

Supporting Information for
“Simultaneous Solvent Screening and Reaction Optimization in Microliter Slugs”

Brandon J. Reizman and Klavs F. Jensen
Department of Chemical Engineering
Novartis-MIT Center for Continuous Manufacturing
Massachusetts Institute of Technology
77 Massachusetts Avenue, Cambridge MA 02139 (USA)

Table of Contents

Table of Contents	2
1. Experimental Method.....	3
1.1. Flow Screening System.....	3
1.2. Preparation of (N-4-methoxybenzyl)-(1R,2R)-(-)-diaminocyclohexane (1)	6
1.3. Automated Reagent Calibration	6
1.4. Test of System Carryover.....	7
2. Optimization Method	7
3. Raw Experimental Data	14
4. Optimization Results.....	17
4.1. Optimization Results after Optimization Step 2	17
4.2. Optimization Results after Optimization Step 3	17
5. NMR Spectra	18
6. Nomenclature.....	21
7. References.....	22

1. Experimental Method

1.1. Flow Screening System

Figure S1 shows a complete schematic of the flow network for the automated screening and optimization system. A tank of nitrogen ($\geq 99.998\%$, Airgas) supplied both the main process flow and provided an inert gas blanket for the reagents stored in 7 mL vials in the liquid handler (Gilson GX-271, 221 x 1.5 x 0.4 mm probe) under a custom-made septum-sealed manifold (Figure S2). All solvents were purchased anhydrous and used as received from Sigma-Aldrich, purged with nitrogen, and introduced under nitrogen into vials under the septum-sealed manifold. 4-methoxybenzyl chloride (Sigma-Aldrich, 98% containing K_2CO_3 as stabilizer) was purged with nitrogen and stored in a vial under the inert gas manifold.

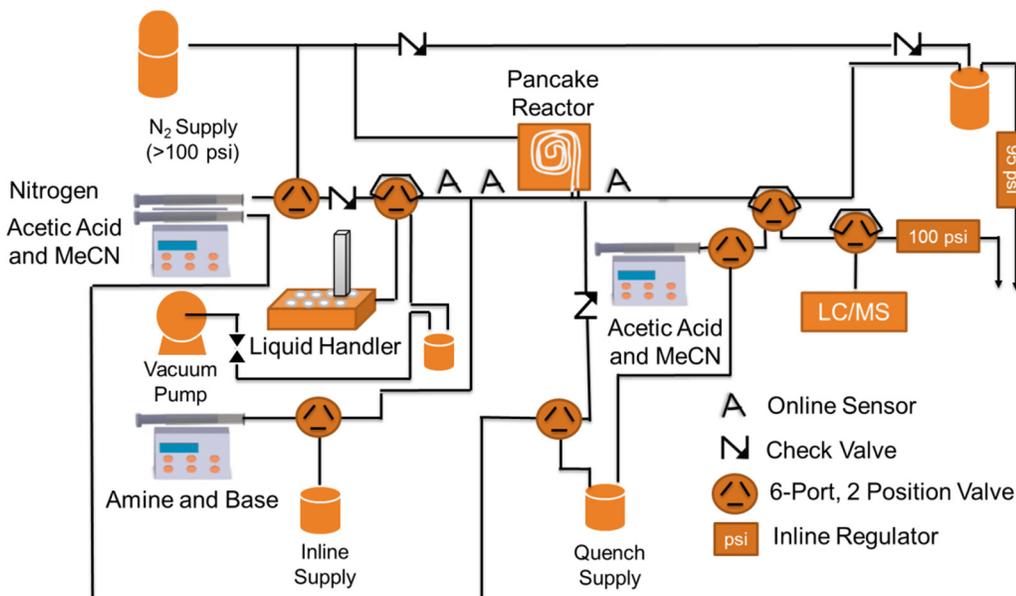


Figure S1. Schematic of automated flow system for alkylation reaction optimization.

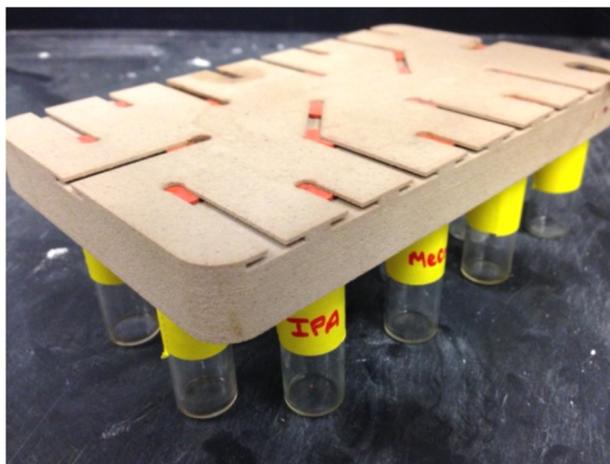


Figure S2. Septum-sealed inert gas manifold for reagent storage under nitrogen atmosphere.

To prepare a slug, the liquid handler (Gilson GX-271) aspirated first a 30 μL volume of nitrogen from an empty vial under the nitrogen manifold, followed by aliquots of a chosen solvent, 4-methoxybenzyl chloride, and the chosen solvent again. To minimize carryover during this process, the liquid handler probe (needle) was dipped in a wash solution of *i*PrOH before each reagent aspiration. 35 μL total liquid volume was nominally aspirated, although the relative volume of 4-methoxybenzyl chloride to solvent was corrected by the density ratio between the two reagents (accounting for the expansion of the nitrogen in the probe during aspiration). Following sample aspiration, the sample was “stirred” three times in the probe under nitrogen by pulling and pushing with the syringe pump 30 μL volume. All reagents were then transferred into a 6 port-2 way injection valve (Cheminert 10S-0503H) containing a 14- μL sample loop. Switching of the sample loop to the inject position created a 14- μL slug.

To minimize reagent carryover in the probe, injection valve, and system, three blank slugs were prepared and introduced to the system prior to introduction of a reagent slug. In sequence, the liquid handler aspirated 20 μL nitrogen, followed by 60 μL water with sample injection, 60 μL THF with sample injection, and finally 60 μL DMF with sample injection. Additionally, the sample injection valve and sample loop were cleared following every injection by pulling a vacuum for 3 seconds on the outlet of the valve (residual liquid was collected in a trap and drained at the end of the optimization). Because acceleration of slugs was observed to occur when trailing slugs entered the heated reactor (a consequence of surface-tension driven thermocapillary flow¹), blank slugs were not prepared and introduced until a reaction slug had traversed a full reactor volume in the system (240 μL).

A 5 mL solution of 2.22 g *trans*-1,2-diaminocyclohexane (Sigma-Aldrich), 1.97 g (1 equivalent) triethylamine (Sigma-Aldrich), and 135 mg naphthalene (Sigma-Aldrich) was prepared and stored under nitrogen. 2.1 μL of this solution was delivered into each reaction slug via syringe pump (Harvard PhD 2000 with 100 μL Hamilton Gastight syringe) through a T-junction (500 μm ID Teflon, Upchurch Scientific) at room temperature. Refractive index sensors (Omron EE-SPX613) were attached to the Teflon tubing before and after the T-junction to correctly time the online reagent injection and to verify that the slug volume was within an acceptable tolerance ($> 12.6 \mu\text{L}$) following the online injection. A network of valves (6 port, 2-way, Rheodyne MXP7960-000) and LabView software (ver. 8.6, National Instruments) allowed both the main nitrogen syringe (Harvard PhD 2000 pump with 8 mL Harvard stainless steel syringe) and the online injection syringe to be refilled automatically when syringe volumes were running low.

For the maximum versatility in our system, we chose to react the slugs in 750 μm FEP tubing. This was critical to the success of our method, as other “inert” tubing materials (stainless steel, PEEK, and even PFA) showed differing degrees of wetting with the solvents of interest, leading to slug degradation along tubing walls and contamination of subsequent slugs. The tubing was inserted into a “pancake” reactor housing, comprising of an aluminum chuck with a 1.6 mm groove for the FEP tubing, a raised lip with an O-ring, and a sheet of polycarbonate which

compressed against the O-ring to allow for pressurization of the reactor to 6.9 bar. With this device, we were able to rapidly heat and cool the reactor tubing between 30°C and 120°C and neutralize nitrogen permeation out of the reactor, allowing for accurately controlled slug flow rates in the range of 1 min to 10 min reaction time. A thermocouple was introduced through the nitrogen supply line and held in place on the aluminum surface by a thin sheet of polycarbonate. The reactor was heated with four 50 W cartridge heaters (McMaster-Carr, two pairs spaced equally on opposite sides of the device). COMSOL (COMSOL, Inc.) simulations revealed that this configuration was sufficient for ensuring less than a 2°C temperature gradient across the reactor tubing of the device at the maximum temperature of 120°C. A PID temperature controller (Omega CN9412) controlled the reactor temperature. Reaction slugs were not introduced into the system unless the reactor temperature was within 1°C of the reaction set point temperature. To conserve overall experimental time, slugs with the same temperature set point in the same optimization step were run consecutively. With the exception of this constraint, all experiments within an optimization step were run in random order.

Downstream of the reactor, slugs were quenched at room temperature with a continuously flowing solution of 10% acetic acid in acetonitrile, delivered via syringe pump (Harvard Apparatus PhD 2000 with 8 mL Harvard stainless steel syringe) through a T-junction (1 mm ID Teflon, Upchurch Scientific). A third refractive index sensor (Omron EE-SPX613) was used downstream of the quench to time the HPLC sampling accurately. Following sampling with a 30 μ L sample loop in a 6-port, 2-way valve (Rheodyne MXP7960-000), the sample was transported via syringe pump (Harvard Apparatus PhD 2000 with 1 mL Hamilton Gastight syringe) in a solution of 10% acetic acid in acetonitrile to a second 6-port, 2-way valve (Agilent G1158A) with a 2 μ L sample loop. An LC/MS (Agilent G1312B binary pump, G1329B ALS, G1316A column compartment, G1365C multi-wavelength detector, 6120 quadrupole MS) method was remotely started with LabView software. The LC flow rate increased from 0.5 mL/min (standby mode) to 3.5 mL/min (required for method), and the sample was injected into the HPLC after 15 s. The sample passed through a C18 guard column (Agilent Poroshell 300SB-C18 5 μ m 2.1 x 12.5 mm), then was heated to 40°C, passed through a T-junction (250 μ m stainless steel Valco), and split by pressure difference between a 1.8 μ m particle diameter column (Agilent Zorbax SB-C18 2.1 x 50 mm) and a 4.6 μ m particle diameter column (Agilent Zorbax SB-C18 2.1 x 50 mm). The sample from the 1.8 μ m particle diameter column was detected by UV at 270 nm and passed to the MS. A suitable HPLC method was found to be 9 min, which included a gradient ramp from 95/5 water/acetonitrile + 1% formic acid to 0/100 acetonitrile + 1% formic acid to 95/5 water/acetonitrile + 1% formic acid. Following UV analysis, the product yield was calculated in MATLAB (ver. R2011a, MathWorks).

Pressure in the system was controlled at 6.9 bar with a nitrogen-regulated Parr bomb, approximately 40 mL in volume. The bomb was drained during refill of the nitrogen and quench syringes by automatically opening a 6-port, 2 way valve (Rheodyne MXT715-000). To minimize

nitrogen loss during regular system operation, 6.6 bar of backpressure (5.2 bar and 1.4 bar backpressure regulators, Upchurch Scientific) was applied to the gas vent of the Parr bomb during regular system operation.

Valve manipulation, HPLC method initiation, and analog input communication with the refractive index sensors were accomplished using a Compact FieldPoint controller from National Instruments (cFP-2020, cFP-RLY-425, cFP-AI-110). The entire system including pumps, liquid handler, temperature control, refractive index monitoring, valving, remote triggering of the HPLC gradient, and MATLAB optimization was controlled with LabView software (ver. 8.6, National Instruments).

1.2. Preparation of (N-4-methoxybenzyl)-(1R,2R)-(-)-diaminocyclohexane (1)

All reagents were used as received. (1R,2R)-(-)-1,2-diaminocyclohexane (410.9 mg, 3.60 mmol, 1 equiv, 98% purity, 99% ee Sigma-Aldrich), triethylamine (358.6 mg, 3.54 mmol, 1 equiv, $\geq 99\%$ Sigma-Aldrich), naphthalene (29.1 mg, 0.23 mmol, 99% Sigma-Aldrich), and DMSO (5 mL, anhydrous $\geq 99.9\%$ Sigma-Aldrich) were stirred in a 25 mL round-bottomed flask. To the flask was added 4-methoxybenzyl chloride (1064.6 mg, 6.80 mmol, 2 equiv, 98% w/ K_2CO_3 as stabilizer Sigma-Aldrich). The reactants were heated to 78°C and stirred for 7.5 min. The reaction was quenched with 4 M aq. NaOH (5 mL) at room temperature, and analysis was taken by HPLC showing 61% yield of **1**. The resulting solution was extracted 5 times with 40 mL ethyl acetate. The collected organic product was then extracted once with 50 mL 4 M NaOH, and the subsequent aqueous product was extracted twice with 25 mL ethyl acetate. The cumulative extracted organic product was then extracted again with 50 mL 4 M NaOH, and the subsequent aqueous product was extracted again twice with 25 mL ethyl acetate. The cumulative organic product was dried with Na_2SO_4 , filtered, and the solvent was removed by rotary evaporation to yield an orange oil. This oil was purified using column chromatography with silica gel. The column mobile phase was increased from DCM/0.1% Et_3N to DCM/8% MeOH/0.1% Et_3N to yield an isolated sample of the product **1**. The solvents were removed by rotary evaporation, and the resulting product was washed 5 times with 5 mL ethyl acetate followed by rotary evaporation to remove excess Et_3N . The product was then dissolved in DCM, precipitated with hexane, filtered, and dried under vacuum, yielding the product **1** (500.9 mg, 2.14 mmol, 59% yield) as a white powder.

Product characterization: ^1H NMR (400 MHz, CDCl_3): δ = 7.36 (d, J = 8.6 Hz, 2H), 6.88 (d, J = 8.7 Hz, 2H), 4.23 (s, 2H), 4.11 – 3.74 (dd, J = 115.5, J = 12.9, 2H), 3.79 (s, 3H), 2.81 (td, J = 11.2, 4.1 Hz, 1H), 2.53 (td, J = 10.9, 3.9 Hz, 1H), 2.18 (dd, J = 12.1 Hz, 2H), 1.75 (dd, J = 25.0, 11.7 Hz, 2H), 1.54 – 1.14 (m, 5H). ^{13}C NMR (101 MHz, CDCl_3): δ = 159.28, 130.12, 129.52, 114.15, 59.67, 55.37, 54.76, 49.57, 31.65, 30.32, 24.68, 24.36. HRMS (ESI) m/z 235.1805 (calculated for $\text{C}_{14}\text{H}_{22}\text{N}_2\text{O}$ 235.1805 $[\text{M}+\text{H}]^+$).

1.3. Automated Reagent Calibration

A solution of 0.542 M (*N*-4-methoxybenzyl)-(1*R*,2*R*)-(-)-diaminocyclohexane and 0.0433 M naphthalene in DMSO and a solution of 0.0474 M naphthalene in DMSO were stored under nitrogen in the liquid handler. Two replicates each of slugs containing 0 M, 0.13 M, 0.26 M, 0.39 M, and 0.52 M (*N*-4-methoxybenzyl)-(1*R*,2*R*)-(-)-diaminocyclohexane were prepared following the same procedure as in the optimization, routed through the screening system (without online injection) at 30°C and a residence time of 5 min, and analyzed by LC/MS. A calibration was constructed based on integrated peak absorbance measurements of the desired product and the internal standard naphthalene at $\lambda = 270$ nm. The calibrated slope was $C_{prod} = 6.98 * C_{naphthalene} * A_{prod} / A_{naphthalene}$ with $R^2 = 0.987$.

1.4. Test of System Carryover

Neat 4-methoxybenzyl chloride and THF were stored under nitrogen in the liquid handler. A slug comprising 1 M 4-methoxybenzyl chloride in THF was prepared following the same procedure as in the optimization, routed through the screening system (without online injection) at 30°C and a residence time of 5 min, and analyzed by LC/MS. Two subsequent slugs were then prepared with only THF as the reagent and routed through the system at the same flow rate and temperature and analyzed by LC/MS. This procedure was repeated 9 times. On average, the LC area for 4-methoxybenzyl chloride was 6578 in the 1 M slugs, 216 in the first subsequent THF slugs, and 116 in the second subsequent THF slugs. The percent carryover was $3.3\% \pm 0.6\%$ in the first THF slugs and $1.8\% \pm 0.6\%$ in the second THF slugs with a maximum carryover in the first THF slugs of 3.9% (twice).

2. Optimization Method

Simultaneous solvent screening and reaction optimization was achieved using a design of experiments (DoE) approach similar to that outlined in Atkinson and Donev.² Continuous variables were first scaled assuming the first-order kinetic relationship:

$$C_{Prod} \propto A_s e^{E_s/RT} C_{0,Diamine} C_{0,MeOBnCl} t_{res} \quad (1)$$

$$\Rightarrow \ln\left(\frac{C_{Prod}}{C_{0,Diamine}}\right) = \ln(Yield) \propto \ln(A_s) + \left(\frac{E_s}{R}\right)T^{-1} + \ln(C_{0,MeOBnCl}) + \ln(t_{res}) \quad (2)$$

$\ln(A_s)$ and (E_s/R) were proposed to be coefficients strictly correlated with the choice of solvent. To correct for inaccuracies in our assumed scaling, additional coefficients were introduced to weigh the $\ln(C_{0,MeOBnCl})$ and $\ln(t_{res})$ terms and account for interactions and quadratic functionality among all continuous variables. The final response surface model to fit was:

$$\hat{b} = \sum_{s=4}^{13} (c_s x_s + a_{s3} x_s x_3) + c_1 x_1 + c_2 x_2 + \sum_{i=1, j=i}^3 a_{ij} x_i x_j \quad (3)$$

where \hat{b} was the response value ($\ln(Yield)$), a_{ij} and c_i were coefficients to fit, and the vector \mathbf{x} was:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_{13} \end{bmatrix} = \begin{bmatrix} 2 \frac{\ln(C_{0,MeOBnCl}) - \ln(C_{0,MeOBnCl})_{\min}}{\ln(C_{0,MeOBnCl})_{\max} - \ln(C_{0,MeOBnCl})_{\min}} - 1 \\ 2 \frac{\ln(t_{res}) - \ln(t_{res})_{\min}}{\ln(t_{res})_{\max} - \ln(t_{res})_{\min}} - 1 \\ 2 \frac{T^{-1} - T^{-1}_{\min}}{T^{-1}_{\max} - T^{-1}_{\min}} - 1 \\ \text{(fraction solvent 1)} \\ \text{(fraction solvent 2)} \\ \vdots \\ \text{(fraction solvent 10)} \end{bmatrix} \quad (4)$$

Naturally \hat{b} was reformulated as:

$$\hat{b} = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{c}^T \mathbf{x}; \quad \mathbf{A} = \begin{bmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_{13}}{2} & 0 & \dots & 0 \\ \frac{a_{12}}{2} & a_{22} & \frac{a_{23}}{2} & 0 & \dots & 0 \\ \frac{a_{13}}{2} & \frac{a_{23}}{2} & a_{33} & \frac{a_{43}}{2} & \dots & \frac{a_{133}}{2} \\ 0 & 0 & \frac{a_{43}}{2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \frac{a_{133}}{2} & 0 & \dots & 0 \end{bmatrix} \quad \mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_{13} \end{bmatrix} \quad (5)$$

which allowed for fast optimization and Hessian calculation in subsequent steps.

To initialize the response surface model, we began experimentation with a 20-experiment fractional factorial design, or 2 experiments for each of 10 solvents. Critical to the selection of these 20 experiments was that the continuous variable experiments for each solvent completed a fractional factorial design—in this case a 2^{3-2} design where high and low values were selected in opposite pairs—and that the sum of all 10 continuous variable designs gave the maximum amount of information—in this case a full 2^3 factorial design with 2 additional replicates. All continuous and discrete variables were randomized so as to minimize the bias in the selection of initial fractional factorial design, and the factorial experiments spanned the entire constrained continuous variable space so as to minimize uncertainty in the yield in future calculations.

From the data collected in 20 experiments, an initial linear response surface was constructed of the form:

$$\hat{\mathbf{b}} = \boldsymbol{\theta}^T \mathbf{X} \quad (6)$$

where \mathbf{X} represented a matrix of the previous (scaled) experiments run ($\mathbf{X} = [\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_3]^T$), $\hat{\mathbf{b}}$ was a vector of model-predicted responses for the 20 experiments, and $\boldsymbol{\theta}$ was a vector of best-fit parameters. The goal of this simplification was to find a quadrant of the experimental space identified by the preliminary fractional factorial design which most likely contained the optimum, and to concentrate future factorial design experiments in that quadrant. This was a more efficient use of real-time information than the traditional central composite design, which would have allowed estimation of quadratic terms only by experimentation over the entire experimental space. To estimate $\boldsymbol{\theta}$ from a vector of the (scaled) responses \mathbf{b} , we had:

$$\boldsymbol{\theta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{b} \quad (7)$$

Optima for each solvent s were then found by solving:

$$\begin{aligned} J_s &= \max_{\mathbf{x}} \boldsymbol{\theta}^T \mathbf{x} & (8) \\ \text{s.t. } x_i &\in [-1, 1] \text{ if } i = 1, \dots, 3 \\ x_i &= 1 \text{ if } i = s \\ x_i &= 0 \text{ if } i > 3 \text{ and } i \neq s \end{aligned}$$

A new set of 20 fractional factorial design experiments was constructed as before, but instead of spanning the full design space these factorial experiments spanned from the midpoint of the continuous variables (0 in this scaling) to the linear maxima \mathbf{x}_s^* .

With a sufficient number of experiments and midpoint experiments, we were able to reformulate \mathbf{X} in terms of scaled linear, interaction, and quadratic variables and solve for the scaling parameter vector $\boldsymbol{\theta}$ as in Equation 7. To estimate the accuracy of a predicted response $\hat{\mathbf{b}}$ at experimental conditions \mathbf{x} required both the sensitivity matrix of experiments already run (given by $(\mathbf{X}^T \mathbf{X})^{-1}$) and the response covariance V_B , which was analogous to the noise in collecting a measurement. The prediction covariance was estimated as:³

$$V_{\hat{\mathbf{b}}} = \mathbf{x}^T (\mathbf{X}^T V_B^{-1} \mathbf{X})^{-1} \mathbf{x} \quad (9)$$

for which an estimate of the scalar V_B was found using:³

$$V_B = \frac{(\boldsymbol{\theta}^T \mathbf{X} - \mathbf{b})^T \mathbf{W} (\boldsymbol{\theta}^T \mathbf{X} - \mathbf{b})}{N_{\text{expts}} - N_{\text{params}}} \quad (10)$$

where \mathbf{W} , a weighting matrix, was taken as the identity matrix and $N_{\text{expts}} - N_{\text{params}}$ was the difference in the number of experiments run and the number of response surface parameters fit. Necessarily N_{expts} had to be greater than N_{params} .

Global optima for the each solvent \mathbf{x}_s^* were identified by converting $\boldsymbol{\theta}$ to \mathbf{A} and \mathbf{c} and solving the quadratic optimization problem:

$$\begin{aligned}
 J_s^* &= \max_{\mathbf{x}} \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{c}^T \mathbf{x} & (11) \\
 \text{s.t. } x_i &\in [-1, 1] \text{ if } i = 1, \dots, 3 \\
 x_i &= 1 \text{ if } i = s \\
 x_i &= 0 \text{ if } i > 3 \text{ and } i \neq s
 \end{aligned}$$

The overall maximum, J^* , corresponded to the maximum over all J_s^* . J_-^* , the lower bound on J^* , was found from a Student's t -distribution around J^* with uncertainty $V_{\hat{B}}$ evaluated at the optimum \mathbf{x}^* .³

$$J_-^* = J^* - \left(V_{\hat{B}}\right)^{1/2} \left(t_{1-\alpha_{conf}, v=N_{expts}-N_{params}}\right) \quad (12)$$

α_{conf} was chosen before experimentation as 0.05, corresponding to a 95% one-sided confidence level on the lower bound of J^* . For all remaining solvents, a paired 2-sample t -test at 95% confidence revealed whether J_s^* was significantly less than the overall optimum J^* :

$$\begin{aligned}
 \text{Null Hypothesis: } H_0 &= J^* - J_s^* = 0 & (13) \\
 H_a &= J^* - J_s^* > 0 \\
 t_{stat} &= \frac{J^* - J_s^*}{V_{\hat{B}}^{1/2}} > t_{crit} = t_{1-\alpha_{conf}, v=N_{expts}-N_{params}} \\
 J^* - \left(V_{\hat{B}}^{1/2}\right) \left(t_{1-\alpha_{conf}, v=N_{expts}-N_{params}}\right) &> J_s^* \\
 J_-^* &> J_s^*
 \end{aligned}$$

(Note this assumed a constant $V_{\hat{B}}$ for all \mathbf{x}_s^* .) Solvents for which the null hypothesis was rejected were fathomed (eliminated) from the ensuing experimental optimization step but were reevaluated at subsequent optimization iterations after updating all response surfaces.

Our goal in subsequent experiments was to minimize the uncertainty on the optimum for each discrete variable, thereby helping reduce the number of discrete variables considered in the optimization system and allowing more experimental emphasis to be focused on the solvents of highest interest. To accomplish this task, we employed a G-optimal experimental design strategy to select experiments where the uncertainty at a proposed optimum could be most greatly minimized. This entailed feeding \mathbf{x}_s^* for each non-fathomed solvent to a constrained optimization program in MATLAB:

$$\begin{aligned}
\min_{\mathbf{x}} V_{\hat{B}} &= (\mathbf{x}_s^*)^T (\mathbf{X}_1^T V_B^{-1} \mathbf{X}_1)^{-1} \mathbf{x}_s^* & (14) \\
s.t. \mathbf{X}_1 &= [\mathbf{X}^T \mathbf{x}^T]^T \\
x_i &\in [-1, 1] \text{ if } i = 1, \dots, 3 \\
x_i &= 1 \text{ if } i = s \\
x_i &= 0 \text{ if } i > 3 \text{ and } i \neq s
\end{aligned}$$

\mathbf{x}_s^* was supplied as the initial guess of experimental conditions. A group of N experiments—one optimal experiment for each unfathomed variable—was selected, those experiments were conducted in the slug flow system, and the optimization was repeated until the termination criterion was satisfied. To minimize downtime, the optimization procedure was performed after experiment $N - 1$, and optimal experiments were added to a queue of the remaining incomplete experiments. We found a suitable termination criterion to be linear convergence of the error in J^* (namely $J^* - J_{-}^*$):

$$\begin{aligned}
\text{let } \Delta J_n^* &= J^* - J_{-}^* \text{ be the error on the optimum through iteration } n & (15) \\
\Delta J_{pred}^* &= \frac{N_{expts} \text{ between } n-2 \text{ and } n}{N_{expts} \text{ between } n-2 \text{ and } n-1} (\Delta J_{n-2}^* - \Delta J_{n-1}^*) + \Delta J_{n-2}^* \\
&\text{if } \frac{|\Delta J_{pred}^* - \Delta J_n^*|}{\Delta J_n^*} < 0.01, \text{ terminate}
\end{aligned}$$

This criterion was independent of the number of remaining unfathomed solvents, which allowed the possibility of multiple similar solvents to be discriminated with the quasi-Newton search and allowed the noise in the optimum to be reduced substantially for the case where only a single solvent remained under consideration after a few G-optimal search iterations.

With sufficient reduction in the error on the optimum J^* (leading to termination of the G-optimal search), a gradient-based quasi-Newton search was conducted for unfathomed solvents, initialized at \mathbf{x}_s^* . The final gradient and line search strategy was employed to test the assumptions of the response surface method for the remaining candidate solvent(s) and identify an optimum within a user-defined tolerance. The quasi-Newton search comprised a gradient estimation and a back-stepping line search, the search direction for which was calculated using the estimated gradient and a BFGS approximation for the Hessian. To estimate a gradient, a randomly chosen 2^{3-1} fractional factorial design was run around \mathbf{x}_s^* at vertex points given by $\mathbf{x}_s^* \pm \Delta \mathbf{x}$, where $\Delta \mathbf{x}$ was a user-defined tolerance on the desired optimization accuracy. For variables where $x_i^* \pm \Delta x_i$ exceeded a constraint, x_i was set equal to the constraint. \mathbf{x}_s^* was also examined experimentally, and a gradient was calculated from all experiments run throughout the optimization within $\mathbf{x}_s^* \pm 1.05^* \Delta \mathbf{x}$. A new estimate of J_s^* was calculated based on the linearized response surface within $\mathbf{x}_s^* \pm \Delta \mathbf{x}$, with uncertainty ΔJ_s^* :

$$\Delta J_s^* = \pm \left(V_{J_s^*}^{1/2} \right) \left(t_{1-\alpha_{conf}, v=N_{expts}-N_{params}} \right) \quad (16)$$

$V_{J_s^*}$ was the prediction covariance estimated over only $\mathbf{x}_s^* \pm 1.05^* \Delta \mathbf{x}$. The confidence parameter α_{conf} was chosen to be 0.16, corresponding to a one-sided single-standard deviation confidence level. The (uphill) search direction \mathbf{p} was evaluated from the gradient \mathbf{g} and approximate Hessian \mathbf{H} :

$$\mathbf{p}_k = \mathbf{H}_k^{-1} \mathbf{g}_k \quad (17)$$

\mathbf{H}^{-1} was initialized as the inverse of Hessian found in the optimization of Equation 11, and an update for \mathbf{H}^{-1} was calculated following the BFGS approximation:⁴

$$\mathbf{H}_{k+1}^{-1} = \mathbf{H}_k^{-1} + \frac{\left[\mathbf{s}_k^T \mathbf{y}_k + \mathbf{y}_k^T \mathbf{H}_k^{-1} \mathbf{y}_k \right] \mathbf{s}_k^T \mathbf{s}_k}{\left(\mathbf{s}_k^T \mathbf{y}_k \right)^2} - \frac{\mathbf{H}_k^{-1} \mathbf{y}_k \mathbf{s}_k^T + \mathbf{s}_k \mathbf{y}_k^T \mathbf{H}_k^{-1}}{\mathbf{s}_k^T \mathbf{y}_k} \quad (18)$$

$$\mathbf{s}_k = -\alpha_k \mathbf{p}_k \quad (19)$$

$$\mathbf{y}_k = \mathbf{g}_{k+1} - \mathbf{g}_k \quad (20)$$

New experiments along a line search were selected with the scaling factor α according to:

$$\mathbf{x}_s = \mathbf{x}_s^* + \alpha \mathbf{p}_k \quad (21)$$

For x_i in violation of a constraint, x_i was chosen equal to the constraint. α was halved iteratively until the subsequent experiment would not be a replicate of the constrained experiment.

The objective function value found at \mathbf{x}_s , B_s , was evaluated statistically against the uncertainty in J_s^* to determine if the line search experiment was a candidate optimum. Similarly to the G-optimal search, a paired t -test was employed with the criterion:

$$J_s^* - \Delta J_s^* > B_s \quad (22)$$

Satisfaction of this criterion implied that \mathbf{x}_s was not a candidate optimum, and hence α was reduced to $\alpha/2$ and a new line search experiment was commenced. In the event that \mathbf{x}_s could not be disregarded as an optimum, a gradient was calculated around \mathbf{x}_s with a randomly-selected 2^{3-1} fractional factorial design. A new candidate optimum, J_s' , was established in the region $\mathbf{x}_s \pm \Delta \mathbf{x}$ with variance $V_{J_s'}$. J_s' was compared to J_s^* with an unpaired 2-sample t -test:

Null Hypothesis: $H_0 = J'_s - J_s^* = 0$

$$H_a = J'_s - J_s^* > 0$$

$$t_{stat} = \frac{J'_s - J_s^*}{\sqrt{V_{J'} + V_{J^*}}} > t_{crit} = t_{1-\alpha_{conf}, \nu} \quad (23)$$

$$\nu = \frac{(V_{J'} + V_{J^*})^2}{\frac{V_{J'}^2}{N_{expts}' - N_{contvariables}} + \frac{V_{J^*}^2}{N_{expts}^* - N_{contvariables}}}$$

Rejection of the null hypothesis established \mathbf{x}_s as the new optimum \mathbf{x}_s , and a new search direction was calculated following Equations 17-20. Failure to reject the null hypothesis resulted in a continuation of the line search with α was reduced to $\alpha/2$. When the step size was small enough such that $\alpha \mathbf{p}_k < \Delta \mathbf{x}$ for all x_i , the optimization terminated and J_s^* was recorded as the optimum for solvent s with uncertainty ΔJ_s^* .

3. Raw Experimental Data

Table S1. Observed yields for conditions screened during first fractional factorial design.

Expt Number	Solvent	t_{res} (s)	T (°C)	$C_{0,\text{MeoBnCl}}$ (M)	Yield
1	<i>i</i> PrOH	60	30.0	0.999	2.4%
2	THF	60	30.0	0.999	1.8%
3	Toluene	600	30.0	0.998	9.7%
4	MeCN	60	30.0	0.206	2.0%
5	DMF	600	30.0	0.206	6.1%
6	DMSO	600	30.0	0.999	52.2%
7	Pyridine	60	30.0	0.206	8.5%
8	DMC	60	30.0	0.206	1.0%
9	DME	600	30.0	1.000	5.4%
10	DCE	600	30.0	0.206	2.7%
11	DMSO	60	120.0	0.206	28.4%
12	<i>i</i> PrOH	600	120.0	0.206	23.5%
13	MeCN	600	120.0	1.000	46.6%
14	Pyridine	600	120.0	0.999	47.2%
15	DMC	600	120.0	0.999	56.8%
16	THF	600	120.0	0.206	27.7%
17	Toluene	60	120.0	0.206	9.3%
18	DMF	60	120.0	0.999	36.4%
19	DME	60	120.0	0.206	8.7%
20	DCE	60	120.0	1.000	44.6%

Table S2. Observed yields for conditions screened during second fractional factorial design.

Expt Number	Solvent	t_{res} (s)	T (°C)	$C_{0,\text{MeoBnCl}}$ (M)	Yield
21	DMC	600	120.0	0.999	45.5%
22	DCE	190	120.0	1.000	53.4%
23	<i>i</i> PrOH	600	120.0	0.444	43.2%
24	MeCN	600	120.0	1.000	36.1%
25	THF	600	120.0	0.444	35.7%
26	DME	190	120.0	0.445	42.3%
27	Pyridine	190	120.0	0.999	51.5%
28	DMSO	190	120.0	0.445	52.2%
29	Toluene	190	120.0	0.444	43.7%
30	DMF	190	120.0	0.444	48.9%
31	DMF	600	69.2	0.999	56.2%
32	DCE	600	69.2	0.444	35.9%
33	THF	190	69.2	0.999	29.9%
34	MeCN	190	69.2	0.444	34.2%
35	Pyridine	600	69.2	0.444	30.5%
36	DME	600	69.2	1.000	41.5%
37	DMC	190	69.2	0.444	16.2%
38	Toluene	600	69.2	0.998	41.9%
39	<i>i</i> PrOH	190	69.2	0.999	33.6%
40	DMSO	600	69.2	0.999	62.9%

Table S3. Observed yields for conditions screened during G-optimal design of experiments.

Expt Number	Solvent	t _{res} (s)	T (°C)	C _{0,MeoBnCl} (M)	Yield
41	DMSO	77	46.4	0.999	36.0%
42	DCE	137	63.6	1.000	24.0%
43	THF	161	120.0	0.999	43.2%
44	Pyridine	94	61.1	0.999	33.0%
45	<i>i</i> PrOH	161	120.0	0.999	48.7%
46	DMF	100	52.1	0.999	31.4%
47	Toluene	91	59.4	0.998	9.9%
48	DMSO	265	90.6	0.999	57.2%
49	DMF	429	120.0	0.999	48.7%
50	DMSO	391	102.8	0.999	59.3%
51	DMSO	323	58.8	0.333	31.4%
52	DCE	326	120.0	0.429	34.8%
53	DMF	229	98.9	0.364	46.1%
54	DMF	600	101.0	0.999	58.6%
55	DCE	419	120.0	1.000	47.7%
56	Pyridine	600	45.0	0.999	44.0%
57	DMSO	600	87.8	0.999	60.7%
58	DMF	431	94.3	0.999	52.3%
59	DMSO	600	85.2	0.999	64.3%
60	Pyridine	455	73.2	0.999	53.1%
61	DMSO	319	74.9	0.999	53.9%
62	Pyridine	340	77.7	0.999	53.5%
63	DMF	332	93.1	0.999	54.4%
64	DMF	340	96.0	0.778	54.9%
65	<i>i</i> PrOH	188	120.0	0.667	23.7%
66	DMSO	514	75.6	0.682	59.4%
67	DMSO	600	79.8	0.999	62.9%
68	DMSO	516	81.1	0.999	62.6%

Table S4. Observed yields for conditions screened during quasi-Newton gradient-based search.

Expt Number	Solvent	t_{res} (s)	T (°C)	$C_{0,\text{MeoBnCl}}$ (M)	Yield
69	10	444	78.1	0.999	61.1%
70	10	459	76.1	0.983	60.7%
71	10	429	76.1	0.999	62.2%
72	10	429	80.1	0.983	61.5%
73	10	459	80.1	0.999	60.4%
74	10	189	76.2	0.999	62.8%
75	10	204	74.2	0.983	56.7%
76	10	174	74.2	0.999	60.6%
77	10	204	78.2	0.999	59.2%
78	10	189	76.2	0.999	54.8%
79	10	174	78.2	0.983	59.8%
80	10	290	77.2	0.999	64.1%
81	10	305	75.2	0.983	65.8%
82	10	275	75.2	0.999	57.1%
83	10	305	79.2	0.999	62.4%
84	10	275	79.2	0.983	59.3%
85	10	290	77.2	0.999	58.3%
86	10	359	77.6	0.999	63.2%
87	10	374	75.6	0.983	61.3%
88	10	344	75.6	0.999	62.5%
89	10	359	77.6	0.999	60.9%
90	10	344	79.6	0.983	60.3%
91	10	374	79.6	0.999	63.8%
92	10	399	77.8	0.999	60.3%
93	10	421	78.0	0.999	61.5%

4. Optimization Results

4.1. Optimization Results after Optimization Step 2

Table S5. Observed maxima and maxima predicted by a linear response surface model through 40 fractional factorial design experiments.

Solvent	Observed Maximum				Predicted Maximum			
	t_{res} (s)	T (°C)	$C_0^{[a]}$	Yield	t_{res} (s)	T (°C)	$C_0^{[a]}$	Yield
DMSO	600	69	1.00	63%	144	64	1.00	197% ^[b]
DMF	600	69	1.00	56%	159	76	1.00	106%
DCE	190	120	1.00	53%	176	89	1.00	85%
Pyridine	190	120	1.00	51%	168	84	1.00	82%
<i>i</i> PrOH	600	120	0.44	43%	216	120	1.00	77%
THF	600	120	0.44	36%	216	120	1.00	77%
Toluene	190	120	0.44	44%	160	77	1.00	77%
DMC	600	120	1.00	57%	216	120	1.00	67% ^[c]
DME	190	120	0.44	42%	170	85	1.00	65% ^[c]
MeCN	600	120	1.00	47%	199	107	1.00	62% ^[c]

^[a]-4-methoxybenzyl chloride initial concentration. ^[b]-Estimated error $\pm 121\%$. ^[c]-Denotes predicted yield less than lower bound on maximum, indicating solvent will not be considered in next iteration of experiments.

Notice that by assuming similar trends among the continuous variables, our method was able to predict higher yields for solvents in other regions of the experimental space that were not considered during the fractional factorial DoE. For that reason, THF, with a maximum observed yield of only 36% in four slugs, ranked above DMC, with a maximum observed yield of 57%.

4.2. Optimization Results after Optimization Step 3

In addition to solvent hydrogen bond basicity, we correlated the results following optimization step to the dielectric constant of the solvent (Figure S3).⁵ Good qualitative agreement was only observed for DMSO, DMF, isopropanol, dimethoxyethane, and toluene.

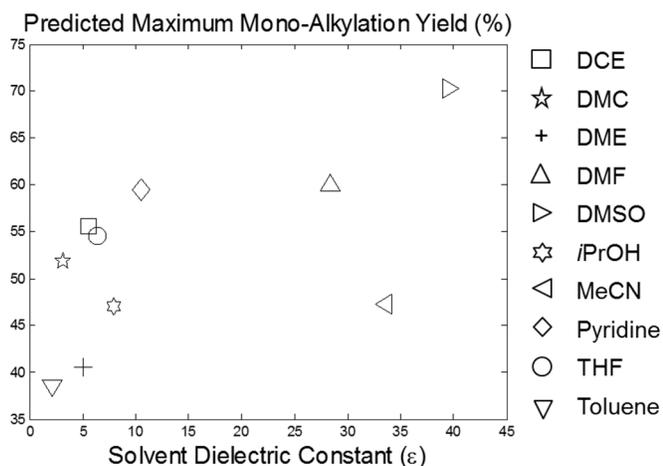
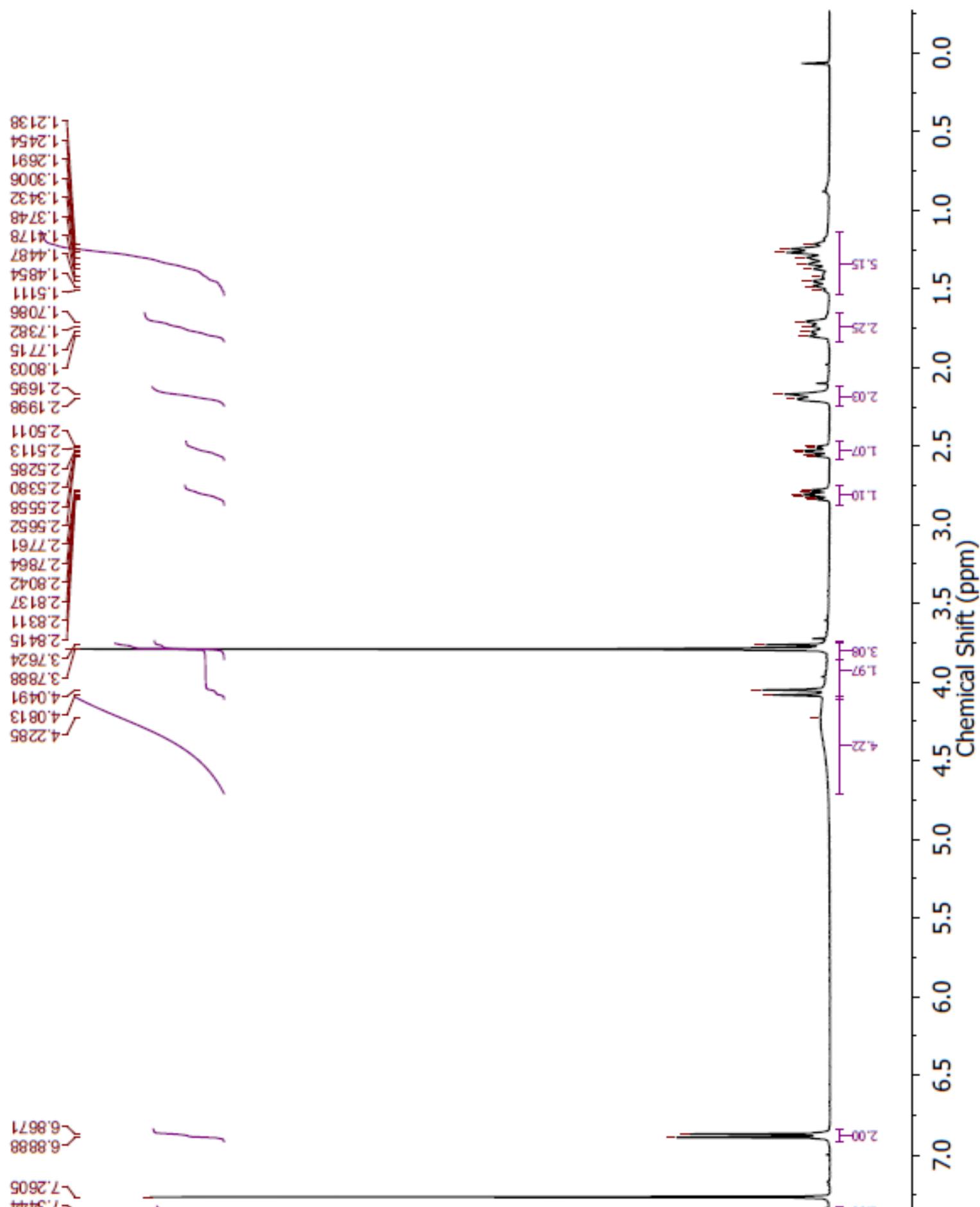
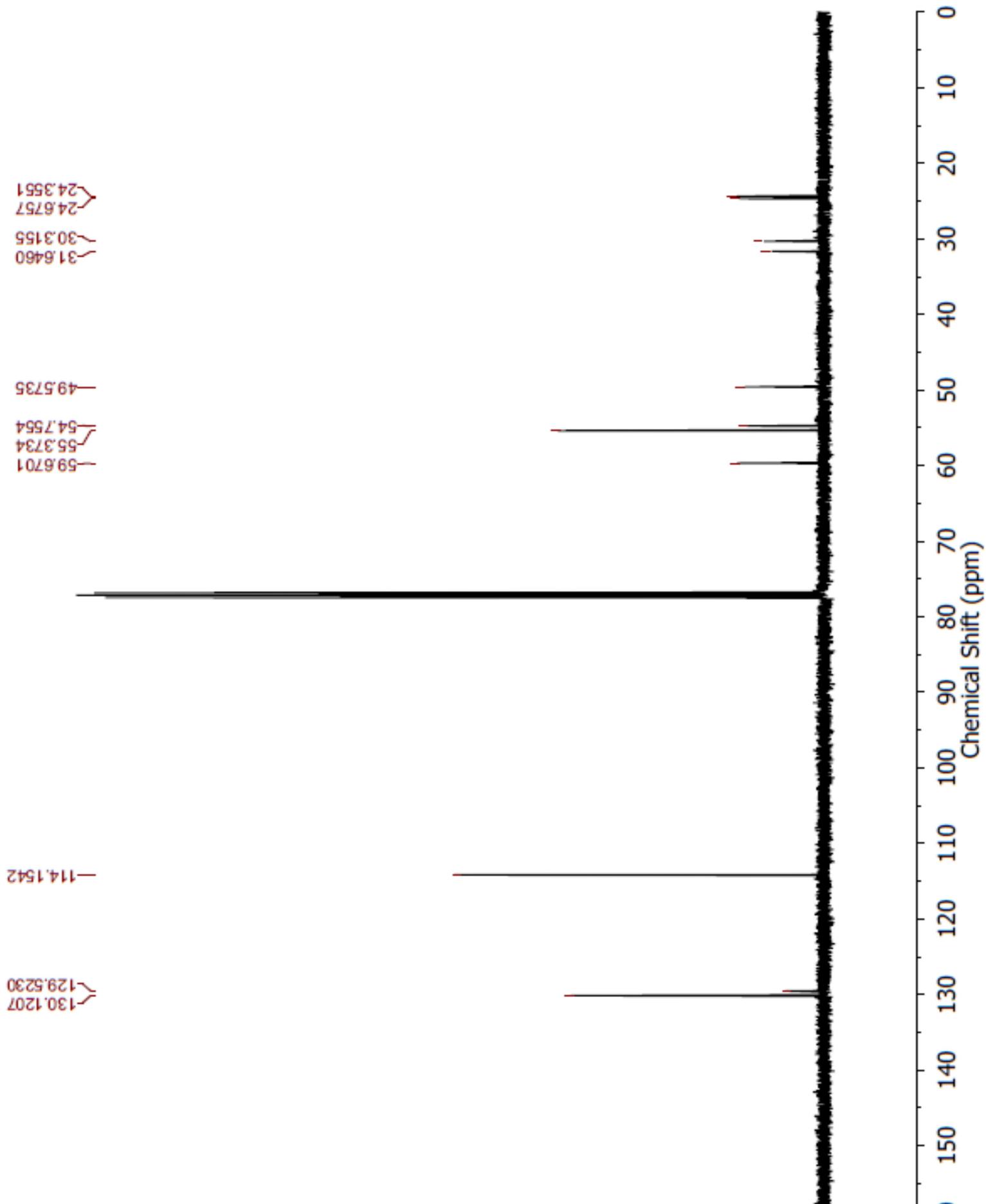


Figure S3. Correlation of maximum predicted yield through step 3 to the solvent dielectric constant.

5. NMR Spectra





6. Nomenclature

Latin	Definition
A	Matrix of quadratic and interaction response surface coefficients
A_s	Pre-exponential factor for solvent s
a_{ij}	Response surface coefficient for continuous variables i and j
a_{si}	Response surface coefficient for solvent s and continuous variable i
B_s	Candidate optimal response for solvent s
b	Vector of measured objective function values (responses) for all experiments
$\hat{\mathbf{b}}$	Vector of predicted objective function values (responses) for all experiments
b	Measured objective function (response) value
\hat{b}	Predicted objective function (response) value
$C_{0,Diamine}$	Initial concentration of <i>trans</i> -1,2-diaminocyclohexane
$C_{0,MeOBnCl}$	Initial concentration of 4-methoxybenzyl chloride
C_{Prod}	Concentration of desired product
c	Vector of linear response surface coefficients
c_i	Response surface coefficient for continuous variable i
c_s	Response surface coefficient for solvent s
E_s	Activation energy for solvent s
\mathbf{g}^k	Gradient at iteration k
\mathbf{H}_k	Hessian at iteration k
H_0	Null hypothesis
H_a	Alternative hypothesis
J^*	Optimal predicted response value (over all solvents)
J_s'	Optimal predicted response value at candidate optimum \mathbf{x}_s^*
J_s^*	Optimal predicted response value for solvent s
J_-^*	Lower bound on J^*
N	Number of experiments (single optimization iteration)
$N_{cont\ variables}$	Number of continuous variables
N_{expts}	Number of experiments (entire optimization)
N_{expts}'	Number of experiments used to calculate J_s'
N_{expts}^*	Number of experiments used to calculate J_s^*
N_{params}	Number of non-zero response surface coefficients (= length of θ)
n	Current experiment number
\mathbf{p}^k	Search direction for line search at iteration k
R	Gas constant
\mathbf{s}^k	Scaled search direction at iteration k
t_{crit}	Critical Student's t statistic for rejection of H_0
t_{res}	Residence time
t_{stat}	Student's t statistic
$t_{\alpha,\nu}$	Inverse Student's t value for $1-\alpha$ confidence and ν degrees of freedom
\mathbf{V}_B	Response covariance matrix
$\mathbf{V}_{\hat{\beta}}$	Prediction covariance matrix
\mathbf{V}_{J^*}	Prediction covariance matrix at the optimum \mathbf{x}^*
$\mathbf{V}_{J'}$	Prediction covariance matrix at candidate optimum \mathbf{x}_s
W	Weighting matrix
X	Matrix of scaled continuous and discrete variables for all experiments

\mathbf{x}	Vector of scaled continuous and discrete variables in a single experiment
\mathbf{x}_s	Vector of continuous and discrete variables for solvent s
\mathbf{x}_s^*	Optimal vector of continuous and discrete variables for solvent s
x_i	Scaled continuous variable i
x_i^*	Scaled continuous variable i in \mathbf{x}^*
x_s	Solvent composition of discrete variable s (1 if solvent s is active, 0 if inactive)
\mathbf{y}^k	Gradient difference at iteration k

Greek	Definition
α	Line search step size
α_{conf}	Rejection confidence level specified for Student's t -test
ΔJ_n^*	Error on the optimal response through experiment n
ΔJ_{pred}^*	Predicted error on the optimal response in next iteration of optimization
ΔJ_s^*	Error on the optimal response for solvent s
$\Delta \mathbf{x}$	Tolerance on the vector of optimal continuous variables
Δx_i	Tolerance on the optimum for continuous variable i
$\boldsymbol{\theta}$	Optimal vector of response surface model coefficients
ν	Degrees of freedom

7. References

1. M. A. Burns, C. H. Mastrangelo, T. S. Sammarco, F. P. Man, J. R. Webster, B. N. Johnsons, B. Foerster, D. Jones, Y. Fields, A. R. Kaiser and D. T. Burke, *Proc. Natl. Acad. Sci.*, 1996, **93**, 5556-5561.
2. A. C. Atkinson and A. N. Donev, *Optimum Experimental Design*, Clarendon Press; Oxford University Press, Oxford, New York, 1992.
3. J. V. Beck and K. J. Arnold, *Parameter Estimation in Engineering and Science*, Wiley, New York, 1977.
4. K. J. Beers, *Numerical Methods for Chemical Engineering: Applications in Matlab*, Cambridge University Press, New York, 2007.
5. *CRC Handbook of Chemistry and Physics*, 95th Edition edn., CRC Press, Cleveland, OH, 2015; M. A. Rivas and T. P. Iglesias, *J. Chem. Thermodynamics*, 2008, **40**, 1120-1130; L. Marcheselli, G. Pistoni, M. Tagliazucchi, L. Tassi and G. Tosi, *J. Chem. & Eng. Data*, 1993, **38**, 204-206.