Electronic Supplementary Material (ESI) for Reaction Chemistry & Engineering.

Supporting information

# Applying quantum mechanics to deconvolute benchtop $^1$H NMR reaction data

Jiayu Zhang,[a] Tristan Maschmeyer,[a] Ben Shapiro,[b] Sunil Babu Paudel,[b] Matthew C. Leclerc,[c]
Jason E. Hein*[a,d,e]

[a] Department of Chemistry, The University of British Columbia, Vancouver, BC V6T 1Z1, Canada.
[b] US Pharmacopeia 12601 Twinbrook Parkway Rockville, MD, 20852, USA
[c] Nanalysis Scientific Corp., 1-4600 5 St NE, Calgary, AB, T2E 7C3, Canada
[d] Department of Chemistry, University of Bergen, N-5007, Norway
[e] Acceleration Consortium, University of Toronto, ON, Canada

*Corresponding author email: jhein@chem.ubc.ca

## Table of Contents

# 1. General methods

a. Chemicals

Tert-Butyl N-allylcarbamate (TBA, **1**, 97%) was purchased from Alfa Aesar. meta-chlorobenzoic acid (mCBA, **3**, 98%) was purchased from AK Scientific; Tetramethylsilane (TMS, 5, $\geq$99%); 3-chloroperbenzoic acid (mCPBA, **2**, $\leq$77%), sodium Sulfate, and dichloromethane (DCM) were purchased from Sigma-Aldrich; Chloroform-D ($CDCl_3$, 99.8%) was purchased from Cambridge Isotope Laboratories.

b. Synthesizing Tert-butyl (oxiran-2-ylmethyl) carbamate

The epoxide product (**4**) was synthesized and purified. The procedures were: the reaction was conducted in 25 ml DCM and was left overnight to reach completion. After that, extra DCM was added (to double the reaction volume) and transferred to a separatory funnel. The solution was washed once with 50 ml of Na2SO3 solution (10% w/v) to quench unreacted mCPBA. The organic layer was subsequently washed with NaHCO3 twice and deionized water once (50 ml each). The final organic layer was concentrated in a rotovap. The concentrated mixture looked oily and yellowish. The epoxide product was then separated on a flash column with normal phase column selected (mobile phase: pet ether + ethyl acetate). The solution is concentrated and left in high vacuum overnight. The final products were confirmed via 1D $^1$H NMR (Figure S13).

# 2. Postprocessing of HF $^1$H NMR data in MNnova

The acquired reaction spectra were processed in the MestReNova (MNova, Mestre Lab Research, spain) software (v.14.2.1). For data acquired on HF NMR spectrometer, the spectra were first referenced based on the added internal standard TMS ($\delta = 0$ ppm); then, exponential apodization (0.3 Hz) was applied to all spectra; Auto phase correction was applied to a single spectrum, then the resulting phase values (PH0 and PH1) were applied to the remaining spectra; Finally, full auto baseline correction was applied to all spectra (Bernstein polynomial with order = 3).

# 3. Postprocessing of Benchtop $^1$H NMR data in MNnova

For data acquired on benchtop NMR spectrometer, steps 1-3 were the same as above. Then, the NMR $^1$H data for the epoxidation reaction acquired on benchtop spectrometer were treated with multipoint baseline correction method. The general strategy was to select one spectrum from the reaction dataset, apply multipoint baseline correction, tune the correction parameters based on visual inspection, and finally apply the correction template to the remaining spectra. The correction parameters used for benchtop experiments were (Figure S1): {Algorithm: Whittaker, RMS calculation span: 300, Smooth: 42}.

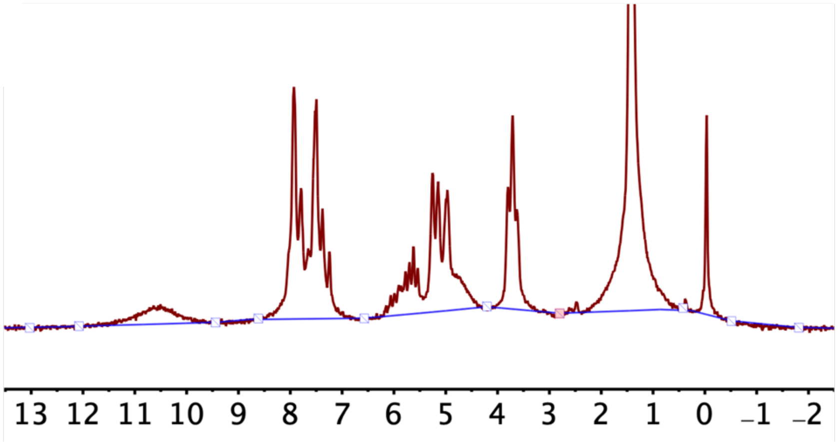
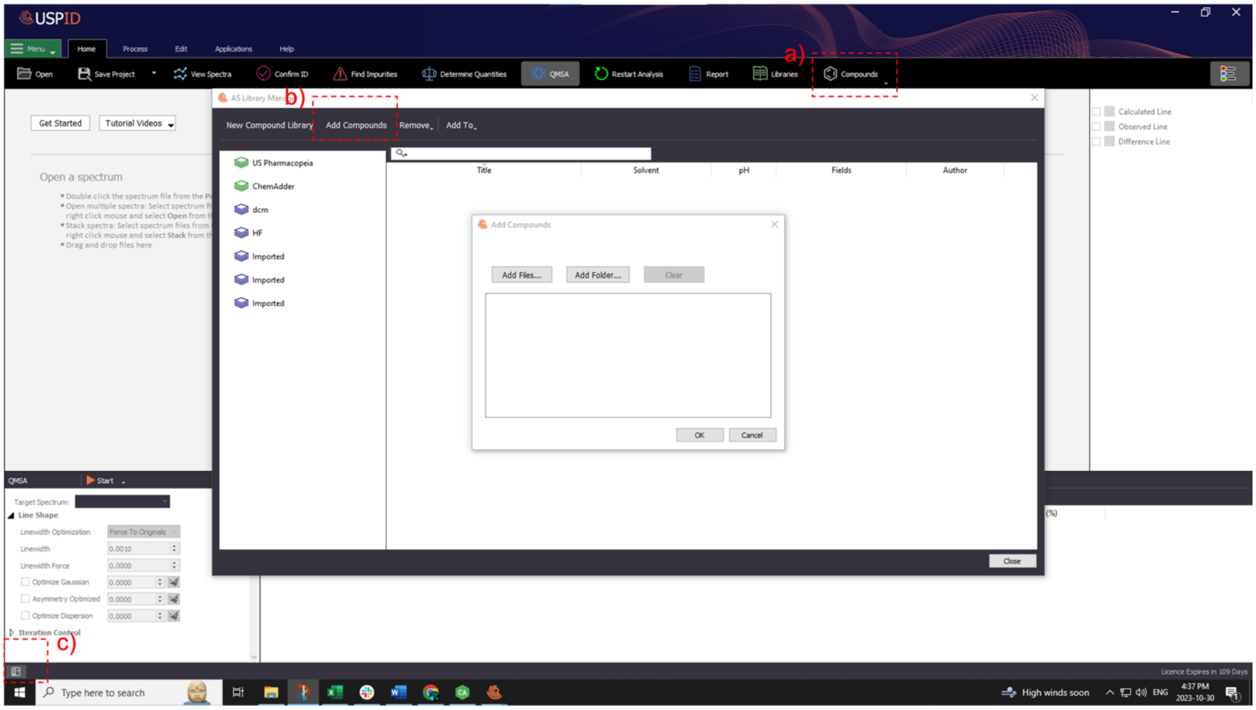


**Figure S1.** Multipoint baseline correction template for epoxidation in $CDCl_3$.

## 4. USP-ID workflow

### a. USP-ID QMSA calculation for a single spectrum

After the generation of model files, those model files will be imported into USP-ID to create a reaction compound library (Figure S2). Clicking the icon in the bottom left will bring up an empty panel where reaction spectra can be loaded by simply dragging the folder into it. After loading all desired spectra to be processed, one spectrum is selected (ideally in the middle of the reaction). The corresponding data will show up in the center panel. Afterwards, the QMSA tool is selected and all relevant model files are loaded (Figure S3b). Individual signal can be adjusted by dragging the tags before QMSA calculation (Figure S3c). After changing QMSA fitting parameters the calculation can be performed by clicking the 'start' otption (Figure S3d). After calculation, the parameter file for this calculation can be saved by clicking the 'Compounds' option in the top panel.

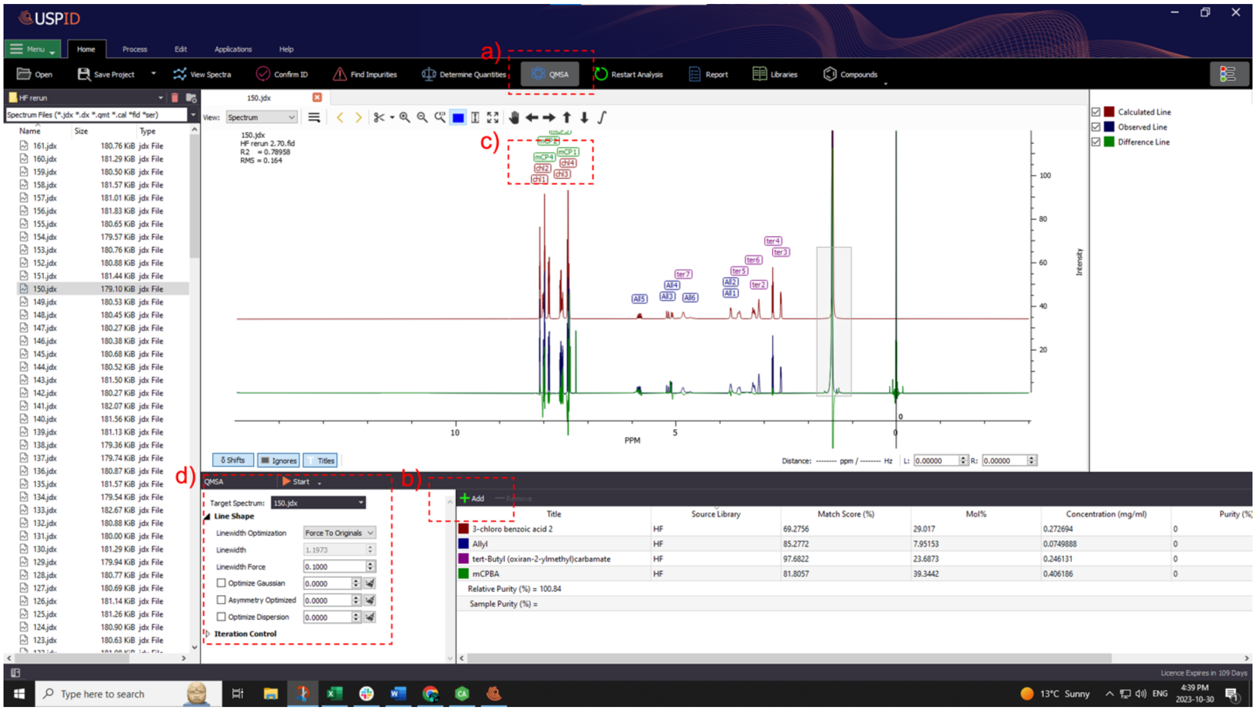


Target Spectrum:

Line Shape

- Linewidth Optimization: Force To Originals
- Linewidth: 0.0010
- Linewidth Force: 0.0000
- ☐ Optimize Gaussian: 0.0000
- ☐ Asymmetry Optimized: 0.0000
- ☐ Optimize Dispersion: 0.0000

**Figure S2.** Compound library generation page in USP-ID.



Spectrum Files (*.jdx *.dx *.qmt *.cal *fid *ser)

| Name | Size | Type |
|---|---|---|
| 161.jdx | 180.76 KiB | jdx File |
| 160.jdx | 181.29 KiB | jdx File |
| 159.jdx | 180.50 KiB | jdx File |
| 158.jdx | 181.57 KiB | jdx File |
| 157.jdx | 181.01 KiB | jdx File |
| 156.jdx | 181.83 KiB | jdx File |
| 155.jdx | 180.65 KiB | jdx File |
| 154.jdx | 179.57 KiB | jdx File |
| 153.jdx | 180.76 KiB | jdx File |
| 152.jdx | 180.88 KiB | jdx File |
| 151.jdx | 181.44 KiB | jdx File |
| 150.jdx | 179.10 KiB | jdx File |
| 149.jdx | 180.53 KiB | jdx File |
| 148.jdx | 180.45 KiB | jdx File |
| 147.jdx | 180.27 KiB | jdx File |
| 146.jdx | 180.38 KiB | jdx File |
| 145.jdx | 180.68 KiB | jdx File |
| 144.jdx | 180.52 KiB | jdx File |
| 143.jdx | 181.50 KiB | jdx File |
| 142.jdx | 180.27 KiB | jdx File |
| 141.jdx | 182.07 KiB | jdx File |
| 140.jdx | 181.56 KiB | jdx File |
| 139.jdx | 181.13 KiB | jdx File |
| 138.jdx | 179.36 KiB | jdx File |
| 137.jdx | 179.74 KiB | jdx File |
| 136.jdx | 180.87 KiB | jdx File |
| 135.jdx | 181.57 KiB | jdx File |
| 134.jdx | 179.54 KiB | jdx File |
| 133.jdx | 182.67 KiB | jdx File |
| 132.jdx | 180.88 KiB | jdx File |
| 131.jdx | 180.00 KiB | jdx File |
| 130.jdx | 181.29 KiB | jdx File |
| 129.jdx | 179.94 KiB | jdx File |
| 128.jdx | 180.77 KiB | jdx File |
| 127.jdx | 180.69 KiB | jdx File |
| 126.jdx | 181.14 KiB | jdx File |
| 125.jdx | 181.26 KiB | jdx File |
| 124.jdx | 180.90 KiB | jdx File |
| 123.jdx | 180.63 KiB | jdx File |

150.jdx
HF rerun 2.70.fid
R2 = 0.78958
RMS = 0.164

Target Spectrum: 150.jdx

Line Shape

- Linewidth Optimization: Force To Originals
- Linewidth: 1.1973
- Linewidth Force: 0.1000
- ☐ Optimize Gaussian: 0.0000
- ☐ Asymmetry Optimized: 0.0000
- ☐ Optimize Dispersion: 0.0000

| Title | Source Library | Match Score (%) | Mol% | Concentration (mg/ml) | Purity (%) |
|---|---|---|---|---|---|
| 🟥 3-chloro benzoic acid 2 | HF | 69.2756 | 29.017 | 0.272694 | 0 |
| 🟦 Allyl | HF | 85.2772 | 7.95153 | 0.0749888 | 0 |
| 🟪 tert-Butyl (oxiran-2-ylmethyl)carbamate | HF | 97.6822 | 23.6873 | 0.246131 | 0 |
| 🟩 mCPBA | HF | 81.8057 | 39.3442 | 0.406186 | 0 |

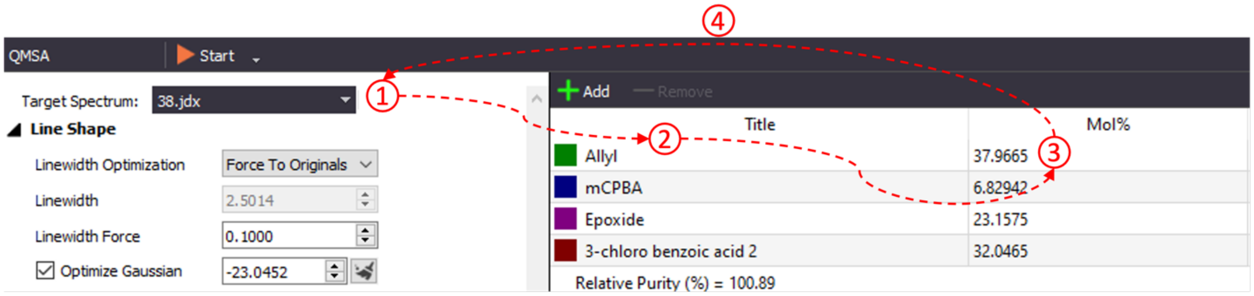
Relative Purity (%) = 100.84

Sample Purity (%) =

**Figure S3.** QMSA calculation page in USP-ID.

b. Automated QMSA calculation for multiple spectra and data extraction

After loading all spectra (which can be done by going back to the data panel, selecting all data, and select stack spectra) into USP-ID calculation panel, click ‘Compounds’ option again to import the parameter file that we just saved. Click yes when the prompt ‘apply to all’ appears. Then, change ‘target spectrum’ to all and click ‘start’. Afterwards, QMSA calculation will be performed to all loaded spectra. To get concentration data after QMSA calculation in USP-ID, the following steps will be taken (Figure S4): 1. Click ‘Target Spectrum’ to select the desired spectrum. 2. Locate the row that contains results for each species. 3. Record down their concentration values in a separate CSV file. 4. Repeat steps 1-3 until all spectra are viewed. Those steps can be automated using the python script discussed in the paper.



Target Spectrum: 38.jdx

Line Shape

Linewidth Optimization: Force To Originals

Linewidth: 2.5014

Linewidth Force: 0.1000

☑ Optimize Gaussian: -23.0452

| Title | Mol% |
|---|---|
| 🟩 Allyl | 37.9665 |
| ■ mCPBA | 6.82942 |
| ■ Epoxide | 23.1575 |
| 🟥 3-chloro benzoic acid 2 | 32.0465 |

Relative Purity (%) = 100.89

**Figure S4.** USP-ID result page after QMSA calculation. The number represents each step taken to get several concentration data points

## 5. QMSA and IHM method developments

a. HF $^1$H NMR epoxidation experiment

Due to the difference in the reaction medium and the environment where model files were created, manual chemical shift adjustments were preformed to the model files for species 1-4 in USP-ID on-the-fly. An example of such adjustments for **2** and **3** are shown in Figure S5a

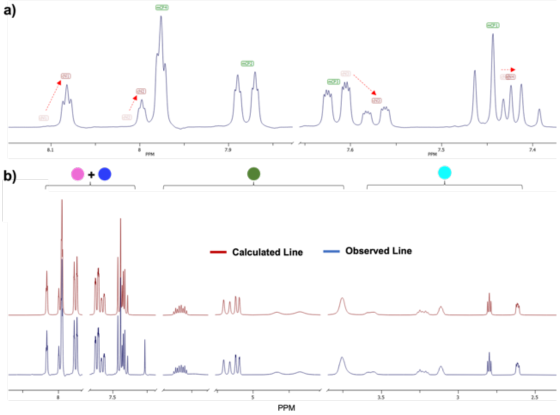


**Figure S5** a) Adjusting the model file for **3** in USP-ID. The transparent red labels represent the stored chemical shift value in the original model file. The arrows indicate the movements of those values to the new locations. b) Comparison of calculated and observed NMR spectrum for epoxidation data acquired on a HF spectrometer after all the mentioned modifications to the fitting steps.

Figure S6 shows the improvements of fitting scores after different modifications to the fitting procedures: b) after model file chemical shift realignments, c) after ignoring t-boc regions, d) after changing fitting mode to ‘Release’.

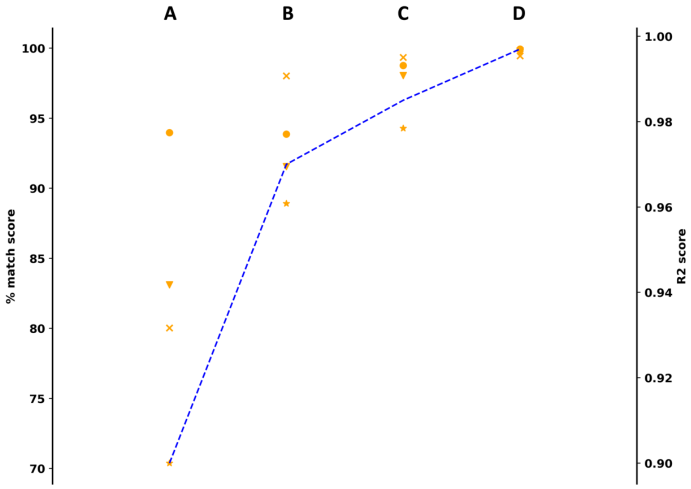


Figure S6 Fitting score improvements after several fitting procedure modifications. The markers for species 1-4 are cross, circle, triangle, and star respectively.

For IHM HF workflow development, the computation time on the representative HF spectrum was extremely long when the fitting mode was set to ‘Maximum interactions’, and the fitting results was poor (Figure S7, top). Computation time and fitting results were improved when the analysis region was reduced to two sub-regions (Figure S7, bottom).

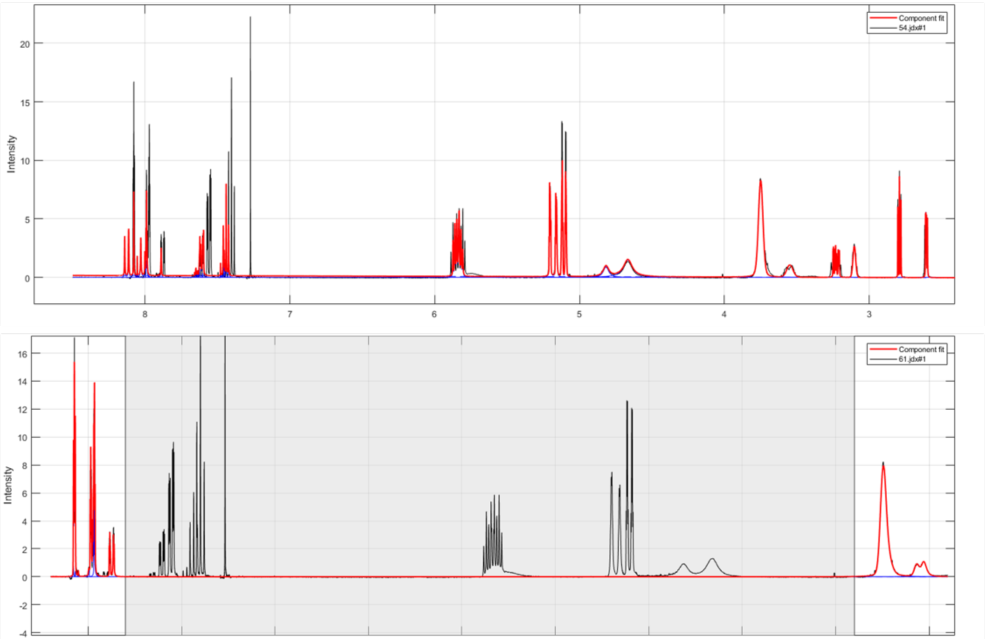


Figure S7. IHM HF representative spectrum fitting results. Top: analysis using full range of the spectrum. Bottom: analysis using reduced range of the spectrum.

### b. Benchtop $^1$H NMR epoxidation experiment

The signals that showed difference in QMSA prediction and observation is the allylic protons highlighted in Figure S6a. Interestingly, the splitting pattern for those protons are different in the model file (obtained on a HF instrument) compared to the pattern shown in the reaction spectra (obtained on a benchtop instrument, Figure S6b). After model file adjustment, the QMSA calculation matched observation better (Figure S6c)

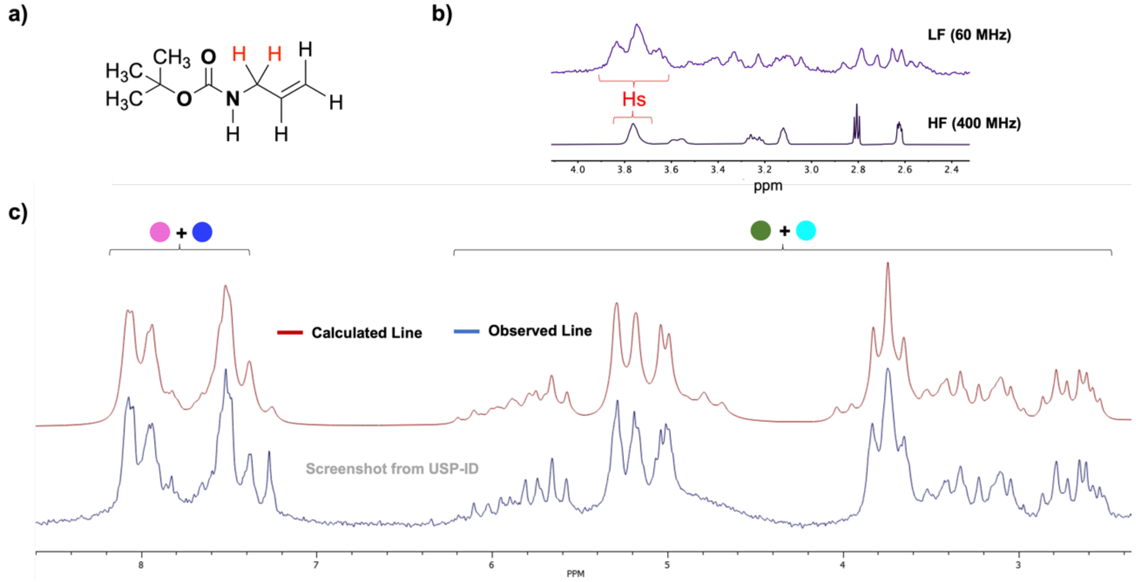


**Figure S*8*.** a) The structure of 1. The allylic protons are highlighted in red. b) Comparison of the allylic proton signals observed on HF (400 MHz) and benchtop (60 MHz) NMR spectrometers. c) Comparison of calculated and observed NMR spectrum for epoxidation data acquired on a benchtop spectrometer.

## 6. QMSA and IHM model files

All model files for QMSA and IHM workflows can be found under the folder 'Supplementary files/model files'. For QMSA, .cadder extension represents pure component model files and .usp extension represents mixture model files. For IHM, .pxm extension represents both the pure component and mixture model files.

## 7. Epoxidation kinetic modelling in Python

The second-order reaction mechanism was used to fit the concentration profiles obtained for the epoxidation experiments:

$$\frac{d[\boldsymbol{Epoxide}]}{dt} = k[\boldsymbol{Olefin}][\boldsymbol{mPCBA}]$$

The rate constant k together with $[1]_0$ and $[2]_0$ was estimated from the data using scipy.odeint for solving differential equations and scipy.optimize.minimize for parameter estimations (Nelder-Mead was used as optimization method as we did not have access to the gradient of those differential equations).

## 8. Preliminary benchtop $^1$H NMR epoxidation data in proteo solvent (DCM)

The same epoxidation reaction was monitoring with the same setup but in DCM (Figure S9). The solvent peak has noticeably larger SNR and is cause baseline distortion. It is envisioned that a custom solvent model file that can address such baseline distortion is crucial to expend the current workflow to include reaction monitoring in proteo solvent.

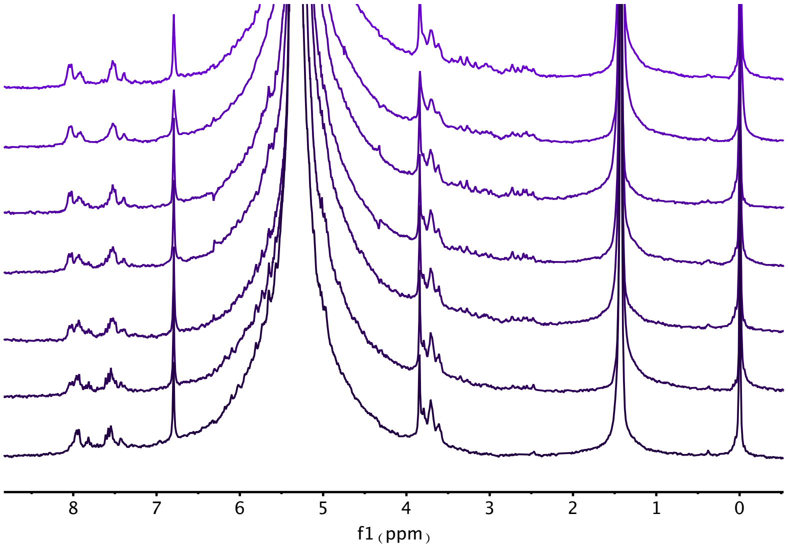


**Figure S9.** Preliminary benchtop 1H NMR epoxidation data in DCM, time increases from bottom to top.

# 9. High field (400 MHz) ¹H reference spectra

## a. Reference ¹H for Tert-Butyl N-allylcarbamate

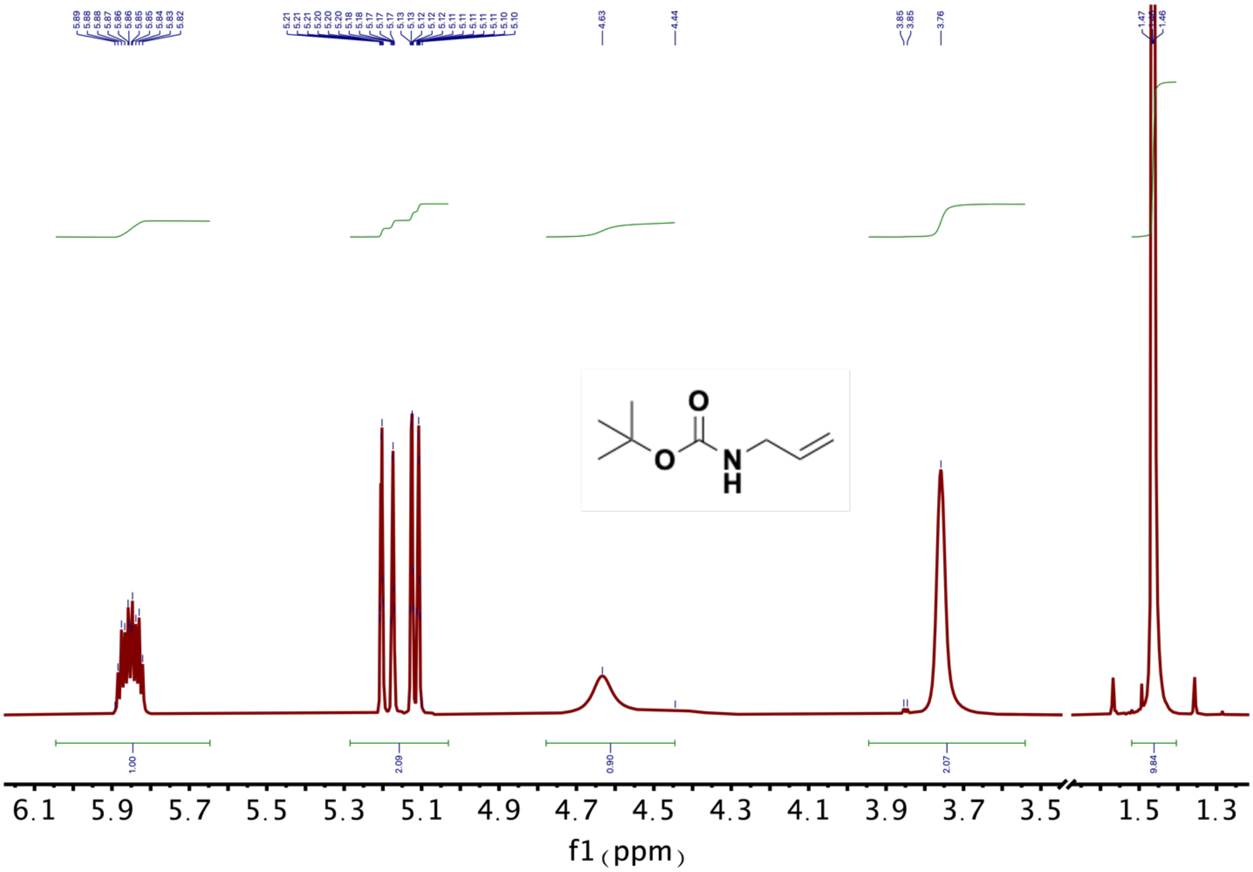


**Figure S10.** ¹H NMR spectrum (400 MHz, $CDCl_3$) of **1**.

b. Reference $^1$H for 3-Chloroperbenzoic acid

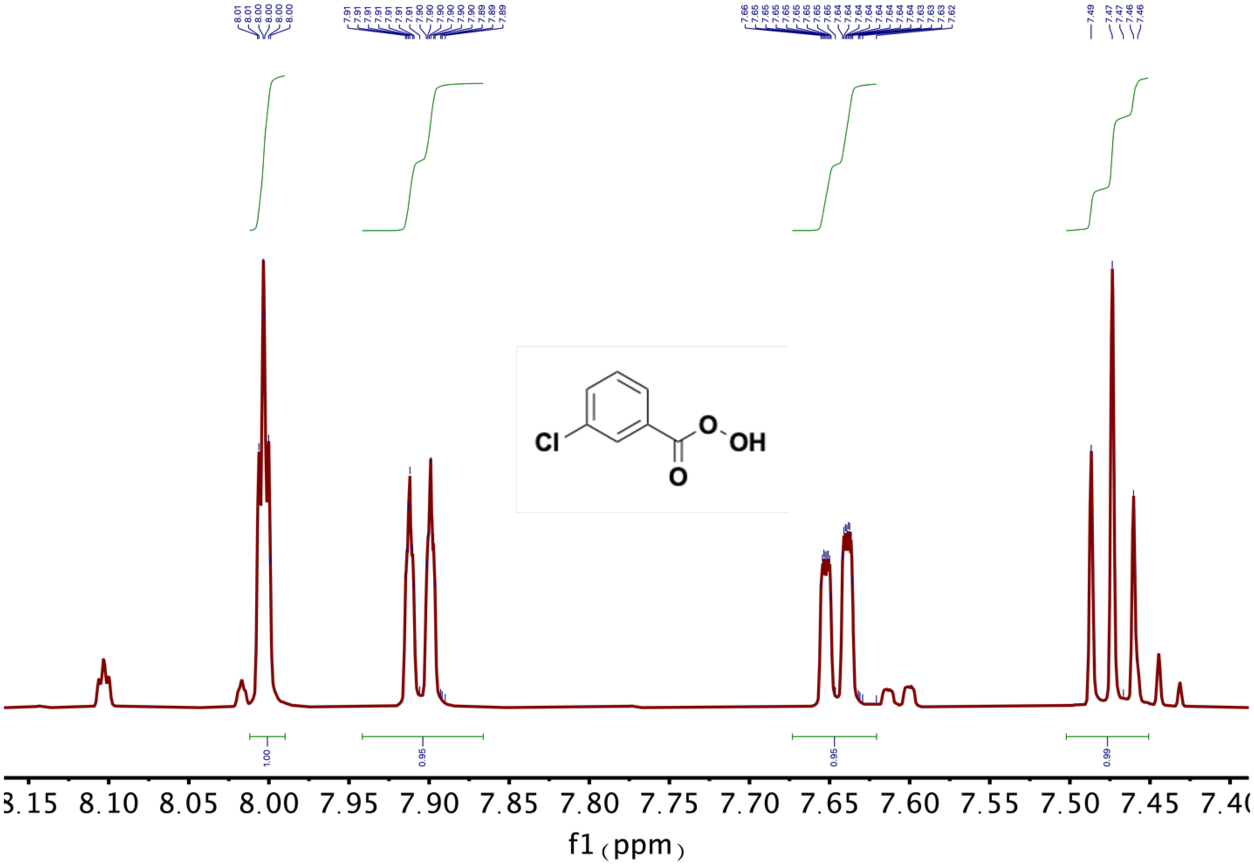


**Figure S11**. $^1$H NMR spectrum (400 MHz, $CDCl_3$) of **2**.

### c. Reference $^1$H for 3-Chlorobenzoic acid

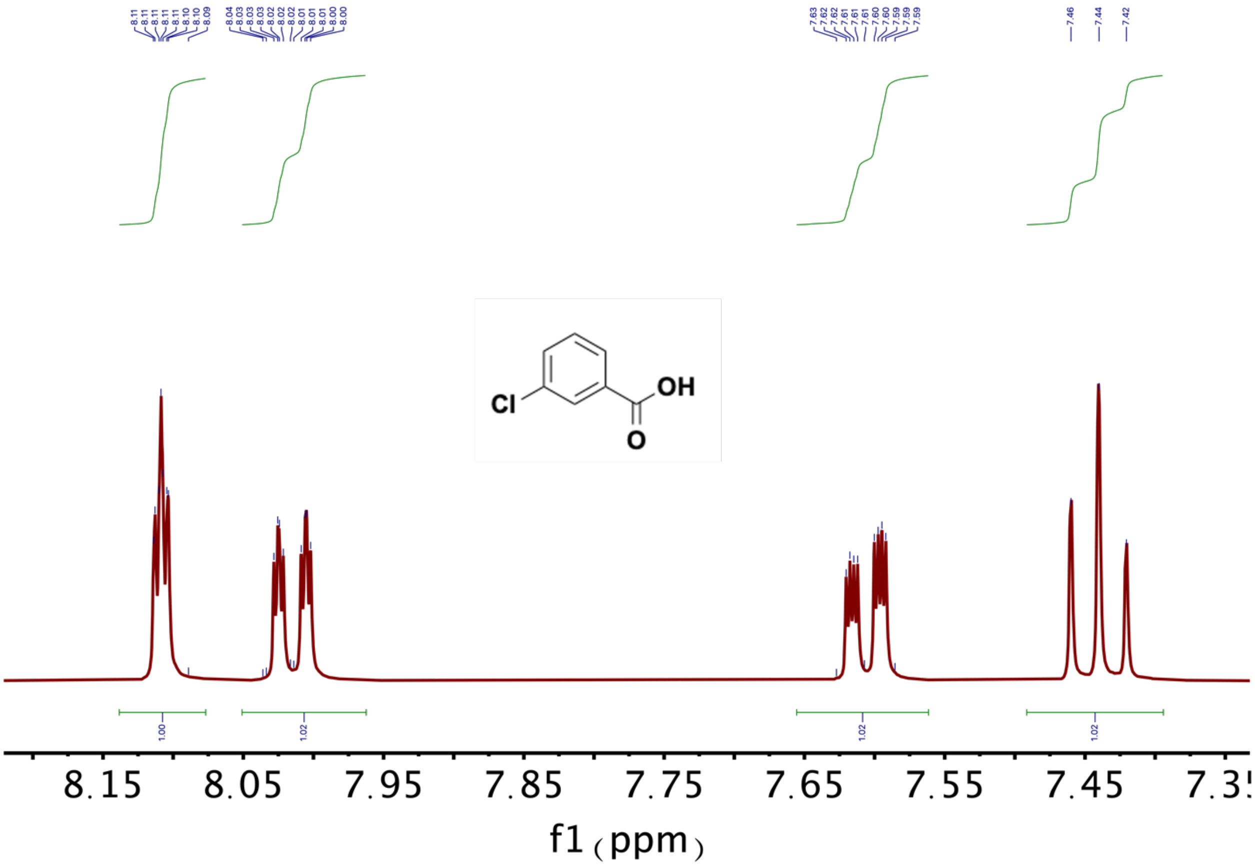


**Figure S12.** $^1$H NMR spectrum (400 MHz, $CDCl_3$) of **3**.

d. Reference $^1$H for Tert-butyl (oxiran-2-ylmethyl) carbamate

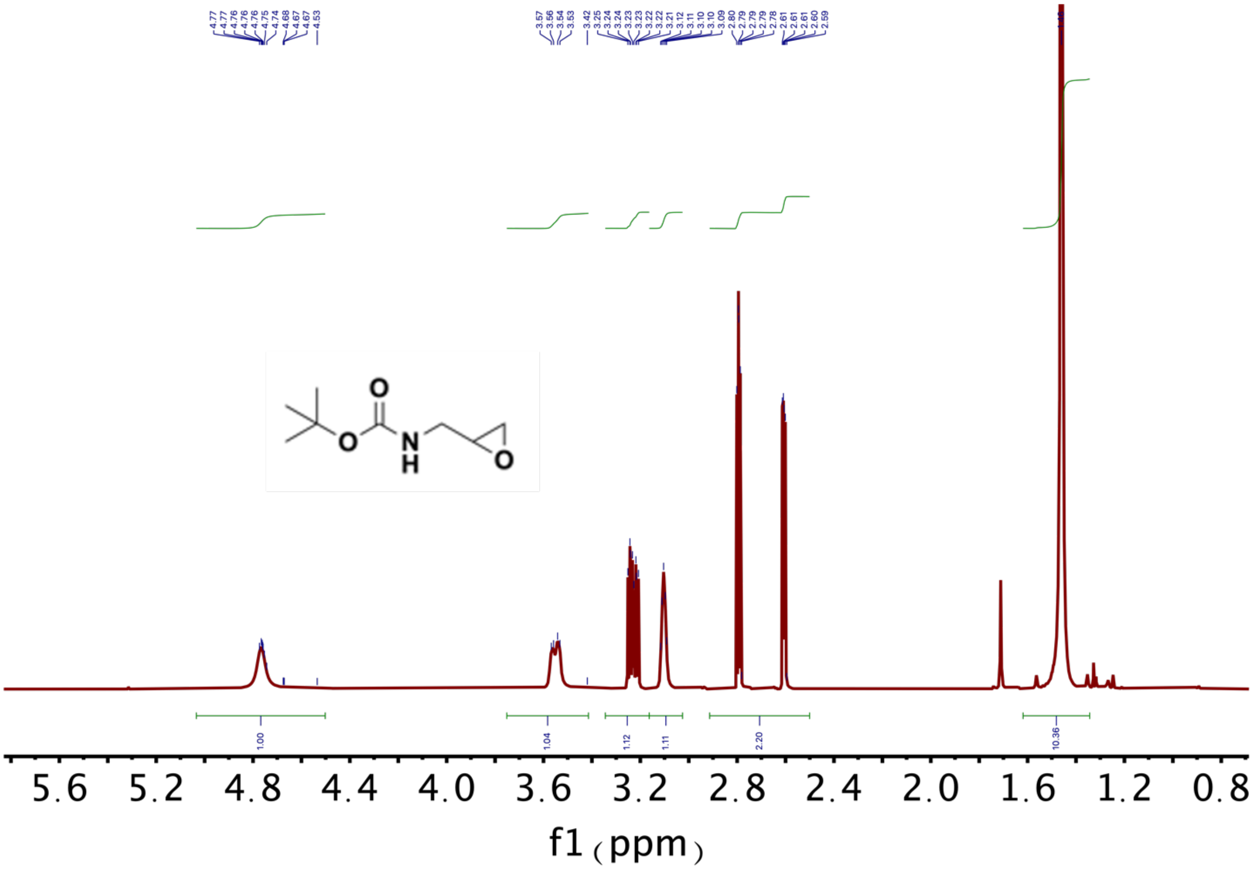


**Figure S13**. $^1$H NMR spectrum (400 MHz, $CDCl_3$) of **4**.