Multivariate analysis of inline benchtop NMR data enables rapid optimization of a complex nitration in flow

Peter Sagmeister\textsuperscript{a,b}, Johannes Poms\textsuperscript{c}, Jason D. Williams\textsuperscript{*a,b} and C. Oliver Kappe\textsuperscript{*a,b}

\textsuperscript{a}Center for Continuous Flow Synthesis and Processing (CCFLOW), Research Center Pharmaceutical Engineering GmbH (RCPE), Inffeldgasse 13, Graz, Austria

\textsuperscript{b}Institute of Chemistry, University of Graz, Heinrichstrasse 28, 8010 Graz, Austria.

\textsuperscript{c}Research Center Pharmaceutical Engineering GmbH (RCPE), Inffeldgasse 13, Graz, Austria

Email: jason.williams@rcpe.at, oliver.kappe@uni-graz.at
Table of Contents

1 General Information ........................................................................................................... 3
  1.1 Materials and Flow Equipment ...................................................................................... 3
  1.2 High Field NMR ........................................................................................................... 3

2 General Flow Configuration .................................................................................................. 3
  2.1 HiTec Zang Control Unit and LabVision Software ....................................................... 6
  2.1.1 HiText Script for High-Pressure Shutdown .............................................................. 6
  2.1.2 HiText Script for Temperature Ramp ...................................................................... 7
  2.2 Reactor Selection ......................................................................................................... 7

3 NMR .................................................................................................................................. 8
  3.1 General Reaction Monitoring NMR .............................................................................. 8
  3.2 Process Integration ....................................................................................................... 8
  3.3 General Overview, Pulse Sequence and Code ............................................................. 9
  3.4 Residence Time Distribution (RTD) Experiments ....................................................... 10
  3.5 Development of Multivariate Models .......................................................................... 12
    3.5.1 Measurements of Training Set and Validation Set Solutions .............................. 12
    3.5.2 MATLAB Script .................................................................................................. 13
    3.5.3 Building of Model .............................................................................................. 13
    3.5.4 Model Robustness / Validation Set .................................................................... 19
  3.6 Monitoring Reaction Optimization by Inline NMR ................................................... 21
    3.6.1 Statistical Approach to Remove Outliers ............................................................ 21
    3.6.2 Experimental Conditions DoE ............................................................................. 22
    3.6.3 Temperature and Pressure data for DoE ............................................................... 22
    3.6.4 Reaction Optimization Data Model Fitting ........................................................ 23
    3.6.5 Experimental Conditions Dynamic Experiment .................................................. 30
    3.6.6 Temperature and Pressure Data from Dynamic Experiment (Slow Ramp) ........ 30
    3.6.7 Offline Analysis of Dynamic Experiment (Slow Ramp) ..................................... 31
    3.6.8 Temperature and Pressure Data of Dynamic Experiment (Fast Ramp) ............. 31
    3.6.9 NMR Results for the Fast Ramp of the Dynamic Experiment ......................... 32

4 UPLC-DAD ........................................................................................................................ 33
  4.1 Chromatographic Conditions ....................................................................................... 33
  4.2 Offline Calibration ......................................................................................................... 34

5 References ......................................................................................................................... 36

6 NMR Spectra 43 MHz ...................................................................................................... 37

7 NMR Spectra 300 MHz ..................................................................................................... 41
1 General Information

1.1 Materials and Flow Equipment

Solvents and chemicals were obtained from commercial suppliers and were used without any further purification unless otherwise noted. In the flow setup, standard PFA tubing (0.8 mm or 1.6 mm i.d.), fittings, T-pieces manufactured from PTFE or PEEK were used as connectors. The back pressure regulator was obtained from Upchurch Scientific.

1.2 High Field NMR

NMR spectra were recorded on a Bruker 300 MHz instrument. $^1$H spectra were recorded at 300 MHz, respectively, with a chemical shift (δ) relative to the methyl group (3.31 ppm) of methanol-$d_4$ expressed in parts per million. The letters s, d, dd, td, t, and m are used to indicate singlet, doublet, doublet of doublets, triplet of doublets, multiplet respectively.

Individual spectra of SA, 3-NSA, 5-NSA and DNSA were recorded in d-4 methanol.

SA:

$^1$H NMR (300 MHz, Methanol-$d_4$) δ 7.85 (dd, $J = 7.9, 1.8$ Hz, 1H), 7.45 (td, $J = 8.7, 7.2, 1.8$ Hz, 1H), 6.95 – 6.84 (m, 2H).

3-NSA:

$^1$H NMR (300 MHz, Methanol-$d_4$) δ 8.18 (dd, $J = 7.8, 1.8$ Hz, 1H), 8.11 (dd, $J = 8.1, 1.8$ Hz, 1H), 7.04 (t, $J = 8.0$ Hz, 1H).

5-NSA:

$^1$H NMR (300 MHz, Methanol-$d_4$) δ 8.72 (d, $J = 2.9$ Hz, 1H), 8.34 (dd, $J = 9.2, 2.9$ Hz, 1H), 7.09 (d, $J = 9.2$ Hz, 1H).

DNSA:

$^1$H NMR (300 MHz, Methanol-$d_4$) δ 8.97 (d, $J = 2.9$ Hz, 1H), 8.94 (d, $J = 2.9$ Hz, 1H).

2 General Flow Configuration

Input solutions were made up with conc. H$_2$SO$_4$ (95 %) in 250 mL or 500 mL volumetric flasks.

0.5 M SA: In a 250 mL volumetric flask 17.264 g of SA was diluted in conc. H$_2$SO$_4$ (95 %).

0.6 M HNO$_3$: A 500 mL volumetric flask was placed in an ice bath and filled with 400 mL conc. H$_2$SO$_4$ (95 %), then 19.2 mL of conc. HNO$_3$ (15.6 M, 68%) was slowly added. After the addition the volumetric flask was removed from the ice bath and filled up to the 500 mL mark with conc. H$_2$SO$_4$ (95 %).

The reaction was performed in a Modular MicroReaction System (MMRS) from Ehrfeld Mikrotechnik, the phase separation was accomplished with an inline phase separator (SEP-10) from Zaiput and the reaction stream was analyzed by a 43 MHz Spinsolve Ultra benchtop NMR from Magritek (Fig. S1). The flow setup is shown in Fig. S1. The HNO$_3$ feed and SA feed were delivered with two SyrDos2 pumps (30 bar valve, 5 mL syringes, pump 1 for SA feed and pump 2 for HNO$_3$ feed) through PFA
tubing (1.6 mm i.d.) to the MMRS system. Both feeds entered the system through 1/8” input connectors (0711-2-0224-F, Hastelloy C-276), followed by pressure sensor modules (0518-1-60x4-F, Hastelloy C-276), coax heat exchangers (0309-4-0004-F, Hastelloy C-276), then temperature sensors (0501-2-1004-X, Hastelloy C-276) to the MMRS system. The two streams were mixed in a temperature controlled cascade mixer 06 (0216-3-0014-F, mixing structure 10 µL, Hastelloy C-276) and delivered through a temperature dividing block to the Lonza FlowPlate Lab (1701-3-0004-F, Hastelloy C-276) where the reaction solution was quenched with premixed water and iPrOAc. The total reactor volume for the nitration was 343 µL. The temperature in the heat exchangers (SA and HNO₃ feeds) and the cascade reactor was controlled by thermostat 1 (Huber, Minisat 240). The water feed was pumped with a SyrDos 2 pump (30 bar valve, 5 mL syringes) through PTFE tubing (1.6 mm i.d.) and a check valve (Upchurch, CV-3321) to the MMRS system where it entered through a 1/8” input connector (0711-2-0224-F, Hastelloy C-276), pressure sensor module (0518-1-60x4-F, Hastelloy C-276) and coax heat exchanger (0309-4-0004-F, Hastelloy C-276) into the FlowPlate Lab. The iPrOAc feed was delivered with a SyrrDos 2 pump (90 bar valve, 2.5 mL syringes) through PTFE tubing (0.8 mm i.d.) and a check valve (Upchurch, CV-3321) to the MMRS system where it was connected with a 1/16” input connector (0711-2-0124-F, Hastelloy C-276) to the FlowPlate Lab. The temperature in heat exchanger and the FlowPlate Lab was controlled by thermostat 2 (Huber, Minisat 240). The FlowPlate Lab was equipped with an LL design Process Plate (1701-4682-HC, LL-Mixer, Nominal Width 0.2 mm, Hastelloy C-22). The iPrOAc stream entered at port 1, the water stream at port 2 and the process stream from the nitration at port 3 and the quenched stream exit the process plate at port 7 (Fig. S2). The quenched stream left the MMRS system through a temperature sensor module (0501-2-1004-X, Hastelloy C-276), 1/8” input connector (0711-2-0224-F, Hastelloy C-276) and a short PTFE tubing (150 mm, 3.2 mm) connected to the Zaiput separator (SEP-10) equipped with a hydrophobic PTFE membrane (0.5 µm, OB-500-S10). The aqueous stream was diverted to a waste bottle and the organic phase stream was delivered through PFA tubing (1/16” o.d., 0.8 mm i.d.) to the benchtop NMR. A backpressure regulator (Upchurch cartridge holder (P-465) equipped with a 2.8 bar (blue, P-761) cartridge) was installed prior the glass flow cell (800 µL internal volume, 550 mm length) for the NMR for safety reasons (Fig. S5). Fractions for offline analysis were collected after the NMR in 15 mL PP tubes. The detailed flow rates and temperature settings for the experiments are listed in section 3.6.
Fig. S1 Photograph of reaction setup in the lab, showing the following equipment: A: Pumps, B: MMRS, C: Thermostat 1, D: Thermostat 2, E: liquid-liquid separator, F: NMR, G: LabManager control box.

Fig. S1 The process plates with the LL design. Inputs are colored in green and outputs in red.
2.1 **HiTec Zang Control Unit and LabVision Software**

The pumps and thermostats were controlled via RS232 interface by a control module (HiTec Zang, LabManager) and its associated software (HiTec Zang, LabVision). Temperature and pressure probes were connected to the HiTec Zang LabManager and the data were recorded in the LabVision software. Data points were collected every 1 second. A HiText script for automatically shutting off the pumps, if the pressure exceeded 15 bar due to a blockage, was running constantly in the background of LabVision Software. The temperature ramp for the dynamic experiment was programmed and executed by a HiText script.

![Schematic of the process in LabVision.](image)

**Fig. S 2** Schematic of the process in LabVision.

2.1.1 **HiText Script for High-Pressure Shutdown**

```plaintext
begin
if !P_1 > 15 then
    !PUMP_1.ON=0
    !PUMP_2.ON=0
    !PUMP_3.ON=0
    !PUMP_4.ON=0
else if !P_2 > 15 then
    !PUMP_1.ON=0
    !PUMP_2.ON=0
    !PUMP_3.ON=0
    !PUMP_4.ON=0
```

else if !P_3 > 15 then
    !PUMP_1.ON=0
    !PUMP_2.ON=0
    !PUMP_3.ON=0
    !PUMP_4.ON=0
end if

goto begin

2.1.2 HiText Script for Temperature Ramp

begin
    if  !THERMOSTAT.TS < 35 then
        !THERMOSTAT.TS = !THERMOSTAT.TS + 0.01
    end if
    wait 1 sec

goto begin

2.2 Reactor Selection

Fig. S 3 Conversion of SA depending on flow rate and different mixers. The Lonza Flow Plate with the LL mixing structure process plate shows for low flow rates (<2.5 mL/min) poorer performance compared to the cascade mixer.
3 NMR

3.1 General Reaction Monitoring NMR

Inline NMR reaction monitoring was accomplished by recording $^1$H spectra using a low field benchtop 43.795 MHz NMR (Magritek, Spinsolve Ultra). Shims were performed with the flow cell filled with iPrOAc and referenced to the methyl group of the acetate moiety of iPrOAc at 1.85 ppm. Typically, a “QUICKSHIM: ALL” was performed in the Spinsolve software (Magritek) and shim values were usually below 0.6 Hz for width at 50 %, below 9 Hz for width at 0.55 % and a signal to noise ratio above 12,000.

3.2 Process Integration

![Diagram of NMR integration](image)

**Fig. S 4** A detailed overview and photograph of the integration of the NMR into the process. The glass flow through cell had an internal volume of 800 µL and a length of 550 mm.
3.3 General Overview, Pulse Sequence and Code

![Flow diagram of script used for inline NMR monitoring.]

**Fig. S 5** Flow diagram of script used for inline NMR monitoring.

**Code for Reaction Monitoring $^1$H Loop**

```python
TimeStampFolder = "c:/ReactionMonitor/XXX"
ProtonPhase = "FirstScan"

# Loop
loop(6000, 00:00:01:000)
  TimeStampFolder = "c:/ReactionMonitor/XXX/1H"
  RunProtocol("1D EXTENDED+", ["Number=1", "RepetitionTime=2", "PulseAngle=90",
  "AcquisitionTime=1.6"])
  wait(00:00:00:000)
endloop

wait(00:00:10:000)
RunMnovaFile("ScriptUtilities/ReactionMonitor.qs", "process", ["1H"])
```
3.4 Residence Time Distribution (RTD) Experiments

RTD experiments were performed in the setup described in section 2. Modifications were as followed instead of the 0.5 M SA solution a 0.2 M nitrobenzene in conc. H₂SO₄ (95 %) (pump 1) was used. Instead of the 0.6 M HNO₃ a conc. H₂SO₄ (95 %) solution was used (pump 2). At the beginning of each experiment pump 1, pump 2, pump 3 and pump 4 were set to the investigated flow rates (Table S1), whereas only pump 2 was switched on and pump 1 switched off. The step was induced by switching on pump 1 and switching off pump 2, vice versa for the step down. Proton spectra were acquired with single scan, a 90° pulse, 3.2 s acquisition time, and 4 s repetition time. The stacked proton spectra were processed (Fourier transform, phasing and baseline correction) in Mestrenova (v11, Mestrelab Research) and the total integral was obtained by integration of peaks between 7.1 and 8.5 ppm (corresponds to 5 aromatic protons of nitrobenzene). The experimental data were fitted with the transfer function Pₛ (Fig. S7). The concentration cᵢₙ and cᵪₒᵤₜ are the Laplace-transformed concentrations at the inlet and outlet of the reactor. The time constants T₁ and T₂ and dead time Tₜ (time delay from starting the stream of tracer until the tracer appears at the NMR) and the scaling factor K were identified using experimental data.

\[
Pₛ = \frac{cₒᵤₜ}{cᵢₙ} = K \cdot \frac{1}{1 + sT₁} \cdot \frac{1}{1 + sT₂} \cdot e^{-sTₜ}
\]

Fig. S 6 Transfer function to fit the experimental data.

| Entry | Pump 1 [mL·min⁻¹] | Pump 2 [mL·min⁻¹] | Pump 3 [mL·min⁻¹] | Pump 4 [mL·min⁻¹] |
|-------|------------------|------------------|------------------|------------------|
| 1     | 0.8              | 0.8              | 4.0              | 0.8              |
| 2     | 1.1              | 1.1              | 5.5              | 1.1              |
| 3     | 1.6              | 1.6              | 8.0              | 1.6              |
| 4     | 2.4              | 2.4              | 12.0             | 2.4              |

Table S 1 Flow rates for each pump for the RTD experiments. Pump 1 was delivering the 0.2 M nitrobenzene in H₂SO₄ feed, pump 2 the H₂SO₄ stream, pump 3 the water feed and pump 4 the iPrOAc stream.

Fig. S 7 Obtained NMR data and fitted curves using the transfer function for the RTD for entry 1.
Obtained NMR data and fitted curves using the transfer function for the RTD for entry 2.

Obtained NMR data and fitted curves using the transfer function for the RTD for entry 3.

Obtained NMR data and fitted curves using the transfer function for the RTD for entry 4.
3.5 Development of Multivariate Models

3.5.1 Measurements of Training Set and Validation Set Solutions

Training and validation set solutions were prepared by weighing the correct amounts of SA, 3-NSA, 5-NSA and DNSA into either a 10 mL or 25 mL volumetric flask and filled up with iPrOAc to the mark (Table S2). The solutions were typically sonicated and stored in the fridge prior usage. The general measuring procedure was as followed. The Knauer Azura HPLC pump, tubing and flow cell were purged with iPrOAc prior to flushing the system with 10 mL of the training or validation solution to avoid cross contamination. Then the solution was circulated through the NMR with the corresponding flow rate of 1.1 mL/min (Fig. S12).

| Entry   | SA [g] | 3-NSA [g] | 5-NSA [g] | DNSA [g] | iPrOAc [mL] | SA [mM] | 3-NSA [mM] | 5-NSA [mM] | DNSA [mM] |
|---------|--------|-----------|-----------|----------|-------------|--------|------------|------------|----------|
| Level_1 | 0      | 0.0466    | 0.045     | 0        | 25          | 0      | 0          | 10.2       | 9.8      | 0.0      |
| Level_2 | 0.6893 | 0         | 0         | 1.1419   | 25          | 199.6  | 0          | 0.0        | 0.0      | 200.2    |
| Level_3 | 0.5223 | 0.1140    | 0.1178    | 0.8557   | 25          | 151.3  | 24.9       | 25.7       | 150.1    |
| Level_4 | 0.3467 | 0.2280    | 0.2257    | 0.5774   | 25          | 100.4  | 49.8       | 49.3       | 101.2    |
| Level_5 | 0.1720 | 0.3396    | 0.3424    | 0.2870   | 25          | 49.8   | 74.2       | 74.8       | 50.3     |
| Level_6 | 0.0885 | 0.4548    | 0.4562    | 0.1400   | 25          | 25.6   | 99.3       | 99.7       | 24.5     |
| Level_7 | 0.0374 | 0.5550    | 0.5550    | 0.0571   | 25          | 10.8   | 121.2      | 121.2      | 10.0     |
| Level_8 | 0.0560 | 0.1288    | 0.1653    | 0.1415   | 10          | 40.5   | 70.3       | 90.3       | 62.0     |
| Level_9 | 0.1610 | 0.1632    | 0.1044    | 0.1877   | 10          | 116.6  | 89.1       | 57.0       | 82.3     |
| Level_10| 0.0329 | 0.1011    | 0.0531    | 0.2559   | 10          | 23.8   | 55.2       | 29.0       | 112.2    |
| Val_1   | 0.1045 | 0.1992    | 0.2048    | 0.1606   | 10          | 75.7   | 108.8      | 111.8      | 70.4     |
| Val_2   | 0.1810 | 0.1520    | 0.0658    | 0.3137   | 10          | 131.0  | 83.0       | 35.9       | 137.5    |
| Val_3   | 0.2367 | 0.0638    | 0.1231    | 0.3966   | 10          | 171.4  | 34.8       | 67.2       | 173.9    |
| Val_4   | 0.1596 | 0.2143    | 0.2126    | 0.2663   | 25          | 46.8   | 46.1       | 46.4       | 46.7     |

Table S2 Overview of the prepared solutions for the training set (Level_1 to Level_10) and validation set (Val_1 to Val_4). Note: Level 2 was not included in 3-NSA and 5-NSA models. Level 1 was not included in the DNSA model.

Fig. S11 Setup for the NMR measurements, showing the NMR, Knauer HPLC pump and vessel with prepared training and validation set solutions.
3.5.2 MATLAB Script

To remove the influence of minor shifts in NMR spectra, a MATLAB script was set up, to find the highest peak between 8.75 and 8.95 ppm (DNSA peak), and set it to 8.89. The script also interpolates the spectra, to remove influence of variations in the recorded points on the ppm scale.

```matlab
% Matlab script to perform alignment and interpolation of NMR data.
% Author: Peter Sagmeister
% Matlab Version: 2017B
% define parameters
ppmstart = 6.0;        %[ppm] set start of ppm scale
ppmEnd = 9.1;          %[ppm] set end of ppm scale
spacing = 0.002;       %[ppm] set spacing of ppm scale
ref = 8.89;            %[ppm] reference number for alignment
peakfind_scale = [8.75 ; 8.95]; %[ppm] set range for finding highest peak

% Reading in text file
fileID = readtable('XXX.txt');
fileID = width(fileID) - 1;
% convert to double
A = table2array(fileID(:,1:wfileID));
%transpose data
nmr_data = A';
sizeX = size(nmr_data);
% define X-variable values
X = nmr_data(1:1,1:sizeX(2));
%Set ppm range and grid for interpolation
Xq = (ppmstart:spacing:ppmEnd);
%define result file
nmr_inpol = Xq;
% loop for interpolation of all experiments in text file
%peak_find_min
[val,id_min] = min(abs(X-peakfind_scale(1)));
Valmin=X(id_min);
%peak_find_max
[val,id_max] = min(abs(X-peakfind_scale(2)));
Valmax=X(id_max);
%loop to go through each spectra
for i = [2:sizeX(1)]
    t = X(1,i:id_min:id_max);
    maxval = 0;
    for ii = 1:length(t)
        find_y = nmr_data(i:i,id_min:id_max);
        if find_y(ii) > maxval
            id = ii;
            maxval = find_y(ii);
        end
    end
    Xspac = ref - t(id);
    Xttemp = X + Xspac;
    Y= nmr_data(1:1:);
    output=interp1(Xtemp,Y,Xq,'linear','extrap');
end
% copy data together
nmr_inpol = [nmr_inpol ; output];

3.5.3 Building of Model

MVA models were built in Simca (Version 16.0.1, Umetrics). The columns for the PLS components are as follows:

Component component index
R2X fraction of X variation modeled in the component
R2X(cum) cumulative R2X up to the specified component
**Eigenvalue** the minimum number of observations (N) and variables (K) multiplied by R²X

**R²Y** fraction of Y variation modeled in the component

**R²Y(cum)** cumulative R²Y up to the specified component

**Q²** overall cross-validated R² for the component

**Limit** critical value of Q² under which the component is insignificant according to CV rule 1

**Q²(cum)** cumulative Q² up to the specified component. Note that unlike R²(cum), Q²(cum) is not additive

**Significance** CV insignificant (NS) or significant according to rule Rz

**Iterations** the number of NIPALS iterations until convergence

The columns for the OPLS components are as follows:

**Component** component index

**R²X** fraction of X variation modeled in the component, using the X model

**R²X(cum)** cumulative R²X up to the specified component

**Eigenvalue** the minimum number of observations (N) and variables (K) multiplied by R²X

**R²** fraction of Y variation modeled by X in the component, using the X model

**R²(cum)** cumulative R² up to the specified component

**Q²** overall cross-validated R² for the component

**Limit** critical value of Q² under which the component is insignificant according to CV rule 1

**Q²(cum)** cumulative Q² up to the specified component. Note that unlike R²X(cum), Q²(cum) is not additive

**R²Y** fraction of Y variation modeled by Y in the component, using the Y model

**R²Y(cum)** cumulative R²Y up to the specified component

**EigenvalueY** the minimum number of observations (N) and variables (M) multiplied by R²Y

**Significance** CV insignificant (NS) or significant according to rule Rz

### 3.5.3.1 Manuscript Table 1, Entry 1

The PLS model consisted of 4 components and was trained with a total of 1320 observations (N) and 1067 variables (K).

| Component | R²X   | R²X(cum) | Eigenvalue | R²Y    | R²Y(cum) | Q²   | Limit | Q²(cum) | Significance | Iterations |
|-----------|-------|----------|------------|--------|----------|------|-------|---------|--------------|------------|
| 0         | Cent. | 0.385    | 0.385      | 411    | 0.677    | 0.677| 0.624 | 0.624   | R1           | 7          |
| 1         | 0.129 | 0.514    | 138        | 0.255  | 0.932    | 0.689| 0.683 | 0.88    | R1           | 4          |
| 2         | 0.0569| 0.571    | 60.7       | 0.0323 | 0.964    | -0.0232| 0     | 0.88    | R2           | 16         |
| 3         | 0.0994| 0.67     | 106        | 0.0108 | 0.975    | 0.0656| 0     | 0.888   | R1           | 6          |
| 4         |       |          |            |        |          |      |       |         |              |            |
3.5.3.2 Manuscript Table 1, Entry 2

SA

The PLS model consisted of 4 components and was trained with a total of 1320 observations (N) and 501 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2Y | R2Y(cum) | Q2 | Limit | Q2(cum) | Significance | Iterations |
|------------|-----|----------|------------|-----|----------|----|-------|----------|--------------|------------|
| 0          |     |          |            |     |          |    |       |          |              |            |
| 1          | 0.62| 0.62     | 310        | 0.947| 0.947    | 0.93| 0     | 0.93     | RN1         | 1          |
| 2          | 0.0948| 0.714  | 47.4       | 0.0401| 0.987    | 0.527| 0     | 0.967    | RN1         | 1          |
| 3          | 0.0572| 0.772  | 28.6       | 0.00731| 0.995    | 0.43 | 0     | 0.981    | N4          | 1          |
| 4          | 0.0235| 0.795  | 11.7       | 0.00294| 0.998    | 0.216| 0     | 0.985    | N4          | 1          |

3-NSA

The PLS model consisted of 4 components and was trained with a total of 1320 observations (N) and 552 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2Y | R2Y(cum) | Q2 | Limit | Q2(cum) | Significance | Iterations |
|------------|-----|----------|------------|-----|----------|----|-------|----------|--------------|------------|
| 0          |     |          |            |     |          |    |       |          |              |            |
| 1          | 0.55| 0.55     | 303        | 0.756| 0.756    | 0.659| 0     | 0.659    | RN1         | 1          |
| 2          | 0.164| 0.714 | 90.1       | 0.225| 0.982    | 0.814| 0     | 0.936    | RN1         | 1          |
| 3          | 0.0753| 0.789 | 41.5       | 0.00741| 0.989    | 0.208| 0     | 0.95     | N4          | 1          |
| 4          | 0.0258| 0.815  | 14.2       | 0.00494| 0.994    | 0.315| 0     | 0.965    | N4          | 1          |

5-NSA

The PLS model consisted of 4 components and was trained with a total of 1320 observations (N) and 701 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2Y | R2Y(cum) | Q2 | Limit | Q2(cum) | Significance | Iterations |
|------------|-----|----------|------------|-----|----------|----|-------|----------|--------------|------------|
| 0          |     |          |            |     |          |    |       |          |              |            |
| 1          | 0.475| 0.475   | 333        | 0.803| 0.803    | 0.719| 0     | 0.719    | RN1         | 1          |
| 2          | 0.156| 0.632   | 110        | 0.159| 0.962    | 0.634| 0     | 0.897    | RN1         | 1          |
| 3          | 0.0681| 0.7    | 47.7       | 0.0157| 0.978    | -0.175| 0     | 0.867    | N5          | 1          |
| 4          | 0.0738| 0.774  | 51.6       | 0.0115| 0.989    | 0.385| 0     | 0.93     | RN1         | 1          |

DNSA

The PLS model consisted of 4 components and was trained with a total of 1320 observations and 67 (N) variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2Y | R2Y(cum) | Q2 | Limit | Q2(cum) | Significance | Iterations |
|------------|-----|----------|------------|-----|----------|----|-------|----------|--------------|------------|
| 0          |     |          |            |     |          |    |       |          |              |            |
| 1          | 0.697| 0.697   | 46.7       | 0.903| 0.903    | 0.877| 0     | 0.877    | RN1         | 1          |
| 2          | 0.132| 0.829   | 8.88       | 0.0806| 0.987    | 0.611| 0     | 0.977    | RN1         | 1          |
| 3          | 0.0963| 0.926 | 6.45       | 0.00563| 0.993    | 0.406| 0     | 0.986    | N4          | 1          |
| 4          | 0.00734| 0.933 | 0.492      | 0.00428| 0.997    | 0.514| 0     | 0.993    | N4          | 1          |
3.5.3.3  Manuscript Table 1, Entry 3

SA

The OPLS model consisted of 1+4+0 components and was trained with a total of 1320 observations (N) 501 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-----------|-----|----------|------------|----|---------|----|-------|---------|----|----------|------------|--------------|
| Model     | 0.637 | 0.998 | 0.995 | 1 |
| Predictive | 0.585 | 0.998 | 0.995 | 1 |
| P1 | 0.585 | 0.585 | 292 | 0.998 | 0.998 | 0.995 | 0.01 | 0.995 | 1 | 1 | 1 | R1 |
| Orthogonal in X (OPLS) | 0.253 | 0 |
| O1 | 0.123 | 0.123 | 61.5 | 0 | 0 | R1 |
| O2 | 0.0619 | 0.165 | 30.9 | 0 | 0 | R1 |
| O3 | 0.0253 | 0.21 | 12.6 | 0 | 0 | NS |
| O4 | 0.0424 | 0.253 | 21.2 | 0 | 0 | NS |

3-NSA

The OPLS model consisted of 1+5+0 components and was trained with a total of 1320 observations (N) 552 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-----------|-----|----------|------------|----|---------|----|-------|---------|----|----------|------------|--------------|
| Model     | 0.862 | 0.996 | 0.99 | 1 |
| Predictive | 0.381 | 0.996 | 0.99 | 1 |
| P1 | 0.381 | 0.381 | 210 | 0.996 | 0.996 | 0.99 | 0.01 | 0.995 | 1 | 1 | 1 | R1 |
| Orthogonal in X (OPLS) | 0.48 | 0 |
| O1 | 0.326 | 0.326 | 180 | 0 | 0 | R1 |
| O2 | 0.0765 | 0.405 | 43.3 | 0 | 0 | R1 |
| O3 | 0.0278 | 0.433 | 15.3 | 0 | 0 | R1 |
| O4 | 0.0374 | 0.47 | 20.6 | 0 | 0 | NS |
| O5 | 0.0103 | 0.48 | 5.68 | 0 | 0 | NS |

5-NSA

The OPLS model consisted of 1+6+0 components and was trained with a total of 1320 observations (N) 701 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-----------|-----|----------|------------|----|---------|----|-------|---------|----|----------|------------|--------------|
| Model     | 0.636 | 0.997 | 0.99 | 1 |
| Predictive | 0.358 | 0.997 | 0.99 | 1 |
| P1 | 0.358 | 0.358 | 251 | 0.997 | 0.997 | 0.99 | 0.01 | 0.995 | 1 | 1 | 1 | R1 |
| Orthogonal in X (OPLS) | 0.478 | 0 |
| O1 | 0.26 | 0.26 | 182 | 0 | 0 | R1 |
| O2 | 0.0737 | 0.334 | 51.6 | 0 | 0 | R1 |
| O3 | 0.0786 | 0.413 | 55 | 0 | 0 | R1 |
| O4 | 0.039 | 0.451 | 27.3 | 0 | 0 | R1 |
| O5 | 0.00731 | 0.459 | 5.11 | 0 | 0 | R1 |
| O6 | 0.0189 | 0.478 | 13.3 | 0 | 0 | NS |

DNSA

The OPLS model consisted of 1+4+0 components and was trained with a total of 1320 observations (N) 67 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-----------|-----|----------|------------|----|---------|----|-------|---------|----|----------|------------|--------------|
| Model     | 0.942 | 0.998 | 0.998 | 1 |
| Predictive | 0.62 | 0.998 | 0.998 | 1 |
| P1 | 0.62 | 0.62 | 41.5 | 0.998 | 0.998 | 0.998 | 0.01 | 0.998 | 1 | 1 | 1 | R1 |
| Orthogonal in X (OPLS) | 0.322 | 0 |
| O1 | 0.202 | 0.202 | 13.5 | 0 | 0 | R1 |
| O2 | 0.1 | 0.302 | 6.73 | 0 | 0 | NS |
| O3 | 0.01 | 0.312 | 0.672 | 0 | 0 | NS |
| O4 | 0.0103 | 0.322 | 0.687 | 0 | 0 | NS |
3.5.3.4  Manuscript Table 1, Entry 4

SA

The OPLS model consisted of 1+4+0 components and was trained with a total of 1218 observations (N) 344 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-----------|-----|----------|------------|----|---------|----|-------|---------|----|---------|-------------|-------------|
| Model     | 0.912 | 0.999 | 0.996 | 1 |
| Predictive | 0.835 | 0.998 | 0.996 | 1 |
| P1        | 0.835 | 0.998 | 0.996 | 1 |

Orthogonal in X (OPLS)

|            | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-------------|-----|----------|------------|----|---------|----|-------|---------|----|---------|-------------|-------------|
| O1          | 0.0543 | 0.0543 | 18.6 | 0 | 0 |
| O2          | 0.0162 | 0.0705 | 5.57 | 0 | 0 |
| O3          | 0.00365 | 0.0743 | 1.32 | 0 | 0 |
| O4          | 0.00297 | 0.0773 | 1.02 | 0 | 0 |

3-NSA

The OPLS model consisted of 1+5+0 components and was trained with a total of 1319 observations (N) 484 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-----------|-----|----------|------------|----|---------|----|-------|---------|----|---------|-------------|-------------|
| Model     | 0.833 | 0.997 | 0.992 | 1 |
| Predictive | 0.426 | 0.997 | 0.992 | 1 |
| P1        | 0.426 | 0.997 | 0.992 | 1 |

Orthogonal in X (OPLS)

|            | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-------------|-----|----------|------------|----|---------|----|-------|---------|----|---------|-------------|-------------|
| O1          | 0.25 | 0.25 | 121 | 0 | 0 |
| O2          | 0.0954 | 0.346 | 46.1 | 0 | 0 |
| O3          | 0.0467 | 0.393 | 22.6 | 0 | 0 |
| O4          | 0.00965 | 0.402 | 4.66 | 0 | 0 |
| O5          | 0.00037 | 0.408 | 2.59 | 0 | 0 |

5-NSA

The OPLS model consisted of 1+6+0 components and was trained with a total of 1319 observations (N) 709 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-----------|-----|----------|------------|----|---------|----|-------|---------|----|---------|-------------|-------------|
| Model     | 0.803 | 0.998 | 0.992 | 1 |
| Predictive | 0.398 | 0.998 | 0.992 | 1 |
| P1        | 0.398 | 0.998 | 0.992 | 1 |

Orthogonal in X (OPLS)

|            | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-------------|-----|----------|------------|----|---------|----|-------|---------|----|---------|-------------|-------------|
| O1          | 0.159 | 0.159 | 113 | 0 | 0 |
| O2          | 0.103 | 0.262 | 72.8 | 0 | 0 |
| O3          | 0.108 | 0.37 | 76.3 | 0 | 0 |
| O4          | 0.024 | 0.394 | 17 | 0 | 0 |
| O5          | 0.00631 | 0.4 | 4.46 | 0 | 0 |
| O6          | 0.00444 | 0.404 | 3.14 | 0 | 0 |

DNSA

The OPLS model consisted of 1+4+0 components and was trained with a total of 1319 observations (N) 102 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-----------|-----|----------|------------|----|---------|----|-------|---------|----|---------|-------------|-------------|
| Model     | 0.96 | 0.999 | 0.998 | 1 |
| Predictive | 0.81 | 0.999 | 0.998 | 1 |
| P1        | 0.81 | 0.999 | 0.998 | 1 |

Orthogonal in X (OPLS)

|            | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-------------|-----|----------|------------|----|---------|----|-------|---------|----|---------|-------------|-------------|
| O1          | 0.114 | 0.114 | 11.5 | 0 | 0 |
| O2          | 0.0218 | 0.136 | 2.2 | 0 | 0 |
| O3          | 0.00951 | 0.145 | 0.961 | 0 | 0 |
| O4          | 0.0046 | 0.15 | 0.465 | 0 | 0 |
3.5.3.5 Manuscript Table 1, Entry 5

SA

The OPLS model consisted of 1+4+0 components and was trained with a total of 1319 observations (N) 344 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-----------|-----|----------|------------|----|---------|----|-------|---------|----|---------|------------|-------------|
| Model     | 0.908 | 0.998 | 0.996 | 1 | Predictive | 0.783 | 0.837 | 0.998 | 0.996 | 0.01 | 0.996 | 1 | 1 | 1 | R1 |
| Orthogonal in X (OPLS) | 0.124 | 0 | 0 | 0 | Q1 | 0.0952 | 0.0952 | 32.6 | 0 | 0 | R1 |
| Orthogonal in X (OPLS) | 0.122 | 0.222 | 107 | 0 | 0 | Q2 | 0.14 | 0.362 | 67.9 | 0 | 0 | R1 |
| Orthogonal in X (OPLS) | 0.0602 | 0.422 | 29.1 | 0 | 0 | Q3 | 0.0115 | 0.434 | 5.54 | 0 | 0 | R1 |
| Orthogonal in X (OPLS) | 0.00289 | 0.124 | 0.99 | 0 | 0 | Q4 | 0.00289 | 0.124 | 0.99 | 0 | 0 | NS |

3-NSA

The OPLS model consisted of 1+5+0 components and was trained with a total of 1220 observations 484 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-----------|-----|----------|------------|----|---------|----|-------|---------|----|---------|------------|-------------|
| Model     | 0.819 | 0.996 | 0.99 | 1 | Predictive | 0.379 | 0.379 | 183 | 0.996 | 0.996 | 0.99 | 0.996 | 1 | 1 | 1 | R1 |
| Orthogonal in X (OPLS) | 0.44 | 0 | 0 | 0 | Q1 | 0.222 | 0.222 | 107 | 0 | 0 | R1 |
| Orthogonal in X (OPLS) | 0.14 | 0.362 | 67.9 | 0 | 0 | Q2 | 0.0602 | 0.422 | 29.1 | 0 | 0 | R1 |
| Orthogonal in X (OPLS) | 0.0115 | 0.434 | 5.54 | 0 | 0 | Q4 | 0.00289 | 0.124 | 0.99 | 0 | 0 | NS |
| Orthogonal in X (OPLS) | 0.00289 | 0.124 | 0.99 | 0 | 0 | Q4 | 0.00289 | 0.124 | 0.99 | 0 | 0 | NS |

5-NSA

The OPLS model consisted of 1+6+0 components and was trained with a total of 1220 observations (N) 709 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-----------|-----|----------|------------|----|---------|----|-------|---------|----|---------|------------|-------------|
| Model     | 0.793 | 0.997 | 0.99 | 1 | Predictive | 0.363 | 0.363 | 257 | 0.997 | 0.997 | 0.99 | 0.996 | 1 | 1 | 1 | R1 |
| Orthogonal in X (OPLS) | 0.43 | 0 | 0 | 0 | Q1 | 0.181 | 0.181 | 128 | 0 | 0 | R1 |
| Orthogonal in X (OPLS) | 0.129 | 0.31 | 91.2 | 0 | 0 | Q2 | 0.0675 | 0.377 | 47.8 | 0 | 0 | R1 |
| Orthogonal in X (OPLS) | 0.0408 | 0.418 | 28.9 | 0 | 0 | Q4 | 0.0075 | 0.425 | 5.31 | 0 | 0 | R1 |
| Orthogonal in X (OPLS) | 0.00246 | 0.43 | 3.15 | 0 | 0 | Q6 | 0.00246 | 0.43 | 3.15 | 0 | 0 | NS |

DNSA

The OPLS model consisted of 1+4+0 components and was trained with a total of 1218 observations (N) 102 variables (K).

| Component | R2X | R2X(cum) | Eigenvalue | R2 | R2(cum) | Q2 | Limit | Q2(cum) | R2Y | R2Y(cum) | EigenvalueY | Significance |
|-----------|-----|----------|------------|----|---------|----|-------|---------|----|---------|------------|-------------|
| Model     | 0.96 | 0.999 | 0.998 | 1 | Predictive | 0.813 | 0.813 | 82.1 | 0.999 | 0.999 | 0.998 | 0.998 | 1 | 1 | 1 | R1 |
| Orthogonal in X (OPLS) | 0.147 | 0 | 0 | 0 | Q1 | 0.11 | 0.11 | 11.1 | 0 | 0 | R1 |
| Orthogonal in X (OPLS) | 0.0236 | 0.134 | 2.39 | 0 | 0 | Q2 | 0.00887 | 0.142 | 0.896 | 0 | 0 | NS |
| Orthogonal in X (OPLS) | 0.00461 | 0.147 | 0.465 | 0 | 0 | Q4 | 0.00461 | 0.147 | 0.465 | 0 | 0 | NS |
### 3.5.3.6 Summary of Root Mean Square Errors of the models

The RMSEE (Root Mean Square Error of Estimation) indicates the fit of the observations to the model. The RMSEcv (Root Mean Square Error of Cross Validation) is an analogous measure, but estimated using the cross-validation procedure. The RMSE_E and RMSE_CV were calculated in Simca.

| Entry | RMSEE_E [mM] | RMSEE_CV [mM] | 3-NSA [mM] | RMSEE_E [mM] | RMSEE_CV [mM] | 5-NSA [mM] | RMSEE_E [mM] | RMSEE_CV [mM] | DNSA [mM] | RMSEE_E [mM] | RMSEE_CV [mM] |
|-------|---------------|----------------|-----------|---------------|----------------|-----------|---------------|----------------|-----------|---------------|----------------|
| 1     | 4.6           | 14.1           | 5.3       | 10.4          | 6.6            | 11.0      | 10.3          | 27.2           |
| 2     | 2.7           | 8.7            | 2.6       | 8.4           | 3.6            | 9.4       | 3.1           | 7.8            |
| 3     | 2.5           | 3.9            | 2.0       | 3.3           | 1.8            | 3.4       | 2.2           | 2.4            |
| 4     | 2.4           | 3.5            | 1.8       | 2.9           | 1.6            | 3.2       | 1.6           | 2.1            |
| 5     | 2.5           | 3.6            | 1.8       | 2.8           | 1.6            | 3.1       | 1.6           | 2.1            |

**Table S3** RMSE for estimation and cross validation for the estimated concentration of each of the four reaction components, through different generations of the multivariate analysis models.

### 3.5.4 Model Robustness / Validation Set

The RMSE_V (Root Mean Square Error of Validation) was calculated by taking the square root of the mean of the squared value of the difference between the expected value (known concentration of the validation set) and the predicted value (predicted concentration of the MVA model).

\[
RMSE_V = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (expected \ value - predicted \ value)^2}
\]

To validate the robustness of the model different flow rates (1.7 mL/min and 2.4 mL/min) and different acquisition times (3.2 s and 6.4 s) were tested on two validation solutions (Table S4). For the different acquisition times the repetition times had to be modified to 4 s and 7 s.
Table S4 Overview of the RMSE with different flow rates and acquisition times.

| Acquisition time [s] | Flow rate [mL·min⁻¹] | SA [mM] | 3-NSA [mM] | 5-NSA [mM] | DNSA [mM] |
|----------------------|-----------------------|---------|------------|------------|----------|
| 1.6                  | 1.1                   | 2.2     | 3.1        | 3.6        | 1.7      |
| 1.6                  | 1.7                   | 2.7     | 2.4        | 3.3        | 2.6      |
| 1.6                  | 2.4                   | 4.0     | 2.1        | 7.1        | 4.1      |
| 3.2                  | 1.1                   | 3.3     | 2.2        | 2.6        | 2.4      |
| 6.4                  | 1.1                   | 2.9     | 2.3        | 2.8        | 2.4      |

| Acquisition time [s] | Flow rate [mL·min⁻¹] | SA [mM] | 3-NSA [mM] | 5-NSA [mM] | DNSA [mM] |
|----------------------|-----------------------|---------|------------|------------|----------|
| 1.6                  | 1.1                   | 2.9     | 1.7        | 2.5        | 1.5      |
| 1.6                  | 1.7                   | 2.8     | 1.9        | 2.6        | 2.5      |
| 1.6                  | 2.4                   | 2.6     | 2.1        | 2.4        | 3.3      |
| 3.2                  | 1.1                   | 2.9     | 2.2        | 2.0        | 1.9      |
| 6.4                  | 1.1                   | 2.9     | 2.6        | 2.5        | 2.4      |

| Acquisition time [s] | Flow rate [mL·min⁻¹] | SA [mM] | 3-NSA [mM] | 5-NSA [mM] | DNSA [mM] |
|----------------------|-----------------------|---------|------------|------------|----------|
| 1.6                  | 1.1                   | 1.8     | 4.9        | 4.0        | 6.8      |
| 1.6                  | 1.7                   | 1.7     | 5.8        | 4.8        | 7.4      |
| 1.6                  | 2.4                   | 6.4     | 4.6        | 7.5        | 13.3     |
| 3.2                  | 1.1                   | 3.0     | 5.5        | 3.8        | 5.3      |
| 6.4                  | 1.1                   | 3.2     | 6.3        | 3.5        | 4.5      |

**Fig. S12** Representative section of NMR spectra recorded at different flow rates to show the peak broadening due to higher flow rates.
3.6 Monitoring Reaction Optimization by Inline NMR

3.6.1 Statistical Approach to Remove Outliers

For the DoE experiments quartiles (Q) and interquartile ranges (IQR) were calculated from each individual data set. For the dynamic experiment the quartiles (Q) and interquartile (IQR) were calculated for each data point from the 18 previous data points. Outliers were detected and removed if the data point exceeded the lower (Q1 – 1.5 IQR) or upper (Q3 + 1.5 IQR) boundary (Fig. S14).

![Fig. S13](image.png)

Fig. S13 (A) raw data obtained from the MVA model for the dynamic experiment. (B) Processed data after statistical outlier detection and removal.
3.6.2 Experimental Conditions DoE

For the three-factor full factorial DoE design the conditions in Table S5 were used. To investigate curvature effects the facial points were added.

| DoE Entry | Pump 1 [mL/min] | Pump 2 [mL/min] | Pump 3 [mL/min] | Pump 4 [mL/min] | RT (reactor) [s] | Equivalence of HN03 | Termostat 1 [°C] | Termostat 2 [°C] |
|-----------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| 1         | 0.66           | 0.44           | 5.5            | 1.1            | 18.7           | 0.8            | 0              | 0              |
| 2         | 0.41           | 0.69           | 5.5            | 1.1            | 18.7           | 2              | 0              | 0              |
| 3         | 1.44           | 0.96           | 12             | 2.4            | 8.6            | 0.8            | 0              | 0              |
| 4         | 0.90           | 1.50           | 12             | 2.4            | 8.6            | 0              | 0              | 0              |
| 5         | 0.66           | 0.44           | 5.5            | 1.1            | 18.7           | 0.8            | 35             | 0              |
| 6         | 0.41           | 0.69           | 5.5            | 1.1            | 18.7           | 2              | 35             | 0              |
| 7         | 1.44           | 0.96           | 12             | 2.4            | 8.6            | 0.8            | 35             | 0              |
| 8         | 0.90           | 1.50           | 12             | 2.4            | 8.6            | 0              | 35             | 0              |
| 9         | 0.81           | 0.94           | 8.75           | 1.75           | 11.8           | 1.4            | 17.5           | 0              |
| 10        | 0.81           | 0.94           | 8.75           | 1.75           | 11.8           | 1.4            | 17.5           | 0              |
| 11        | 0.81           | 0.94           | 8.75           | 1.75           | 11.8           | 1.4            | 17.5           | 0              |
| 12        | 0.81           | 0.94           | 8.75           | 1.75           | 11.8           | 1.4            | 0              | 0              |
| 13        | 0.81           | 0.94           | 8.75           | 1.75           | 11.8           | 1.4            | 35             | 0              |
| 14        | 1.05           | 0.70           | 8.75           | 1.75           | 11.8           | 0.8            | 17.5           | 0              |
| 15        | 0.66           | 1.09           | 8.75           | 1.75           | 11.8           | 2              | 17.5           | 0              |
| 16        | 0.51           | 0.59           | 5.5            | 1.1            | 18.7           | 1.4            | 17.5           | 0              |
| 17        | 1.11           | 1.29           | 12             | 2.4            | 8.6            | 1.4            | 17.5           | 0              |

Table S5. Experimental conditions for the performed DoE.

3.6.3 Temperature and Pressure data for DoE

| Entry | Total flow rate [mL/min] | Thermostat 1 Set value [°C] | Equiv HNO3 | P_1 Mean [bar] | P_1 STD [bar] | P_2 Mean [bar] | P_2 STD [bar] | P_3 Mean [bar] | P_3 STD [bar] |
|-------|-------------------------|----------------------------|------------|----------------|----------------|----------------|----------------|----------------|----------------|
| 1     | 7.70                    | 0                          | 0.8        | ±2.5           | ±0.1           | ±2.6           | ±0.1           | ±2.4           | ±0.1           |
| 2     | 7.70                    | 0                          | 2          | ±2.4           | ±0.0           | ±2.4           | ±0.0           | ±2.4           | ±0.0           |
| 3     | 16.80                   | 0                          | 0.8        | ±6.1           | ±0.1           | ±6.1           | ±0.1           | ±6.1           | ±0.1           |
| 4     | 16.80                   | 0                          | 2          | ±5.9           | ±0.0           | ±5.9           | ±0.0           | ±5.9           | ±0.1           |
| 5     | 7.70                    | 35                         | 0.8        | ±2.4           | ±0.0           | ±2.4           | ±0.0           | ±2.4           | ±0.0           |
| 6     | 7.70                    | 35                         | 2          | ±2.4           | ±0.0           | ±2.4           | ±0.0           | ±2.4           | ±0.0           |
| 7     | 16.80                   | 35                         | 0.8        | ±5.8           | ±0.1           | ±5.9           | ±0.1           | ±6.0           | ±0.1           |
| 8     | 16.80                   | 35                         | 2          | ±5.8           | ±0.1           | ±5.8           | ±0.1           | ±6.0           | ±0.1           |
| 9     | 12.25                   | 17.5                       | 1.4        | ±3.9           | ±0.1           | ±3.9           | ±0.1           | ±3.9           | ±0.1           |
| 10    | 12.25                   | 17.5                       | 1.4        | ±4.0           | ±0.1           | ±4.0           | ±0.0           | ±4.0           | ±0.0           |
| 11    | 12.25                   | 17.5                       | 1.4        | ±3.9           | ±0.0           | ±3.9           | ±0.0           | ±3.9           | ±0.0           |
| 12    | 12.25                   | 0                          | 1.4        | ±4.0           | ±0.0           | ±4.0           | ±0.0           | ±4.0           | ±0.1           |
| 13    | 12.25                   | 35                         | 1.4        | ±3.9           | ±0.1           | ±3.9           | ±0.0           | ±4.0           | ±0.0           |
| 14    | 12.25                   | 17.5                       | 0.8        | ±4.2           | ±0.1           | ±4.2           | ±0.1           | ±4.2           | ±0.0           |
| 15    | 12.25                   | 17.5                       | 2          | ±3.8           | ±0.1           | ±3.8           | ±0.1           | ±3.9           | ±0.1           |
| 16    | 7.70                    | 17.5                       | 1.4        | ±2.4           | ±0.0           | ±2.4           | ±0.0           | ±2.4           | ±0.0           |
| 17    | 16.80                   | 17.5                       | 1.4        | ±5.8           | ±0.1           | ±5.8           | ±0.1           | ±6.0           | ±0.1           |

Table S6. Temperature and pressure data obtained during the DoE experiments.
### Table S7. Temperature data obtained during the DoE experiments from the temperature probes in the process stream (T_1, T_2 and T_3) and probes in the cooling cycle (T_4 before and T_5 after Lonza FlowPlate Lab).

| Entry | Total flow rate [mL/min] | Thermostat 1 Set value [°C] | Thermistor 1 Mean STD Mean | Thermistor 2 Mean STD Mean | Thermistor 3 Mean STD Mean | Thermistor 4 Mean STD Mean | Thermistor 5 Mean STD Mean | ΔT_4, std | S
|-------|--------------------------|-----------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|-----------|--------|
|       |                          |                             | T_1 [°C]                  | T_2 [°C]                  | T_3 [°C]                  | T_4 [°C]                  | T_5 [°C]                  |           |        |
| 1     | 7.70                     | 0.0                         | 0.8                       | 2.51                      | 0.01                     | 3.03                      | 0.01                     | 3.40                  | 0.08   |
| 2     | 7.70                     | 0.0                         | 2.0                       | 2.44                      | 0.01                     | 3.10                      | <0.01                    | 3.43                  | 0.10   |
| 3     | 16.80                    | 0.0                         | 0.8                       | 2.21                      | 0.01                     | 2.82                      | 0.01                     | 4.86                  | 0.07   |
| 4     | 16.80                    | 0.0                         | 2.0                       | 2.22                      | 0.01                     | 3.03                      | <0.01                    | 4.84                  | 0.08   |
| 5     | 7.70                     | 35.0                        | 0.8                       | 33.67                     | 0.01                     | 33.75                     | 0.02                     | 4.08                  | <0.01  |
| 6     | 7.70                     | 35.0                        | 2.0                       | 33.73                     | 0.01                     | 33.70                     | 0.01                     | 4.06                  | <0.01  |
| 7     | 16.80                    | 35.0                        | 0.8                       | 33.82                     | 0.01                     | 33.86                     | 0.01                     | 5.39                  | <0.01  |
| 8     | 16.80                    | 35.0                        | 2.0                       | 33.92                     | <0.01                    | 33.81                     | <0.01                    | 5.33                  | <0.01  |
| 9     | 12.25                    | 17.5                        | 1.4                       | 17.98                     | 0.01                     | 18.28                     | <0.01                    | 4.25                  | <0.01  |
| 10    | 12.25                    | 17.5                        | 1.4                       | 18.06                     | 0.00                     | 18.38                     | <0.01                    | 4.36                  | 0.05   |
| 11    | 12.25                    | 17.5                        | 1.4                       | 18.05                     | 0.00                     | 18.38                     | <0.01                    | 4.31                  | <0.01  |
| 12    | 12.25                    | 0.0                         | 1.4                       | 2.35                      | 0.01                     | 3.01                      | 0.01                     | 3.95                  | 0.11   |
| 13    | 12.25                    | 35.0                        | 1.4                       | 33.81                     | 0.01                     | 33.79                     | <0.01                    | 4.49                  | 0.05   |
| 14    | 12.25                    | 17.5                        | 0.8                       | 17.91                     | 0.01                     | 18.21                     | 0.01                     | 4.19                  | <0.01  |
| 15    | 12.25                    | 17.5                        | 2.0                       | 18.02                     | 0.01                     | 18.34                     | <0.01                    | 4.29                  | 0.12   |
| 16    | 7.70                     | 17.5                        | 1.4                       | 18.03                     | 0.01                     | 18.33                     | 0.01                     | 3.85                  | 0.10   |
| 17    | 16.80                    | 17.5                        | 1.4                       | 18.03                     | <0.01                    | 18.35                     | 0.01                     | 5.26                  | 0.08   |

#### Table S8 Summary of fit for all models. R^2 is a measure of how well the model fits the experimental data points. Q^2 measures how well the model predicts future data (should be greater than 0.1 for a significant model and greater than 0.5 for a good model). Reproducibility is a measure of experimental error. Model validity can be low (negative) in good models due to very good replicates.

### 3.6.4 Reaction Optimization Data Model Fitting

The data for the optimization were fitted in MODDE (version 12.1, Umetrics). The data were imported from Excel. Models were fitted for the NMR and UPLC data for 

- **SA**: 3-NSA, 5-NSA and DNSA by using multiple linear regression (MLR). The histogram plot of SA in the NMR or UPLC data set did not show a “bell shaped” normal distribution, therefore the response was log transformed. Models were fit by using main, square and interaction terms.

---

| Design                  | Data Set | X Label | R2   | Q2     | Model validity | Reproducibility |
|-------------------------|----------|---------|------|--------|----------------|-----------------|
| three-factor full factorial DoE | NMR | SA      | 0.747 | 0.642  | -0.128         | 0.996           |
|                         |          | 3-NSA   | 0.850 | 0.601  | -0.163         | 0.997           |
|                         |          | 5-NSA   | 0.933 | 0.822  | -0.034         | 0.998           |
|                         |          | DNSA    | 0.998 | 0.990  | 0.348          | 0.999           |
| Face Centered Design    | NMR | SA      | 0.903 | 0.849  | 0.193          | 0.995           |
|                         |          | 3-NSA   | 0.969 | 0.883  | 0.300          | 0.996           |
|                         |          | 5-NSA   | 0.960 | 0.903  | 0.216          | 0.997           |
|                         |          | DNSA    | 0.995 | 0.984  | 0.337          | 0.999           |
| Face Centered Design    | UPLC | SA      | 0.919 | 0.884  | 0.996          | 0.675           |
|                         |          | 3-NSA   | 0.928 | 0.737  | -0.054         | 0.998           |
|                         |          | 5-NSA   | 0.977 | 0.939  | 0.310          | 0.998           |
|                         |          | DNSA    | 0.993 | 0.982  | 0.632          | 0.997           |
Fig. S 14 Coefficients and terms for the three-factor full factorial DoE design with NMR data set after non-significant terms were removed.

Fig. S 15 Coefficients and terms for the face centered DoE design with NMR data set after non-significant terms were removed.

Fig. S 16 Coefficients and terms for the face centered DoE design with UPLC data set after non-significant terms were removed.
Fig. S17 DoE surfaces obtained for three-factor full factorial DoE design with NMR data set.
Fig. S18 DoE surfaces obtained for the face centered DoE design with NMR data set.
Fig. S19 DoE surfaces obtained for the face centered DoE design with UPLC data set.
**Fig. S 20** Face centered DoE surface generated from the UPLC data set for 1.1 mL/min flow rate subtracted from the face centered DoE surface from the NMR data set for 1.1 mL/min flow rate.

**Table S 9.** Results of the DoE, measuring the distribution of SA, 3-NSA, 5-NSA and DNSA with NMR. Each percentage value is the mean of roughly 100 measuring points.
Table S 10. Results of the DoE, measuring the distribution of SA, 3-NSA, 5-NSA and DNSA with UPLC. Each percentage value is the mean of 3 measuring points.

| Entry | SA UPLC | 3-NSA UPLC | 5-NSA UPLC | DNSA UPLC |
|-------|---------|-------------|-------------|-----------|
|       | Mean    | STD         | Mean        | STD       | Mean    | STD     |
| 1     | 18.6%   | ±0.2%       | 38.9%       | ±0.2%     | 41.6%   | ±0.2%   | 0.9%    | ±0.0%    |
| 2     | 0.0%    | ±0.0%       | 44.0%       | ±0.1%     | 37.7%   | ±0.1%   | 18.2%   | ±0.0%    |
| 3     | 3.7%    | ±1.1%       | 46.8%       | ±0.6%     | 48.0%   | ±0.6%   | 1.5%    | ±0.0%    |
| 4     | 0.0%    | ±0.0%       | 46.6%       | ±0.3%     | 44.8%   | ±1.0%   | 8.7%    | ±1.3%    |
| 5     | 13.3%   | ±0.0%       | 41.9%       | ±0.1%     | 39.9%   | ±0.0%   | 5.0%    | ±0.1%    |
| 6     | 0.0%    | ±0.0%       | 27.5%       | ±0.1%     | 7.3%    | ±0.1%   | 65.2%   | ±0.2%    |
| 7     | 11.8%   | ±0.3%       | 42.9%       | ±0.2%     | 41.0%   | ±0.1%   | 4.3%    | ±0.0%    |
| 8     | 0.0%    | ±0.0%       | 32.7%       | ±0.1%     | 11.6%   | ±0.1%   | 55.8%   | ±0.2%    |
| 9     | 0.0%    | ±0.0%       | 45.4%       | ±0.0%     | 37.4%   | ±0.1%   | 17.3%   | ±0.1%    |
| 10    | 0.7%    | ±1.0%       | 45.1%       | ±0.1%     | 37.8%   | ±0.9%   | 16.3%   | ±1.9%    |
| 11    | 0.0%    | ±0.0%       | 44.9%       | ±0.1%     | 36.7%   | ±0.1%   | 18.4%   | ±0.0%    |
| 12    | 0.0%    | ±0.0%       | 47.0%       | ±0.1%     | 46.2%   | ±0.0%   | 6.7%    | ±0.0%    |
| 13    | 0.0%    | ±0.0%       | 41.3%       | ±0.0%     | 25.3%   | ±0.1%   | 33.4%   | ±0.1%    |
| 14    | 4.2%    | ±1.0%       | 46.8%       | ±0.5%     | 44.9%   | ±0.4%   | 4.2%    | ±0.1%    |
| 15    | 0.0%    | ±0.0%       | 41.5%       | ±0.3%     | 28.4%   | ±0.7%   | 30.1%   | ±1.0%    |
| 16    | 0.0%    | ±0.0%       | 44.0%       | ±0.0%     | 35.0%   | ±0.0%   | 21.0%   | ±0.1%    |
| 17    | 0.0%    | ±0.0%       | 44.6%       | ±0.1%     | 38.1%   | ±0.6%   | 17.3%   | ±0.7%    |

Fig. S 21 Comparison of the DoE data points obtained by NMR and by UPLC.
3.6.5 Experimental Conditions Dynamic Experiment

Experimental conditions are given in Table S11 for the dynamic experiments. In case of the slow ramp the HiText Script was used to control thermostat 1 and in case of the fast ramp the temperature was immediately changed.

| Experiment    | Pump 1 [mL/min] | Pump 2 [mL/min] | Pump 3 [mL/min] | Pump 4 [mL/min] | RT (reactor) [s] | Equivalence of HNO3 | Thermostat 1 [°C] | Thermostat 2 [°C] |
|---------------|-----------------|-----------------|-----------------|-----------------|-----------------|---------------------|------------------|------------------|
| Slow Ramp     | 0.47            | 0.63            | 5.5             | 1.1             | 18.71           | 1.61                | 0                | 35               |
| Fast Ramp     | 0.47            | 0.63            | 5.5             | 1.1             | 18.71           | 1.61                | 35               | 0                |

Table S 11. Experimental conditions for the dynamic experiments.

3.6.6 Temperature and Pressure Data from Dynamic Experiment (Slow Ramp)

Fig. S 22 Obtained temperature (T_1, T_2 and T_3) and pressure (P_1, P_2 and P_3) data for the dynamic experiment with the slow temperature ramp.
3.6.7 Offline Analysis of Dynamic Experiment (Slow Ramp)

Fig. S23 Comparison of the obtained inline NMR and offline UPLC measurements for the dynamic experiment with the slow temperature ramp.

3.6.8 Temperature and Pressure Data of Dynamic Experiment (Fast Ramp)

Fig. S24 Obtained temperature (T_1, T_2 and T_3) and pressure (P_1, P_2 and P_3) data for the dynamic experiment with the fast temperature ramp.
3.6.9 NMR Results for the Fast Ramp of the Dynamic Experiment

Fig. S25 Comparison of the obtained inline NMR and offline UPLC measurements for the dynamic experiment with the slow temperature ramp. The NMR results are adjusted to the temperature based on the RT.
4 UPLC-DAD

4.1 Chromatographic Conditions

The UPLC-DAD (Shimadzu Nexera X2) was comprised of a degassing unit (DGU-20A), two solvent delivery units (LC-30AD), a thermostated autosampler (SIL-30AD), thermostated column oven (CTO-20AC), diode array detector (SPD-M30A) and a control unit (CBM-20A). The analysis was carried out on a Phenomenex Kinetex Biphenyl reversed-phase analytical column (100 × 2.1 mm, particle size 1.7 μm, pore size 100 Å) at 45 °C using mobile phase A (MeOH) and B (aqueous phosphate buffer 25 mM, pH 6.25) at a flow rate of 0.5 mL/min. The phosphate buffer was freshly prepared on a daily basis by dissolving KH₂PO₄ (1360 mg) and K₂HPO₄ (435 mg) in HPLC grade H₂O (500 mL). Compounds were eluted with following gradient elution: starting with 3 % B, increasing to 40 % B in 2.5 min, holding at 40 % B for 0.75 min, decreasing to 3 % B in 0.01 min and equilibrating the column with 3 % B for 3.74 min. UPLC-DAD calibration curves were measured by injecting 0.8 µL of the calibration solutions and peak detection at a wavelength 210 nm for SA and at wavelength 338 nm for 3-NSA, 5-NSA and DNSA.

![UPLC-DAD Chromatograms](image)

**Fig. S 26** Representative UPLC–DAD chromatograms of the analysis of the calibration mixture at the wavelength of 210 nm and 338 nm.
4.2 Offline Calibration

**Fig. S 27** Calibration curve for offline UPLC-DAD analysis of SA at a wavelength of 210 nm.

**Fig. S 28** Calibration curve for offline UPLC-DAD analysis of 3-NSA at a wavelength of 338 nm.
**Fig. S 29** Calibration curve for offline UPLC-DAD analysis of 5-NSA at a wavelength of 338 nm.

**Fig. S 30** Calibration curve for offline UPLC-DAD analysis of DNSA at a wavelength of 338 nm.
5 References

S1 J. Kruisz, J. Rehrl, S. Sacher, I. Aigner, M. Horn and J. G. Khinast, *Int. J. Pharm.*, 2017, **528**, 334–344.
6  NMR Spectra 43 MHz
NMR Spectra 300 MHz
