Supplementary information

Contents

S1 Batch synthesis

Unless otherwise stated, materials were purchased from commercial sources and used as received without further purification. Automated flash column chromatography was performed using a Teledyne ISCO CombiFlash NextGen 100 system using pre-packed silica columns with a gradient of petroleum ethers/ethyl acetate.

S1.1 3,5-dimethyl-1-phenyl-*1H*-pyrazole (4a)

Acetyl acetone (0.102 g, 1 mmol) and ethanol (5 mL) were added to a 25 mL round bottomed flask containing a stirrer bar with a temperature probe and reflux condenser attached. Phenyl hydrazine (0.109 g, 1 mmol) was added to the reactor and the reaction was heated to 70 °C. After five hours the solvent was evaporated *in vacuo* to afford the *title compound* as an orange oil (0.140 g, 0.813 mmol, 81%).

S1.2 3,5-diethyl-1-phenyl-*1H*-pyrazole (4b)

Heptane-3,5-dione (0.126 g, 1 mmol) and ethanol (5 mL) were added to a 25 mL round bottomed flask containing a stirrer bar with a temperature probe and reflux condenser attached. Phenyl hydrazine (0.109 g, 1 mmol) was added to the reactor and the reaction was heated to 70 °C. After five hours the solvent was evaporated *in vacuo* to afford the *title compound* as an orange oil (0.150 g, 0.714 mmol, 71%).

S1.3 2,8-dimethylnonane-4,6-dione (1d)

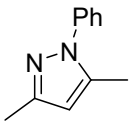
Sodium hydride (60% in mineral oil, 0.4530 g, 11.3 mmol) was loaded into a 10 mL round bottomed flask containing a stirrer bar. The reactor was then filled with nitrogen gas, by use of a Schlenk line, and toluene (2 mL) was added. A mixture of ethyl iso valerate (0.6027 g, 4.63 mmol) and 4-methyl-2-pentanone (0.4530 g, 4.63 mmol) was added via a manual syringe over 5 minutes at room temperature. The reaction was heated to 60 °C and left stirring for 18 hours.

After 18 hours, the reaction mixture was diluted with diethyl ether (10 mL) and quenched with saturated ammonium chloride solution (10 mL) and then 1 M HCl (10 mL). The organic and aqueous layers were separated, and the aqueous layer washed with diethyl ether (2 x 10 mL). The organic layers were combined, washed with water (10 mL) and brine (10 mL) and dried over sodium sulphate. The mixture was then filtered and evaporated under reduced pressure to afford the *title compound* as a yellow oil (0.2992 g, 1.63 mmol, 35%).

S2 Compound characterisation

^{1}H and ^{13}C NMR spectra were recorded on Bruker AVANCE III HD spectrometers (400 MHz and 500 MHz for ^{1}H NMR; 101 MHz and 126 MHz for ^{13}C NMR). ^{1}H NMR chemical shifts (δ_H) and ^{13}C NMR chemical shifts (δ_C) are quoted in parts per million (ppm) downfield of tetramethylsilane (TMS) and reported relative to residual solvent peaks ($CDCl_3$: δ_H = 7.26 ppm and δ_C = 77.16 ppm; CD_3OD: δ_H = 3.31 ppm and δ_C = 49.00 ppm).

S2.1 3,5-dimethyl-1-phenyl-*1H*-pyrazole



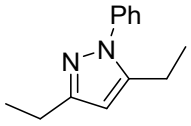
4a

δ_H (400 MHz, $CDCl_3$) 2.23 (3 H, d, J = 0.8 Hz), 2.25 (3 H, s), 5.94 (1 H, s), 7.24–7.31 (1 H, m), 7.37 (2 H, s), 7.38 (2 H, d, J = 0.8 Hz);
δ_H (400 MHz, CD_3OD) 2.24 (3 H, s), 2.26 (3 H, s), 6.07 (1 H, s), 7.35–7.55 (5 H, m);

δ_C (101 MHz, $CDCl_3$) 12.32, 13.47, 106.90, 124.70, 127.19, 128.95, 139.35q, 139.87q, 148.88q;
δ_C (101 MHz, CD_3OD) 12.10, 13.16, 107.77, 126.38, 129.07, 130.29, 140.80q, 141.72q, 150.27q;

HRMS ($APCI^+$) found 173.1076 $C_{11}H_{13}N_2^+$ $[M+H]^+$ requires 173.1073.

S2.2 3,5-diethyl-1-phenyl-*1H*-pyrazole



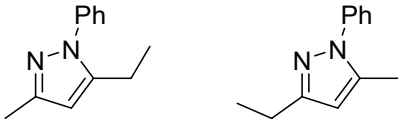
4b

δ_H (400 MHz, CD_3OD) 1.17 (3 H, t, J = 7.6 Hz), 1.26 (3 H, t, J = 7.6 Hz), 2.57-2.68 (4 H, m), 6.14 (1 H, s), 7.37–7.58 (5 H, m);

δ_C (101 MHz, CD_3OD) 13.47, 14.41, 20.47, 22.18, 104.29, 126.88, 129.30, 130.32, 140.91q, 147.95q, 156.39q;

HRMS ($APCI^+$) found 201.1376 $C_{13}H_{17}N_2^+$ $[M+H]^+$ requires 201.1386

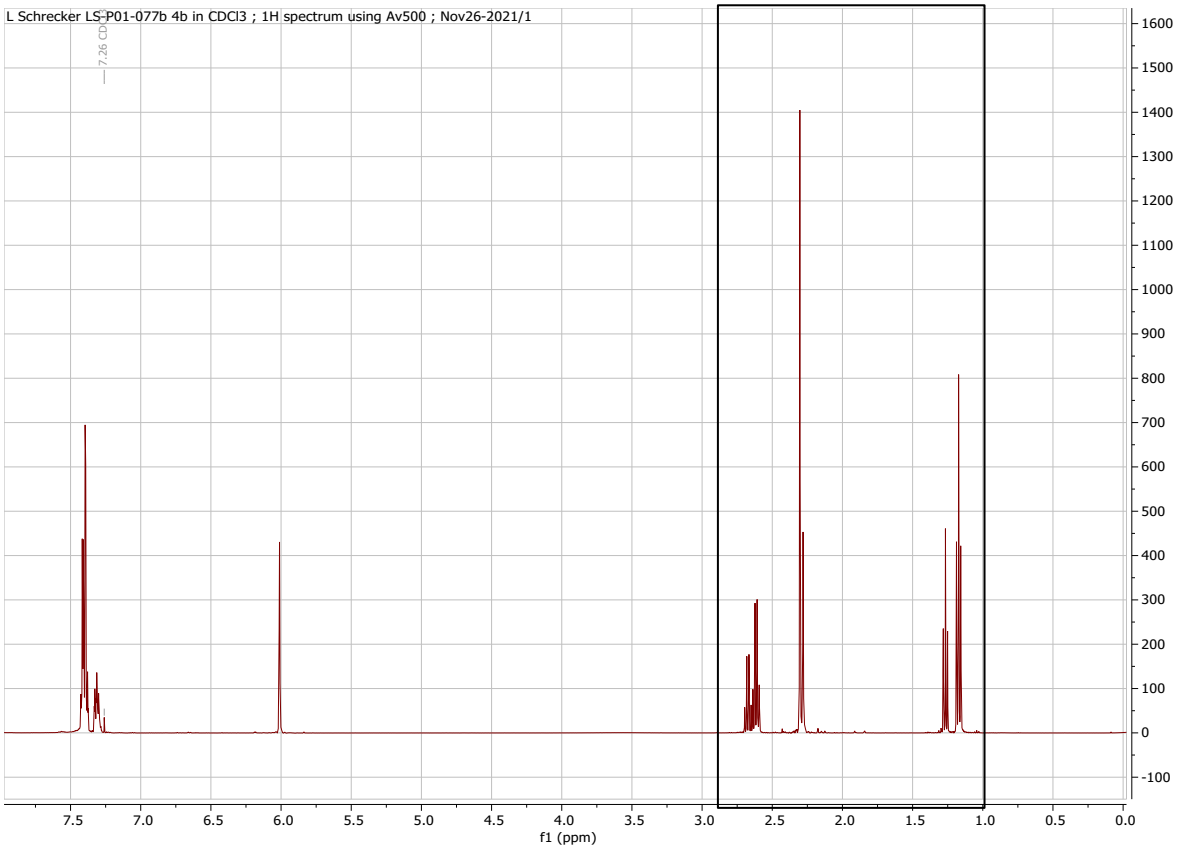
S2.3 Regioisomeric mixture of 5-ethyl-3-methyl-1-phenyl-*1H*-pyrazole (4f) and 5-methyl-3-ethyl-1-phenyl-*1H*-pyrazole (4f')

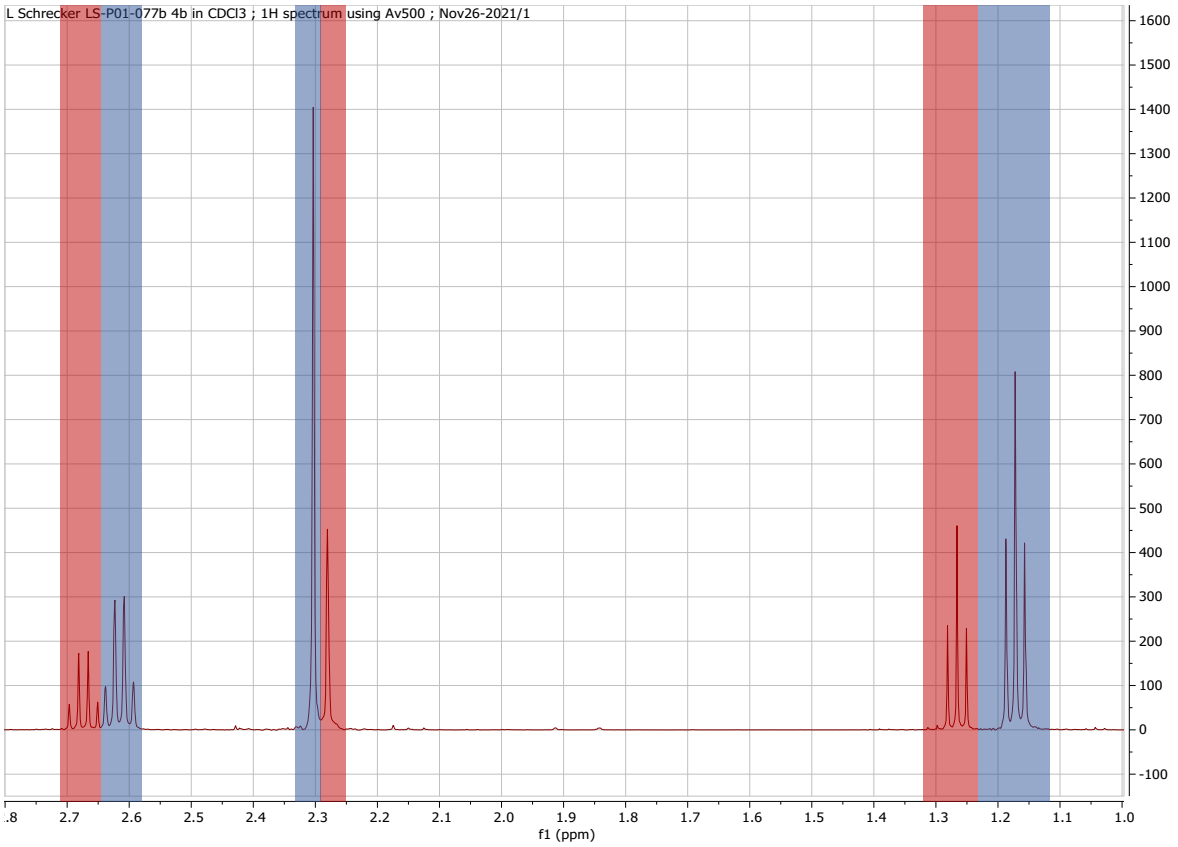


4f **4f'**

Characterised as a mixture: shifts in blue correspond to **4f**, shifts in red correspond to **4f'** and all other peaks correspond to overlapping peaks between both compounds. The corresponding peaks for each regioisomers were identified by relative integrations and assigned by selective nOe exciting at 7.44 ppm. This shift corresponds to a region containing phenyl proton signals of both regioisomers.

δ_H (400 MHz, $CDCl_3$) 1.18 (7.5 H, t, J = 7.6 Hz), 1.28 (3.0 H, t, J = 7.6 Hz), 2.29 (3.1 H, d, J = 0.8 Hz), 2.31 (7.3 H, s), 2.62 (5.0 H, qd, J = 7.6, 0.8 Hz), 2.68 (2.0 H, q, J = 7.6 Hz), 6.02 (3.5 H, s), 6.73-7.48 (17.8 H, m)


L Schrecker LS-P01-077b 4b in CDCl3 ; 1H spectrum using Av500 ; Nov26-2021/1
7.26 CDCl3
f1 (ppm)


L Schrecker LS-P01-077b 4b in CDCl3 ; 1H spectrum using Av500 ; Nov26-2021/1
f1 (ppm)

δ_C (101 MHz, $CDCl_3$) 12.39, 13.13, 13.52, 13.90, 19.68, 21.44, 104.79, 105.32, 112.05, 119.15, 124.69, 125.09, 127.10, 127.34, 128.90, 128.92, 129.07, 139.15, 139.96, 145.83, 148.83, 154.93

HRMS (APCI$^+$) found 187.1220 $C_{12}H_{15}N_2^+$ $[M+H]^+$ requires 187.1230

S3 Flow system setup

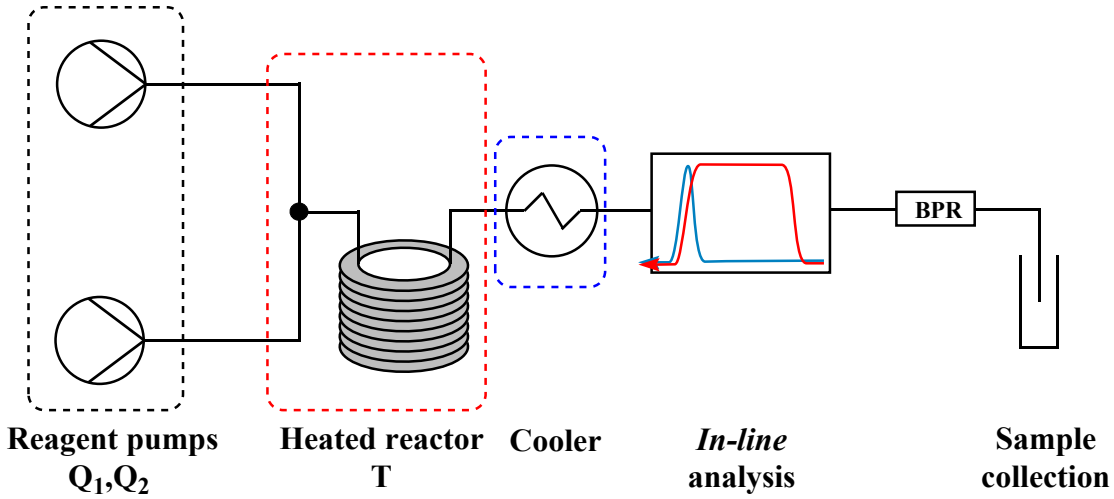


Figure S1: Schematic of the flow system for the control of individual flow rate (Q_1 and Q_2) and temperature (T), with in-line analytics.

S3.1 Flow system overview

A flow system was constructed prioritising the accurate control of parameters needed for transient flow methodology to be developed. The system consisted of reagent and solvent solutions in bottles which contain the inlet tubing with in-line filters attached. The inlet tubing lines were connected to two Gilson 305 HPLC pumps. The outlets of these pumps were connected to two 1 m lengths of stainless steel tubing which fed into a Valco T-piece stainless steel mixer.

The last 0.38 m length of each mixer inlet line and the mixer itself is submerged in an oil bath, heated and stirred by a stirrer hotplate. The outlet of the mixer is connected to a 5.10 m (4.13 mL) tubular reactor constructed from stainless steel tubing also submerged in the oil bath. This is then connected by a stainless steel HPLC style union on the oil bath's surface through an insulating PTFE sheet to a length of stainless steel tubing (14 cm) which passes through a custom built cooling system consisting of an aluminium block and a Peltier assembly.

The tubing is connected to the bottom fitting of the ReactIR flow cell (10 µL) and then a 10 cm length of 1 mm ID PTFE tubing connects this to a 8.1 bar BPR. This is connected at the outlet to a further 15 cm length of 1 mm ID PTFE tubing which can pass into a waste container or into a vial for collection.

S3.2 Pumps

The pumps used are two Gilson 305 HPLC pumps with a 10 mL WSC and 10 mL SC pump heads respectively. These are connected by a GSIOC cable so that the pumps can control each other as a master/slave system through inbuilt Gilson firmware. In order to maintain the pumps

in working order and confirm their accuracy, the check valves are cleaned and sonicated in methanol regularly, and the cumulative flow rate of the system is confirmed at different flow rates and at different times during experimentation.

S3.3 Tubing and fittings

The pump inlet tubing is 1/8" OD, 2 mm ID PTFE tubing with a volume of 1.6 mL and 1.2 mL for pumps A and B respectively. The stainless steel tubing in the system is 1 mm ID, 1/16" OD from Thames Restek UK Ltd. and the T-piece is a VALCO T-piece mixer. The outlet of the back pressure regulator is connected to 1/16" OD, 1 mm ID PTFE tubing.

The fittings to the IR flow cell and connecting the BPR are ¼-28 flat-bottom flangeless ferrules. The inlet and outlet fittings to the Gilson pumps are standard Gilson 30X fittings. All other fittings are HPLC style fittings.

S3.4 Heating and cooling

Heating is performed by an IKA Plate (RCT digital) stirrer hotplate with a silicon oil bath containing a magnetic stirrer bar and with a thermocouple attached for PID feedback control. Cooling is performed by a 60 W Peltier thermo-electric cooler module and heatsink assembly from PiHut with a custom cooling block milled to fit 1/16th inch tubing by the Advanced HackSpace, Imperial College London.

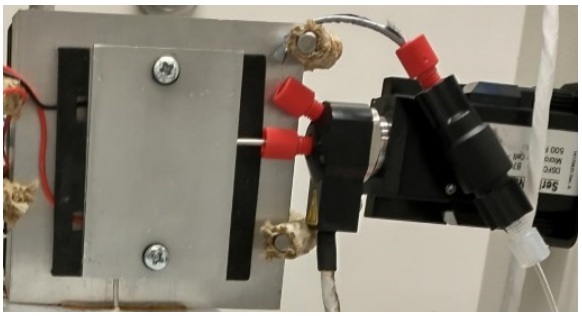


Figure S2: The cooling system used in the flow system consisting of a 60 W Peltier assembly connected to a custom aluminium block containing a cutting which fits to 1/16th inch stainless steel tubing. Insulation is used to cover the outside of the Peltier cooling side in order to improve the efficiency of the device by preventing direct heat transfer back across the heat junction.

S3.5 *In-line* FTIR

The in-line IR spectroscopic data was collected using a Mettler Toledo ReactIR 15 equipped with a 10 µL DS Micro Flow Cell with a DComp (Diamond) probe tip. The spectrometer was set to scan between 3000 cm^{-1} and 650 cm^{-1} using a resolution setting of 8 and a gain setting of 1x.

We employed a 15 second scan rate in order to obtain good time resolution whilst maintain good signal to noise. Diagnostic peaks were identified for pyrazole **4a** and **4b** based on previous batch work and corroborated by individual injections and mixed injections of the reaction components, and by off-line HPLC. The optimal peaks were found in the 2nd derivative of the collected IR spectra as the height to a two point baseline. For **4a**: between 1508-1496 cm^{-1} to a two point baseline at 1514 and 1495 cm^{-1}; for **4b**: between 1508-1502 cm^{-1} to a two point baseline 1516-1495 cm^{-1}. These peaks were picked to allow maximum sensitivity for

concentration of the pyrazole product whilst being independent of effects from other peaks corresponding to the reactants.

Calibration of the relevant peaks was performed at the start of each experimental run by injection of known concentration samples ranging in concentration from 0.236 – 0.028 M (prepared in volumetric flasks via serial dilution of three independent stock solutions) into the IR flow cell with at least four point calibration for each experimental run. Often samples were also injected at the end of an experimental run to confirm that no drift had occurred in the selected peaks. At the end of each experimental run the flow cell was flushed with ethanol and then cleaned with ethanol, water, acetone and petroleum ethers, drying with compressed air between solvents**.**

S3.6 System pressure

The back pressure regulator used was an Upchurch Scientific 250 psi BPR adjusted to 8.1 bar as measured for ethanol on the Vapourtec R series system. This is adequate pressure to keep all solvents used in this work in solution phase for the temperatures used.

S4 Confirmation of plug flow assumption

For transient flow methodology to generate meaningful data, it is important that the flow system is well understood and thus an important step is determining the residence time distribution of the system.

The residence time distribution (RTD) was investigated by pumping an ethanolic solution containing pyrazole **4a** through one pump and ethanol through the other pump. By achieving a steady state at a ratio of flow rates higher than that being investigated and then performing a step change in the ethanolic pump to a lower flow rate, a change in concentration of the pyrazole **4a** could be observed via in-line ReactIR after the given residence time. A sigmoidal change in pyrazole concentration was observed which can be fitted to a dispersion function in order to calculate the Péclet number (Pe) for this system and hence a description of the extent of the deviation of the system from a Plug Flow Reactor (PFR) (Equation S1). [O. Levenspiel, *Chemical Reaction Engineering, 3rd Edition*, 1998.]

[Equation S1]

$$\frac{dF}{dt} = \frac{e^{-\frac{(1-t)^2}{4X}}}{\sqrt{12.566\,X}}$$

where

$$X = \frac{1}{Pe}$$

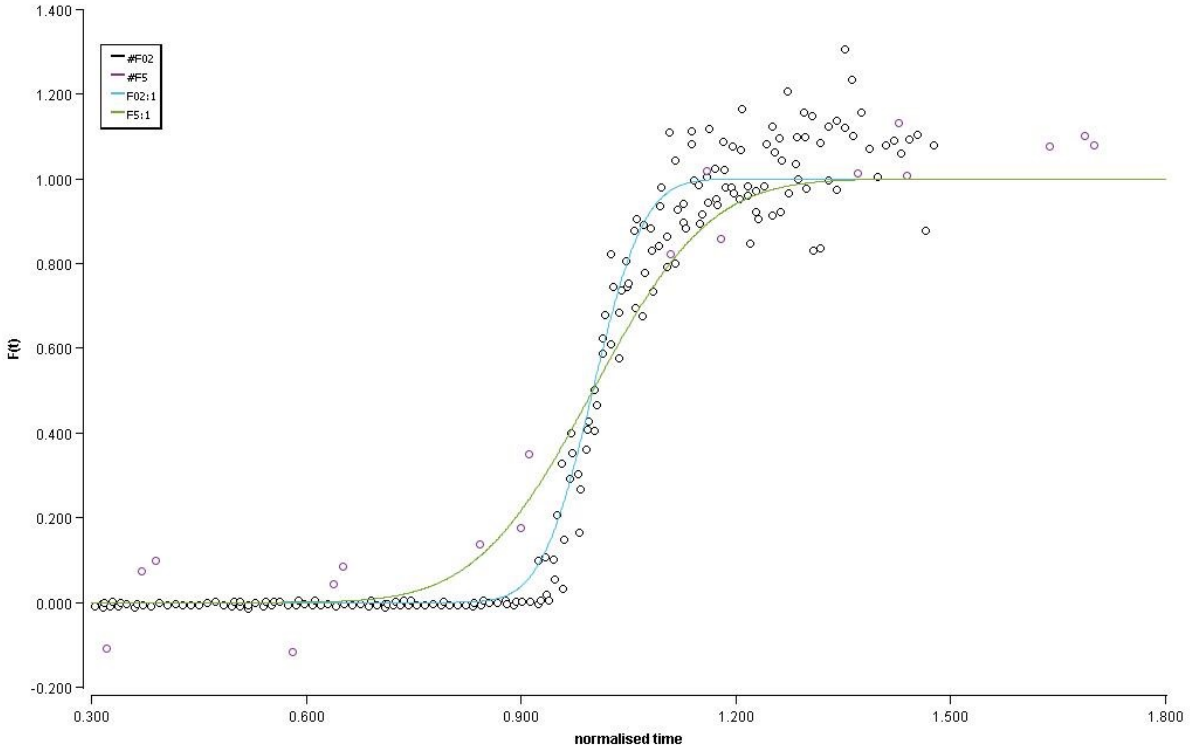


Figure S3: Fitting of Equation S1 in Berkeley Madonna to data used to determine the Péclet number for our system under reaction conditions at our standard initial cumulative flow rate (Q_0, 0.2 mL min^{-1}, #F02) and final cumulative flow rate (Q_{end}, 5 mL min^{-1}, #F5).

This work confirmed that at reaction temperatures the mean residence time was as expected, and the system produces minimal deviations from an ideal PFR at relevant flow rates. The Péclet number was experimentally determined for the minimum and maximum flow rates of our system (0.2 and 5 mL min^{-1}) and was found to satisfy the boundary for minimal deviation from an ideal PFR to be assumed (Pe at 0.2 = 690, Pe at 5 = 129).

S5 General transient flow methods

S5.1 Derivation of deconvolutional maths for step change

Referring to Fig. 1b: a step change of flow rate from an initial cumulative flow rate (Q_0) to a final cumulative flow rate (Q_{end}) performed instantaneously at t = 0 for a system consisting of a reactor of volume V_R followed by a dead volume of volume V_D before an in-line analytical device.

Residence time is equal to the time a differential volume element exits the reactor (t_f) minus the time it enters the reactor (t_i).

[Equation S2]

$$\tau = t_f - t_i$$

Generally the residence time can be determined by integrating the flow rate function over t_f and t_i and equate this to the reactor volume. In the case of a cumulative flow rate step change, the flow rate function has two forms for different regions:

[Equation S3]

$$V_R = \int_{t_i}^{t_f} Q(t)dt = \int_{t_i}^{0} Q_0 dt + \int_{0}^{t_f} Q_{end} dt$$

$$V_R = -Q_0 t_i + Q_{end} t_f$$

Substituting Equation S2 gives:

[Equation S4]

$$\tau = \frac{V_R - (Q_{end} - Q_0)t_f}{Q_0}$$

Due to the dead volume between the exit of the reactor at time t_f and measurement of the volume element by the in-line FTIR at time t_m we must also integrate again:

[Equation S5]

$$V_D = \int_{t_f}^{t_m} Q_{end} dt = Q_{end} t_m - Q_{end} t_f$$

$$t_f = t_m - \frac{V_D}{Q_{end}}$$

Substituting this into Equation S4 gives:

[Equation S6]

$$\tau = \frac{V_R - (Q_{end} - Q_0)(t_m - \frac{V_D}{Q_{end}})}{Q_0}$$

S5.2 General flow experiment setup

The IR was calibrated by the method described in section S3.5. The pumps were then primed with absolute ethanol, the fluidic connection reconnected to the IR flow cell, and the system filled with absolute ethanol. The reactor was then heated to reaction temperature.

A solution of phenyl hydrazine in absolute ethanol and a solution of diketone in absolute ethanol were freshly prepared by weighing out the appropriate amount of the reagents into a volumetric flask and diluting with the requisite volume of absolute ethanol. The solutions were separately both rigorously mixed and transferred into bottles attached to the inlet lines for pump 1 and pump 2 respectively. The pumps were then run at 0.8 mL min^{-1} for two minutes each in order to prime the inlet lines with the reagent solutions. The oil bath is allowed to warm to 70 °C and the Peltier assembly is turned on and allowed to equilibrate.

S5.3 Standard transient residence time ramp experiment (General Method A)

A standard transient residence time ramp is performed by programming the Gilson pumps to

begin at 5 mL min^{-1} at a given %B ratio (flow rate ratio of pumps 1 and 2) for 3 minutes at which time the cumulative flow rate changes to 0.2 mL min^{-1} for a further 21 minutes. This can be directly linked to other transient flow rate experiments. Once all experiments in an experimental run are finished, the system is flushed with ethanol through both pumps for at least 10 reactor volumes.

time / min	flow rate / mL min^{-1}
0	5
2.99	5
3	0.2
24.99	0.2

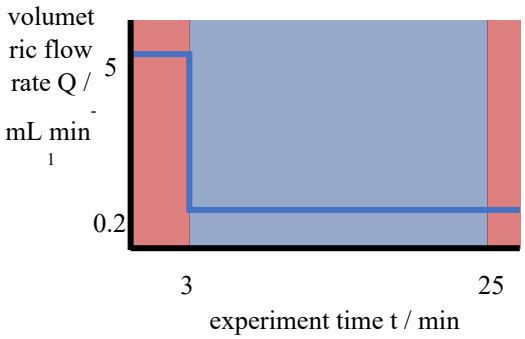


Figure S4: Cumulative volumetric flow rate method for a standard transient residence time ramp experiment (General Method A).

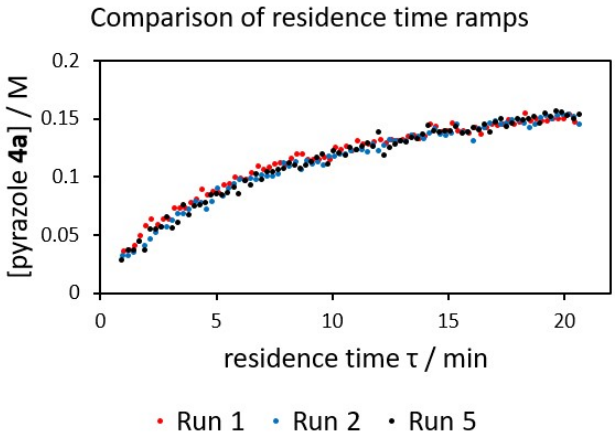


Figure S5: Comparison of the consistency of General Method A performed at 50%B across three experimental runs (experimental run 1, 2 and 5) for the reaction of diketone **1a** and phenyl hydrazine.

S5.4 Standard transient reactant stoichiometry ramp experiment (General Method B)

A standard transient reactant stoichiometry ramp is performed by programming the Gilson pumps to pump at a given cumulative flow rate at a given %B ratio (flow rate ratio of pumps 1 and 2) for 5-10 minutes. The %B is then set to ramp over a given time to the final %B ratio which is then held until a complete reactor volume has exited the reactor.

time / min	flow rate / mL min^{-1}	%B
0	0.2	40

10		40
20		60
50	0.2	60

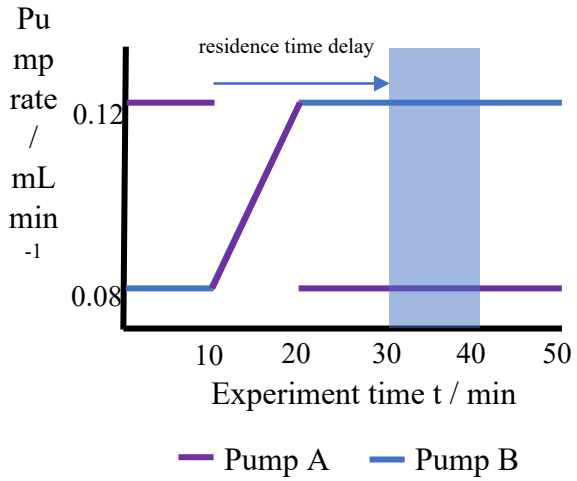


Figure S6: Cumulative volumetric flow rate method for a standard transient reactant stoichiometry ramp experiment (General Method B).

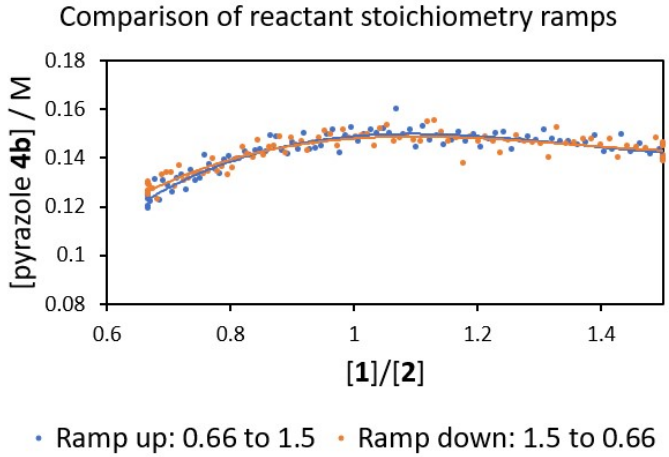


Figure S7: Comparison of the consistency of General Method B when ramping reactant stoichiometry in opposite directions. These experiments were performed at a cumulative flow rate of 0.2 mL min^{-1} in experimental run 6 for the reaction of diketone **1b** and phenyl hydrazine.

As with the transient flow rate methodology, multiple different reactant stoichiometry ramps were performed in order to confirm the robustness of the method and determine the most efficient method.

It was found that concentration ratios of **2**:**1a** from 0.24 M:0.16 M (40%B) to 0.16 M:0.24 M (60%B) could be investigated at the maximum residence time of 20.75 minutes in only 50 minutes of experiment time. The usage of material could also be limited to less than 1.5 reactor volumes of material by using ethanol as a carrying fluid to push the remainder of the reaction out of the reactor.

S5.5 Standard transient bi-variate ramp experiment (General Method C)

A standard transient bi-variate ramp (residence time and reactant stoichiometry) is performed

by programming the Gilson pumps to pump at a cumulative flow rate of 5 mL min^{-1} at a given %B ratio (flow rate ratio of pumps 1 and 2) for 5-10 minutes. The %B is then set to ramp over a given time to the final %B ratio which is then held until a complete reactor volume has exited the reactor.

time / min	flow rate / mL min^{-1}	%B
0	5	40
3	5	40
3.82	5	60
3.83	0.2	60
27	0.2	60

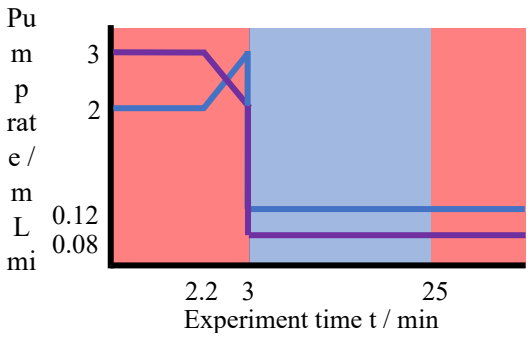


Figure S8: Cumulative volumetric flow rate method for a standard transient bi-variate ramp experiment (General Method C).

Reactor volume is 4.1347 mL (510 cm reactor length) therefore at 5c it takes 0.8269 min to prime. So for combined ramps we need to prime reactor at 5c as normal but at 40%B or 60%B for 3 min and then do a 0.82 min ramp from 40%B to 60%B or 60%B to 40%B which then can be flushed out over the next 23 min at 0.2c. (21.2 min max residence including dead volume).

S6 Combined experimental runs

S6.1 Experimental run 1

General Method A was performed at 50%B using 0.4 M solutions of diketone **1a** and phenyl hydrazine. Then a series of steady state points were collected.

Compound	volume / mL	concentration / M	desired mass / g	actual mass / g	Pump
diketone **1a**	100	0.4	4.004	4.0018	B
phenyl hydrazine	100	0.4	4.328	4.3296	A

General method A.

time / min	flow rate / mLmin-1	%B
0	5	50
2.99	5	

3	0.2	
27	0.2	50

Manual steady state points were collected at 0.2, 0.4, 1, 2, and 5 mL min^{-1} after leaving the system to equilibrate at each of these flow rates.

These data were used for: Fig. 1, repeat comparison (Fig S5)

S6.2 Experimental run 2

Three General Method A runs were performed linked together at 50, 40 and 60%B using 0.4 M solutions of diketone **1a** and phenyl hydrazine. Diluted versions of these 0.4 M solutions were then used for two further General Method A runs to collect same excess data.

Compound	volume / mL	concentration / M	desired mass / g	actual mass / g	Pump
diketone **1a**	50	0.4	2.002	2.0063	B
phenyl hydrazine	50	0.4	2.164	2.1640	A

Three General Method A linked together.

time / min	flow rate / mL min^{-1}	%B
0	5	50
2.99	5	
3	0.2	
24.99	0.2	50
25	5	40
27.99	5	
28	0.2	
49.99	0.2	40
50	5	60
52.99	5	
53	0.2	
74.99	0.2	60

Same excess solutions were made as 14 mL solutions (the required volume for a single General Method A run is 10 mL). They were prepared using a digital pipette to accurately dilute 10 mL of 0.4 M diketone **1a** and phenyl hydrazine stock solutions with 4 mL of ethanol (0.2863 and 0.2859 M respectively). A second set of same excess solution were prepared in the same manner but diluting 8 mL of 0.4 M diketone **1a** and phenyl hydrazine stock solutions with 6 mL of ethanol (0.2290 and 0.2287 M respectively).

General Method A was then run at 50%B for the first set of same excess solutions (0.29 M) and the second set of same excess solutions (0.23 M).

These data were used for: repeat comparison (Fig S5), Fig. 3a, Fig. 6a and b

S6.3 Experimental run 3

A General Method A was performed at 50%B using 0.2 M solutions of diketone **1a** and phenyl hydrazine. A modified version of General Method A was then performed using the same solutions in which tow step changes in cumulative flow rate were performed from 5 to 1 to 0.2 mL min^{-1}. Two more General Method A runs were performed using doped versions of these solution: a run in which 0.2 M pyrazole **1a** was included and a run in which both 0.2 M pyrazole **1a** and 0.4 M water were included.

Doping experiments only (and only two of those).

Compound	volume / mL	concentration / M	desired mass / g	actual mass / g	Pump
diketone **1a**	100	0.2	2.002	2.0028	B
phenyl hydrazine	100	0.2	2.164	2.1648	A

General Method A three separate times.

time / min	flow rate / mLmin-1	%B
0	5	50
2.99	5	
3	0.2	
24.99	0.2	50

Modified General Method A to investigate "incubation period"

time / min	flow rate / mLmin-1	%B
0	5	50
2.99	5	
3	1	
9.99	1	
10	0.2	
32	0.2	50

To 25 mL of the 0.2 M phenyl hydrazine solution 0.8605 g (actually 0.8599 g) of pyrazole **4a** was added to produce a solution of 0.2 M phenyl hydrazine and 0.2 M pyrazole **4a.**

To 25 mL of the 0.2 M diketone **1a** solution 0.182 g of water was added to produce a solution of 0.2 M diketone **1a** and 0.4 M water.

These data were used for: Fig. 3a, Fig. 6a and b

S6.4 Experimental run 4

Three General Method A runs were performed linked together at 50, 40 and 60%B using 0.4 M solutions of diketone **1b** and phenyl hydrazine. Diluted versions of these 0.4 M solutions were then used for two further General Method A runs to collect same excess data.

Compound	volume / mL	concentration / M	desired mass / g	actual mass / g	Pump
diketone **1b**	50	0.4	2.5634	2.5696	B
phenyl hydrazine	50	0.4	2.164	2.1617	A

Three General Method A linked together.

time / min	flow rate / mLmin-1	%B
0	5	50
2.99	5	
3	0.2	
24.99	0.2	50
25	5	40
27.99	5	
28	0.2	
49.99	0.2	40
50	5	60
52.99	5	
53	0.2	
74.99	0.2	60

Same excess solutions were made as 14 mL solutions (the required volume for a single General Method A run is 10 mL). They were prepared using a digital pipette to accurately dilute 10 mL of 0.4 M diketone **1b** and phenyl hydrazine stock solutions with 4 mL of ethanol (0.2864 and 0.2856 M respectively). A second set of same excess solution were prepared in the same manner as 11.55 mL solutions. The second set diluted 6.6 mL of 0.4 M diketone **1b** and phenyl hydrazine stock solutions with 4.95 mL of ethanol respectively (0.2291 and 0.2285 M respectively) giving a similar dilution factor to those in Experimental run 2 for diketone **1a**.

General Method A was then run at 50%B for the first set of same excess solutions (0.29 M) and the second set of same excess solutions (0.23 M).

These data were used for: Fig. 6

S6.5 Experimental run 5

Three General Method A runs were performed linked together at 50, 40 and 60%B using 0.4 M solutions of diketone **1a** and phenyl hydrazine. These were also linked to three General Method B runs performed from 40 to 60%B at cumulative flow rates of 0.2, 0.3 and 0.4 mL min^{-1}. Subsequently two General Method C runs were performed linked together, the first from 40 to 60%B and the second from 60 to 40%B.

Compound	volume / mL	concentration / M	desired mass / g	actual mass / g	Pump
diketone **1a**	100	0.4	4.004	4.0015	B
phenyl hydrazine	100	0.4	4.328	4.3246	A

Three General Method A and three General Method B linked together.

Program	time / min	flow rate / mLmin-1	%B
(3)	0	5	50
	2.99	5	
	3	0.2	

	24.99	0.2	50
	25	5	46
	27.99	5	
	28	0.2	
	49.99	0.2	46
	50	5	54
	52.99	5	
	53	0.2	
	74.99	0.2	54
	75		40
(4)	0	0.2	40
	10		40
	20		60
	49.99	0.2	60
	50	0.3	40
	60		40
	70		60
	94.99	0.3	60
	95	0.4	40
(5)	0	0.4	40
	5		40
	15		60
	30	0.4	60
TOTAL	**200 minutes**	**A: 45.6 mL**	**B: 48.1 mL**

Then two General Method C linked together.

Program	time / min	flow rate / mLmin-1	%B
(6)	0	5	40
	3	5	40
	3.82	5	60
	3.83	0.2	60
	26.99	0.2	60
	27	5	60
	30	5	60
	30.82	5	40
	30.83	0.2	40
	54	0.2	40

These data were used for: repeat comparison (Fig S5), Fig. 3d, Fig. 6

S6.6 Experimental run 6

A General Method A was performed at 50%B using 0.4 M solutions of diketone **1b** and phenyl hydrazine. Then were two General Method B runs linked together were performed at a cumulative flow rate of 0.2 mL min^{-1} from 40 to 60%B and 60 to 40%B respectively. Another General Method B run was then performed at 0.4 mL min^{-1} from 40 to 60%B.

Compound	volume / mL	concentration / M	desired mass / g	actual mass / g	Pump
diketone **1b**	25	0.4	1.2817	1.2811	B
phenyl hydrazine	25	0.4	1.082	1.0797	A

General Method A once, then two General Method B linked at 0.2 mL min^{-1}.

time / min	flow rate / mLmin-1	%B
0	0.2	40
10		40
30		60
60		60
80		40
110	0.2	40

Then General Method B at 0.4 mL min^{-1}.

time / min	flow rate / mLmin-1	%B
0	0.4	40
5		40
15		60
30	0.4	60

These data were used for: Fig. 6

S7 HPLC-MS methods and experiments

LC-MS analysis was performed using an Agilent 1260 Infinity II instrument, a diode array UV-Vis Detector and an InfinityLab LC/MSD mass spectrometer performing positive ES-API MS.

S7.1 Initial intermediate identification method for R = Me

An initial vial reaction, performed using 6 equivalents of phenyl hydrazine and 1 equivalent of diketone **1a** neat, and subsequent samples taken from various reactant stoichiometry ramps and residence time ramps for diketone **1a** were analysed using an Agilent Zorbax SB-C18 (50 mm x 4.6 mm x 3.5 µm) column at 40 °C with acetonitrile (solvent A) and 20 mM aqueous ammonium formate (solvent B). Beginning at 60%B for 0.5 min, followed by a 2.5 min ramp to 95%B, then a 2 minute isocratic hold before returning to the initial conditions.

This method clearly showed separate peaks with m/z values corresponding to di-addition intermediate **5a** (3.181 min) and pyrazole **4a** (2.077 min) in the 6 equivalent neat excess experiment:

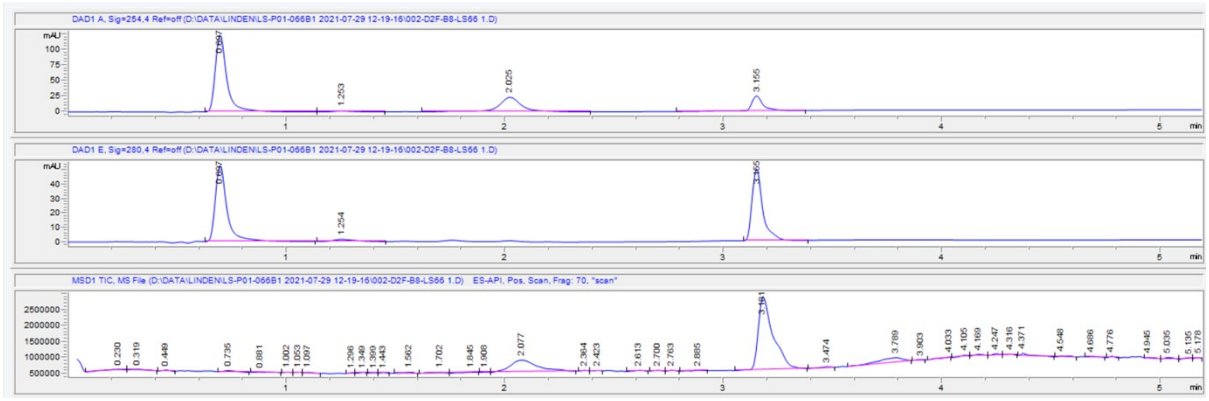


Figure S9: Separation and identification of di-addition intermediate **5a** using LC-MS on a neat reaction at room temperature between 1 equivalent of diketone **1a** and 6 equivalents of phenyl hydrazine. The figure consists of the diode array trace at 254 nm and 280 nm, and the TIC (total ion chromatogram) from the mass spectrometer.

In subsequent experimental runs without such extreme excesses, peaks with m/z values corresponding to mono-addition intermediate **3a** (1.735 min), di-addition intermediate **5a** (2.087 min) and pyrazole **4a** (3.215 min) could be clearly identified. For example, at the end of transient ramp at 0.2 mL min^{-1} and 60%B:

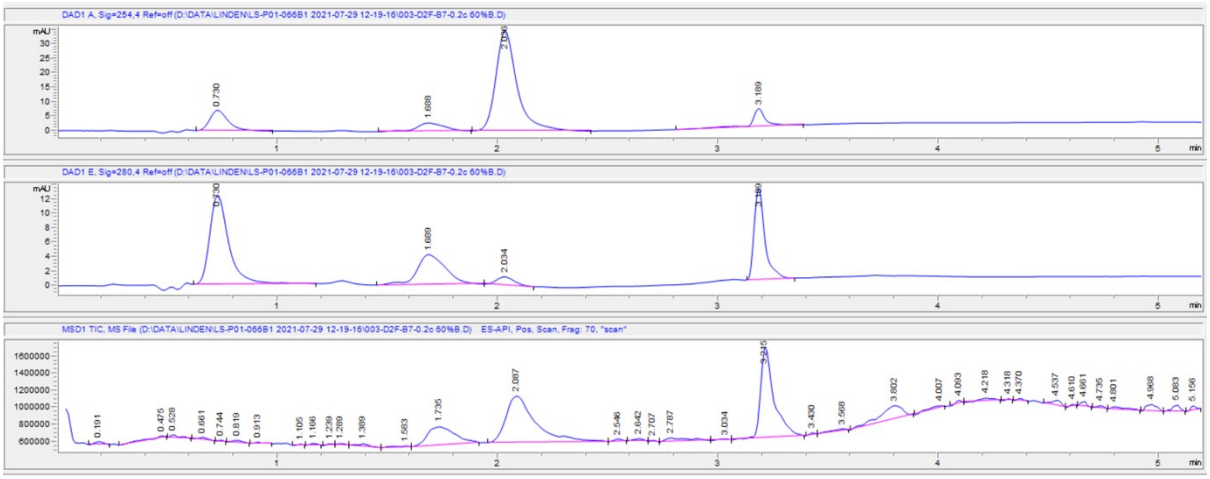


Figure S10: Separation and identification of mono-addition intermediate **3a** and di-addition intermediate **5a** using LC-MS used in a variety of experimental runs for the reaction of diketone **1a** and phenyl hydrazine. The figure consists of the diode array trace at 254 nm and 280 nm, and the TIC from the mass spectrometer.

S7.2 Method for identification of unsymmetric diketone reaction intermediates

The reactions performed using a 0.4 M stock solution of hexane-2,4-dione **1f** and 0.4 M stock solution of phenyl hydrazine at 70 °C and a cumulative flow rate of 0.2 mL min^{-1} at 40 and 60%B respectively were analysed using an Agilent Zorbax SB-C18 (50 mm x 4.6 mm x 3.5 µm) column at 40 °C with acetonitrile (solvent A) and 20 mM aqueous ammonium formate (solvent B). Beginning at 50%B for 0.5 min, followed by a 3 min ramp to 65%B, then a 0.5 minute isocratic hold, followed by a 1 min ramp to 95%B, then a 0.5 min isocratic hold before returning to the initial conditions.

This method allowed separation of peaks with m/z values corresponding to mono-addition intermediate **3f** (1.298 min), di-addition intermediate **5f** (3.760 min) and either or both pyrazole products **4f**/**4f'** as one peak (1.496 min) despite further attempts to separate the regioisomeric pyrazoles using LC-MS.

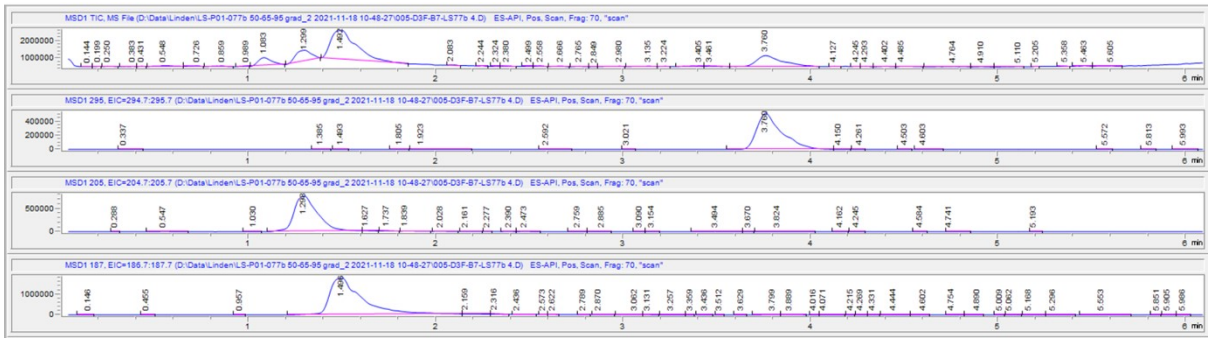


Figure S11: Separation and identification of mono-addition intermediate **3f** and di-addition intermediate **5f** using LC-MS while attempting to separate regioisomeric pyrazoles **4f** and **4f'** for the reaction of diketone **1f** and phenyl hydrazine. The figure consists of the TIC from the mass spectrometer, followed by the EIC (extracted ion chromatogram) for the di-addition intermediate **5f** (m/z = 295.2), mono-addition intermediate **3f** (m/z = 205.2) and pyrazole products **4f**/**4f'** (m/z = 187.2).

After removing the solvent from the solvent from the reactions using a SmartEvaporator, it was found the extract ion chromatograms for intermediates **3f** and **5f** fell to nothing:

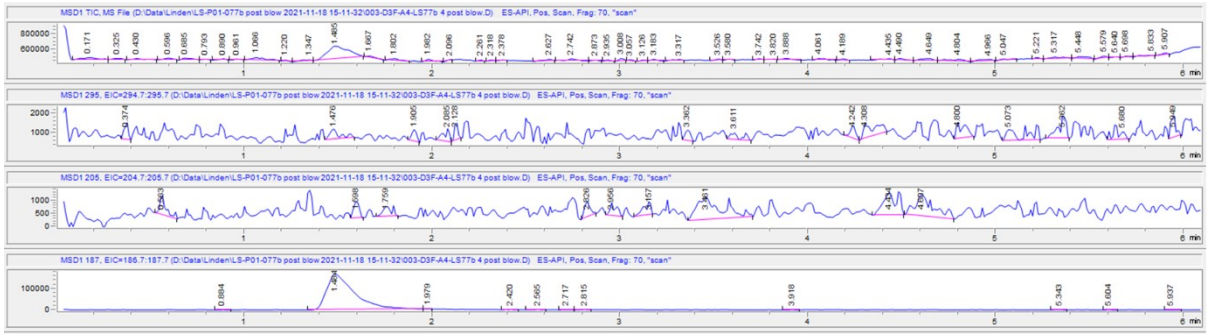


Figure S12: Separation and identification of mono-addition intermediate **3f** and di-addition intermediate **5f** using LC-MS while attempting to separate regioisomeric pyrazoles **4f** and **4f'** for the reaction of diketone **1f** and phenyl hydrazine analysed after concentration by removal of reactant solvent *in vacuo*. The figure consists of the TIC from the mass spectrometer, followed by the EIC (extracted ion chromatogram) for the di-addition intermediate **5f** (m/z = 295.2), mono-addition intermediate **3f** (m/z = 205.2) and pyrazole products **4f**/**4f'** (m/z = 187.2).

S7.3 Intermediate identification method for range of alkyl substituents

Vial reactions performed at 0.2 M in ethanol at room temperature between phenyl hydrazine and a range of diketones sampled at 10 minutes were analysed using a Phenomenex Gemini NX-C18 (50 mm x 2.0 mm x 3.0 µm) column at 40 °C with acetonitrile (solvent A) and 20 mM aqueous ammonium formate (solvent B). Beginning at 65%B for 0.5 min, followed by a 3 min ramp to 95%B, then a 2 minute isocratic hold before returning to the initial conditions.

This method allowed identification of the presence of separate peaks with m/z values corresponding to mono-addition intermediate **3**, di-addition intermediate **5** and pyrazole **4** across the range of alkyl substituents. All peaks were present in R =Me, Et and ^{i}Bu (**a**,**b**,**d**), peaks for mono-addition intermediate **3** and pyrazole **4** were present when R = ^{i}Pr (**c**) and no peaks were identified corresponding to any intermediates or products when R = ^{t}Bu (**e**) as identified by extracted ion chromatograms.

S8 NMR experiments

S8.1 Reaction monitoring for intermediates

Initiation of the reaction in an NMR tube at 0.2 M concentration of phenyl hydrazine and diketone **1a** or **1b** in CD_3OD at room temperature revealed peaks not accounted for by the

starting diketone, phenyl hydrazine or the respective pyrazole product. NMR experiments were performed within 5 minutes of the reactive sample being prepared.

For the reaction of diketone **1a** clear additional peaks in the alkyl region could be observed in both the ^{1}H and ^{13}C NMR. The number of new peaks requires multiple intermediates to be present, however identification of individual intermediates proved challenging due to their low concentration and transient nature.

For the reaction of diketone **1b** clear additional peaks in the alkyl region were present in ^{13}C NMR and further corroborated by DEPT-135 NMR experiments. These experiments allowed the identification of at least four new CH_2 environments. This suggests once again that multiple intermediates are present in this reaction mixture.

For both the reaction of diketone **1a** and **1b** all intermediate peaks disappeared when the sample was left to stand overnight.

R = Me

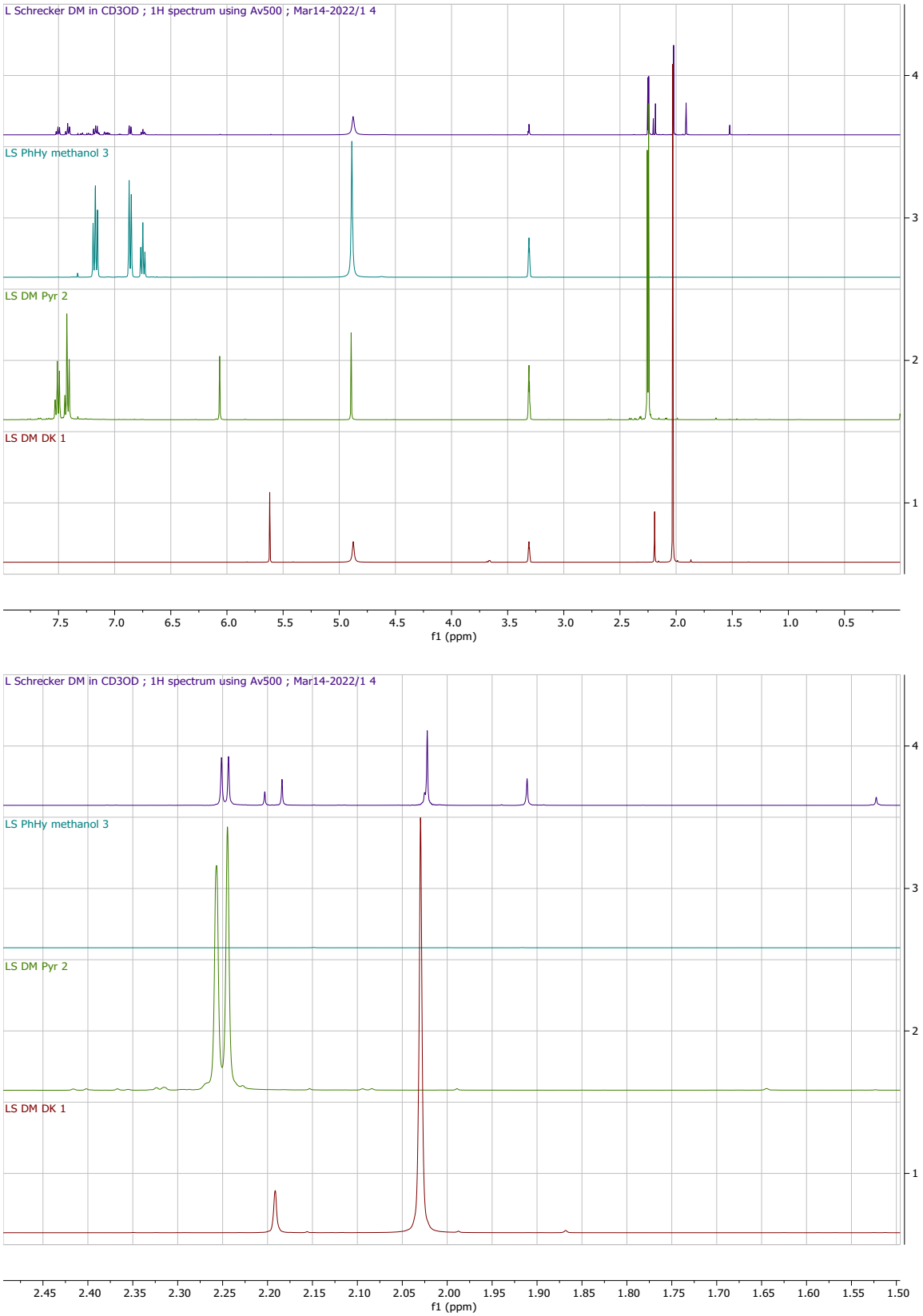


Figure S13: ^{1}H NMR spectra taken under 5 minutes from the initiation of the reaction of diketone **1a** and phenyl hydrazine in CD_3OD with comparative spectra for phenyl hydrazine, pyrazole **4a** and diketone **1a**. Clear intermediate peak formation can be seen, especially in the alkyl region.

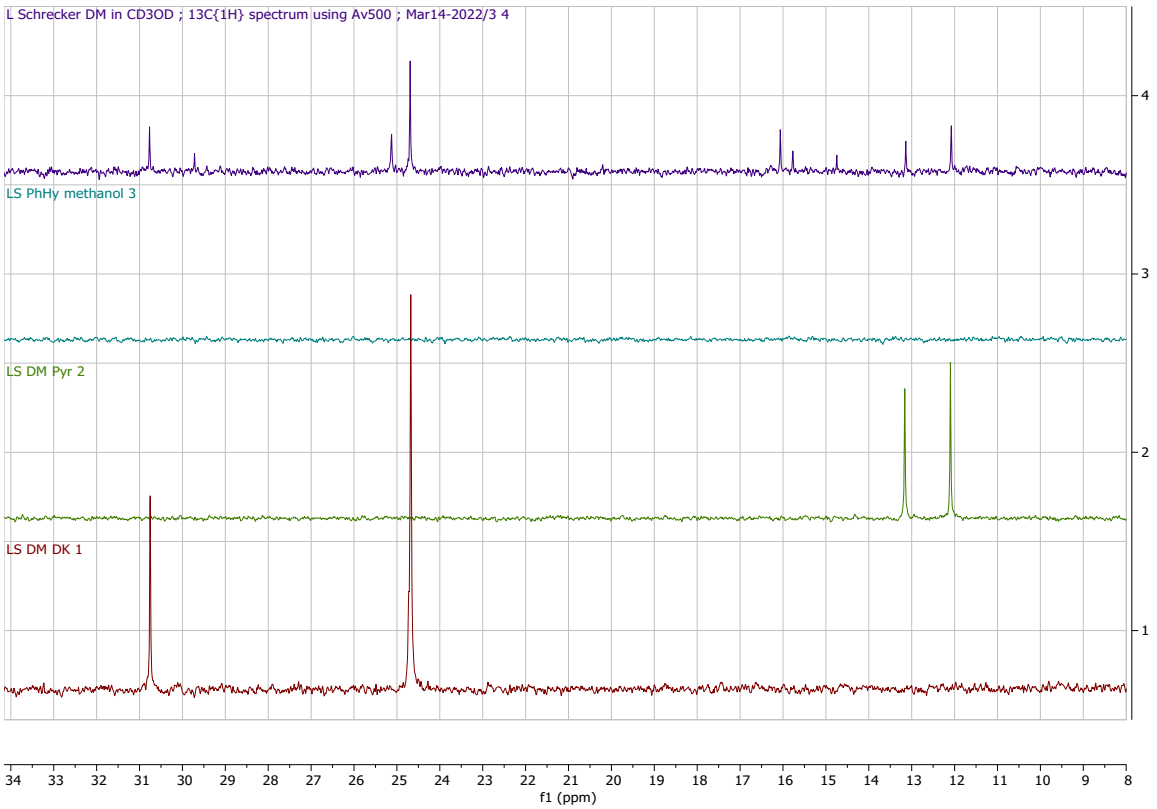


Figure S14: ^{13}C NMR spectra taken under 5 minutes from the initiation of the reaction of diketone **1a** and phenyl hydrazine in CD_3OD with comparative spectra for phenyl hydrazine, pyrazole **4a** and diketone **1a**. Clear intermediate peak formation can be seen, especially in the alkyl region.

R = Et

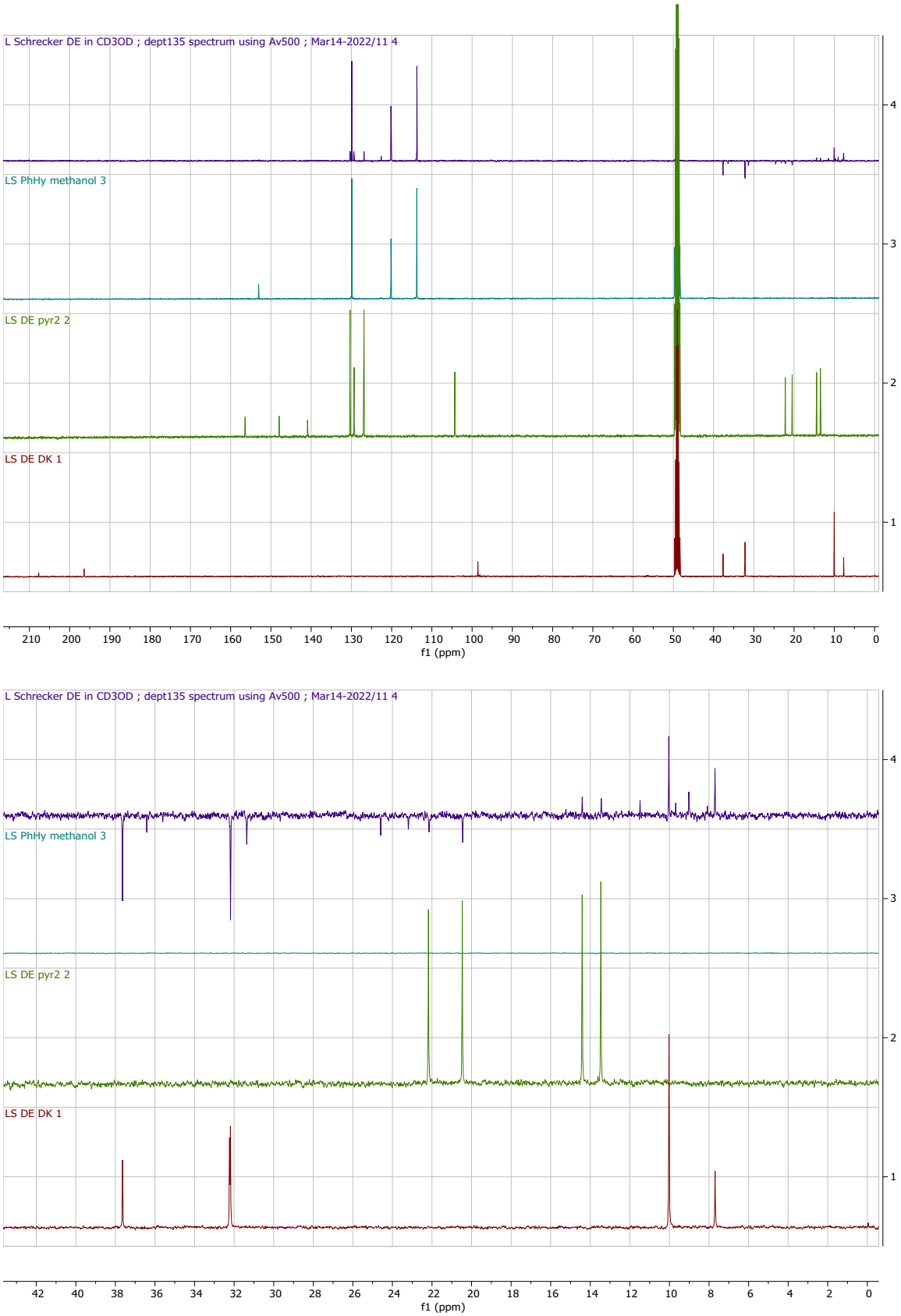


Figure S15: DEPT-135 NMR spectra taken under 5 minutes from the initiation of the reaction of diketone **1b** and phenyl hydrazine in CD_3OD with comparative ^{13}C NMR spectra for phenyl hydrazine, pyrazole **4a** and diketone **1a**. Clear intermediate peak formation can be seen, especially in the alkyl region.

Four new CH_2 peaks present as compared to diketone **1b**, pyrazole **4b** and phenyl hydrazine. This suggests multiple intermediates.

S8.2 Lack of interaction between diketone 1a and pyrazole 4a

^{1}H NMR spectra taken of diketone **1a**, pyrazole **4a**, and a combination of both together in CD_3OD showed no interaction was happening between the two species in solution and thus that both a tandem acid-base catalysed mechanism using a diketone and a pyrazole was plausible.

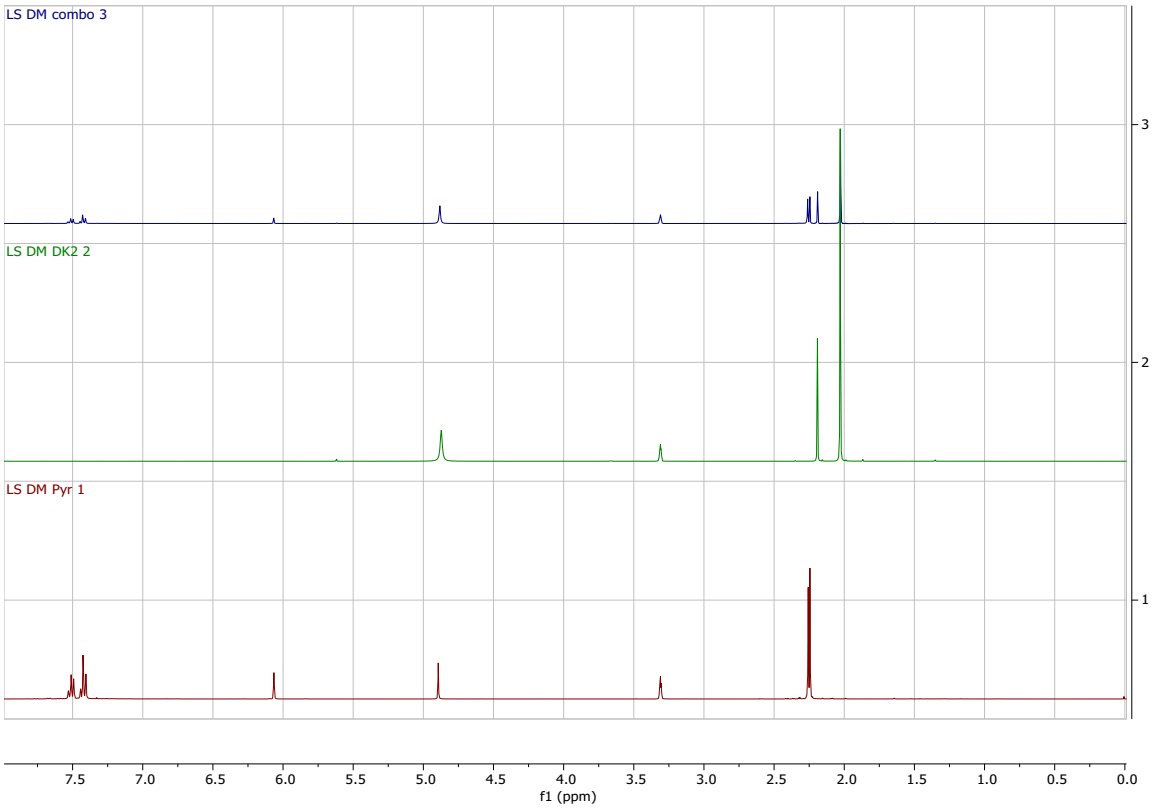


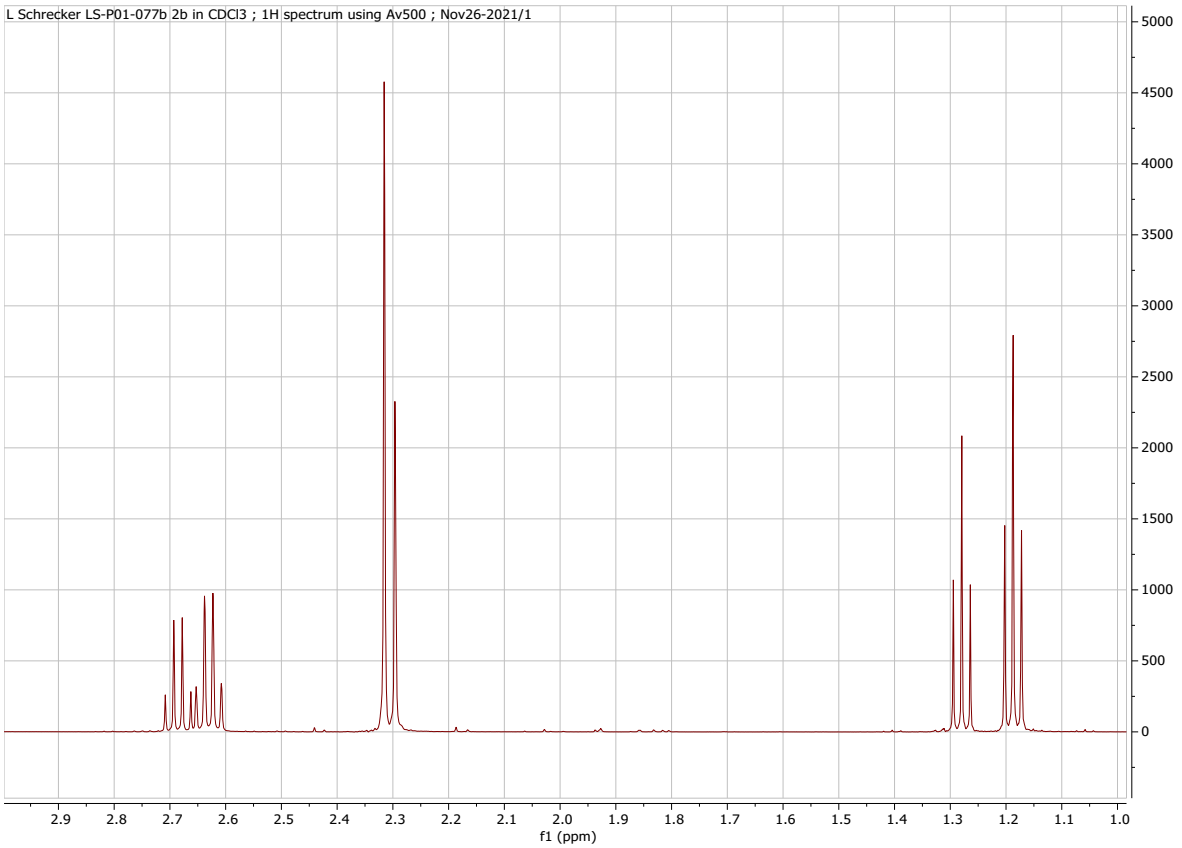
Figure S16: ^{1}H NMR spectra of a combination of diketone **1a** and pyrazole **4a** in CD_3OD with comparative spectra for diketone **1a** and pyrazole **4a** all taken after the samples were allowed to equilibrate in CD_3OD.

S8.3 Major regioisomer identification

^{1}H NMR was performed on the regioisomeric mixtures of pyrazoles **4f** and **4f'** produced from the different excess reactions of phenyl hydrazine with unsymmetric diketone **1f** in excesses of 1.5:1 equivalents. These experiments produced different ratios of pyrazoles **4f** and **4f'** but did not allow us to identify which isomer was which. The shift of the phenyl protons in the pyrazole products were then be used as the centre for excitation for selective excitation nOe. Excitation at 7.44 ppm revealed enhancement of the same signals in both samples as expected and only one of each signal for the methyl and ethyl signals as desired. The enhanced peaks are those which are closest to the excited phenyl protons and therefore correspond to the alkyl group which sits on the same side as the phenyl ring. Enhancement of the neighbouring ethyl CH_2 environment in pyrazole **4f** was 0.7%, with the corresponding ethyl CH_3 enhanced by 0.1% and that of the neighbouring CH_3 in pyrazole **4f'** was 0.5%. As a comparison, enhancement of the 4 position proton on the pyrazole ring was 0.1% and all other environments from the alkyl groups the opposite side of the pyrazole ring were less than 0.1%.

Excess of phenyl hydrazine 1.5 equivalents:

^{1}H NMR



Selective excitation nOe NMR

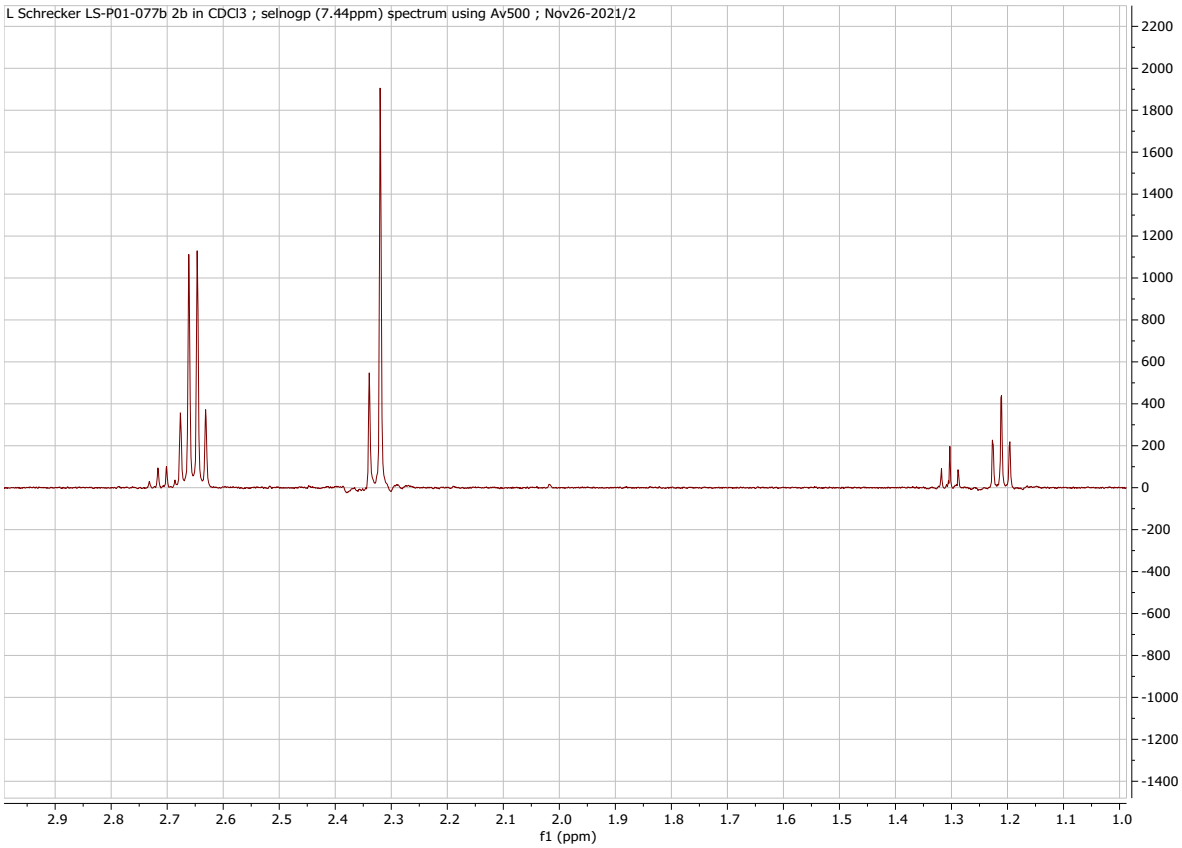
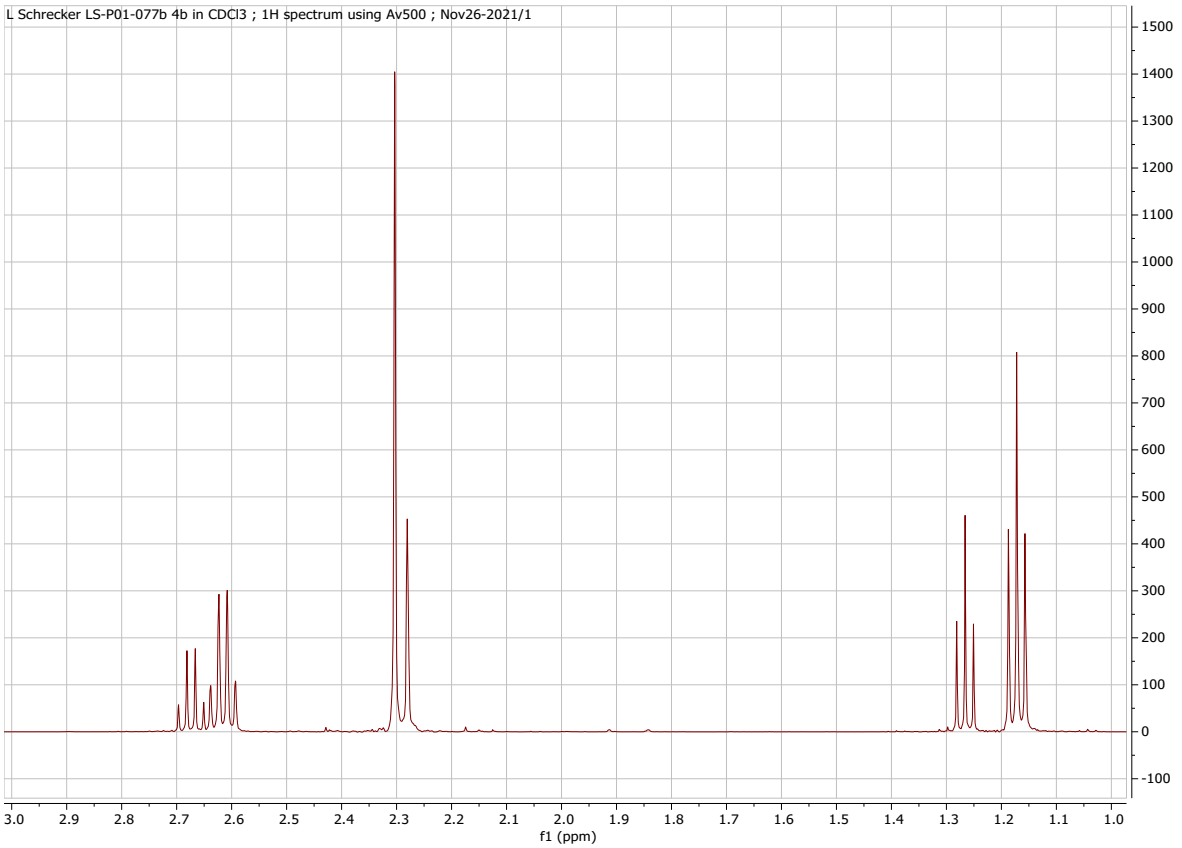


Figure S17: ^{1}H NMR and selective excitation nOe spectra in $CDCl_3$ of the isolated product mixture of the reaction of 1 equivalent of diketone **1f** and 1.5 equivalents of phenyl hydrazine.

Excess of diketone 1f 1.5 equivalents:

^{1}H NMR



Selective excitation nOe

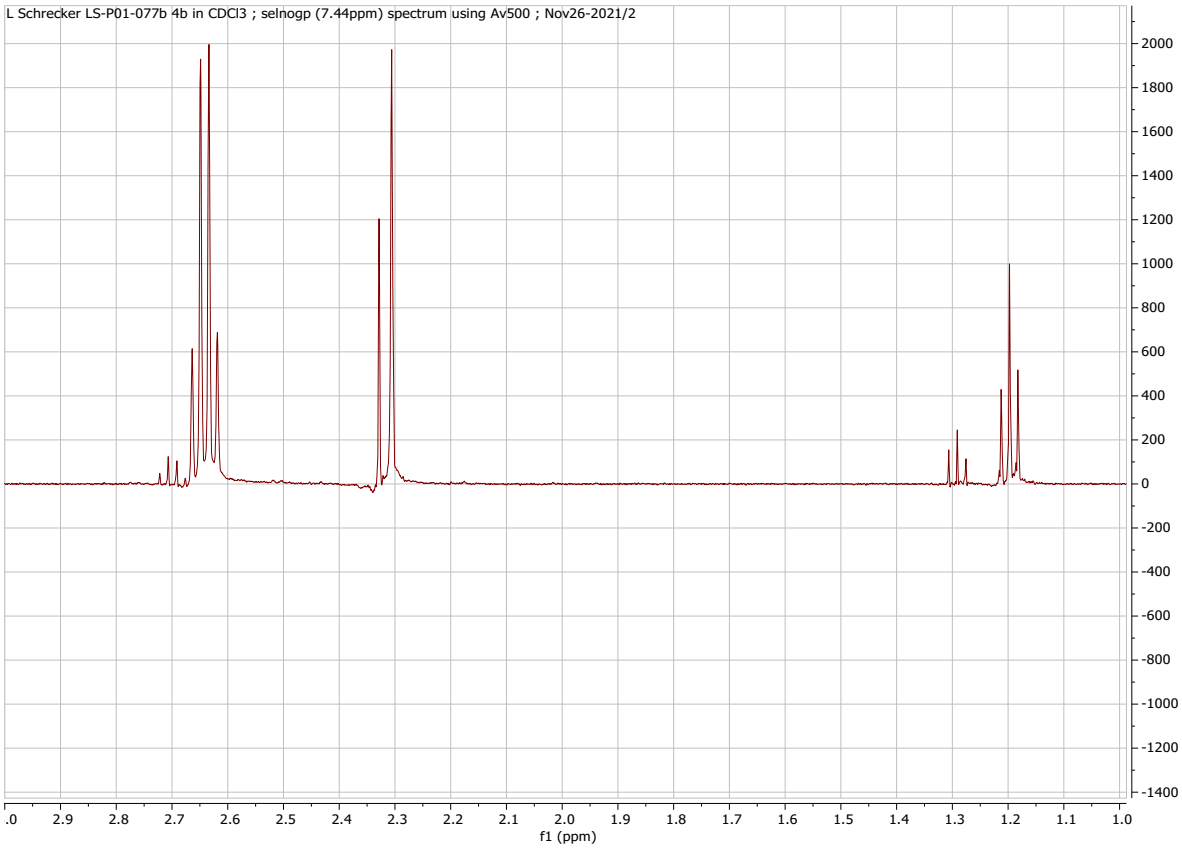
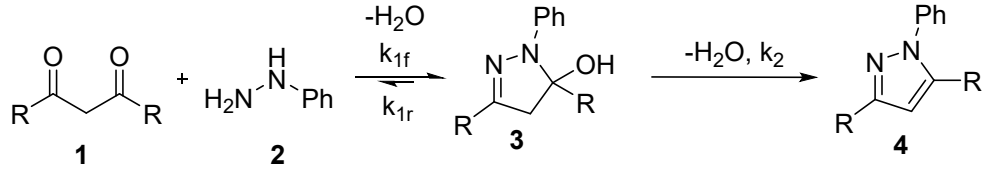


Figure S18: ^{1}H NMR and selective excitation nOe spectra in $CDCl_3$ of the isolated product mixture of the reaction of 1.5 equivalents of diketone **1f** and 1 equivalent of phenyl hydrazine.

S9 Berkeley Madonna modelling

The ODE solver Berkeley Madonna was used to construct microkinetic models which treated rate constants as fitting values to be optimised when fitting to data collected from experiments A-G for diketone **1a** or A-E for diketone **1b** as outlined in Table 2. Fitting was achieved using the multi-curve fitting feature primarily using the solver Runge-Kutta 4, but transferring to the solver Rosenbrock (stiff) once the model complexity increased due to the better performance of this solver on highly coupled series of ODEs as was the case in our microkinetic model.

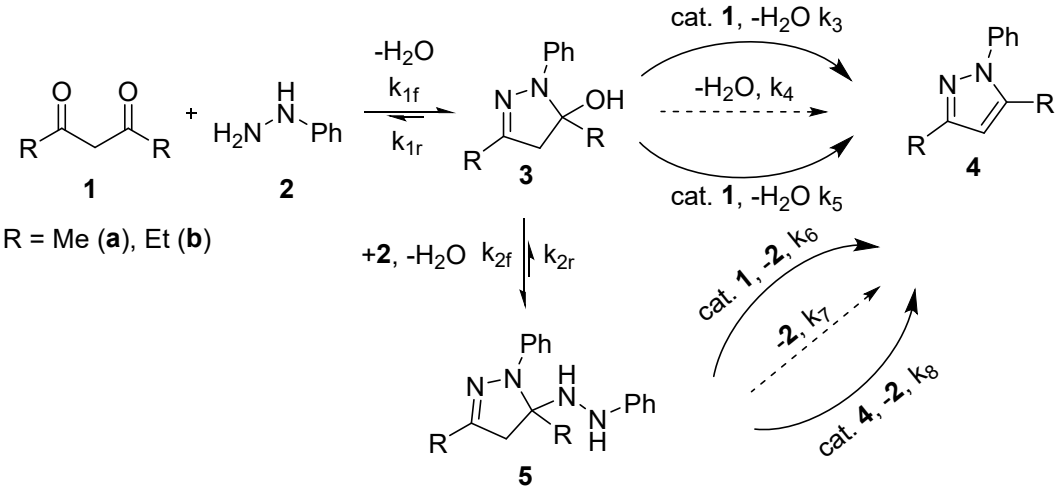
Model building began with a simple linear model **A** *via* intermediate **3** widely accepted in previous work:



R = Me (**a**), Et (**b**)

Figure S19: Scheme of the initial simple linear model (Model A) involving 5 species and 3 fitting parameters.

Fitting of this model to the available data for either diketone **1a** or diketone **1b** was poor both quantitatively (RMSE>0.3) and qualitatively (visually misaligned). Evidence from NMR experiments, LC-MS experiments, doping experiments, and reactant stoichiometry ramps suggested other pathways that would improve the accuracy of the model. Model **B** was constructed by individually introducing these pathways gradually improving model accuracy but still qualitatively performing poorly. This suggested to us that the model was overfitting due to the number of kinetic parameters fitted and thus that adding more parameters would inherently improve the quantitative fit without necessarily improving the visual fit.



R = Me (**a**), Et (**b**)

Figure S20: Scheme of the more complex model (Model B) involving 6 species and 10 fitting parameters developed from insight into the reaction from other experiments.

Changing from model **B** *via* intermediate **3** to a more competitive divergent model *via* intermediate **6** (model **C**) produced promising results with both good quantitative and qualitative fit (RMSE<0.05).

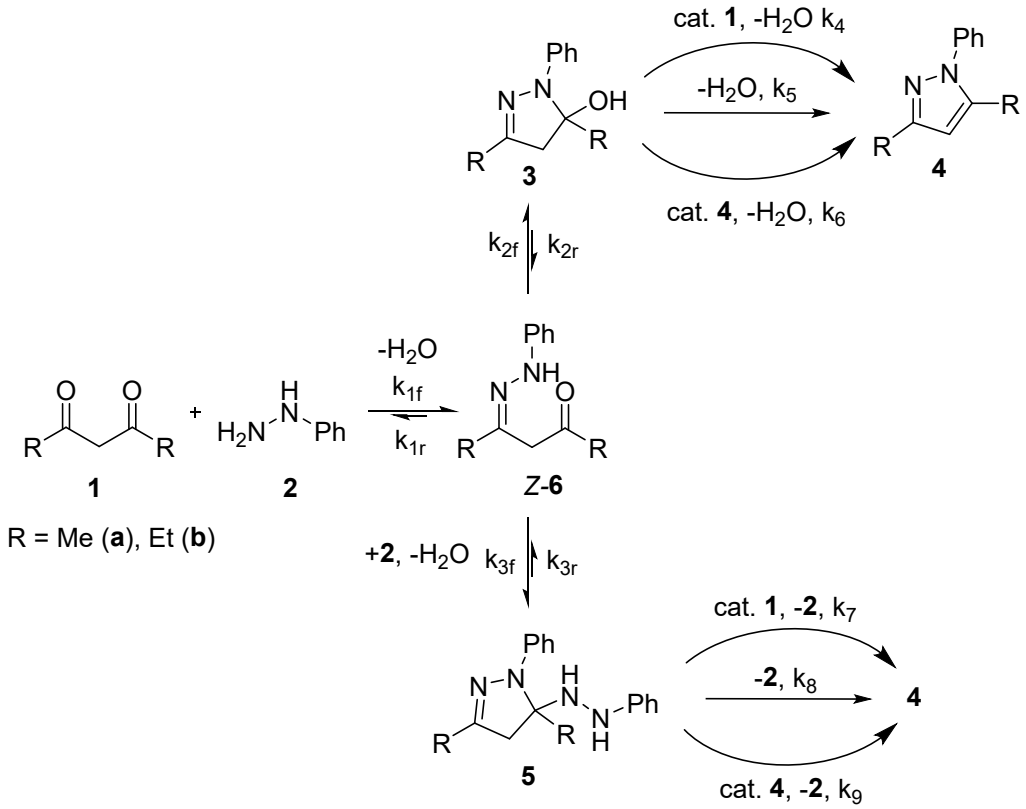


Figure S21: Scheme of the competitive divergent model (Model C) involving 7 species and 12 fitting parameters.

Cutting down the number of fitted rate constants gives a more accurate fit of the ones remaining and reduces the risk of the model over fitting. A simpler functional model is also preferable as it has greater interpretability by showing only the dominant relevant pathways in the kinetics of the system. With the aim to simplify the model whilst retaining or improving its good fit we developed an iterative fitting drop-out (IFDO) method. This method involves first determining a cut-off of rate constant value which is kinetically irrelevant, in our case $k<1 \times 10^{-5}$. The model is fitted using all available kinetic parameters and any parameters that are below the predetermined cut-off are set to 0 and removed from the list of fitting parameters. The model is then refitted with this shortened list of kinetic parameters with any remaining kinetic parameters set initially to the values determined in the previous iteration of model fitting. This method was repeated until no further kinetic parameters are dropped and the RMSE does not improve in a new iteration.

It was found that this method retained some insensitive parameters and could be further improved by resetting k_4-k_6 and k_7-k_9 to the mean value for each group and then reoptimizing again to further improve the RMSE score and qualitative fit. This produced consistent fits for both diketone **1a** and **1b** with the same active reaction pathways as described in figure 6 and table 3.

An example of this iterative method for diketone **1a** beginning with model **C** where red routes indicate dropped pathways:

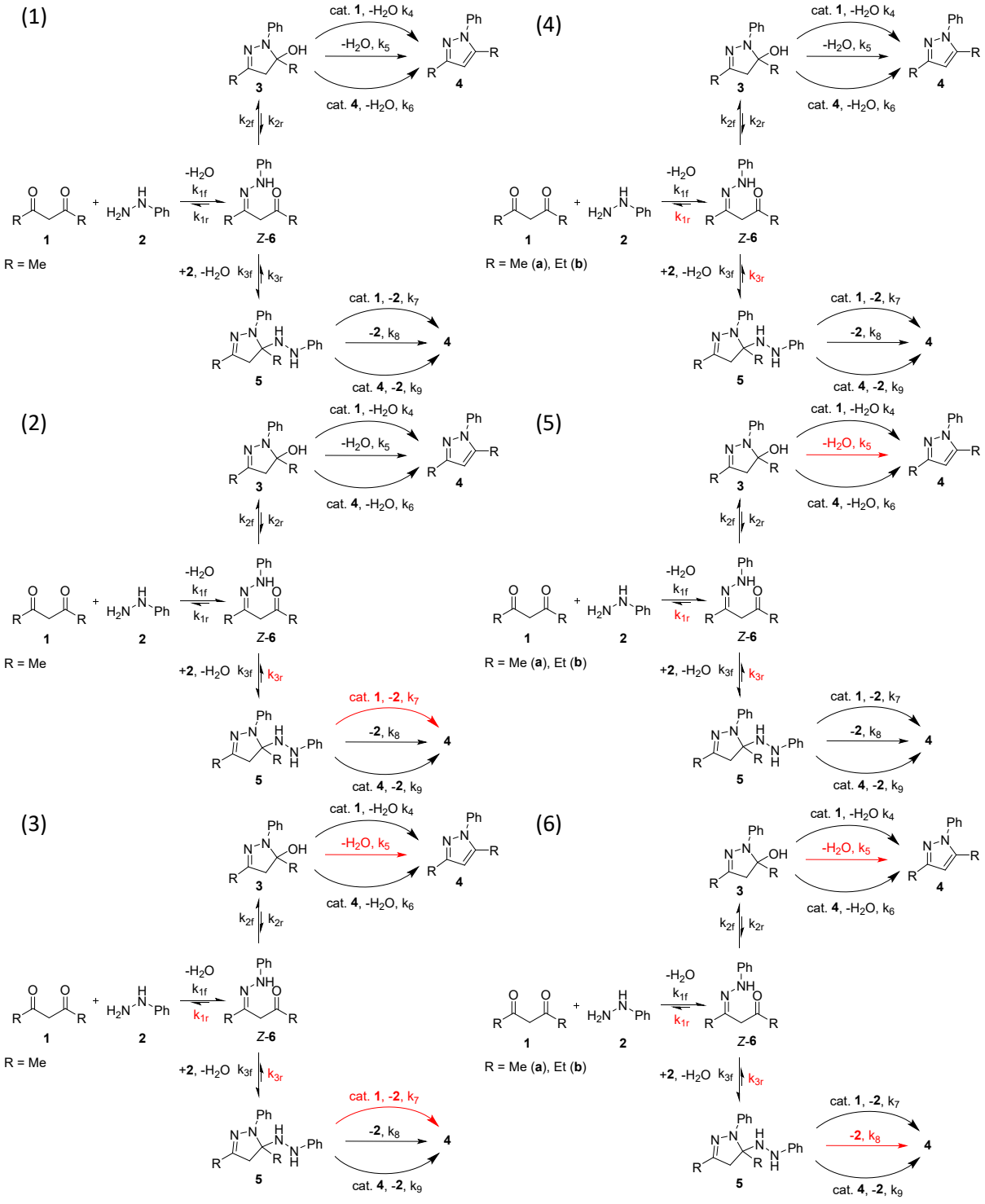


Figure S22: An example of the progression of model C (1) when fitted to data for the reaction of diketone **1a** using experiments A-G (Table 2) through our IFDO (iterative fitting drop-out) method. This leads to the dropping and refitting of the equilibrium steps (K_1-K_3) to produce two irreversible steps k_{1f} and k_{3f} and one equilibrium K_2. The later irreversible steps (k_4-k_9) are then given the mean value of their respective group (k_4-k_6 and k_7-k_9) (4) before performing IFDO again to reach a stable simple functional model (6).

The final fitted models and values for diketone **1a** and **1b** were then solved again using these fitted parameters *via* different solvers (Runge-Kutta 4 and Runge-Kutta 2) to confirm that the model was independent of the solver used and the results were not an artifact of the methods used by the solvers.

S10 MATLAB Surface plotting

3D visualisation and microkinetic model response surface generation was performed in MATLAB R2021b. The response surface was generated by creating a function containing the reactions and series of ODEs as used in Berkeley Madonna which takes time, reaction component concentrations and kinetic rate constants as in put parameters. The kinetic parameters are then inputted and the function solved using ode45 (a solver for coupled ODEs) for a range of starting concentrations of diketone **1** and phenyl hydrazine from [**1**]/[**2**]=0.66 to 1.5 in intervals of 0.02 at a time resolution of 0.0021 minutes. This could then be plotted using the function surf. The collected data was plotted in series over this surface using the function plot3 to generate the graphs shown in Fig. 8 as interactable 3D graphs in order to qualitatively determine the quality of model fit to independent data (Fig. S23). An interaction plot containing surfaces and data for both diketone **1a** and **1b** could also be generated from similar code (Fig. S24).

It is worth noting that ODE solver ode45 solves coupled ODEs in a different manner to Rosenbrock (stiff) as used in Berkeley Madonna and thus acts as another method of showing the solver independent nature of our microkinetic model solution.

All code used is available on https://github.com/LindenSchrecker

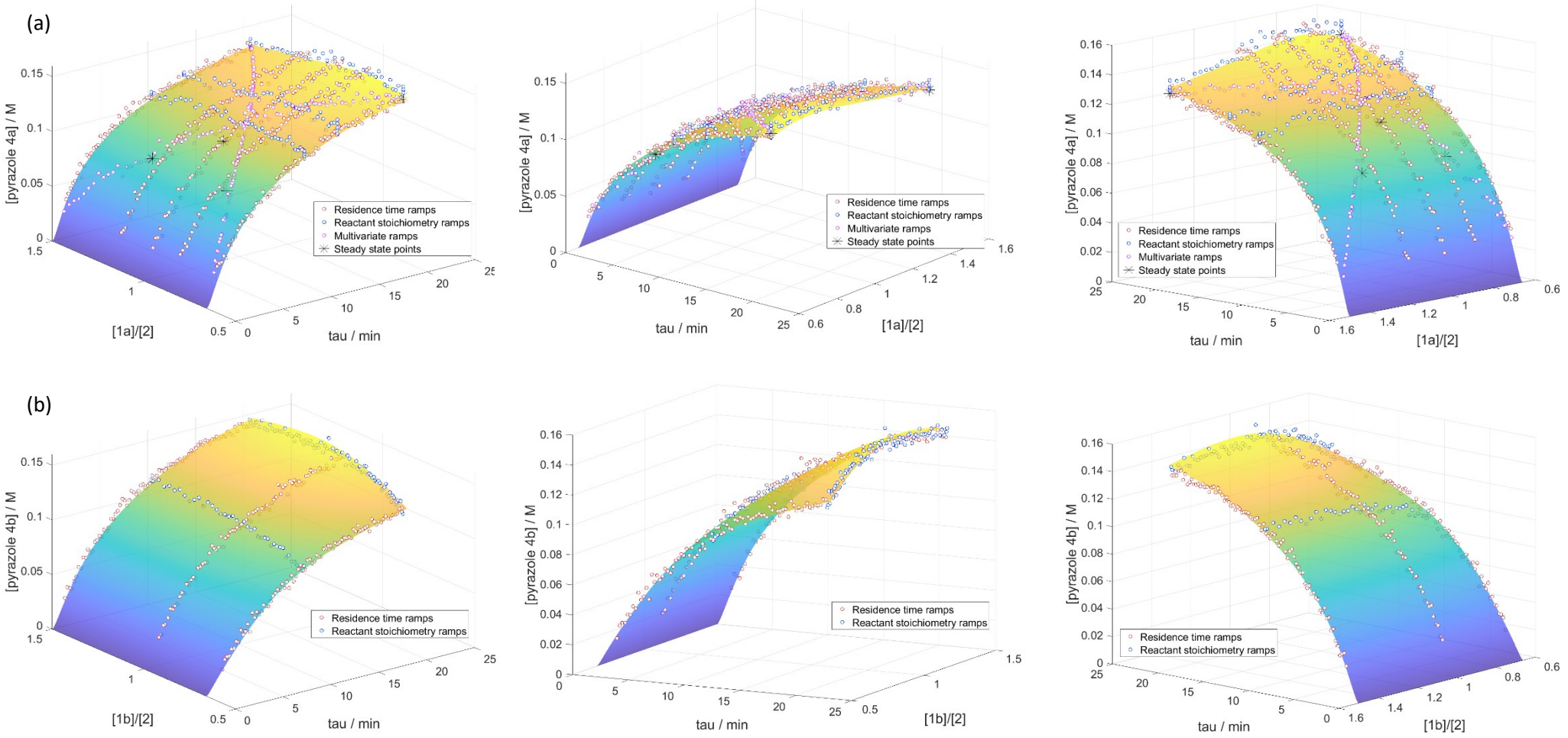


Figure S23: Multiple angles of the 3D visualised response surfaces for the reaction of phenyl hydrazine with (a) diketone **1a**; (b) diketone **1b**.

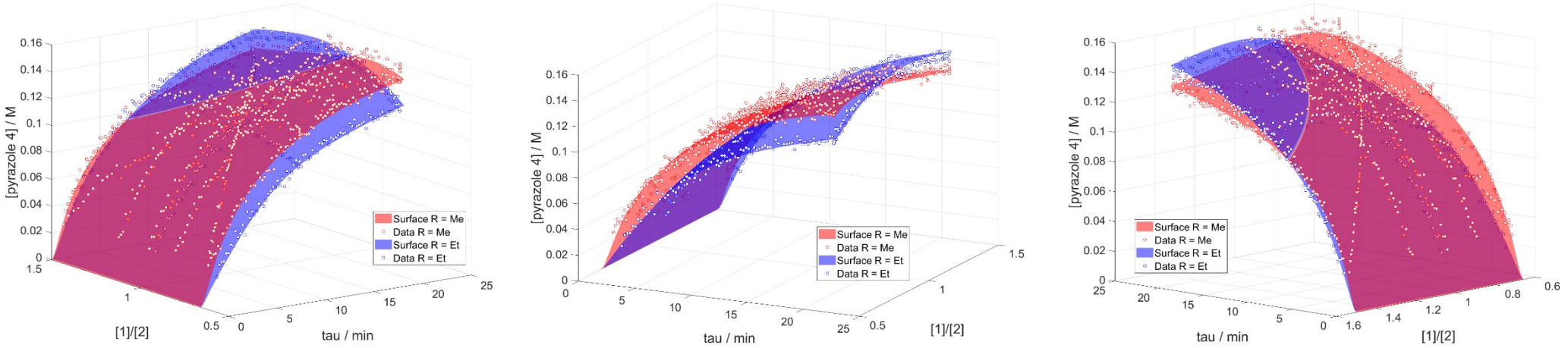


Figure S24: Multiple angles of the 3D visualised interaction plot of surfaces for diketone **1a** and **1b**.