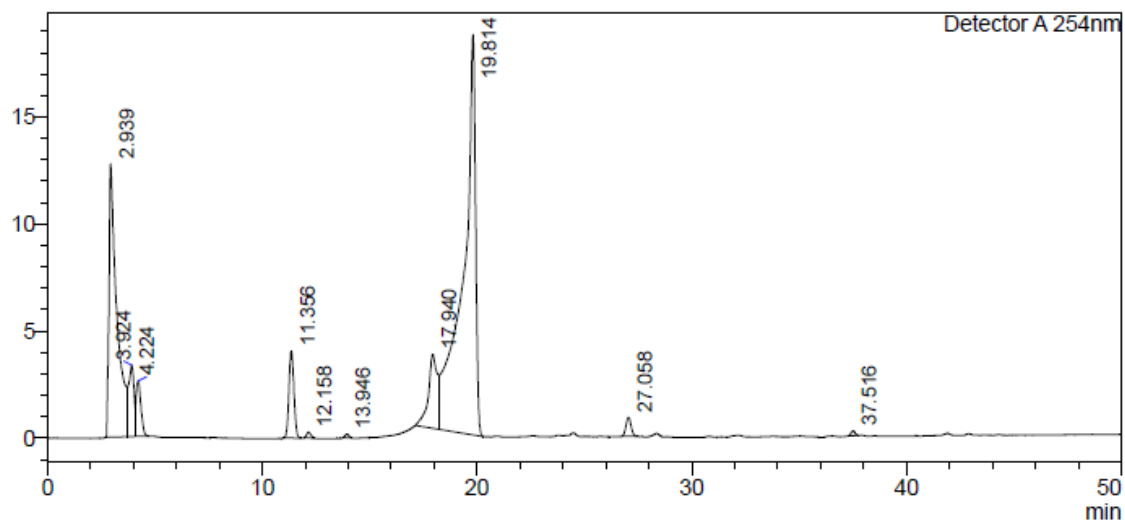
mV



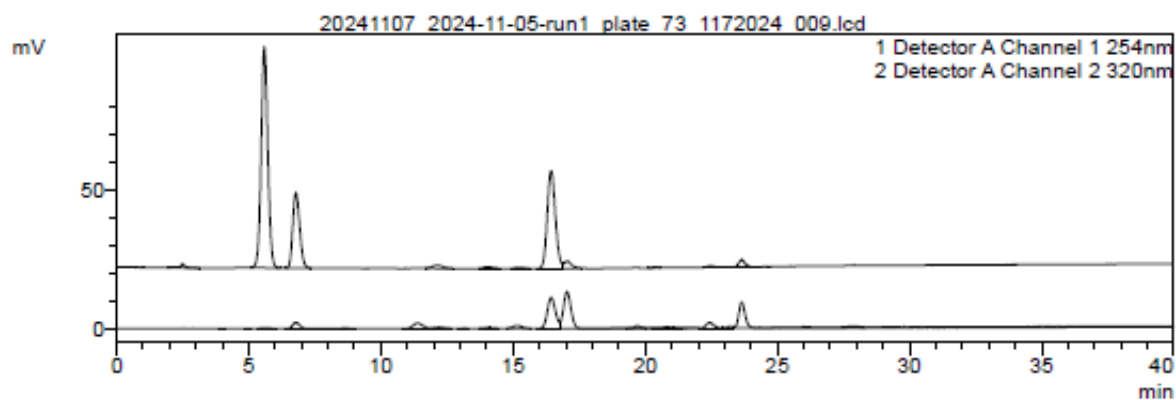
HPLC chromatogram of sample 20 from

Table S10

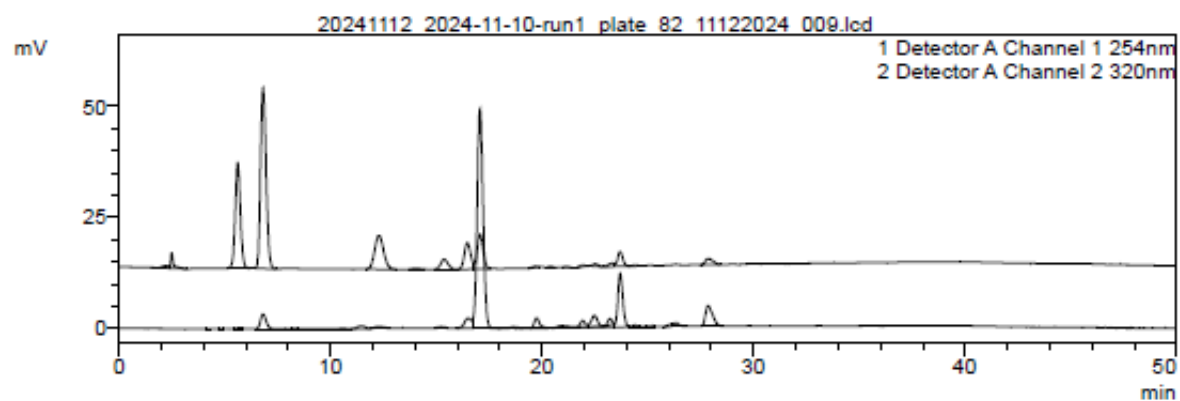
mV



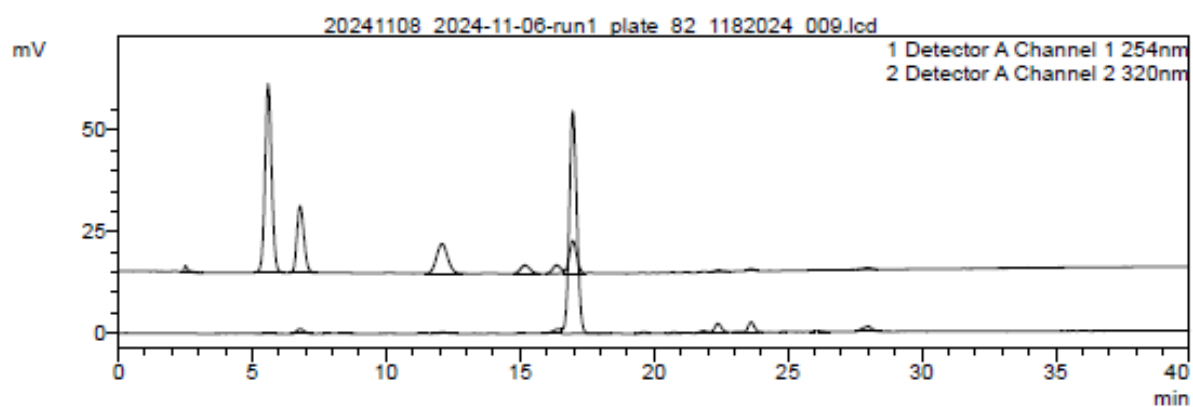
HPLC chromatogram for condition 46, from Table S12



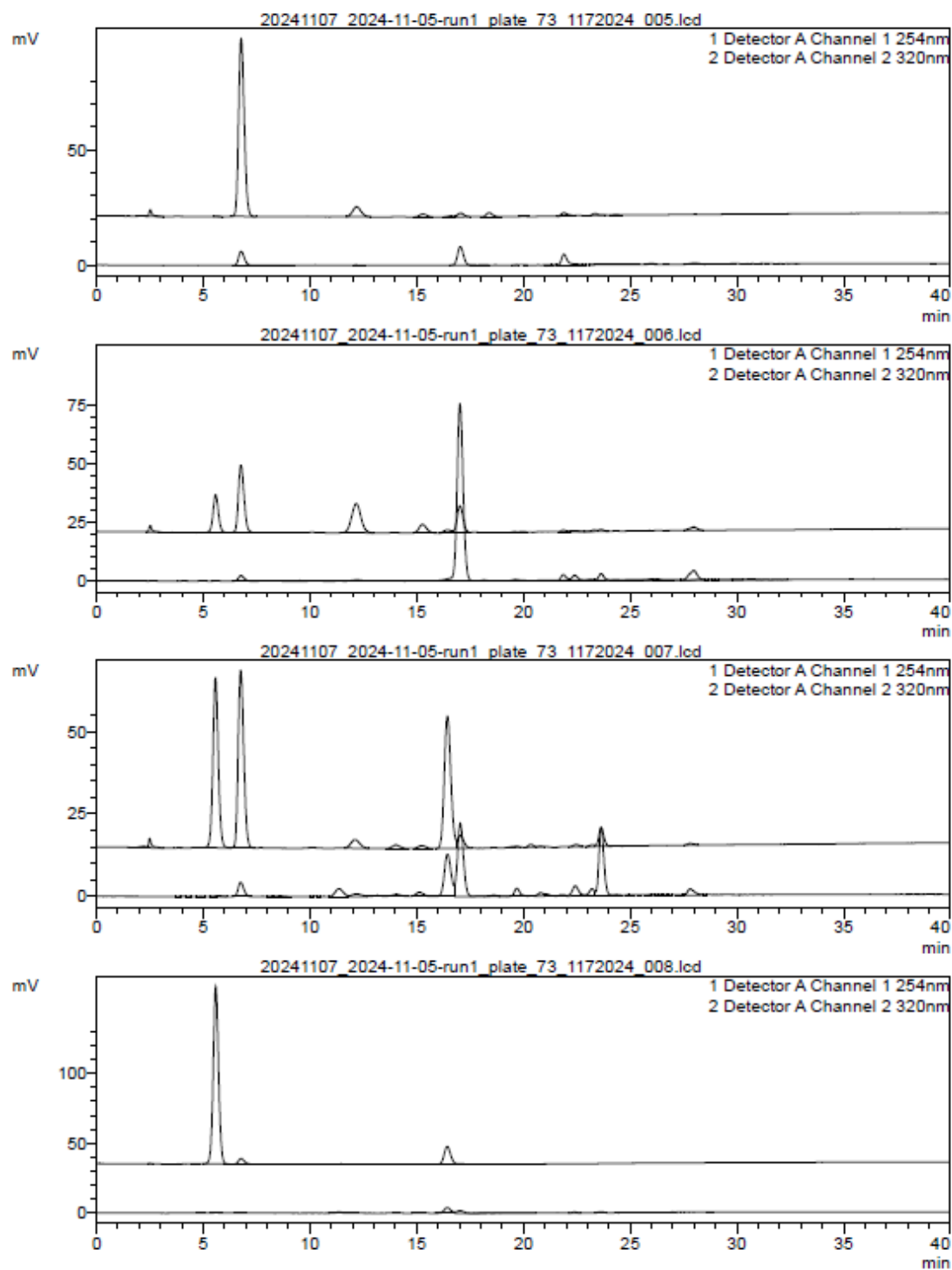
HPLC chromatogram for condition 49, from Table S12.



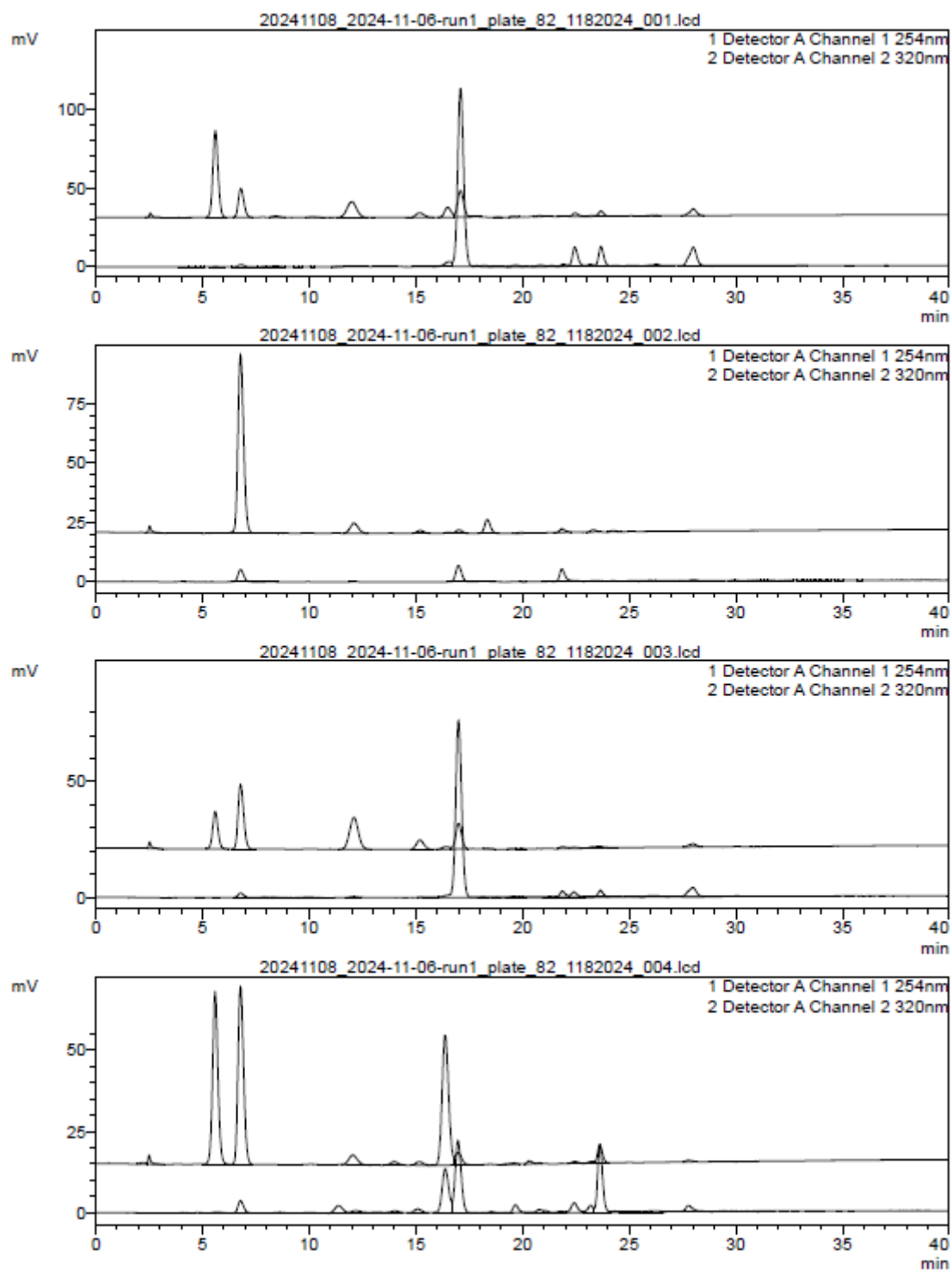
HPLC chromatogram for condition 64, from Table S12



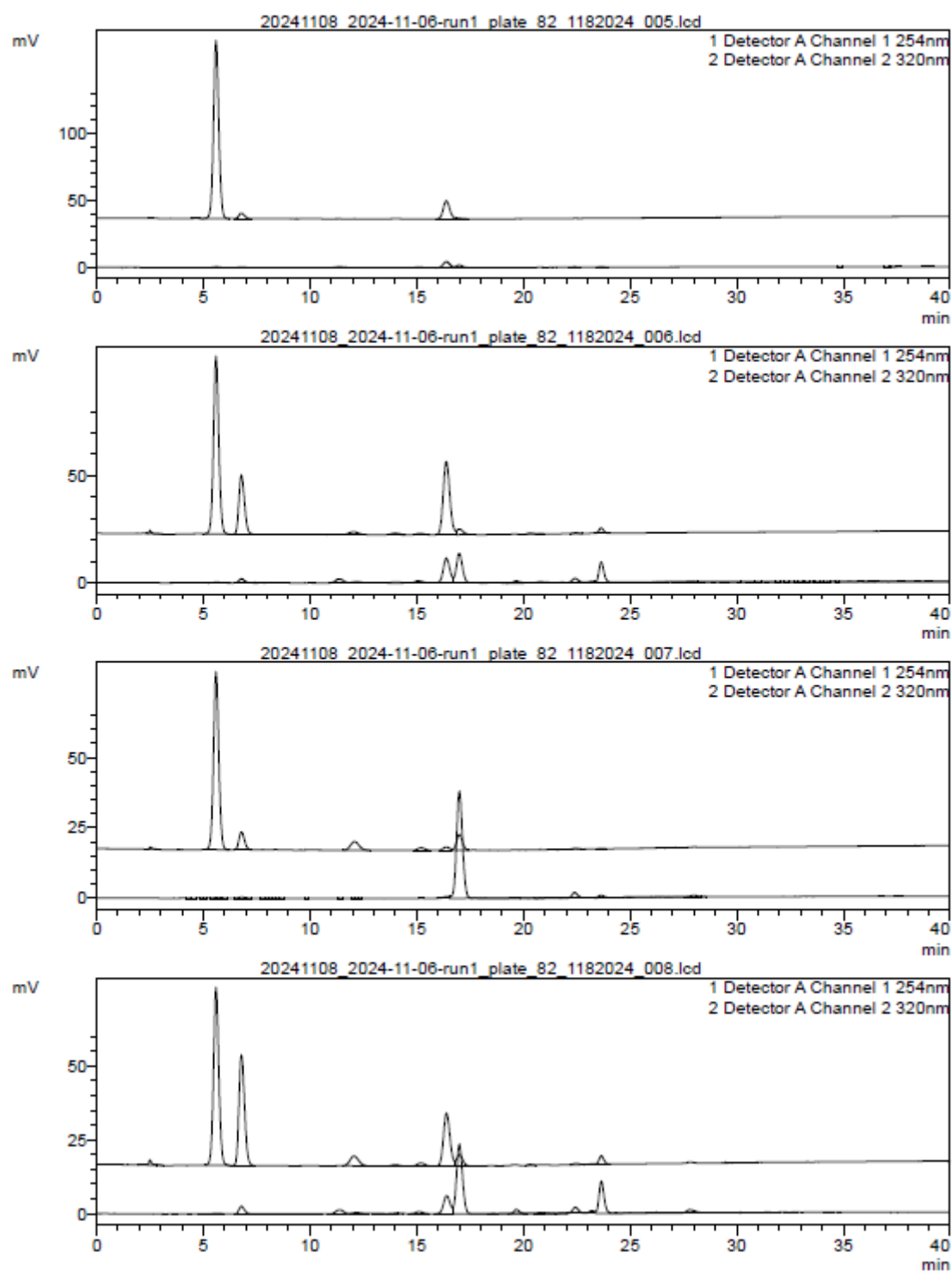
HPLC chromatogram for conditions 31, 91, 16, 4, from Table S12, top to bottom



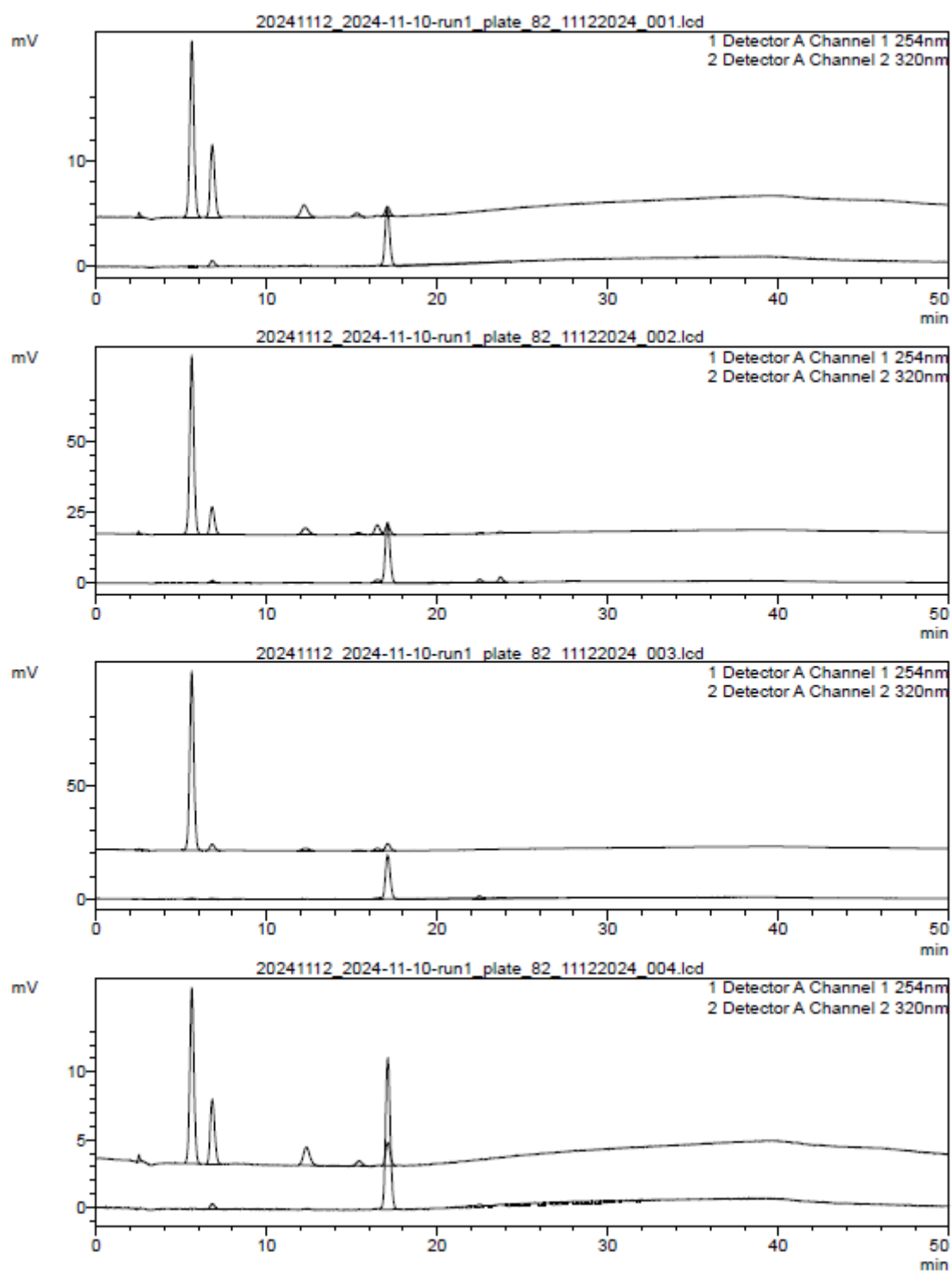
HPLC chromatogram for conditions 34, 31, 91, 16, from Table S12



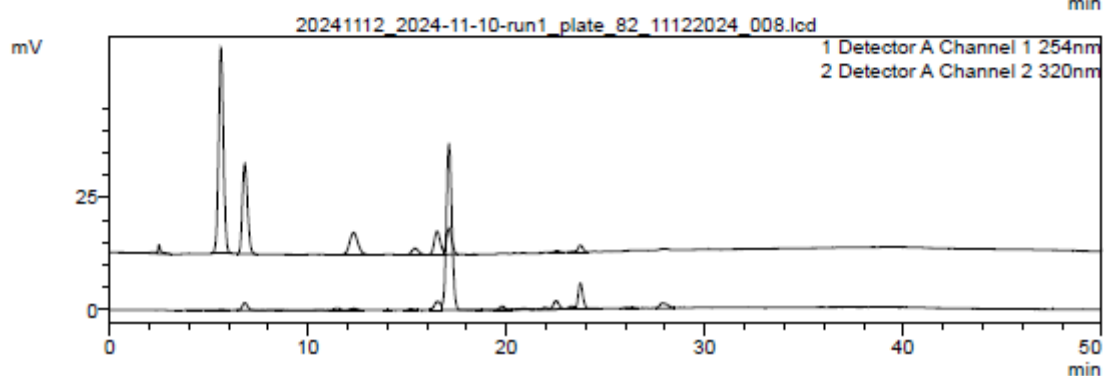
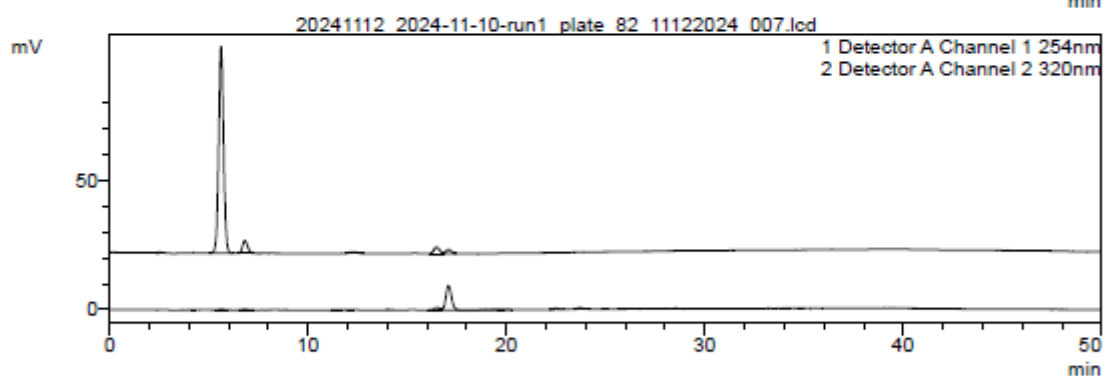
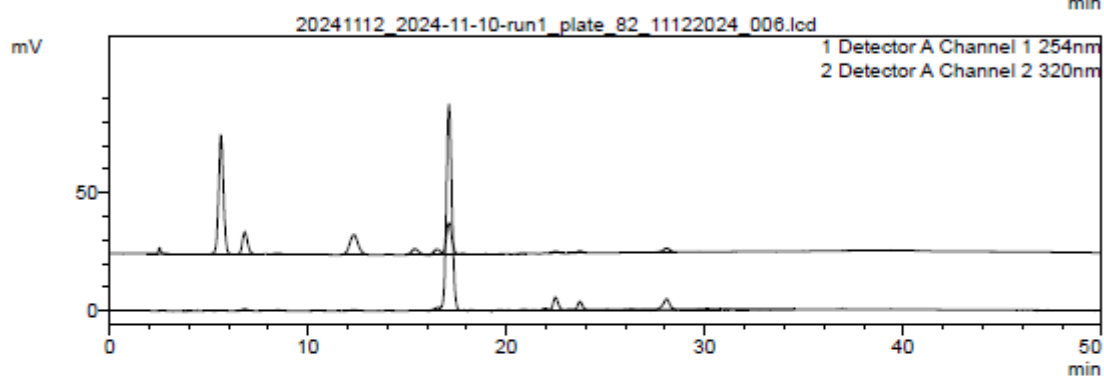
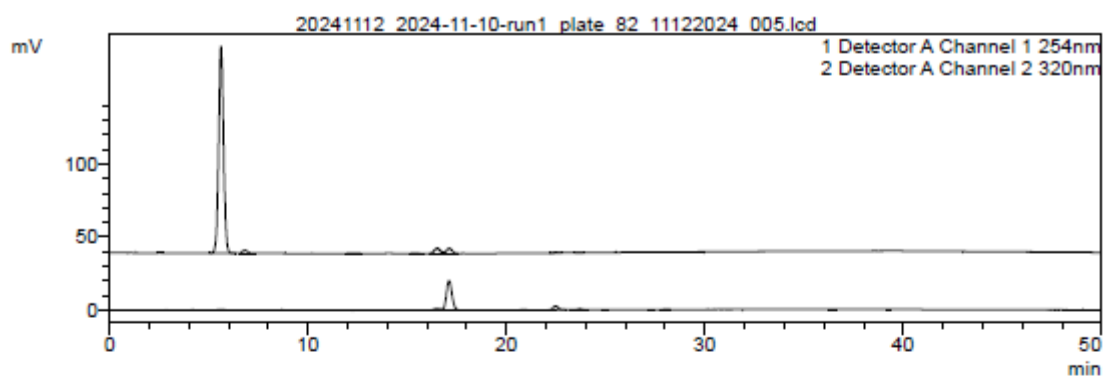
HPLC chromatogram for conditions 4, 46, 25, 10, from Table S12



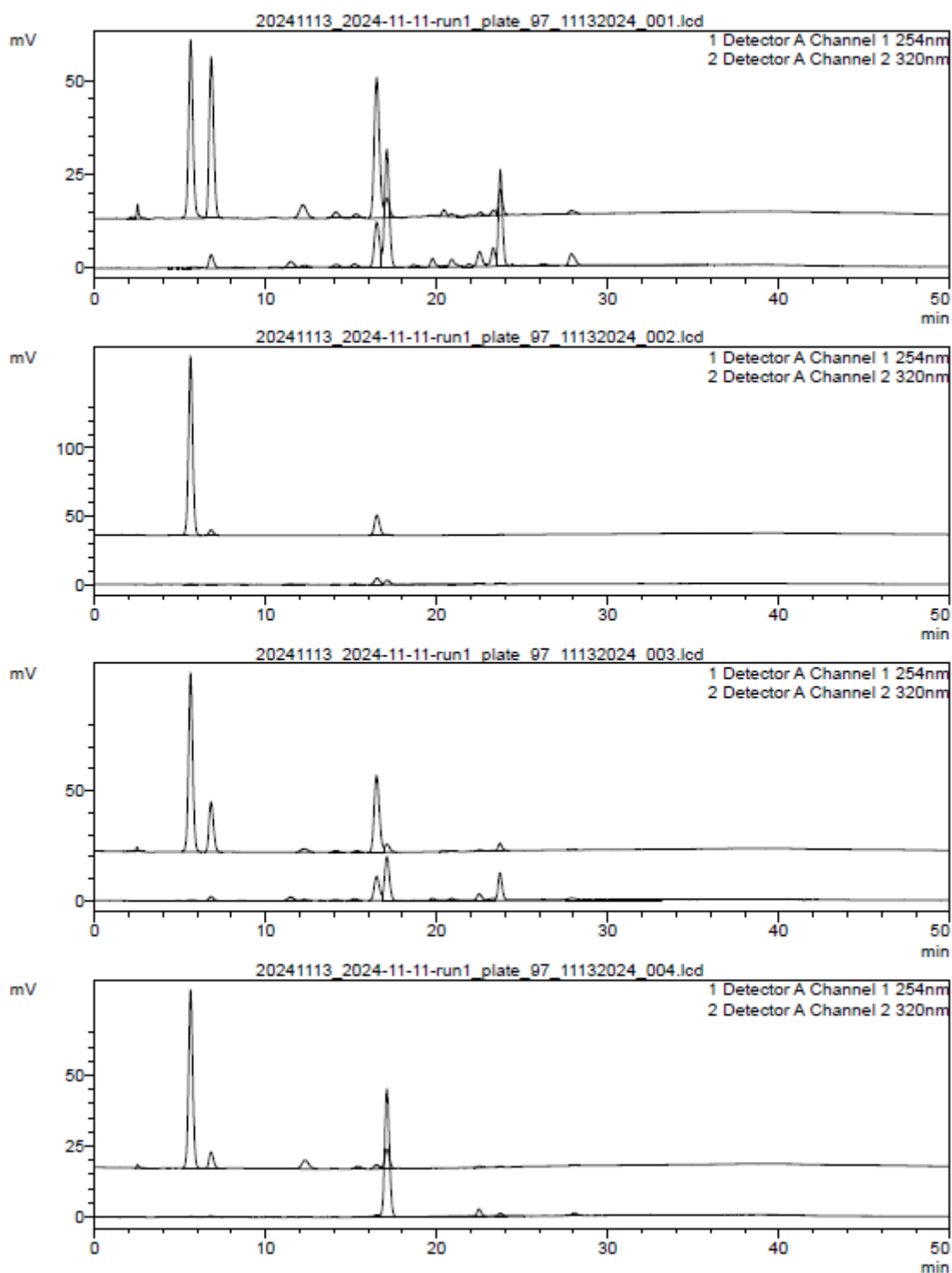
HPLC chromatogram for conditions 1, 7, 55, 19, from Table S12



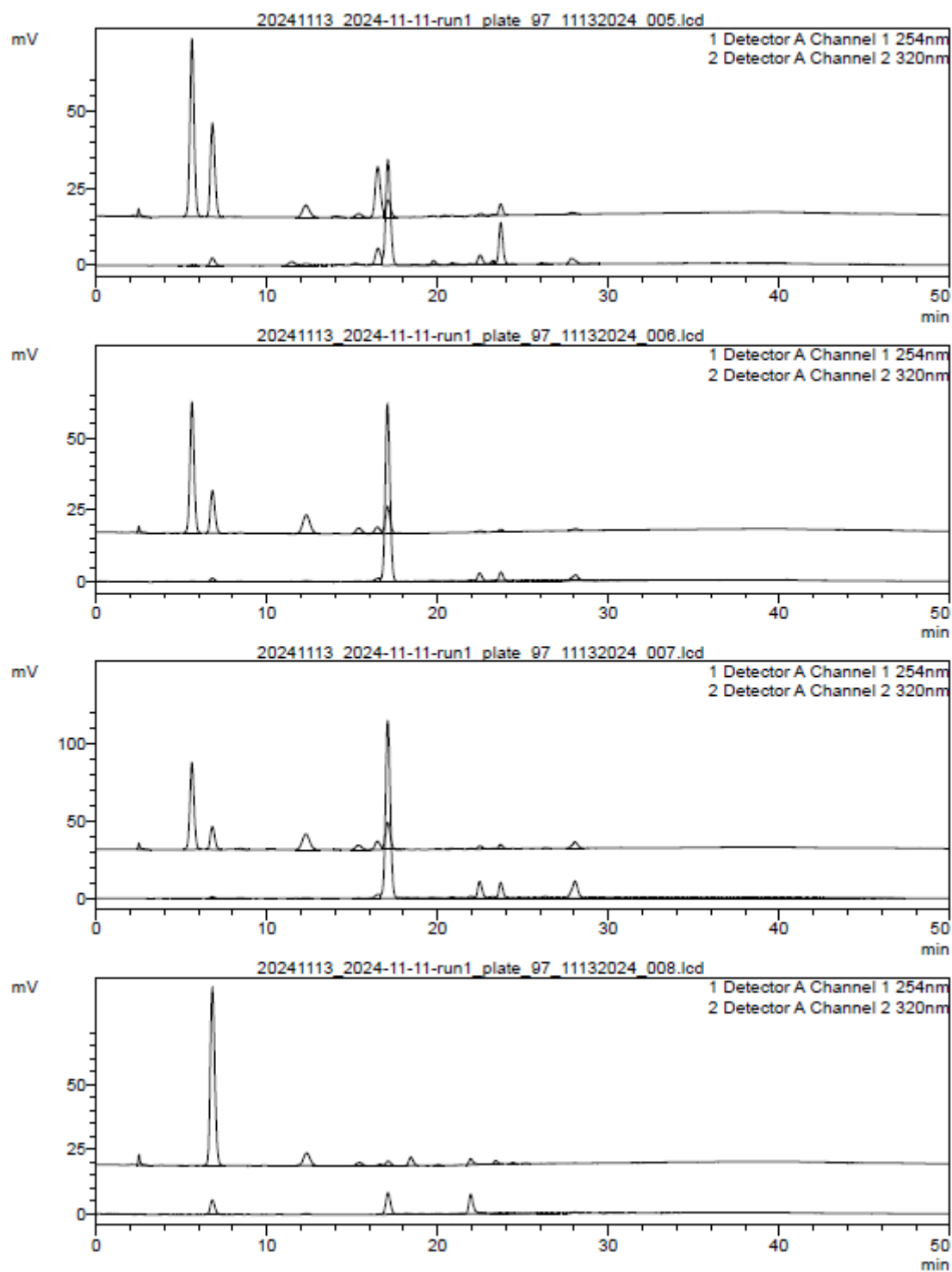
HPLC chromatogram for conditions 58, 28, 37, 43, from Table S12



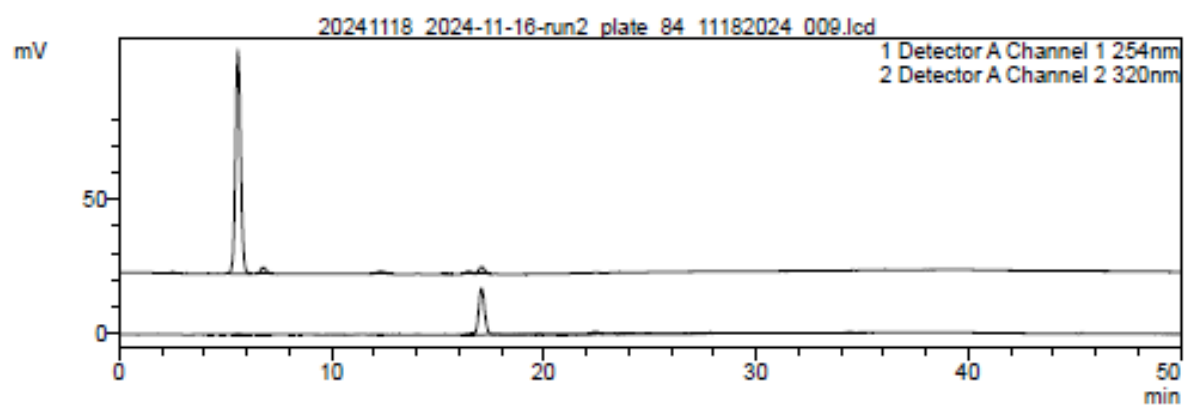
HPLC chromatogram for conditions 16, 4, 46, 25, from Table S12



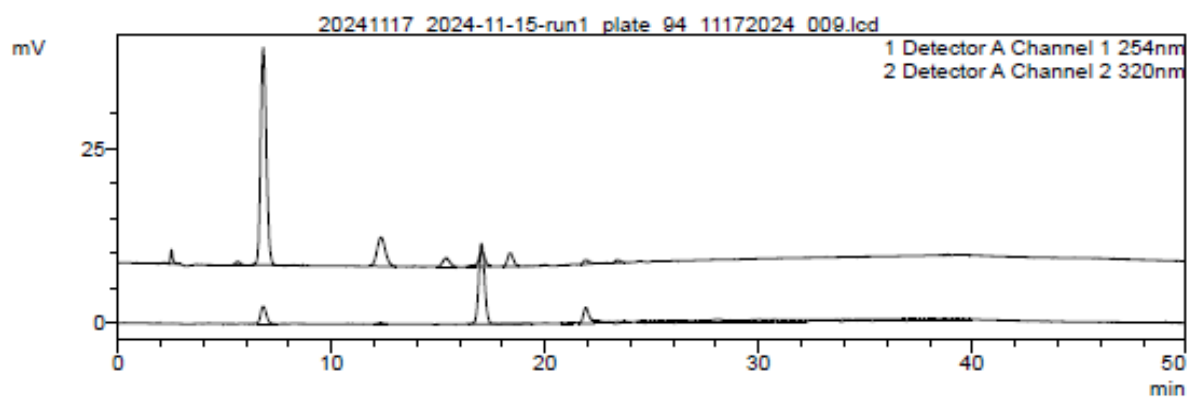
HPLC chromatogram for conditions 10, 64, 34, 31, from Table S12



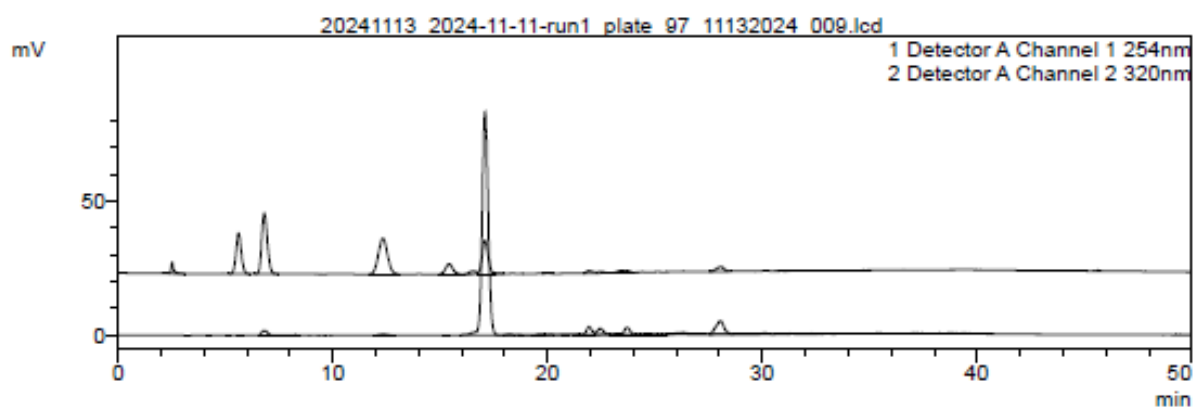
HPLC chromatogram for condition 55, from Table S12



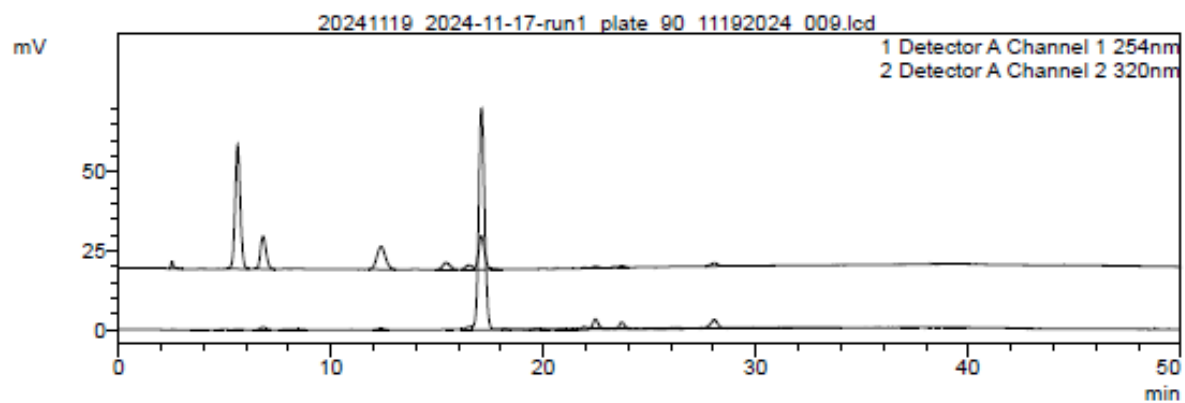
HPLC chromatogram for condition 61, from Table S12.



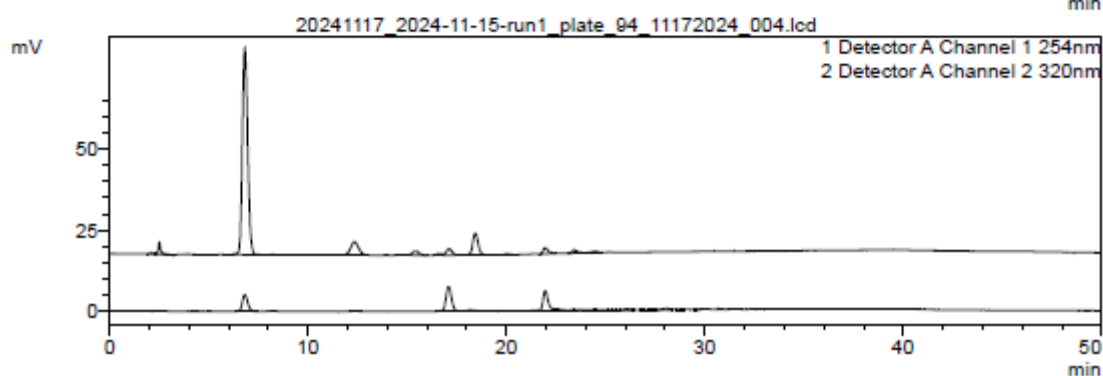
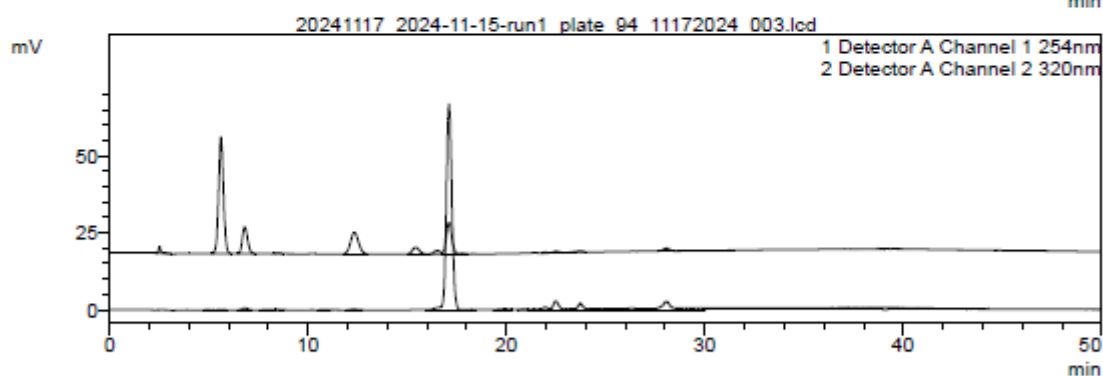
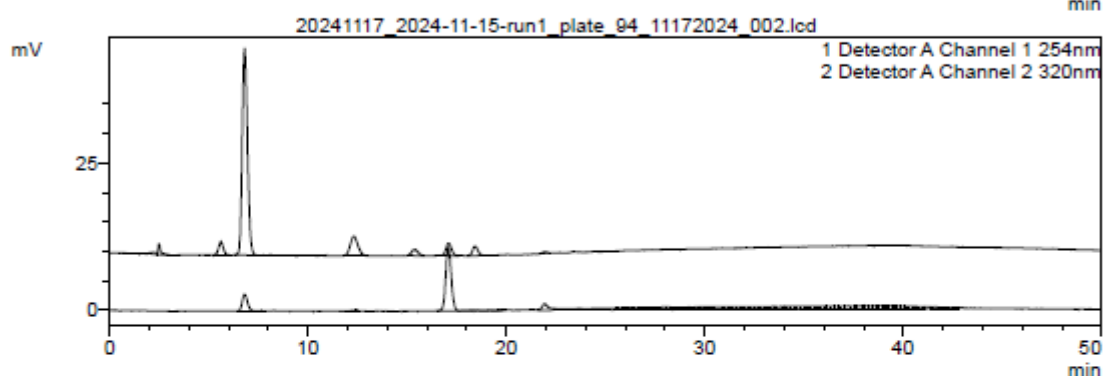
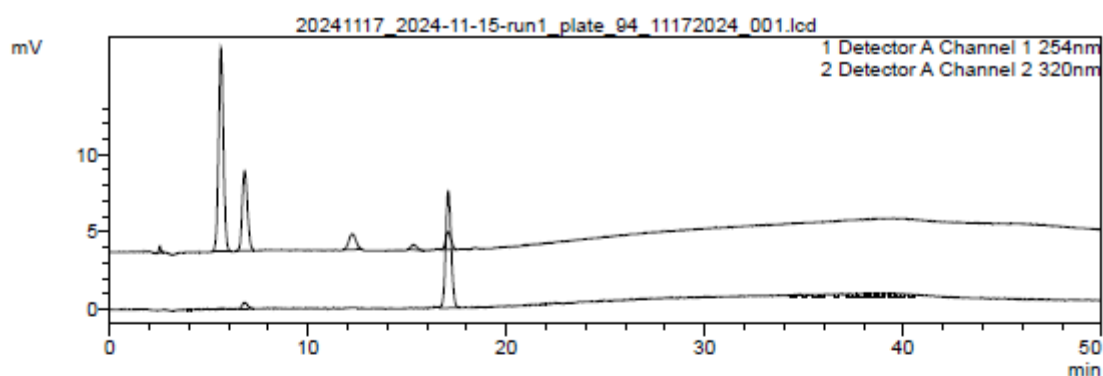
HPLC chromatogram for condition 91, from Table S12



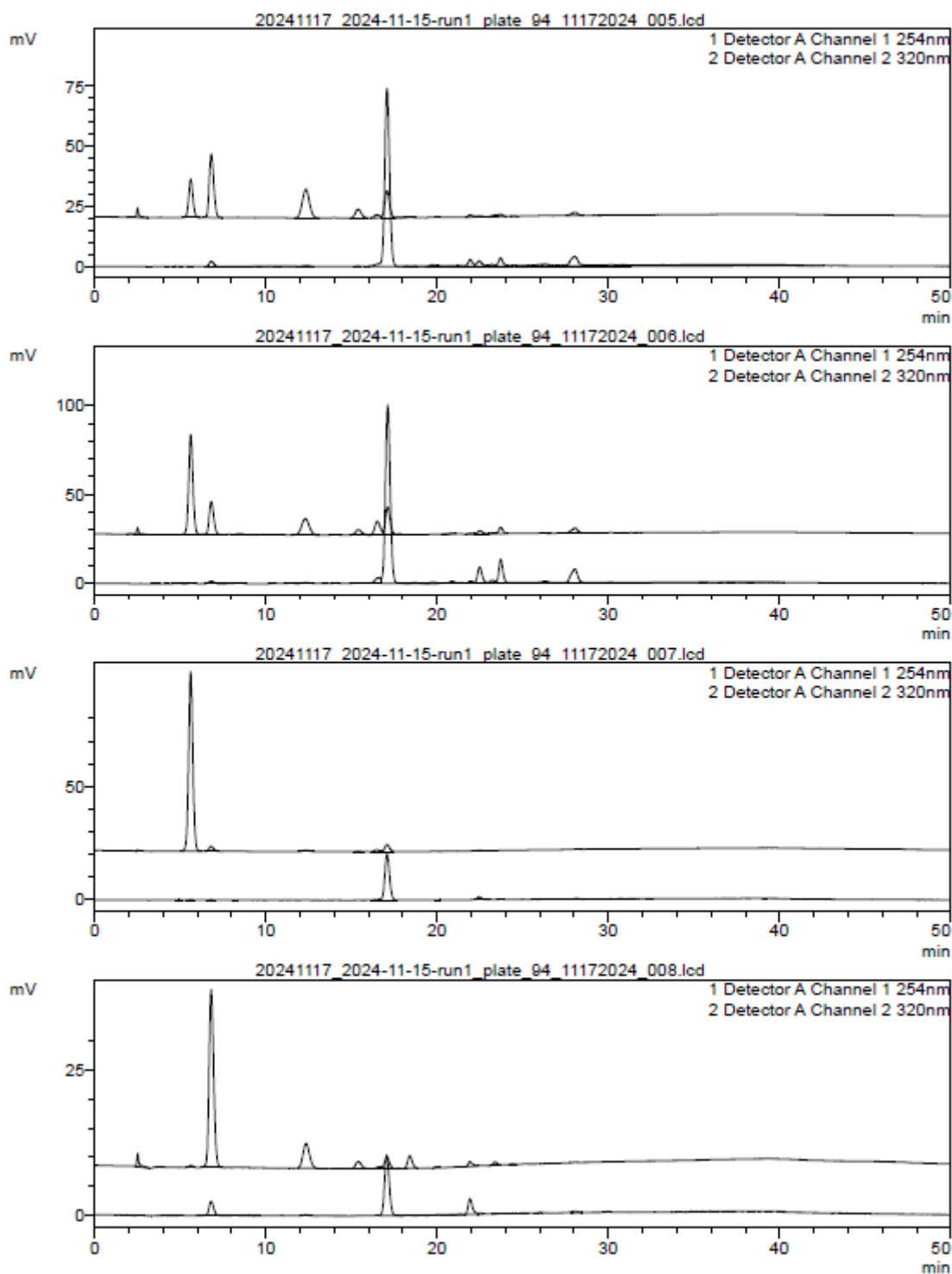
HPLC chromatogram for condition 85, from Table S12



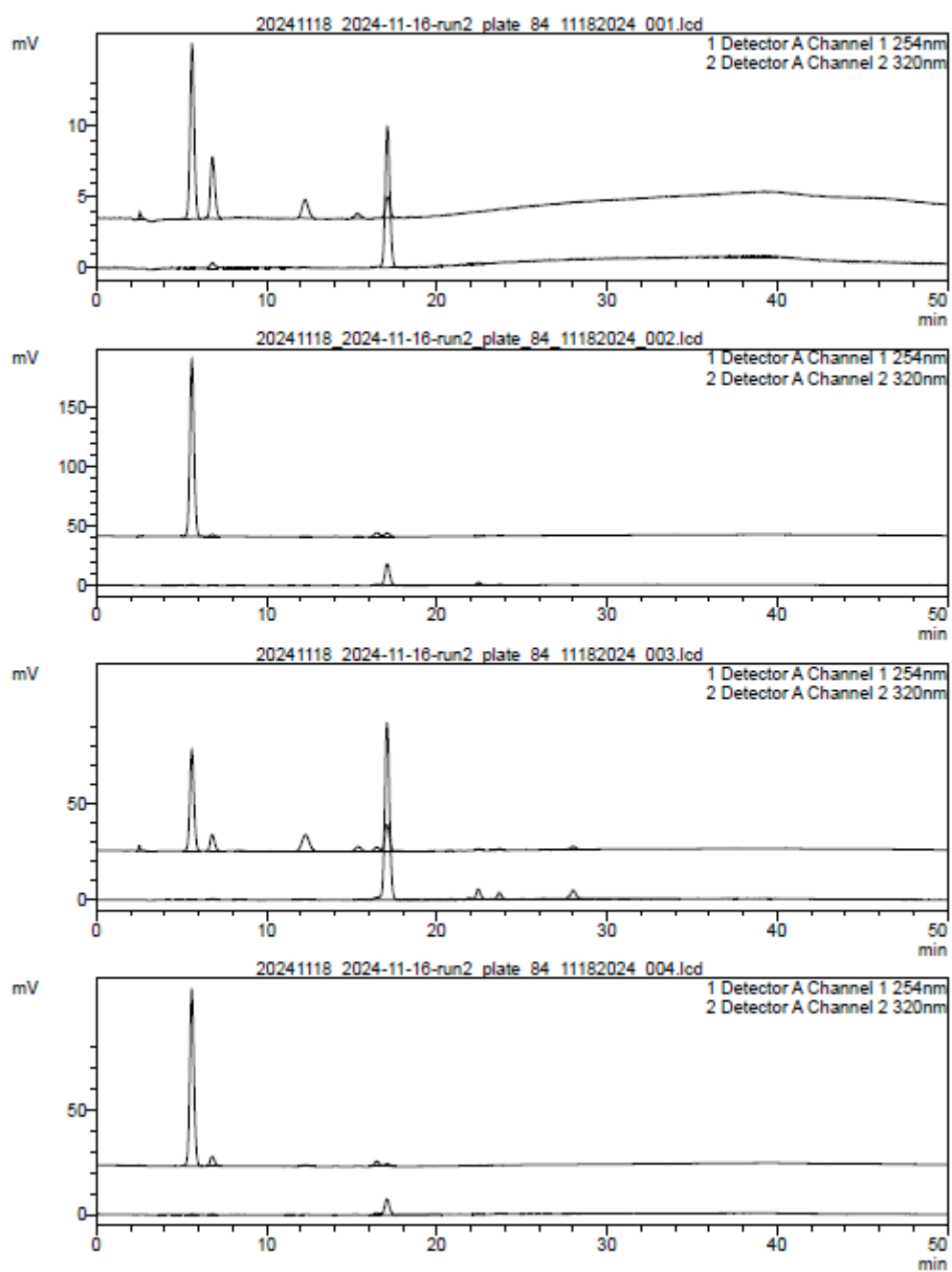
HPLC chromatogram for conditions 52, 40, 85, 70, from Table S12



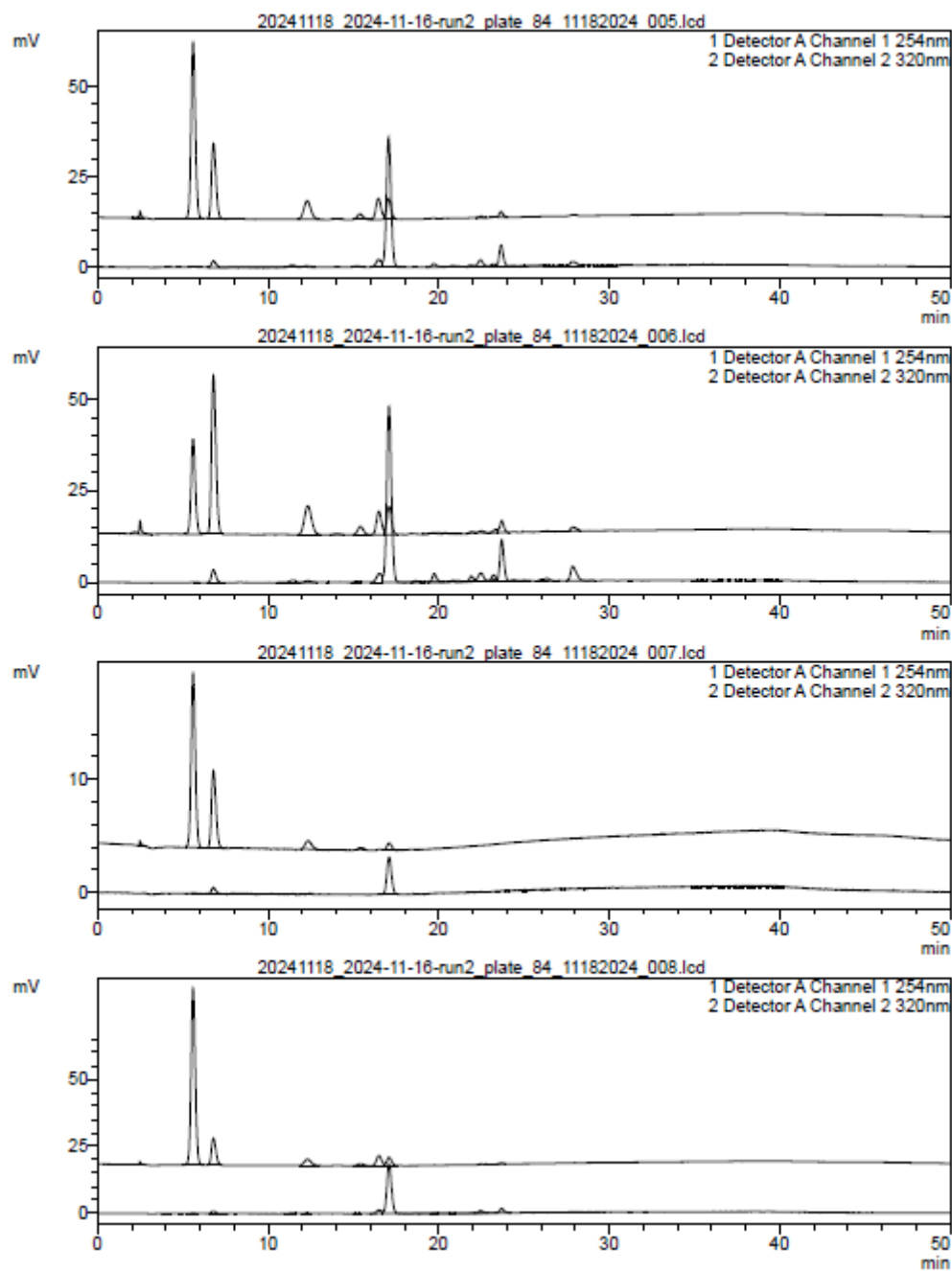
HPLC chromatogram for conditions 73, 76, 79, 82, from Table S12



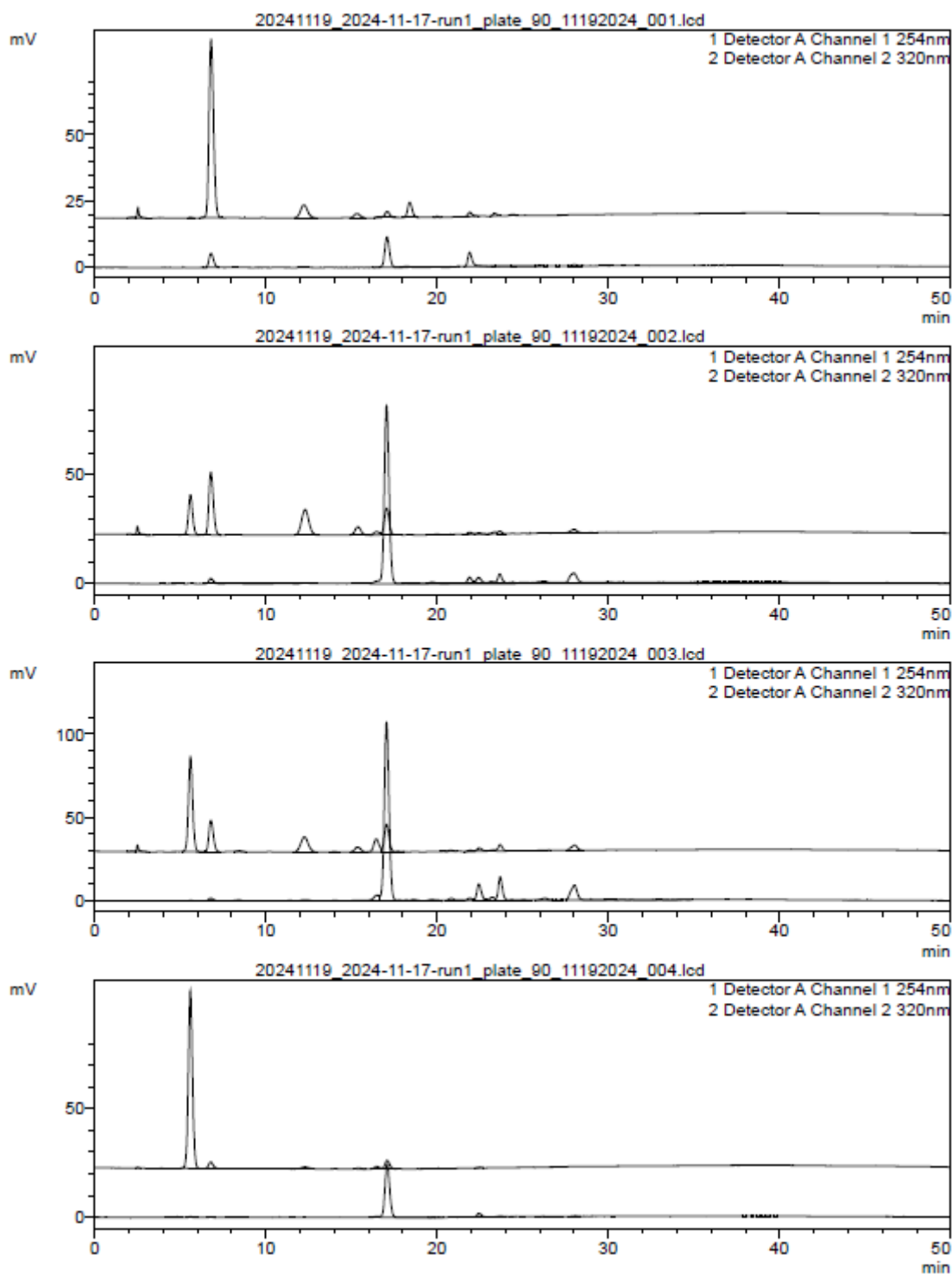
HPLC chromatogram for conditions 19, 58, 28, 37, from Table S12



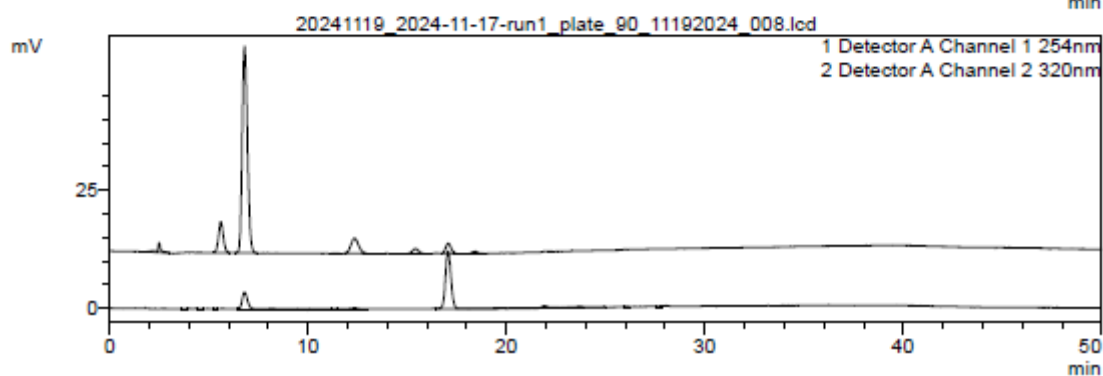
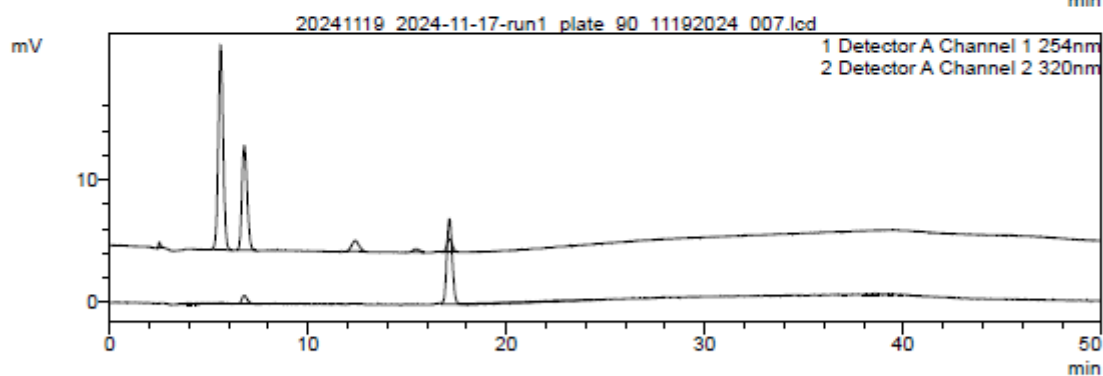
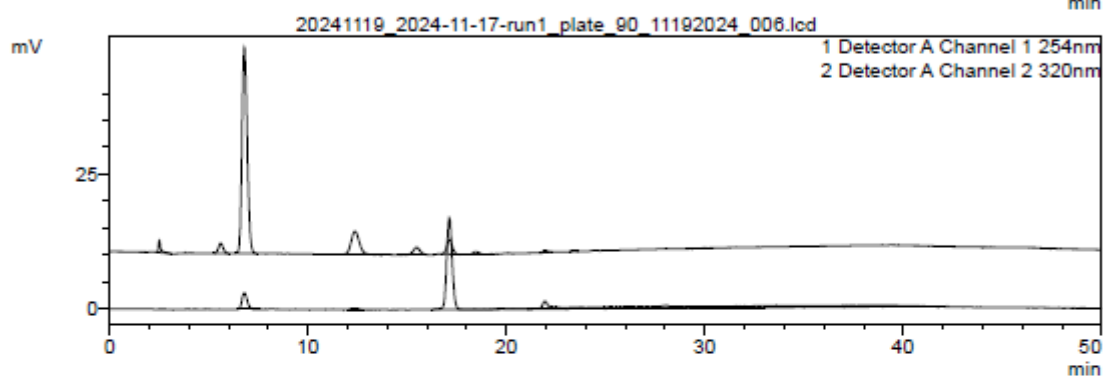
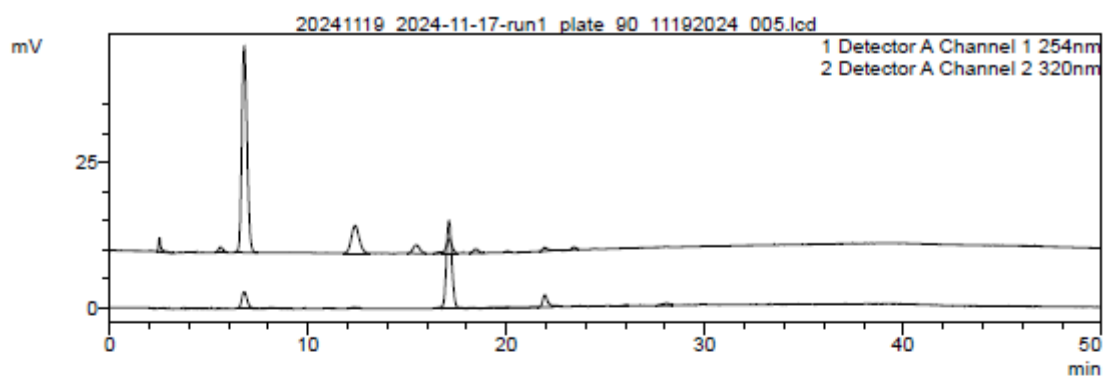
HPLC chromatogram for conditions 43, 49, 1, 7, from Table S12



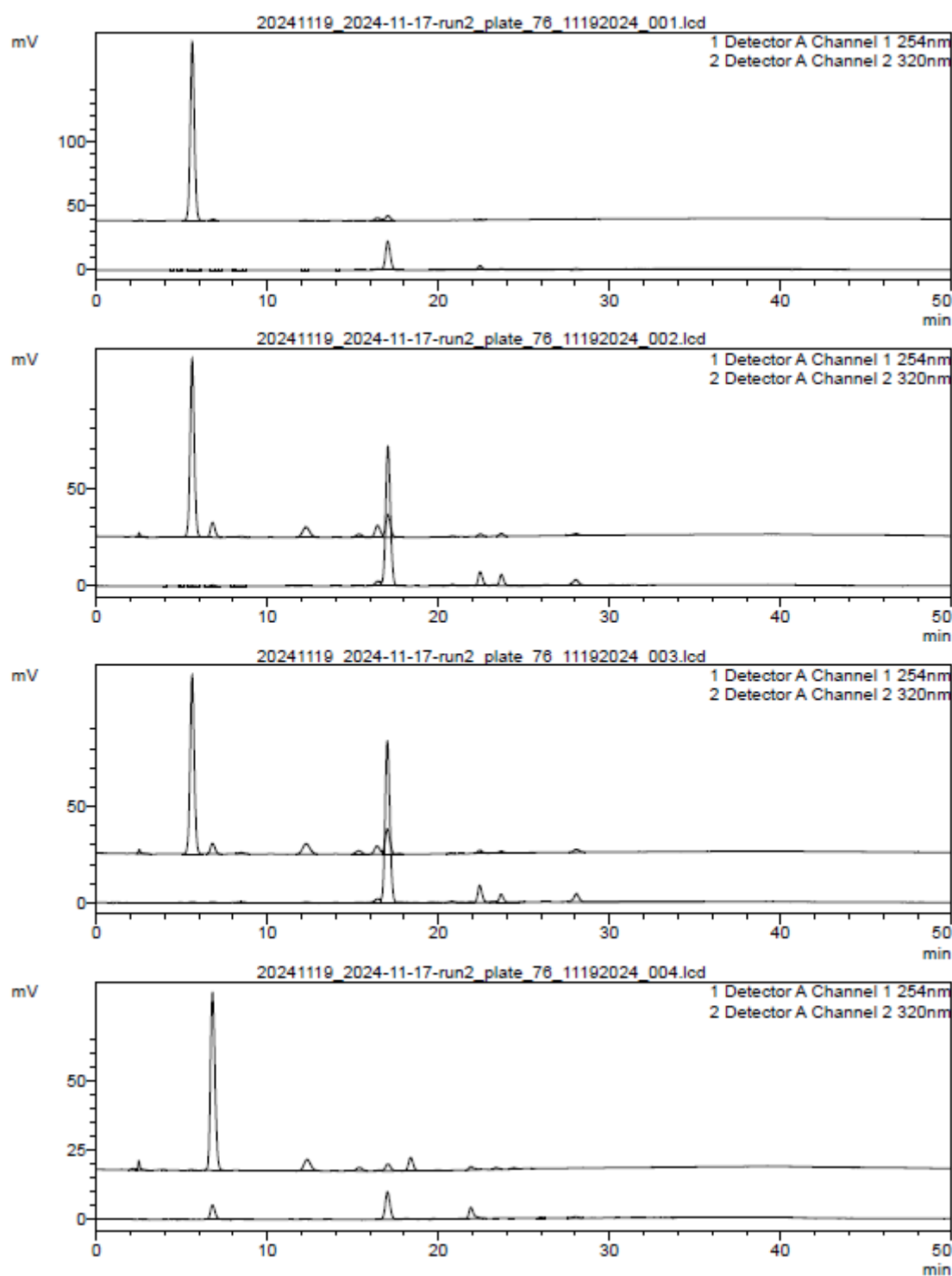
HPLC chromatogram for conditions 70, 73, 76, 79, from Table S12



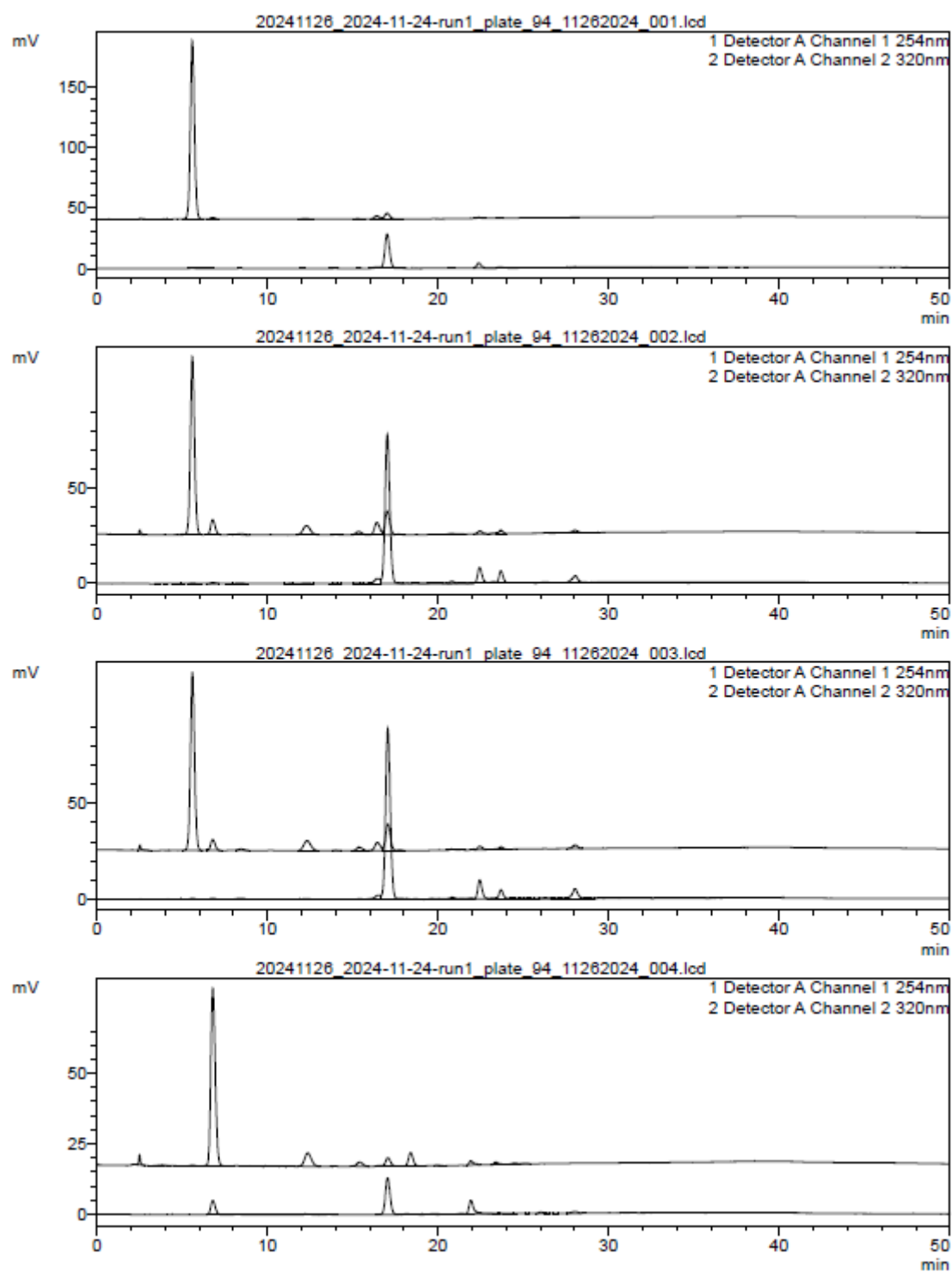
HPLC chromatogram for conditions 82, 61, 52, 40, from Table S12



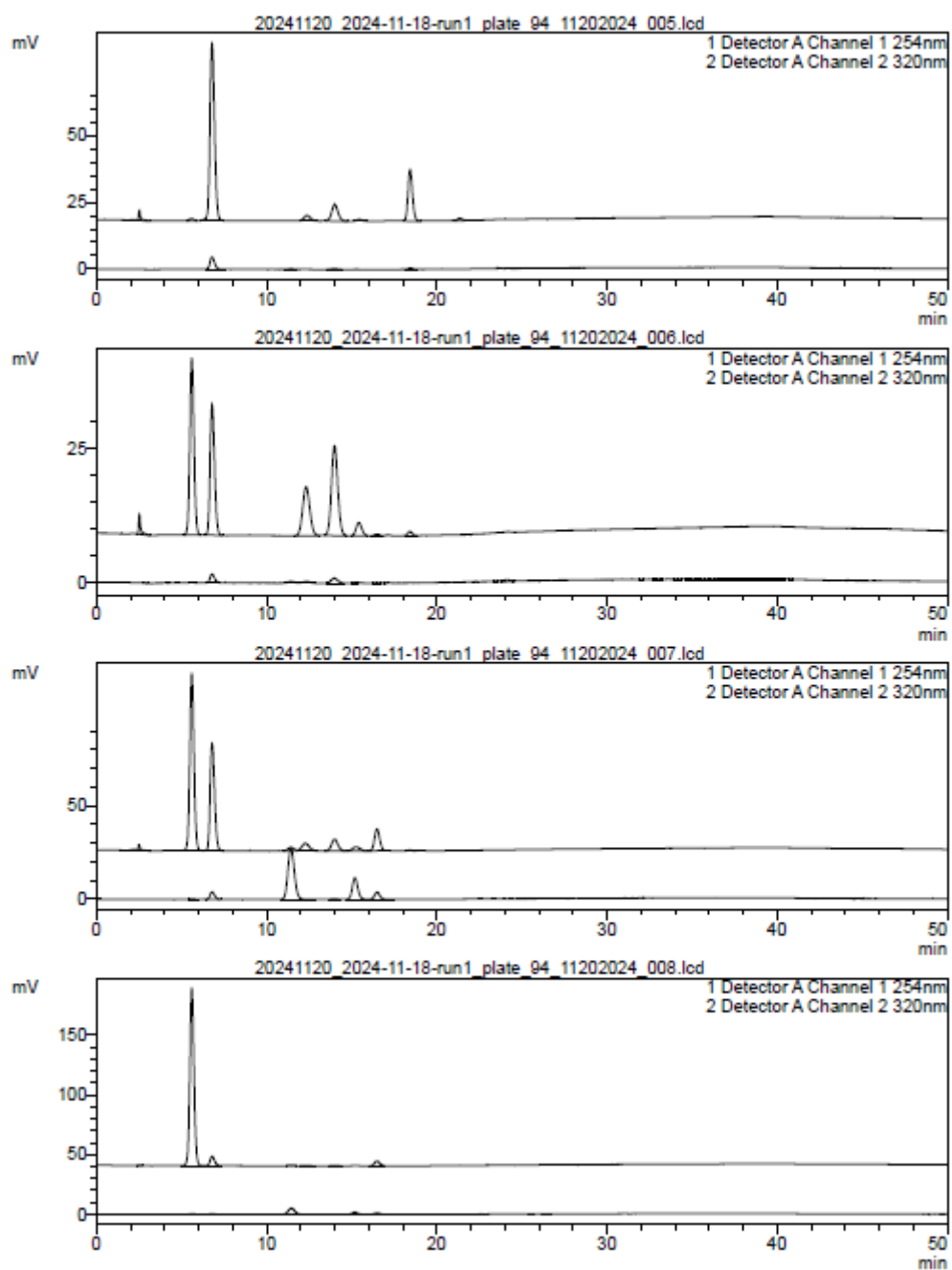
HPLC chromatogram for conditions 22, 67, 88, 13, from Table S12



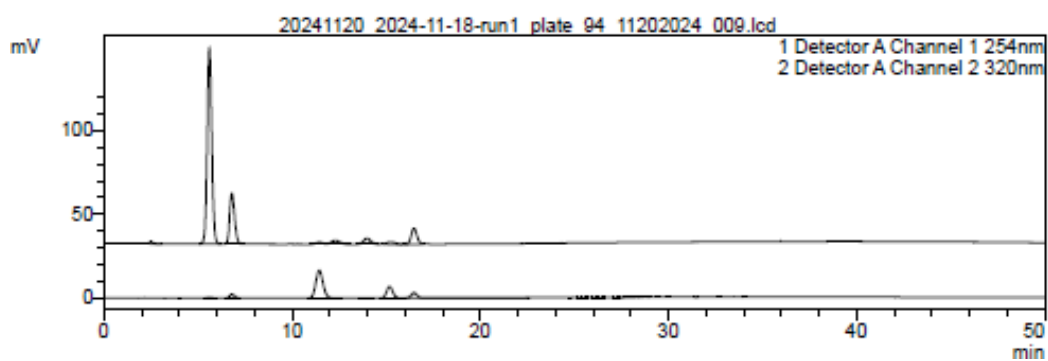
HPLC chromatogram for conditions 22, 67, 88, 13, from Table S12



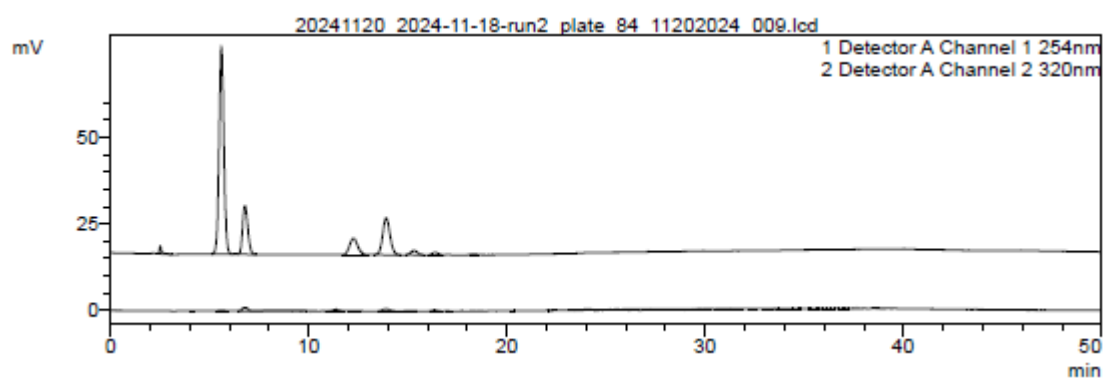
HPLC chromatogram for conditions 31, 91, 16, 4, from Table S13



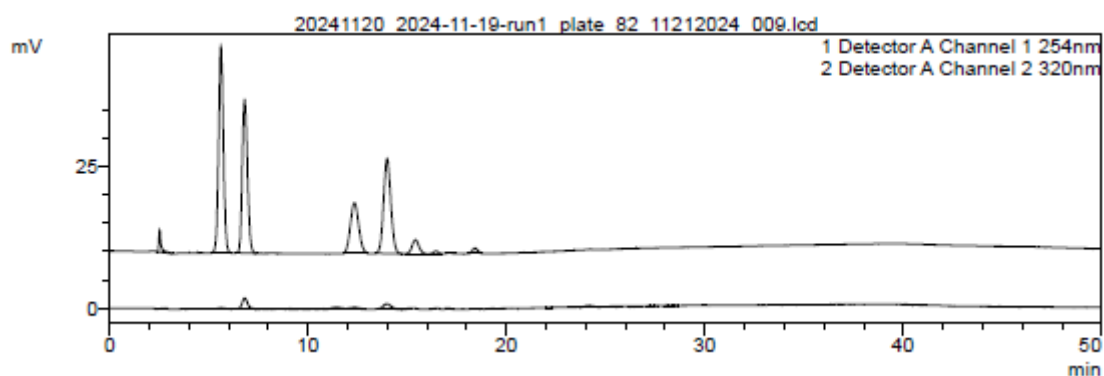
HPLC chromatogram for condition 46, from Table S13



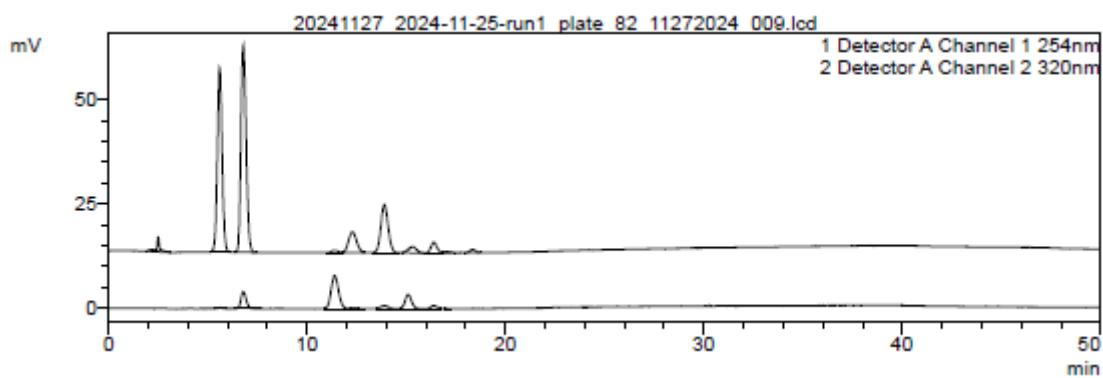
HPLC chromatogram for condition 64, from Table S13



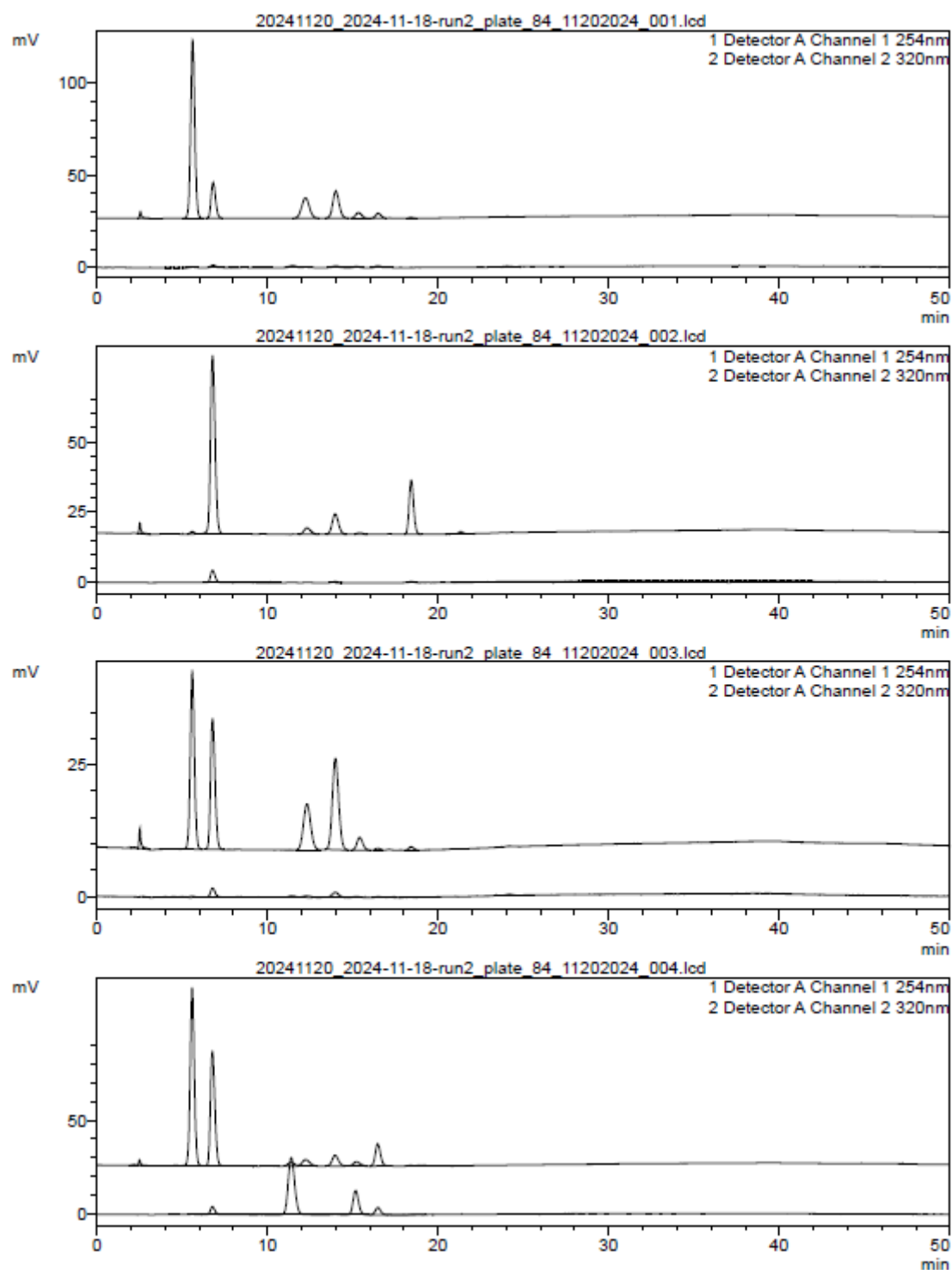
HPLC chromatogram for condition 91, from Table S13



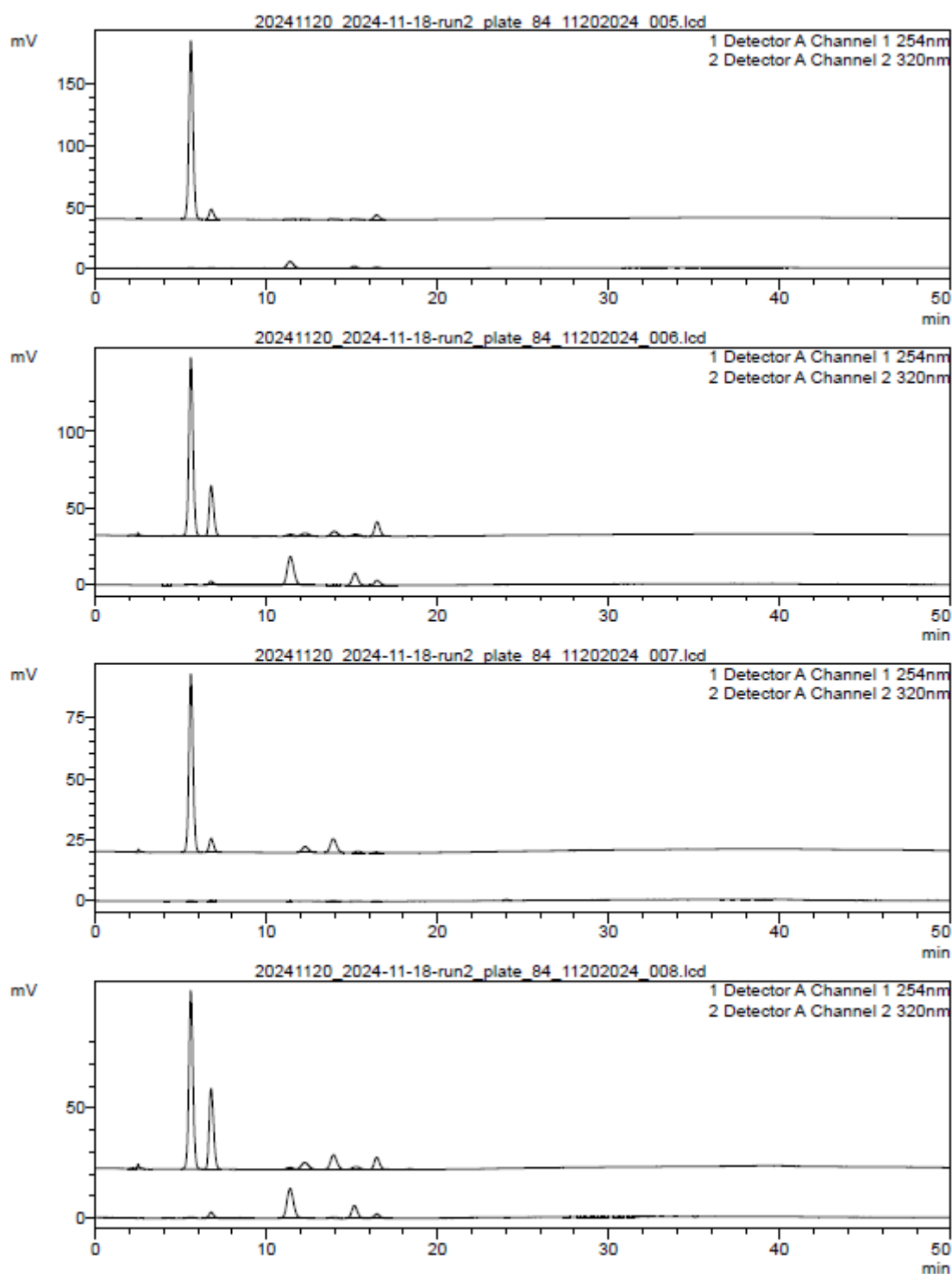
HPLC chromatogram for condition 49, from Table S13



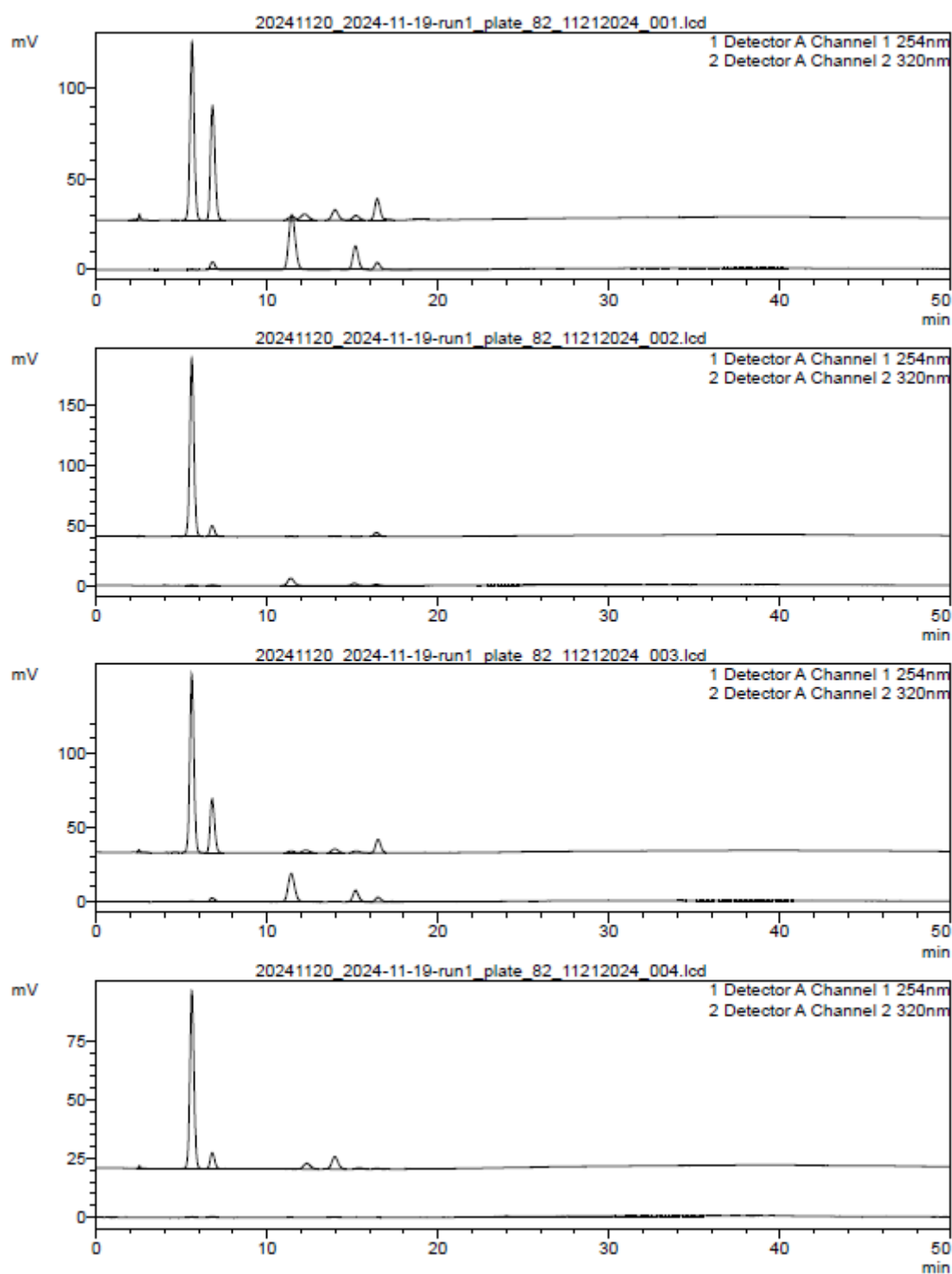
HPLC chromatogram for conditions 34, 31, 91, 16, from Table S13



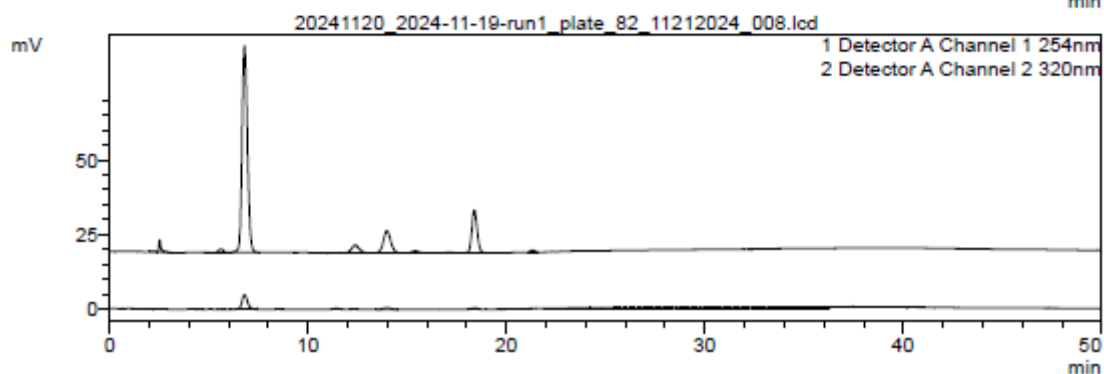
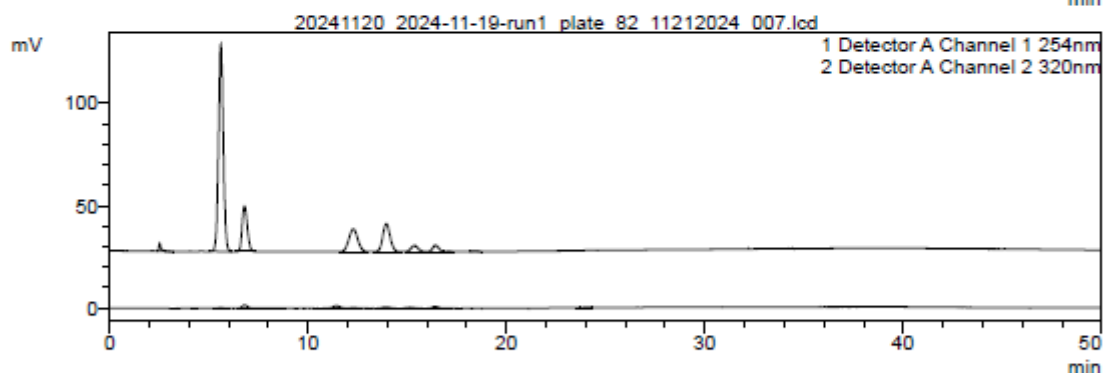
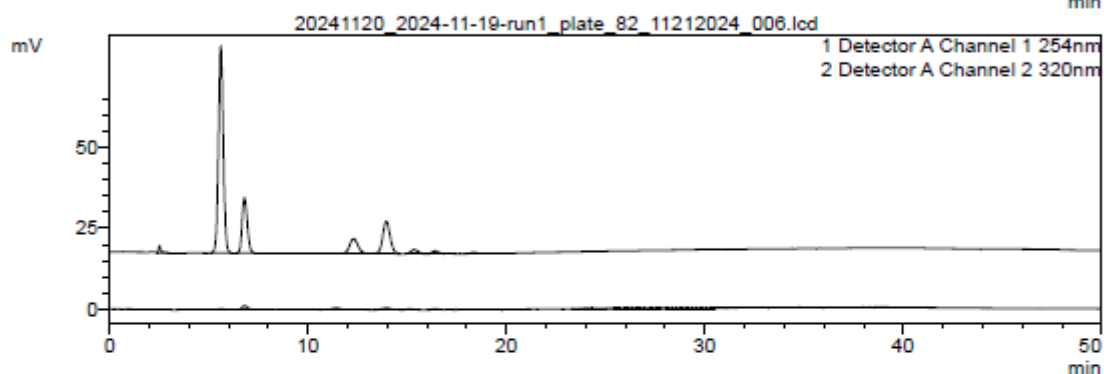
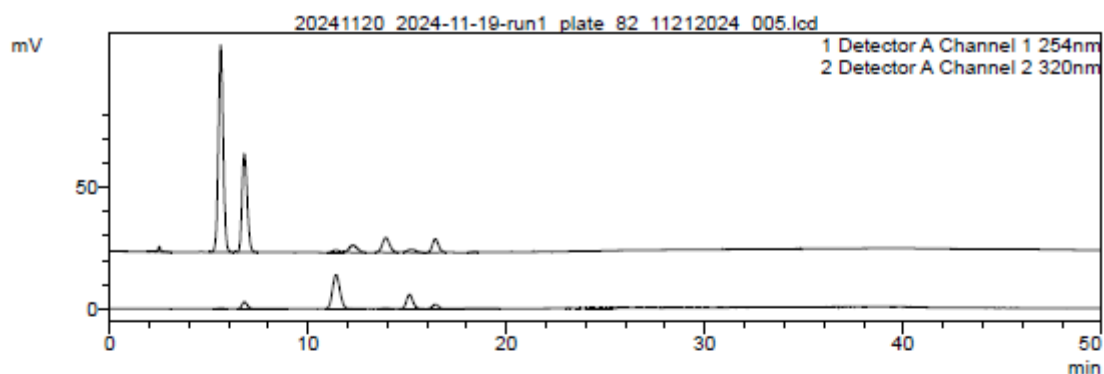
HPLC chromatogram for conditions 4, 46, 25, 10, from Table S13



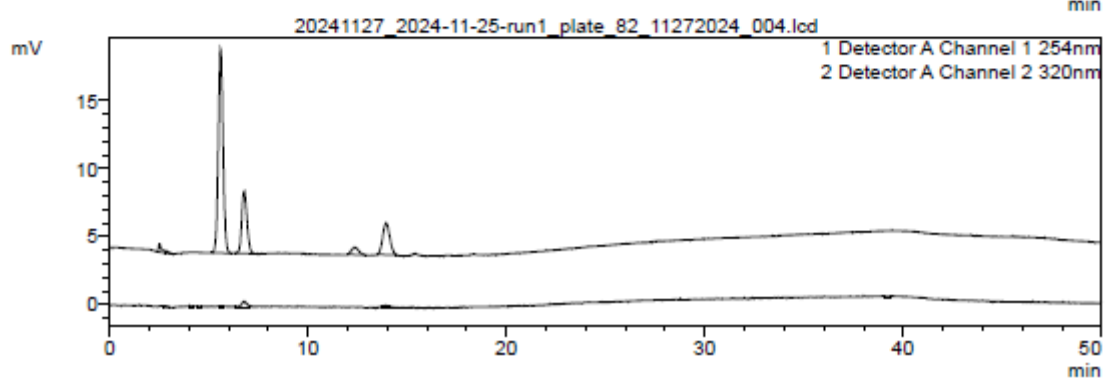
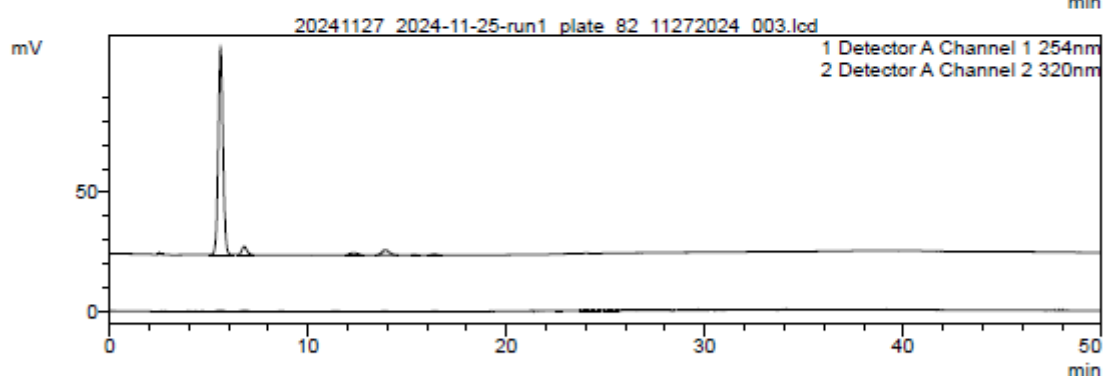
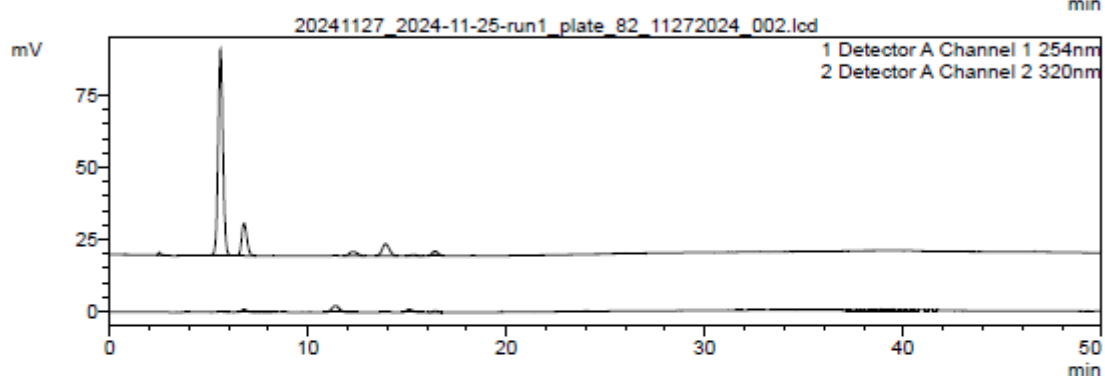
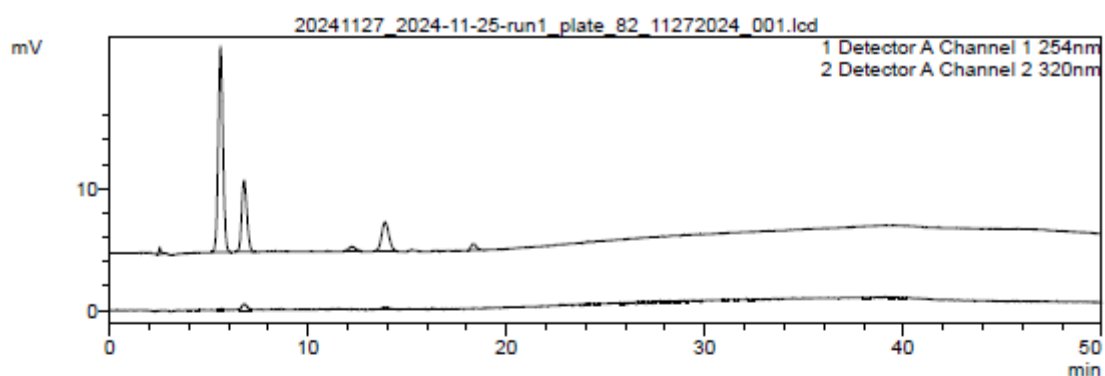
HPLC chromatogram for conditions 16, 4, 46, 25, from Table S13



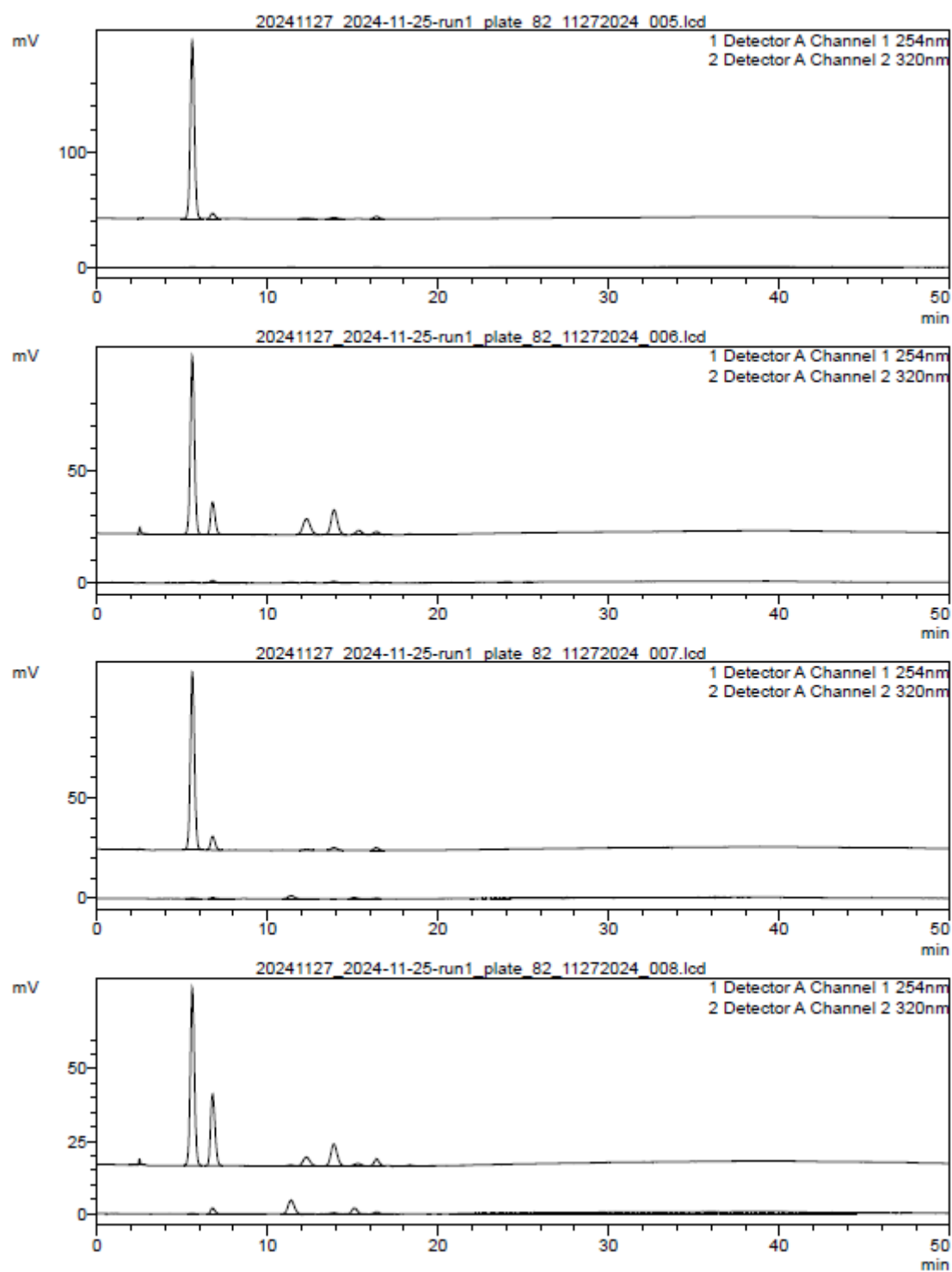
HPLC chromatogram for conditions 10, 64, 34, 31, from Table S13



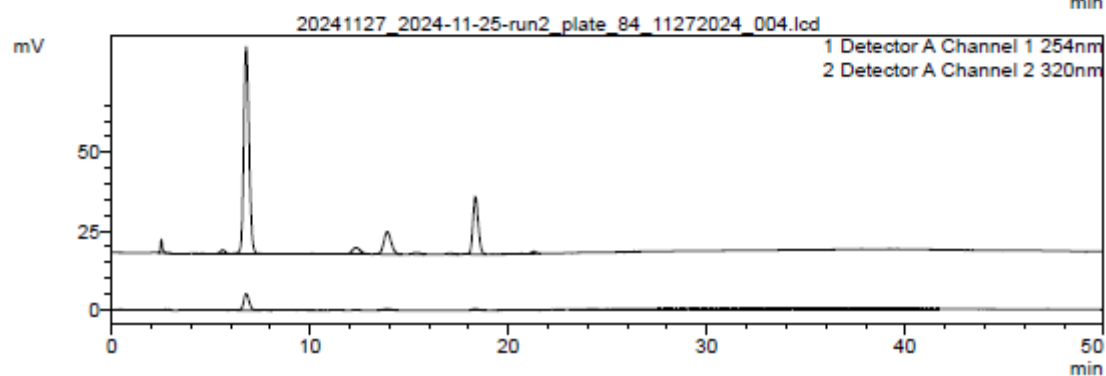
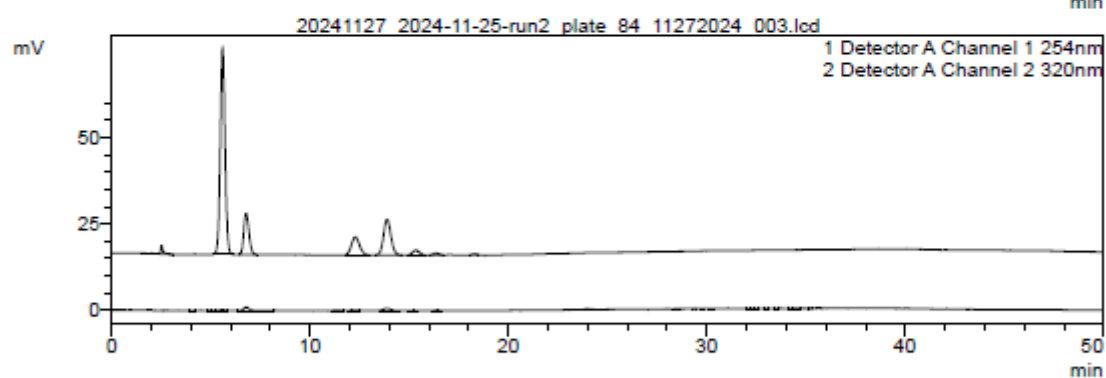
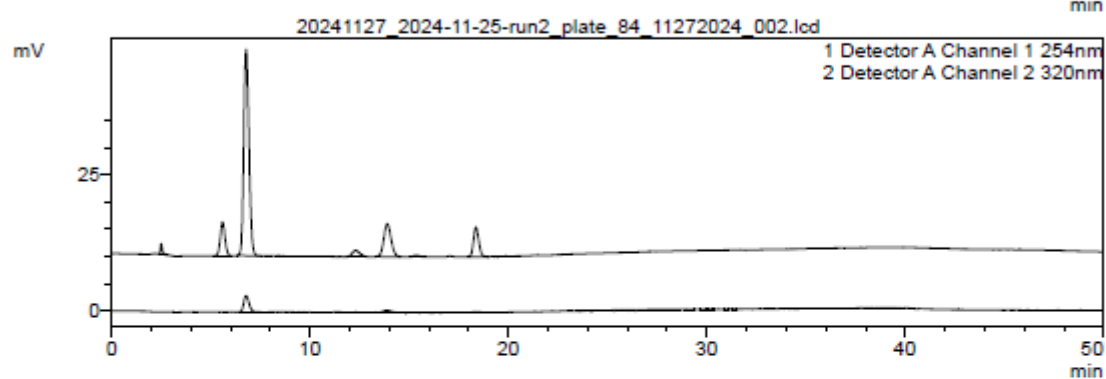
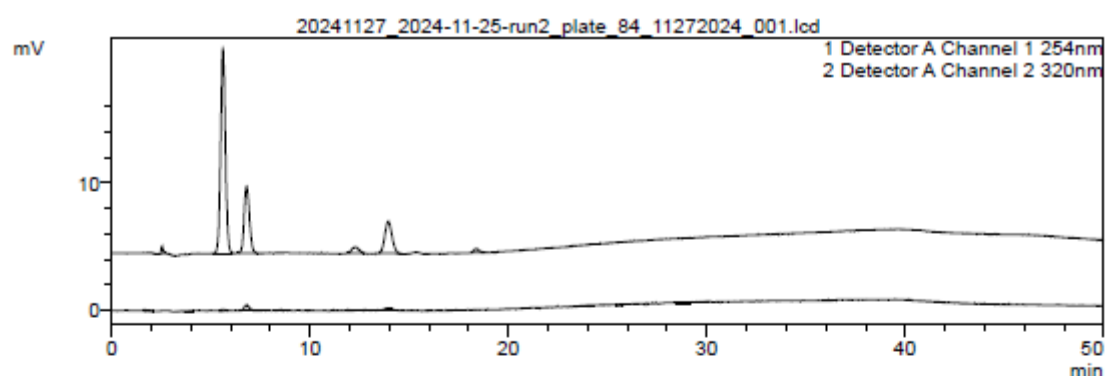
HPLC chromatogram for conditions 1, 7, 55, 19, from Table S13



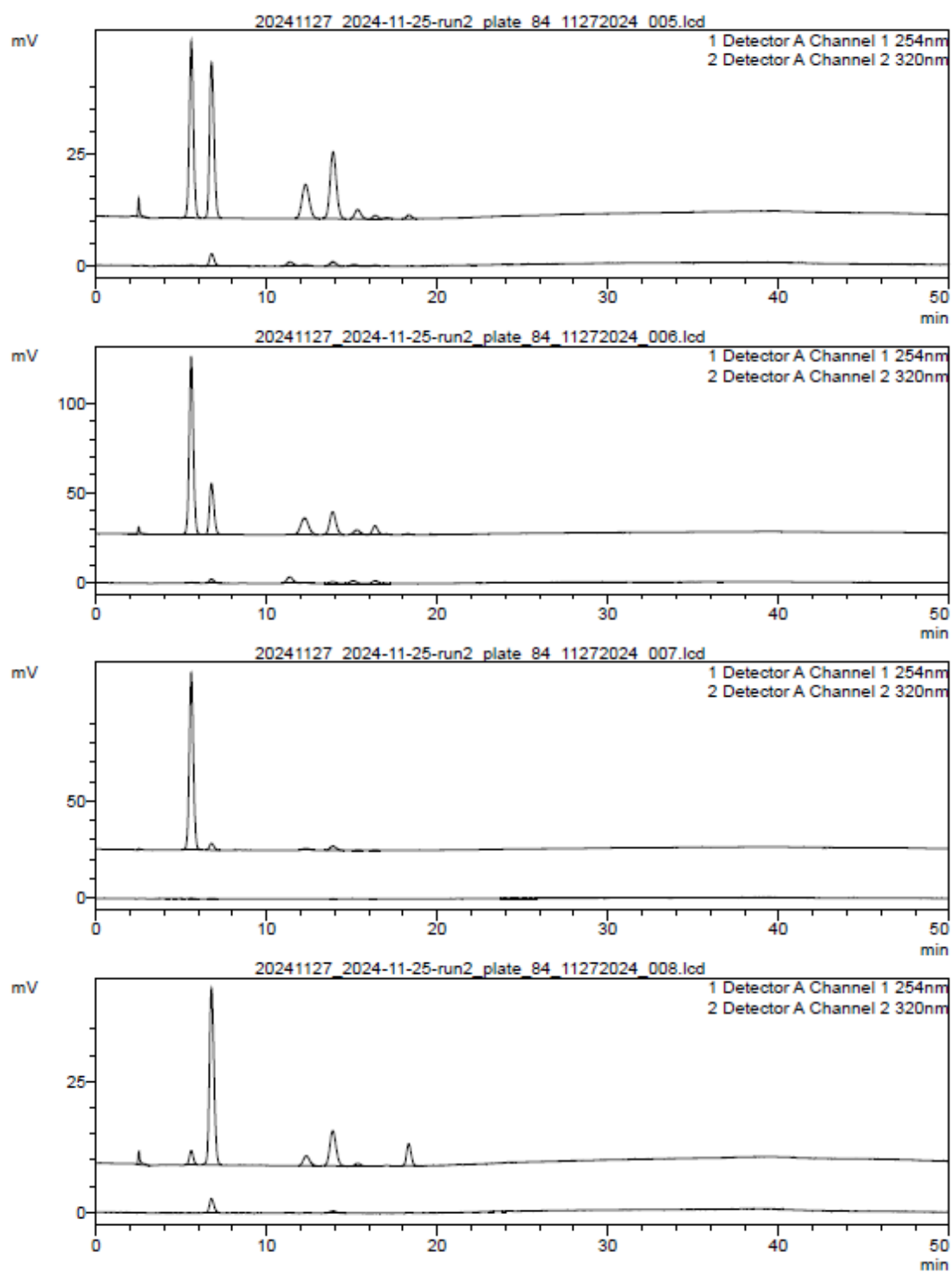
HPLC chromatogram for conditions 58, 28, 37, 43, from Table S13



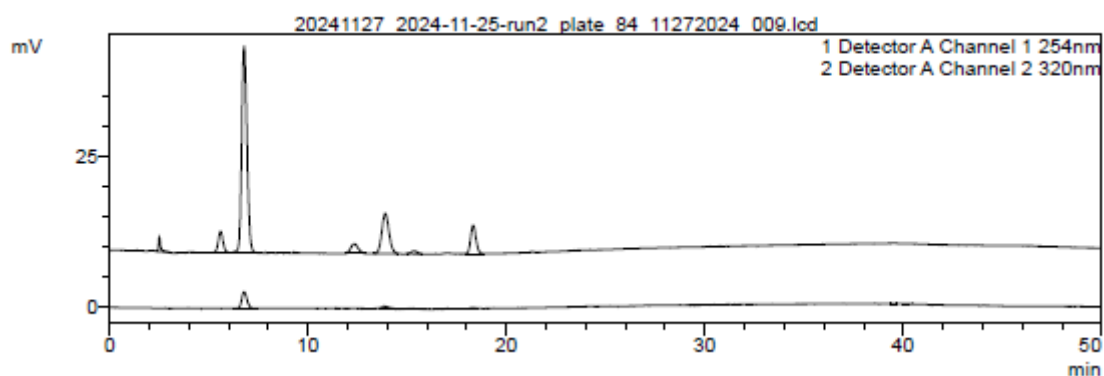
HPLC chromatogram for conditions 52, 40, 85, 70, from Table S13



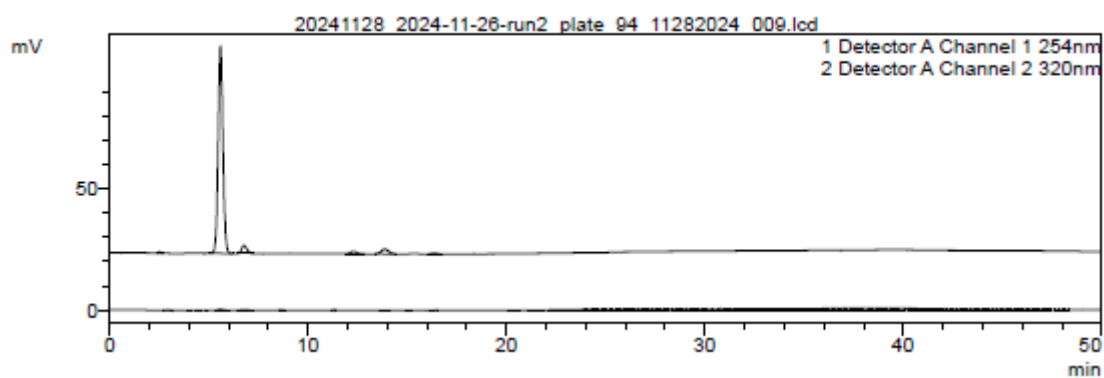
HPLC chromatogram for conditions 73, 76, 79, 82, from Table S13



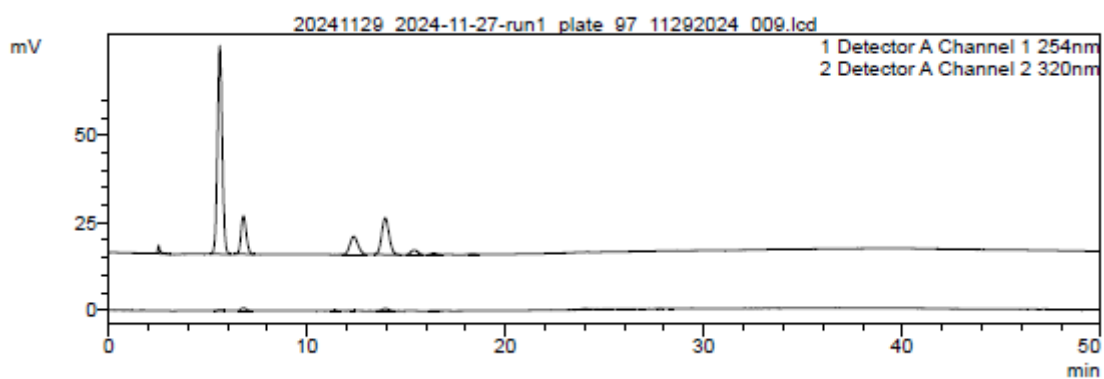
HPLC chromatogram for condition 61, from Table S13



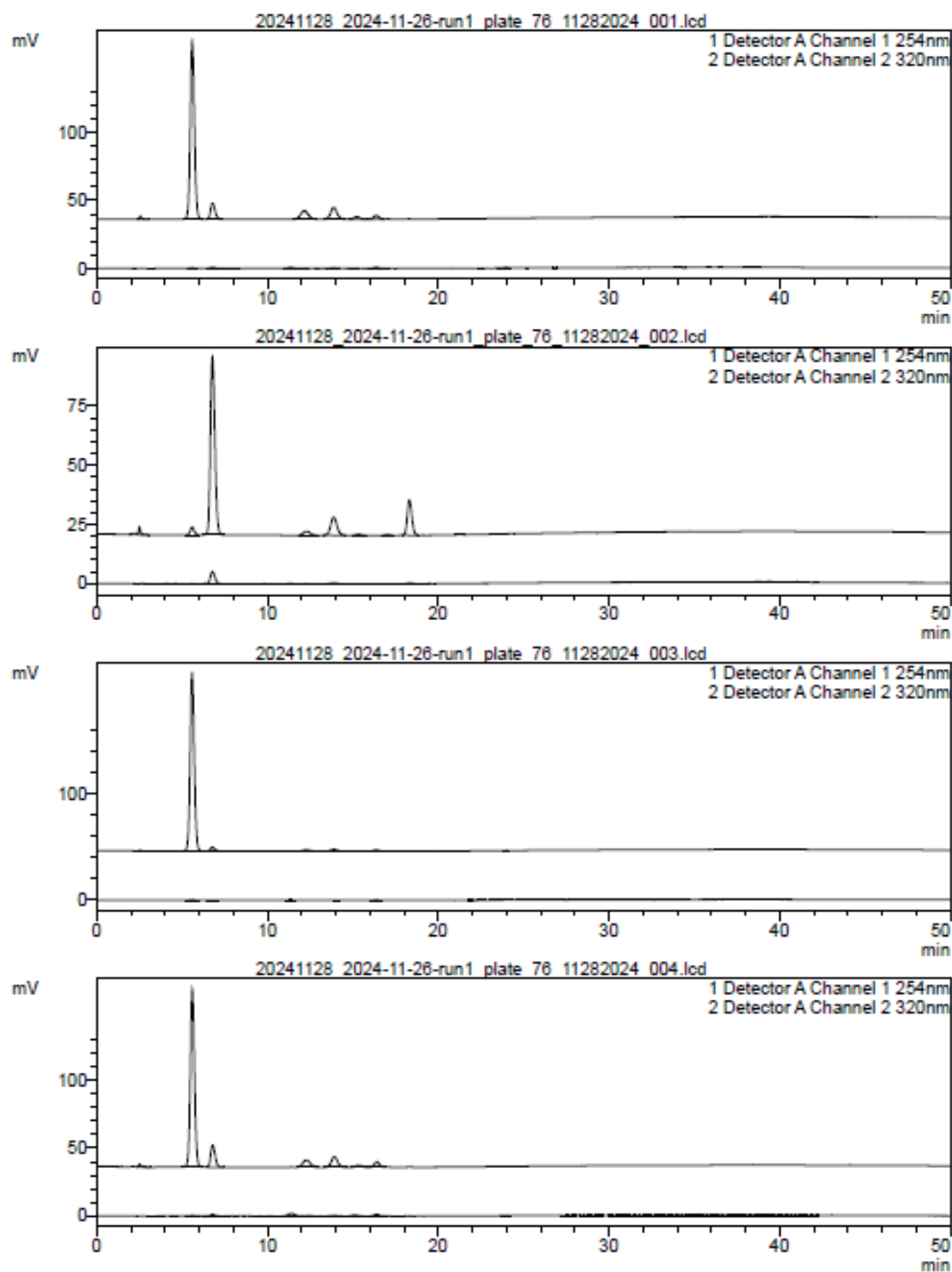
HPLC chromatogram for conditions 55, from Table S13



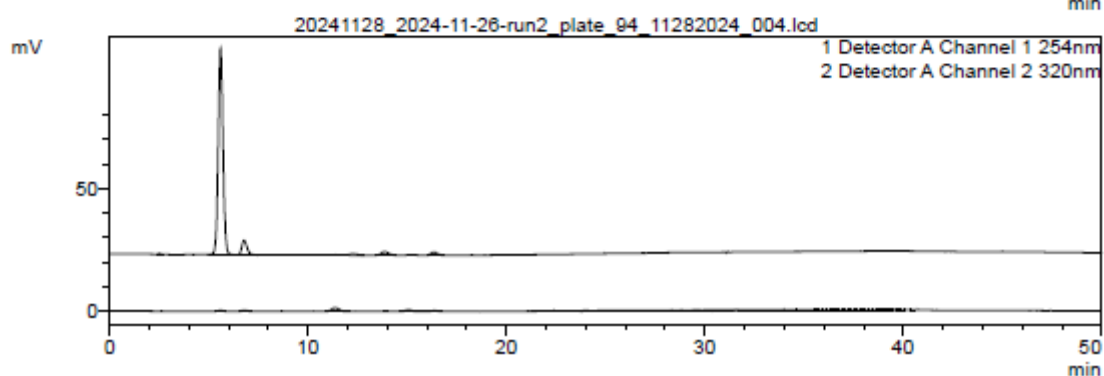
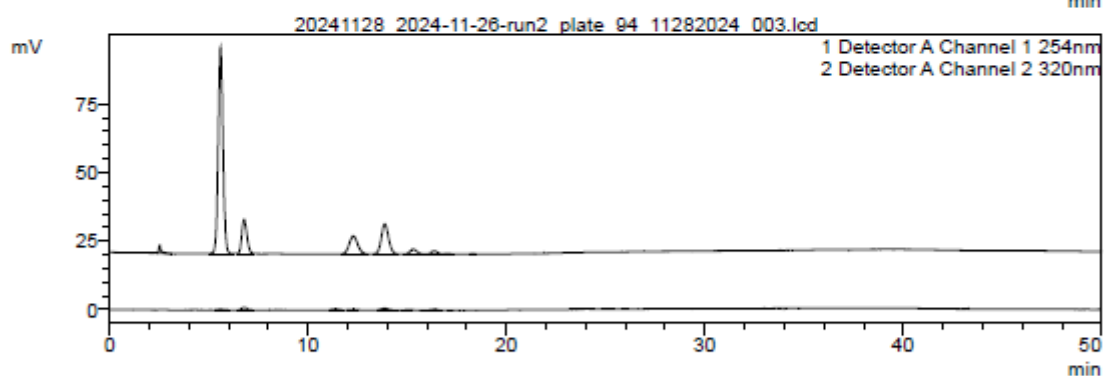
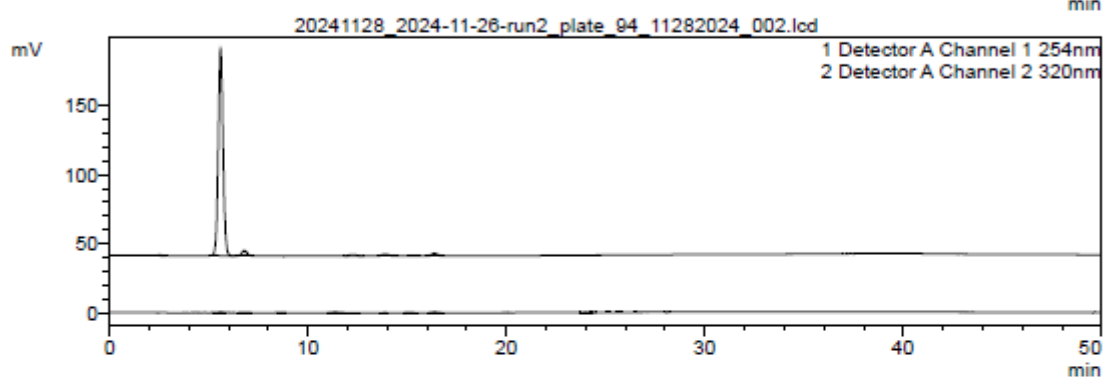
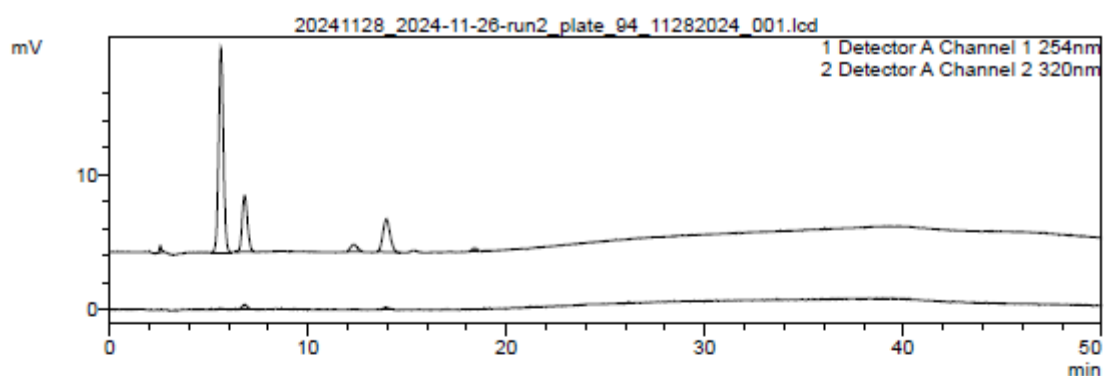
HPLC chromatogram for condition 85, from Table S13



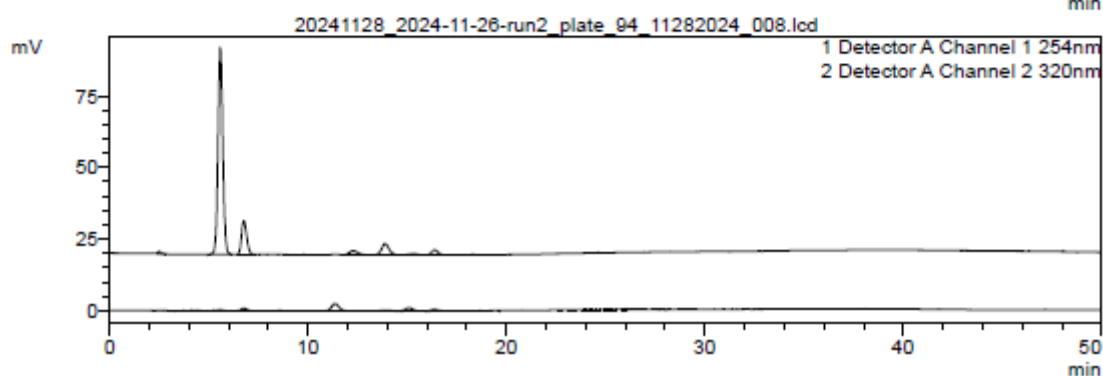
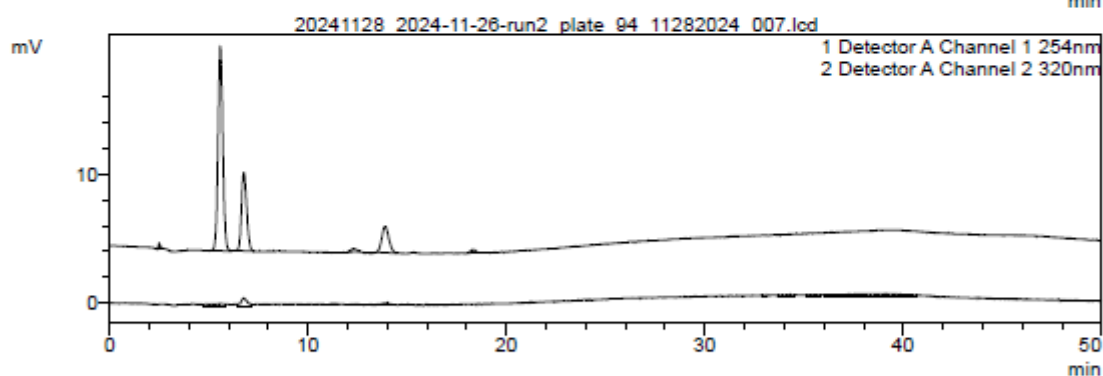
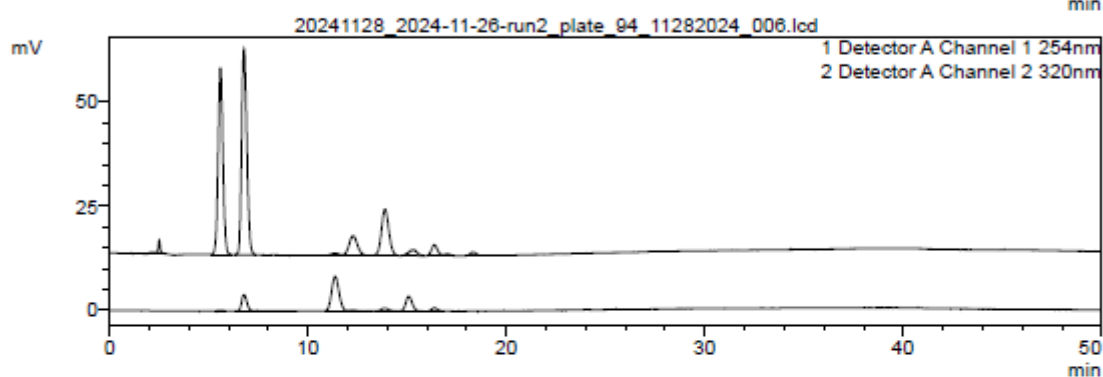
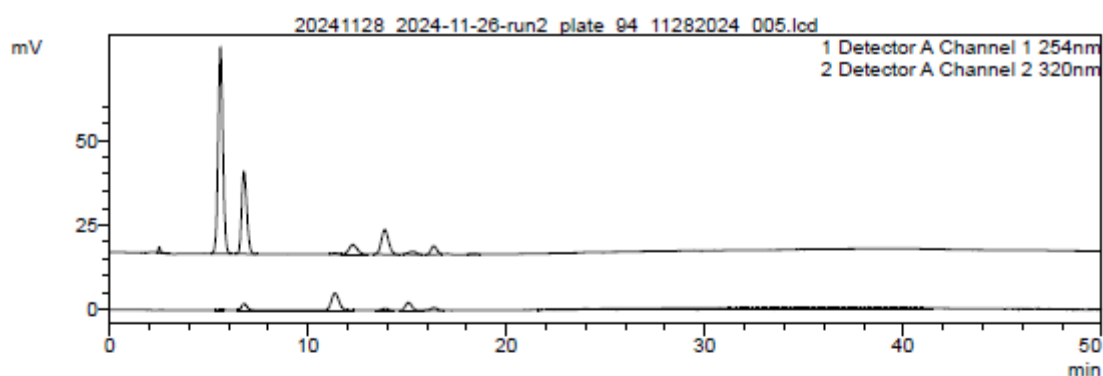
HPLC chromatograms for conditions 88, 13, 22, 67, from Table S13



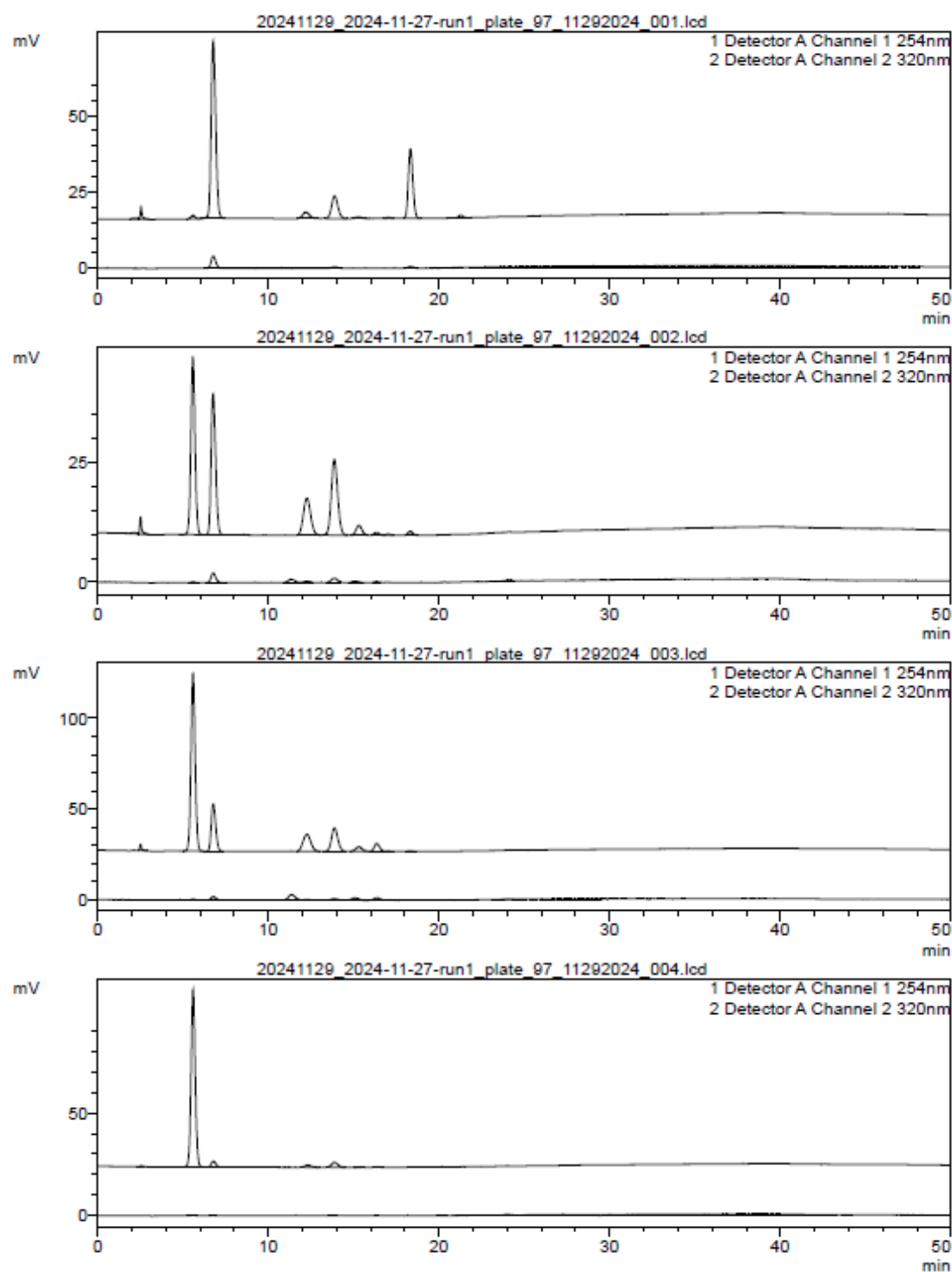
HPLC chromatogram for conditions 19, 58, 28, 37, from Table S13



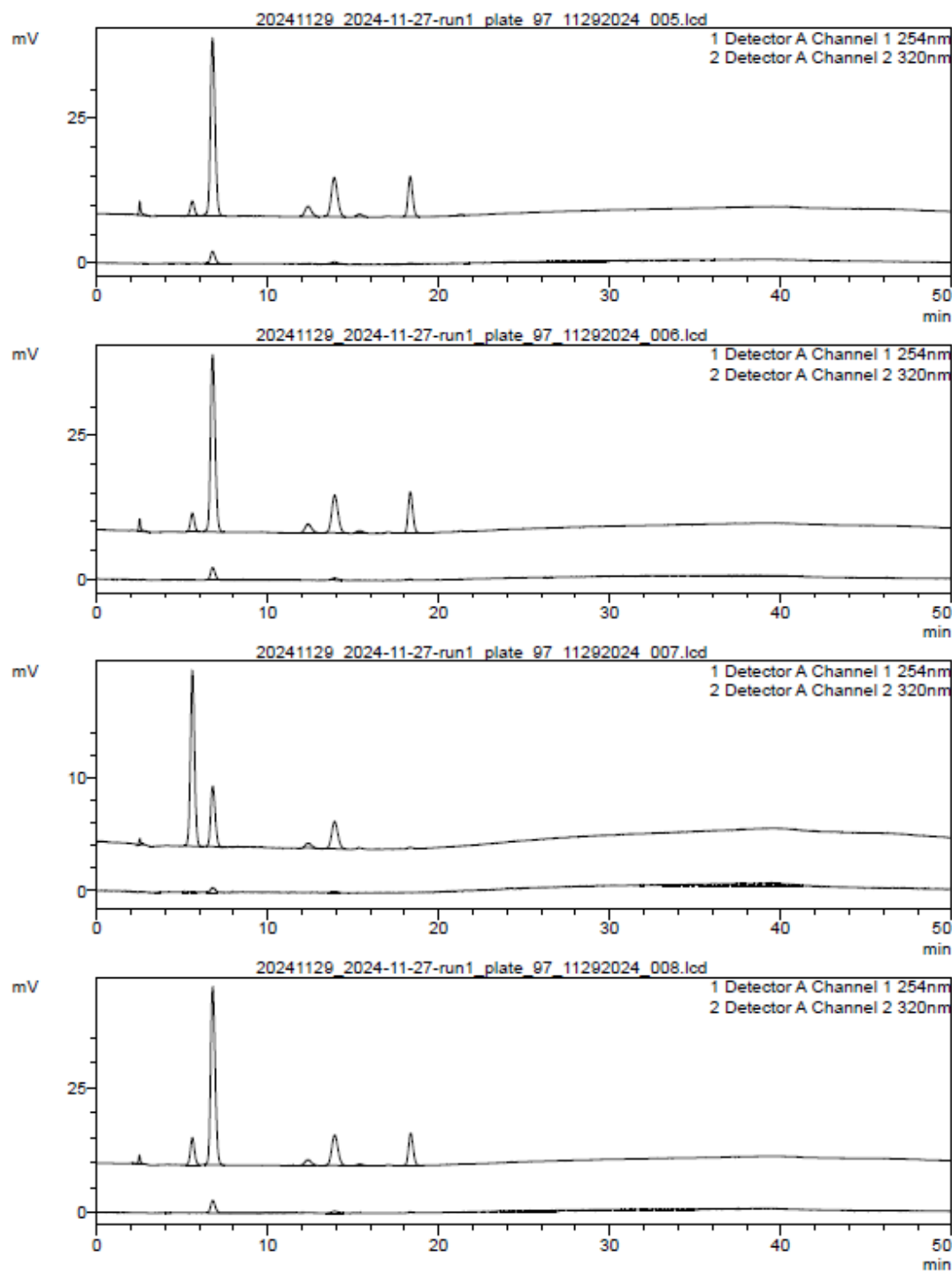
HPLC chromatogram for conditions 43, 49, 1, 7, from Table S13



HPLC chromatogram for conditions 70, 73, 76, 79, from Table S13



HPLC chromatogram for conditions 82, 61, 52, 40, from Table S13



HPLC chromatogram for conditions 22, 67, 88, 13, from Table S13

