

Efficient copper-catalyzed amination of DNA-conjugated aryl iodides under mild aqueous conditions

Supporting information:

Table of Contents

SI I Materials and methods.....	2
SI II Analysis and purification of oligonucleotides-conjugates.....	2
Structure of the standard model oligonucleotide-amine	2
Synthesis and Characterization of Iodo-aromatics DNA conjugates 1a-1i.....	3
SI III Synthesis and Characterization of structural analogs of DMPAO L11-L21.....	14
SI IV Protocol and analysis for the screening of known ligands on 1a (main text Figure 1)	39
Effect of the addition of sodium ascorbate for the reactions using L8 (DMPAO) as ligand and characterization of the products 3a-f	39
Preliminary evaluation of DNA degradation.....	44
SI V Protocol and analysis for the screening of DMPAO analogs on 1a and 1b (main text Figure 2)	46
Selected screening results and comparison of the efficiency of L8 and L15	46
SI VI Protocol and analysis for the optimization of the concentration of L15 and DMSO proportion on the reaction between 1a and 1b and amines 2b-j(main text Figure 3).....	50
Effect of the concentration of L15 and cosolvent proportion on the proportion of Ligand L15 <i>N</i> - arylation side reaction using 1a as substrate.	51
SI VII Protocol and analysis for the optimization of the DMSO proportion, concentration of L15, amine, and base for the reaction between 1b and amine 2f	52
SI VIII Protocol and analysis for the scope determination on 1b-1i (main text Figure 4)	56
SI IX Protocol and analysis for the evaluation of the stability of the DNA tags under reaction conditions 1	57
SI X Confirmation the structure of DNA conjugates 3a, 3h and 4h by synthesis using commercial reagents	59

Abbreviations.

ACN: Acetonitrile

DIC: Diisopropylcarbodiimide

DMF: Dimethylformamide

DMPAO: 2-((2,6-Dimethylphenyl)amino)-2-oxoacetic Acid

DCM: Dichloromethane

HFIP: Hexafluoroisopropanol

HOAt: 1-Hydroxy-7-azabenzotriazole

MeOH: Methanol

TEA: Triethylamine

Rt: Retention time

SI I Materials and methods

All reagents and solvents were purchased from ABCR, ChemBridge, Chemcia Scientific, Trans World Chemicals, Sigma-Aldrich, Fluka, Enamine, TCI, and Alfa Aesar at the highest commercial quality and used without further purification unless stated otherwise.

Oligonucleotides were purchased from Integrated DNA Technologies (Synthesis scale: 5 μ moles, Standard desalting) and used without additional purification. Derivatized oligonucleotides **1a-i** were purified by reverse-phase preparative HPLC.

SI II Analysis and purification of oligonucleotides-conjugates

Structure of the standard model oligonucleotide-amine

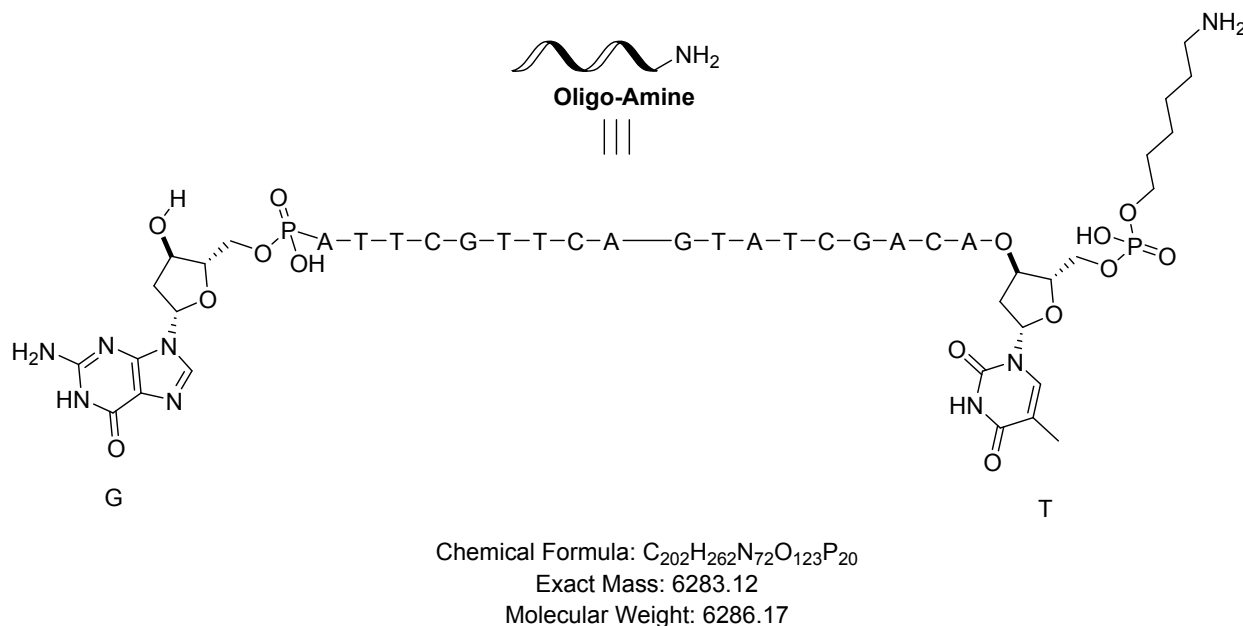


Figure SI II1: Molecular structure of the **Oligo-Amine** starting material.

The sequence of this model oligonucleotide (5'-TACAGCTATGACTTGCTTAG-3') was randomized.

Synthesis and Characterization of DNA-conjugated aryl halides 1a-1i

The acylation protocol on immobilized DNA was adapted from Harbury *et al.*¹

Representative acylation protocol for the synthesis of 1a:

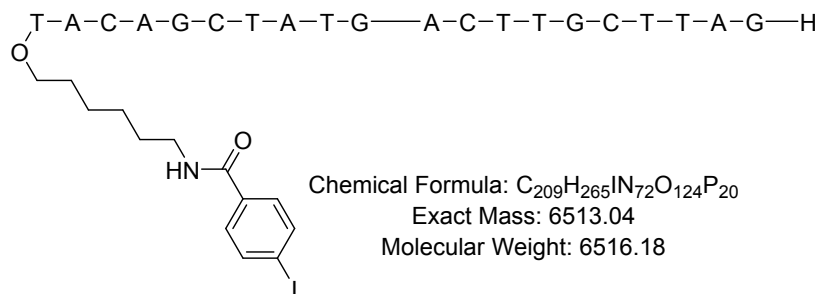


Figure SI II2: Molecular structure of the starting material **1a**.

5000 μ l of a suspension of DEAE sepharose (DEAE SephadexTM A-25, GE Healthcare 17-0170-02) was washed with 10ml of DNA bind solution (10 mM Acetic acid in water+0.005% triton-X100).

5000ul of a solution of **Oligo-Amine** as a solution (0.5 nmole/ μ l) (3.14 mg, 0.5 μ mol) was then added to the resin together with 5ml of DNA bind solution. This suspension was agitated for 2min, followed by filtration and washing with water (10ml) and MeOH (10ml). The resin was then washed for 2 min with 10ml of a 5% 4-methylpiperidine in MeOH to neutralize and remove any remaining acetic acid.

A solution of activated acid was prepared by mixing HOAt (CAS 39968-33-7)(100mM in MeOH) (3000 μ l, 300 μ mol), 4-iodobenzoic acid (100mM in DMSO) (3000 μ l, 300 μ mol) and DIC (46.7 μ l, 300 μ mol). This solution was agitated 30 seconds before addition to the DEAE sepharose resin. The resin was agitated 15min with the activated acid and this process repeated once.

A few beads were then incubated in 65 μ l of DNA elute solution (1.5M NaCl, 50mM pH8 phosphate buffer in water+0.005% triton-X100) and the released DNA purified by P6 chromatography (Micro Bio-SpinTM P-6 Gel Columns) before HPLC-TOF analysis.

Workup and Purification

At this stage the reaction was complete, and the DNA was released by treatment with 2ml of DNA elute solution (1.5M NaCl, 50mM phosphate buffer pH8, 0.005% triton-X) for 30min, followed by 3 treatments with 2000 μ l of the same solution for 15 min.

The mixture was concentrated using a 3K MWCO Ultrafiltration device (PALL, Nanosep 3K-30K Omega Centrifugal Device, 5mL, 5000G, 40°C, 30 min).

The concentrated sample was diluted to 4 ml with water and concentrated again (5000G, 40°C, 30 min) (twice).

The final solution was purified by reverse phase HPLC (Column Xbridge[®] Prep C18 5 μ m OBDTM 19x50mm, Solvent A: 100mM HFIP 86mM TEA, Solvent B: MeOH/ACN 1/1). The fractions were combined and concentrated using a 3K MWCO centrifugal device (5000G, 40C, 30 min) and adjusted to 4ml to give a solution containing 445nmoles of pure product **1a** (89%) as determined by UV absorbance at 260nm.

Analysis for 1a

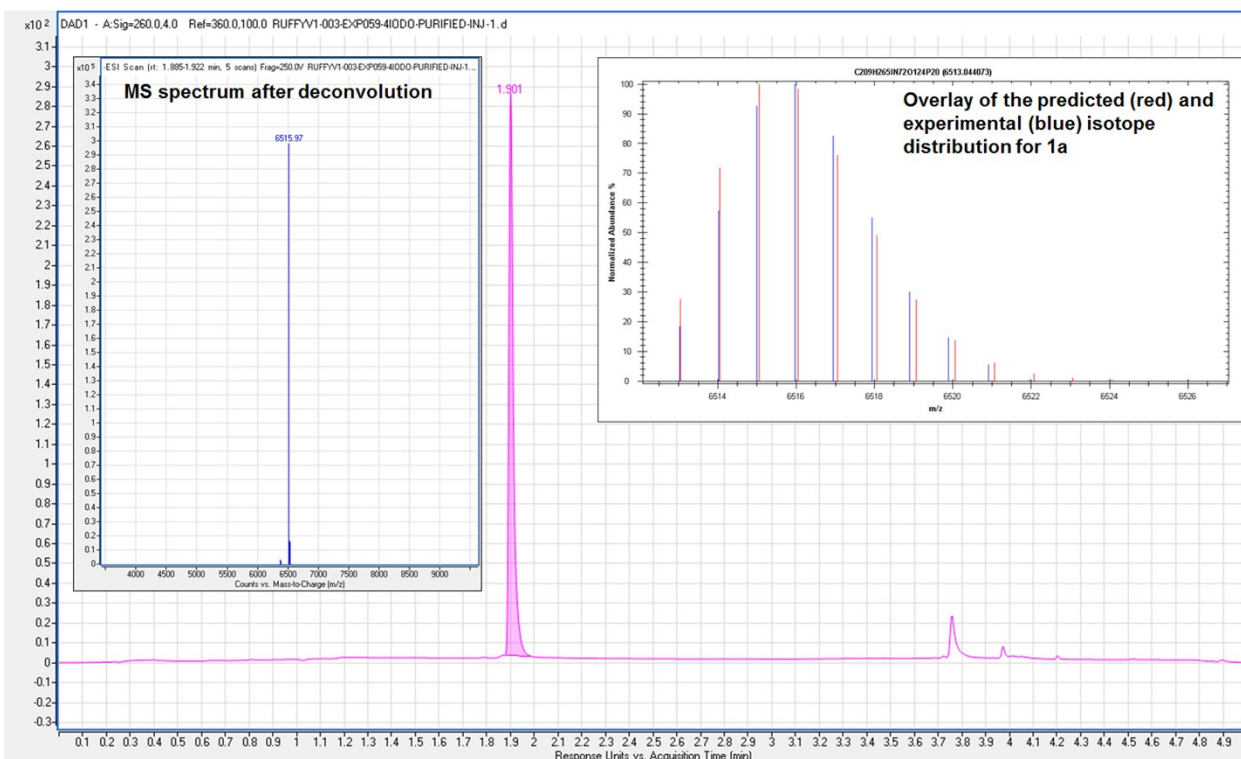


Figure SI II3: UPLC-TOF analysis of 1a.

- **UPLC/MS** Agilent 6200 series TOF/6500 series Q-TOF, Column ACQUITY BEH OST C18 2.1x50 mm 1.7 μm at 60°C; Flow 0.6 mL/min, Gradient: 15-30%B in 2.7min; 30-85%B in 0.95min; 85-100%B in 0.05min; 100% for 0.8min with A= water + 10mM TEA + 200mM HFIP and B= MeOH, Rt= 1.90 min, m/z =6515.97 amu (M after deconvolution with Agilent MassHunter Qualitative Analysis 6.00)(Calc. 6516.05), interpreted as compatible with the structure of the final compound 1a.

The predicted isotope distribution (ChemBioDraw Ultra 14.00) of the product (red) was overlapped with the experimental isotope distribution (Blue) using the Isotope Distribution Calculator add-on from the Agilent MassHunter 6.00 software suite.

Synthesis of 1b

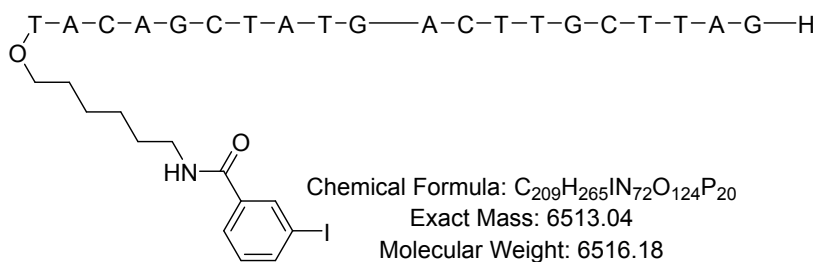


Figure SI II4: Molecular structure of the starting material 1b.

For the synthesis of **1b** the standard acylation was performed using 3-iodobenzoic acid (Fluka CAS 618-51-9) (100mM in DMF) (3000 µl, 300 µmol).

Analysis for 1b

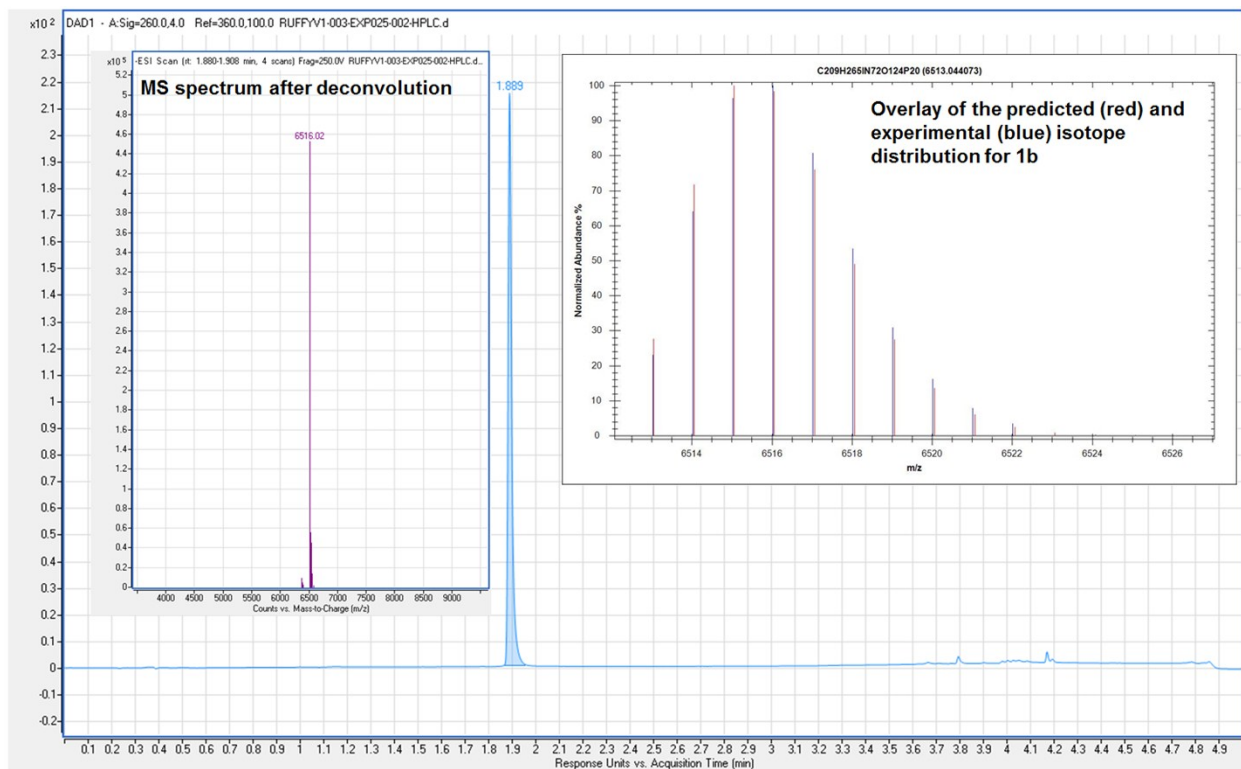


Figure SI II5: UPLC-TOF analysis of **1b**.

- **UPLC/MS** 6200 series TOF/6500 series Q-TOF, Column ACQUITY BEH OST C18 2.1x50 mm 1.7 µm at 60°C.; Flow 0.6 mL/min, Gradient: 15-30%B in 2.7min; 30-85%B in 0.95min; 85-100%B in 0.05min; 100% for 0.8min with A= water + 10mM TEA + 200mM HFIP and B= MeOH, Rt= 1.889 min, m/z =6516.0 amu (M deconvoluted), interpreted as compatible with the structure of the final compound **1b** (M calculated 6516.05).

Synthesis of 1c

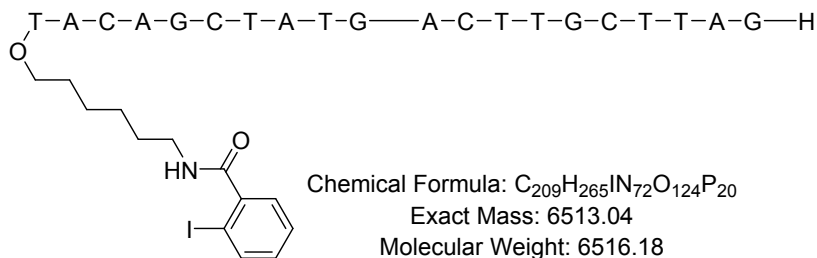


Figure SI II6: Molecular structure of the starting material **1c**.

For the synthesis of **1c** the standard acylation was performed using 2-iodobenzoic acid (Fluka CAS 88-67-5).

Analysis for **1c**

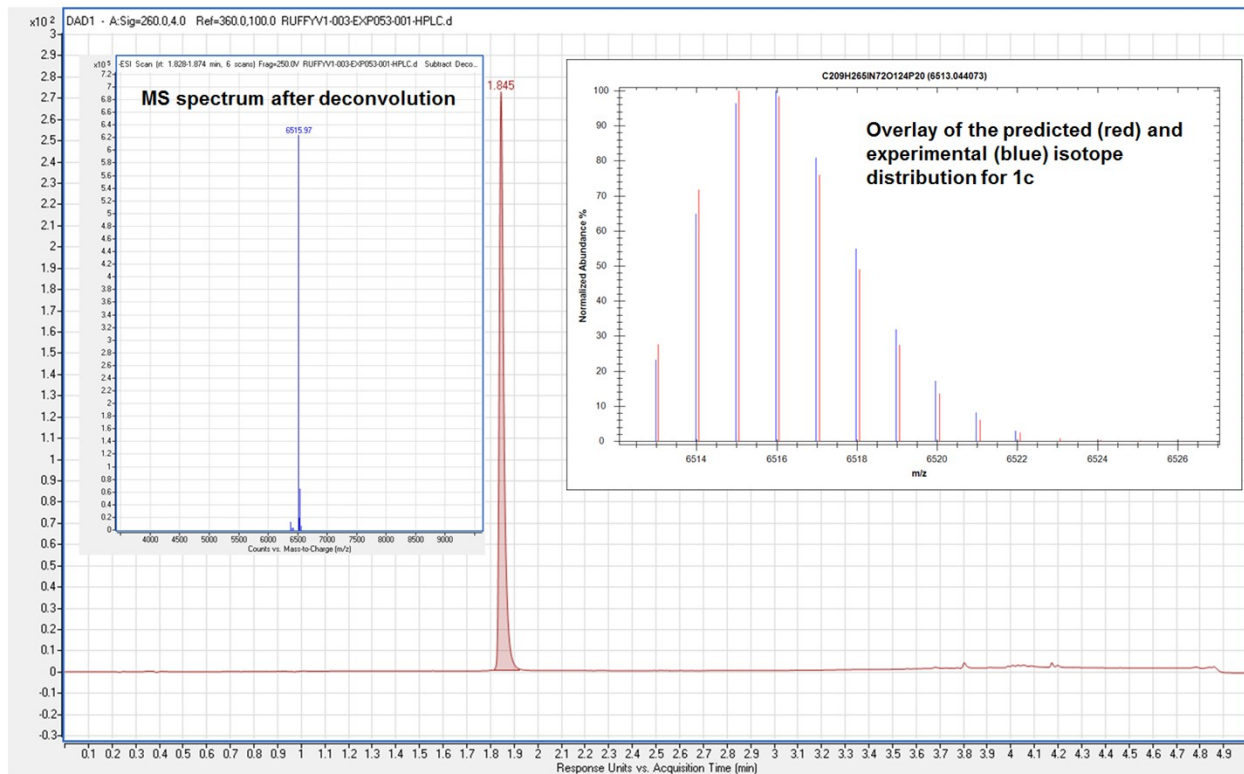


Figure SI II7: UPLC-TOF analysis of **1c**.

- **UPLC/MS** Agilent 6200 series TOF/6500 series Q-TOF, Column ACQUITY BEH OST C18 2.1x50 mm 1.7 μ m at 60°C.; Flow 0.6 mL/min, Gradient: 15-30%B in 2.7min; 30-85%B in 0.95min; 85-100%B in 0.05min; 100% for 0.8min with A= water + 10mM TEA + 200mM HFIP and B= MeOH, R_t = 1.845 min, m/z =6515.97 amu (M after deconvolution with Agilent MassHunter Qualitative Analysis 6.00)(Calc. 6516.05), interpreted as compatible with the structure of the final compound **1c**.

The predicted isotope distribution (ChemBioDraw Ultra 14.00) of the product (red) was overlapped with the experimental isotope distribution (Blue) using the Isotope Distribution Calculator add-on from the Agilent MassHunter 6.00 software suite.

Synthesis of 1d

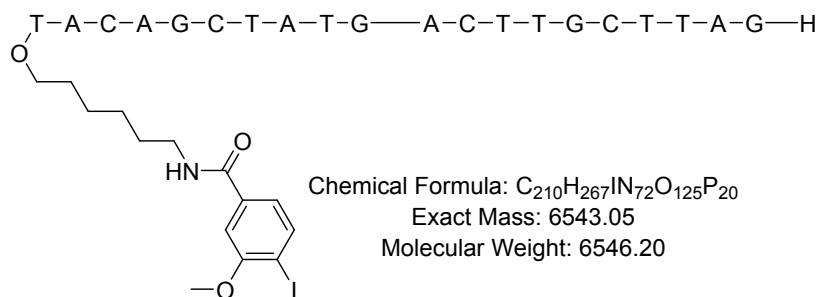


Figure SI II8: Molecular structure of the starting material 1d.

For the synthesis of 1d the standard acylation was performed using 4-iodo-3-methoxybenzoic acid (Chemcia Scientific CAS 282087-44-9).

Analysis for 1d

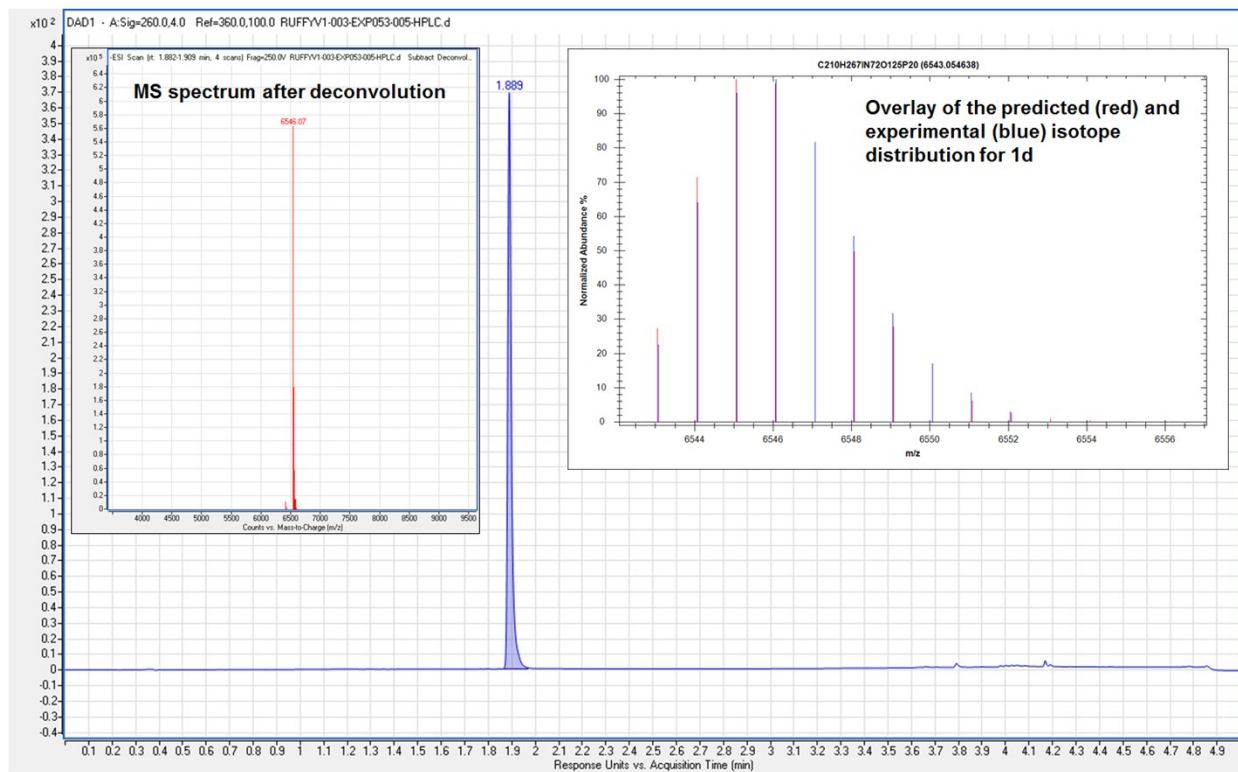


Figure SI II9: UPLC-TOF analysis of 1d.

- **UPLC/MS** Agilent 6200 series TOF/6500 series Q-TOF, Column ACQUITY BEH OST C18 2.1x50 mm 1.7 μ m at 60°C.; Flow 0.6 mL/min, Gradient: 15-30%B in 2.7min; 30-85%B in 0.95min; 85-100%B in 0.05min; 100% for 0.8min with A= water + 10mM TEA + 200mM HFIP and B= MeOH, Rt= 1.889 min, m/z

=6546.07 amu (M after deconvolution with Agilent MassHunter Qualitative Analysis 6.00)(Calc. 6546.06), interpreted as compatible with the structure of the final compound **1d**.

The predicted isotope distribution (ChemBioDraw Ultra 14.00) of the product (red) was overlapped with the experimental isotope distribution (Blue) using the Isotope Distribution Calculator add-on from the Agilent MassHunter 6.00 software suite.

Synthesis of **1e**

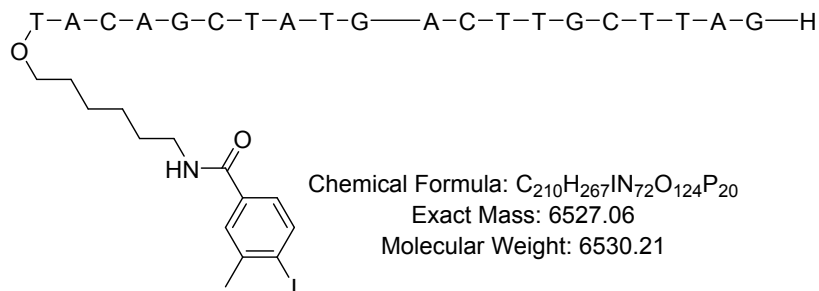


Figure SI II10: Molecular structure of the starting material **1e**.

For the synthesis of **1e** the standard acylation was performed 4-iodo-3-methylbenzoic acid (Trans World Chemicals I1188-J).

Analysis for **1e**

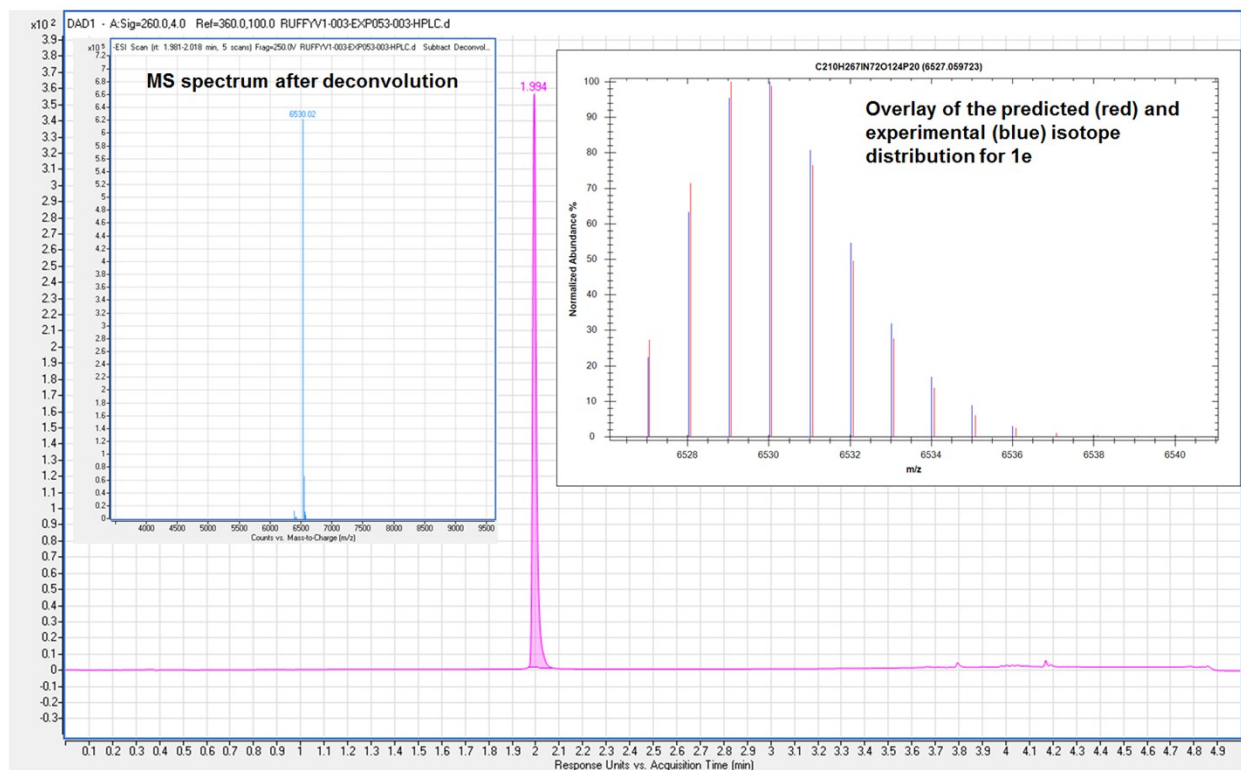


Figure SI II11: UPLC-TOF analysis of **1e**.

- **UPLC/MS** Agilent 6200 series TOF/6500 series Q-TOF, Column ACQUITY BEH OST C18 2.1x50 mm 1.7 μ m at 60°C; Flow 0.6 mL/min, Gradient: 15-30%B in 2.7min; 30-85%B in 0.95min; 85-100%B in 0.05min; 100% for 0.8min with A= water + 10mM TEA + 200mM HFIP and B= MeOH, Rt= 1.994 min, m/z =6530.02 amu (M after deconvolution with Agilent MassHunter Qualitative Analysis 6.00)(Calc. 6530.06), interpreted as compatible with the structure of the final compound **1e**.

The predicted isotope distribution (ChemBioDraw Ultra 14.00) of the product (red) was overlapped with the experimental isotope distribution (Blue) using the Isotope Distribution Calculator add-on from the Agilent MassHunter 6.00 software suite.

Synthesis of **1f**

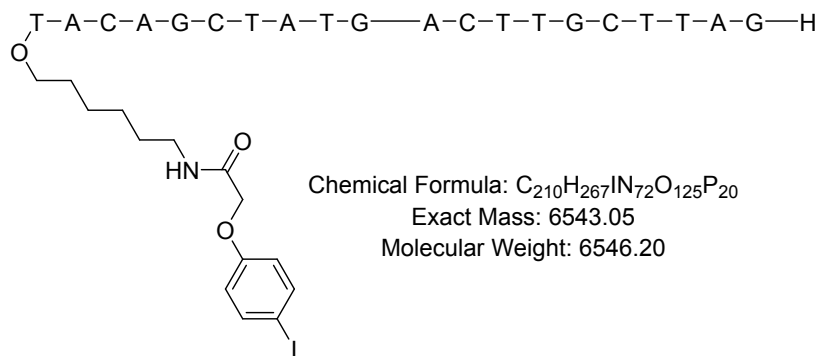


Figure SI II12: Molecular structure of the starting material **1f**.

For the synthesis of **1f** the standard acylation was performed using 4-iodophenoxyacetic acid (Aldrich CAS 1878-94-0).

Analysis for **1f**

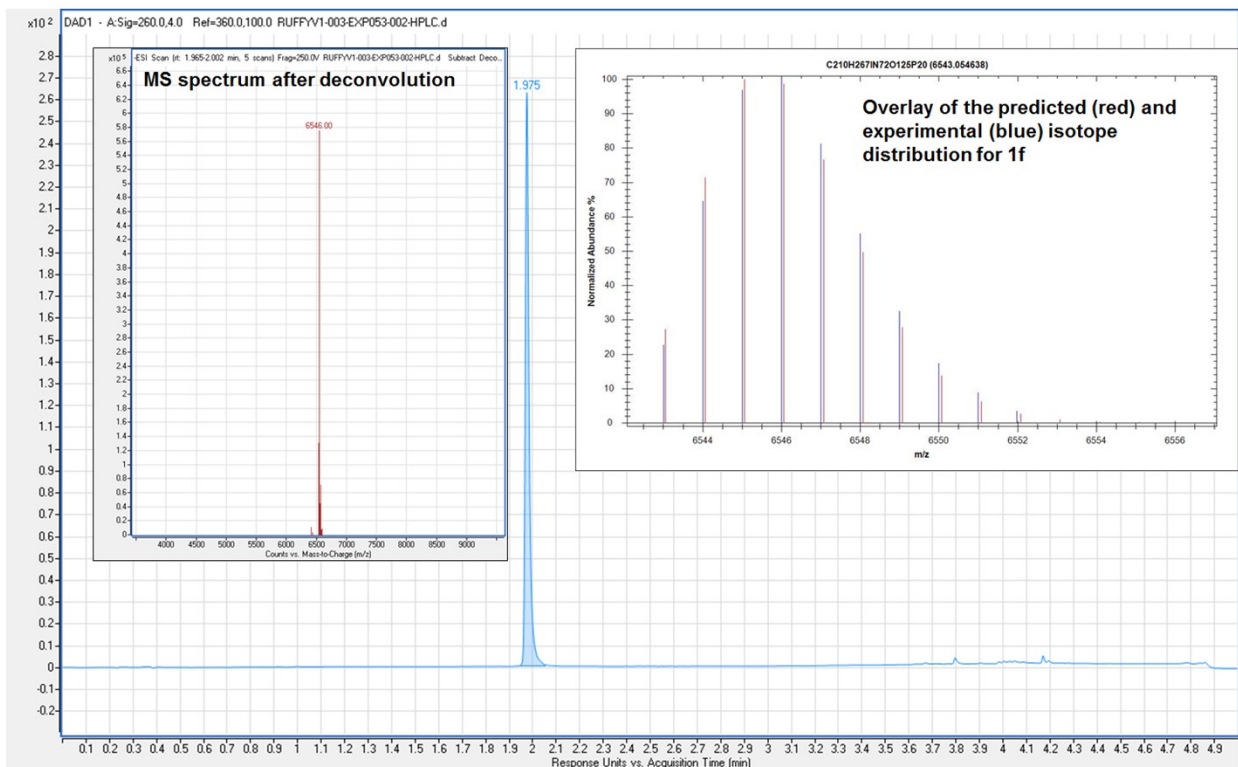


Figure SI II13: UPLC-TOF analysis of **1f**.

- **UPLC/MS** Agilent 6200 series TOF/6500 series Q-TOF, Column ACQUITY BEH OST C18 2.1x50 mm 1.7 um at 60°C.; Flow 0.6 mL/min, Gradient: 15-30%B in 2.7min; 30-85%B in 0.95min; 85-100%B in 0.05min; 100% for 0.8min with A= water + 10mM TEA + 200mM HFIP and B= MeOH, Rt= 1.975 min, m/z =6546.00 amu (M after deconvolution with Agilent MassHunter Qualitative Analysis 6.00)(Calc. 6546.06), interpreted as compatible with the structure of the final compound **1f**.

The predicted isotope distribution (ChemBioDraw Ultra 14.00) of the product (red) was overlapped with the experimental isotope distribution (Blue) using the Isotope Distribution Calculator add-on from the Agilent MassHunter 6.00 software suite.

Synthesis of **1g**

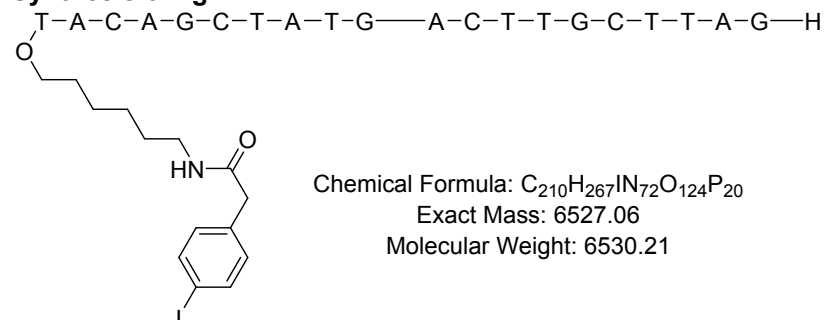


Figure SI II14: Molecular structure of the starting material **1g**.

For the synthesis of **1g** the standard acylation was performed using 4-iodophenylacetic acid (Alfa Aesar CAS 1798-06-7).

Analysis for **1g**

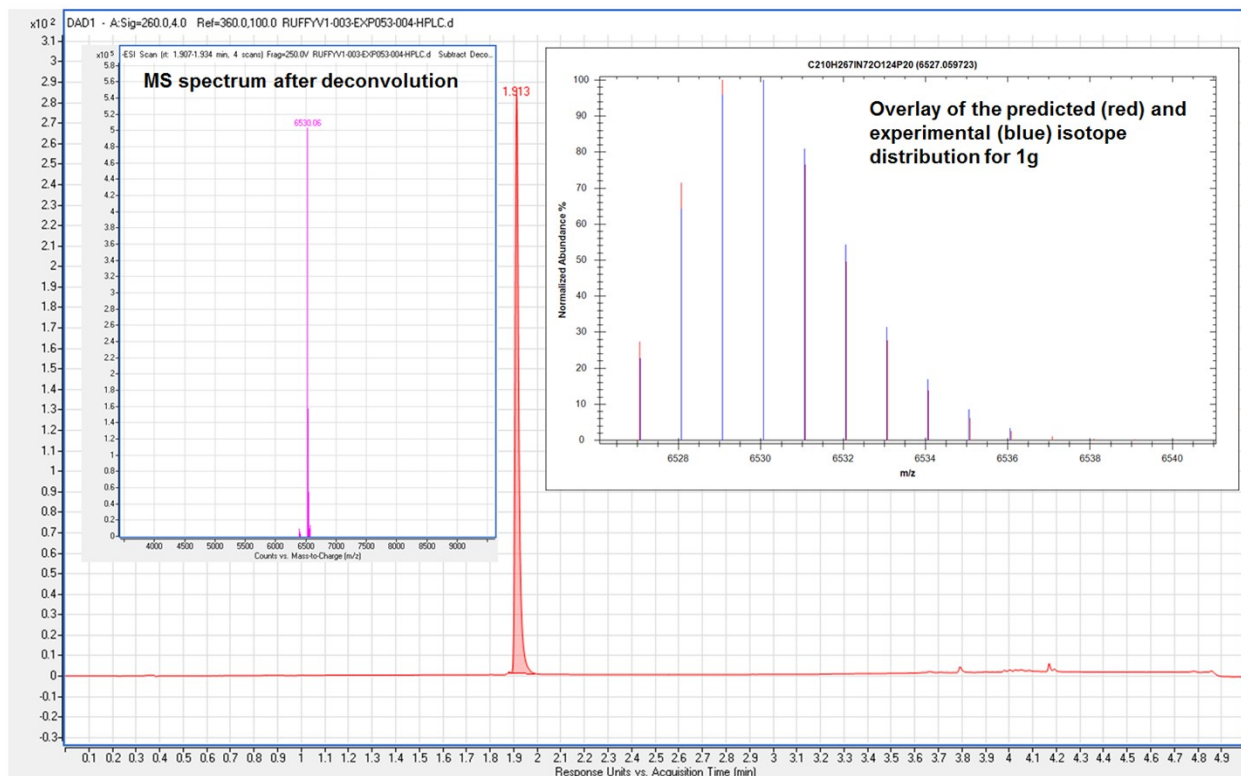


Figure SI II15: UPLC-TOF analysis of **1g**.

- **UPLC/MS** Agilent 6200 series TOF/6500 series Q-TOF, Column ACQUITY BEH OST C18 2.1x50 mm 1.7 μ m at 60°C.; Flow 0.6 mL/min, Gradient: 15-30%B in 2.7min; 30-85%B in 0.95min; 85-100%B in 0.05min; 100% for 0.8min with A= water + 10mM TEA + 200mM HFIP and B= MeOH, Rt= 1.913 min, m/z =6530.06 amu (M after deconvolution with Agilent MassHunter Qualitative Analysis 6.00)(Calc. 6530.06), interpreted as compatible with the structure of the final compound **1g**.

The predicted isotope distribution (ChemBioDraw Ultra 14.00) of the product (red) was overlapped with the experimental isotope distribution (Blue) using the Isotope Distribution Calculator add-on from the Agilent MassHunter 6.00 software suite.

Synthesis of 1h

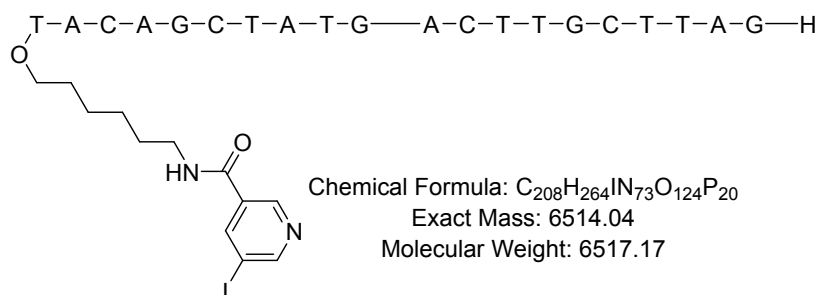


Figure SI II16: Molecular structure of the starting material 1h.

For the synthesis of 1h the standard acylation was performed using 5-iodopyridine-3-carboxylic acid (Ark- Pharm. CAS 15366-65-1).

Analysis for 1h

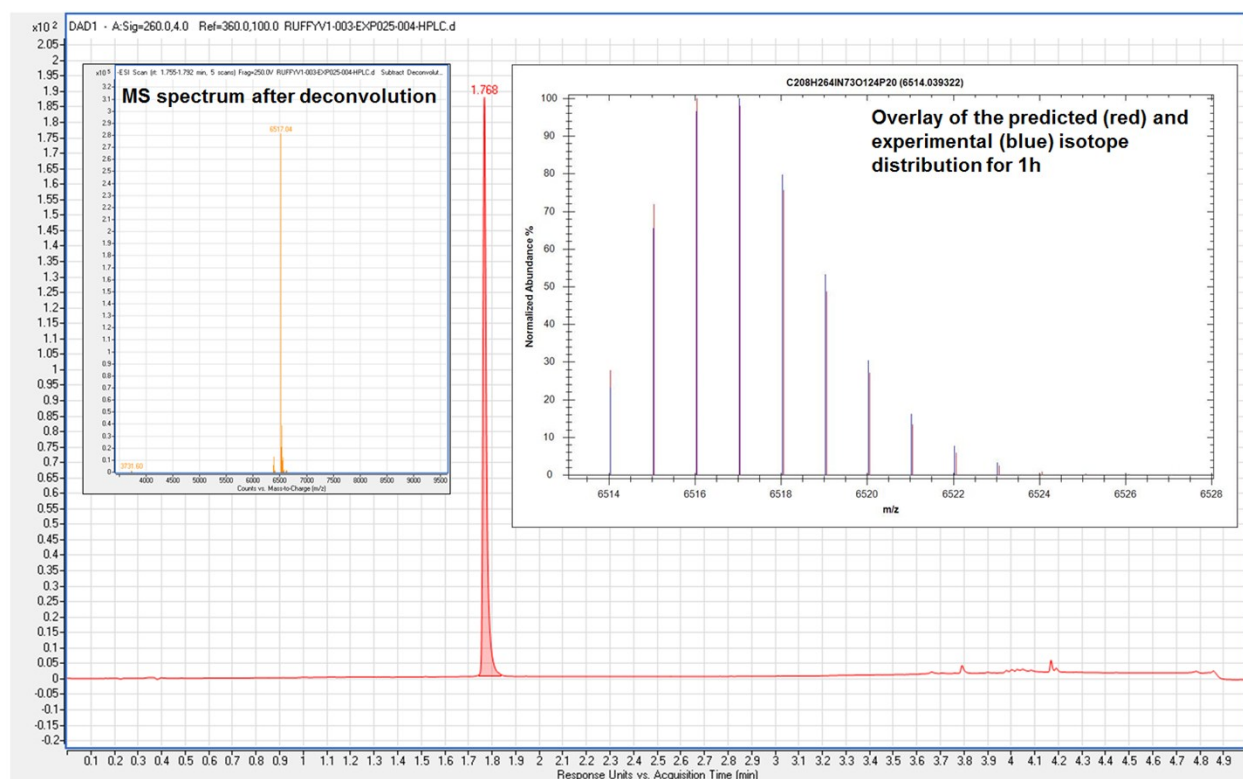


Figure SI II17: UPLC-TOF analysis of 1h.

- **UPLC/MS** Agilent 6200 series TOF/6500 series Q-TOF, Column ACQUITY BEH OST C18 2.1x50 mm 1.7 μ m at 60°C.; Flow 0.6 mL/min, Gradient: 15-30%B in 2.7min; 30-85%B in 0.95min; 85-100%B in 0.05min; 100% for 0.8min with A= water + 10mM TEA + 200mM HFIP and B= MeOH, Rt= 1.768 min, m/z

=6517.00 amu (M after deconvolution with Agilent MassHunter Qualitative Analysis 6.00)(Calc. 6517.04), interpreted as compatible with the structure of the final compound **1h**.

The predicted isotope distribution (ChemBioDraw Ultra 14.00) of the product (red) was overlapped with the experimental isotope distribution (Blue) using the Isotope Distribution Calculator add-on from the Agilent MassHunter 6.00 software suite.

Synthesis of **1i**

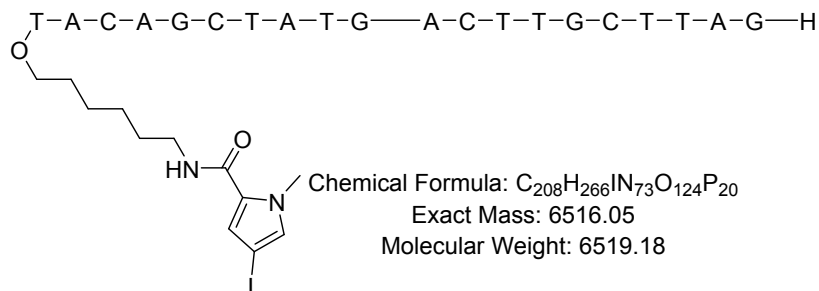


Figure SI II18: Molecular structure of the starting material **1i**.

For the synthesis of **1i** the standard acylation was performed using 4-iodo-1-methyl-1H-pyrrole-2-carboxylic acid (BIONET FG-0709).

Analysis for **1i**

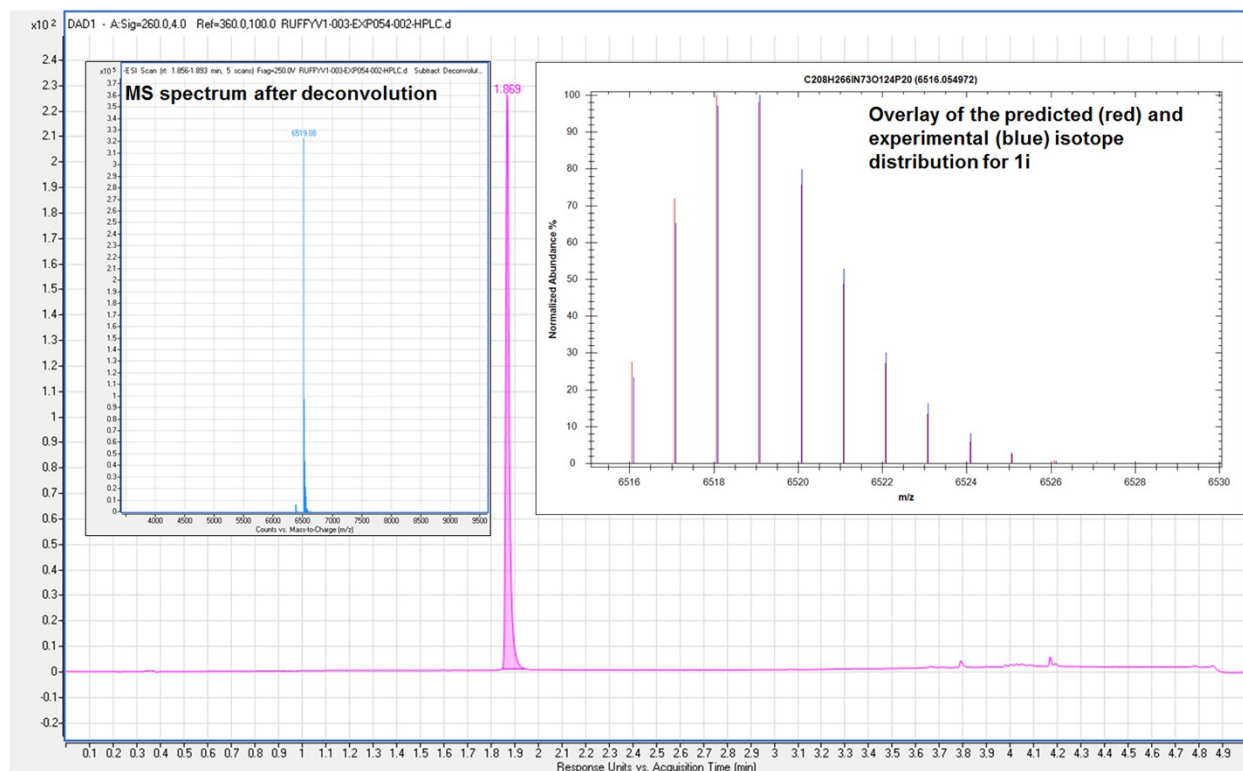


Figure SI II19: UPLC-TOF analysis of **1i**.

- **UPLC/MS** Agilent 6200 series TOF/6500 series Q-TOF, Column ACQUITY BEH OST C18 2.1x50 mm 1.7 μ m at 60°C.; Flow 0.6 mL/min, Gradient: 15-30%B in 2.7min; 30-85%B in 0.95min; 85-100%B in 0.05min; 100% for 0.8min with A= water + 10mM TEA + 200mM HFIP and B= MeOH, Rt= 1.869 min, m/z =6519.03 amu (M after deconvolution with Agilent MassHunter Qualitative Analysis 6.00)(Calc. 6519.06), interpreted as compatible with the structure of the final compound **1i**.

The predicted isotope distribution (ChemBioDraw Ultra 14.00) of the product (red) was overlapped with the experimental isotope distribution (Blue) using the Isotope Distribution Calculator add-on from the Agilent MassHunter 6.00 software suite.

SI III Synthesis and Characterization of structural analogs of DMPAO L11-L21

The synthesis of **L11-L21** was adapted from the synthesis of **L8** (DMPAO) described by Ma *et al.*²

Standard protocol for the synthesis of **L11**

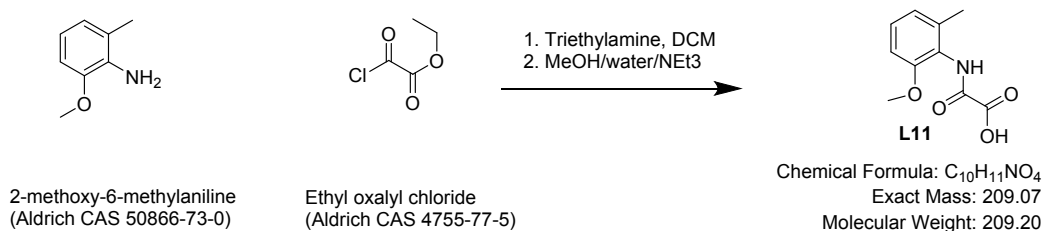


Figure SI III1 : Synthesis of **L11**.

2-methoxy-6-methylaniline (Aldrich CAS 50866-73-0) (0.896 g, 6.53 mmol) was dissolved in CH₂Cl₂ (volume: 12 mL) under argon. Triethylamine (1.820 mL, 13.06 mmol) was then added to this solution followed by ethyl oxalyl chloride (Aldrich CAS 4755-77-5) (0.802 mL, 7.18 mmol). The reaction was stirred at room temperature for 2 hr.

Workup

The reaction was diluted with 100ml of DCM. The organic phase was washed with water and brine. The organic phase was dried on Na₂SO₄, filtered and concentrated in vacuo.

The crude product was saponified by dissolving it in MeOH/H₂O/NEt₃ 50/20/10 and stirring at room temperature until complete hydrolysis (3 hours)

After 3 hours the pH of the solution was adjusted to 2 by adding concentrated HCl. The reaction was diluted with DCM and washed 3 times with water. The organic phase was dried on Na₂SO₄, filtered and concentrated in vacuo.

Purification

The crude product was recrystallized from cyclohexane/ethyl acetate to yield the expected product **L11** (Yield: 47%).

Analysis of L11

^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 14.02 (s, 1H), 9.91 (s, 1H), 7.19 (t, $J = 7.9$ Hz, 1H), 6.88 (dd, $J = 22.8, 7.9$ Hz, 2H), 3.74 (s, 3H), 2.22 (s, 0H), 2.12 (s, 3H).

^{13}C NMR (101 MHz, $\text{DMSO-}d_6$) δ 162.12, 157.00, 154.76, 136.40, 127.77, 123.58, 121.89, 109.25, 55.59, 39.52, 17.70.

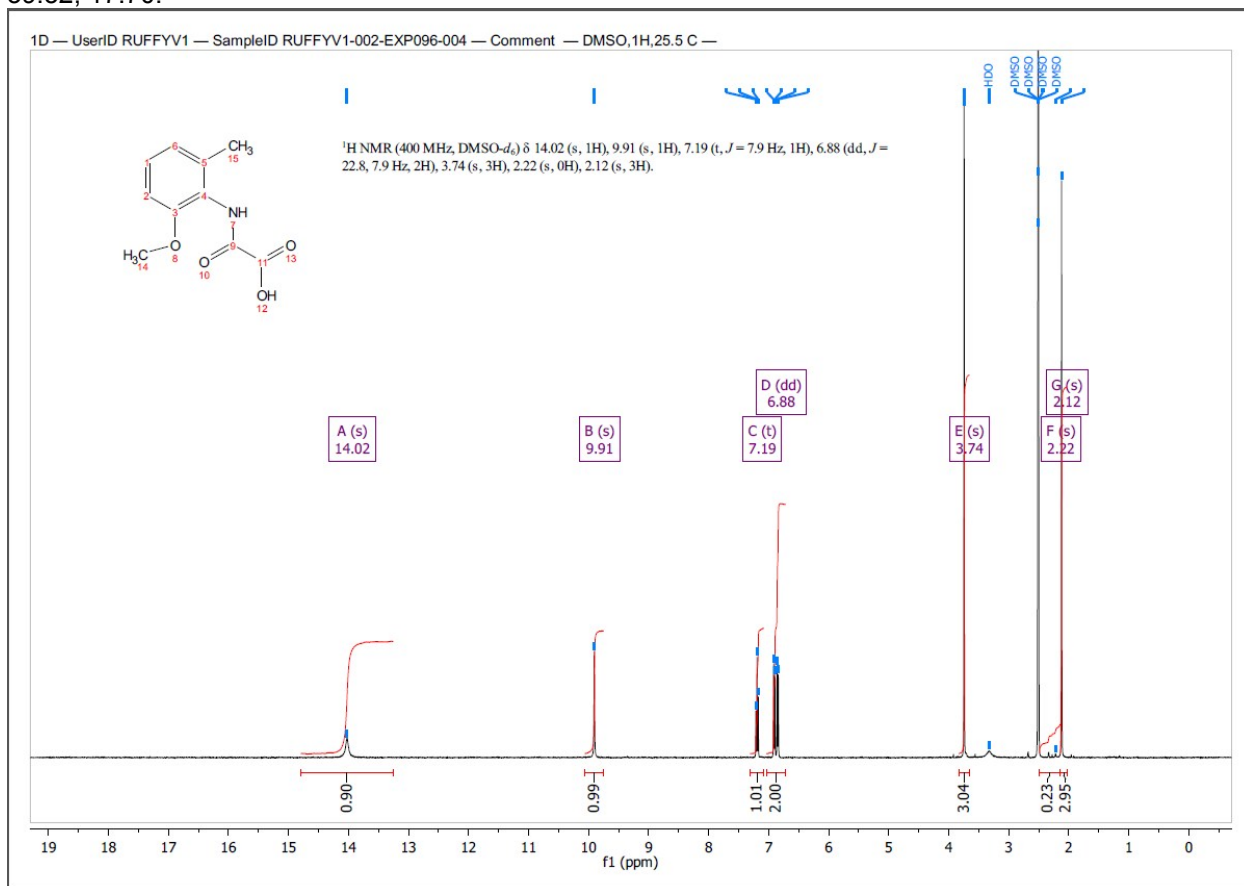


Figure SI III2: ^1H NMR analysis of L11.

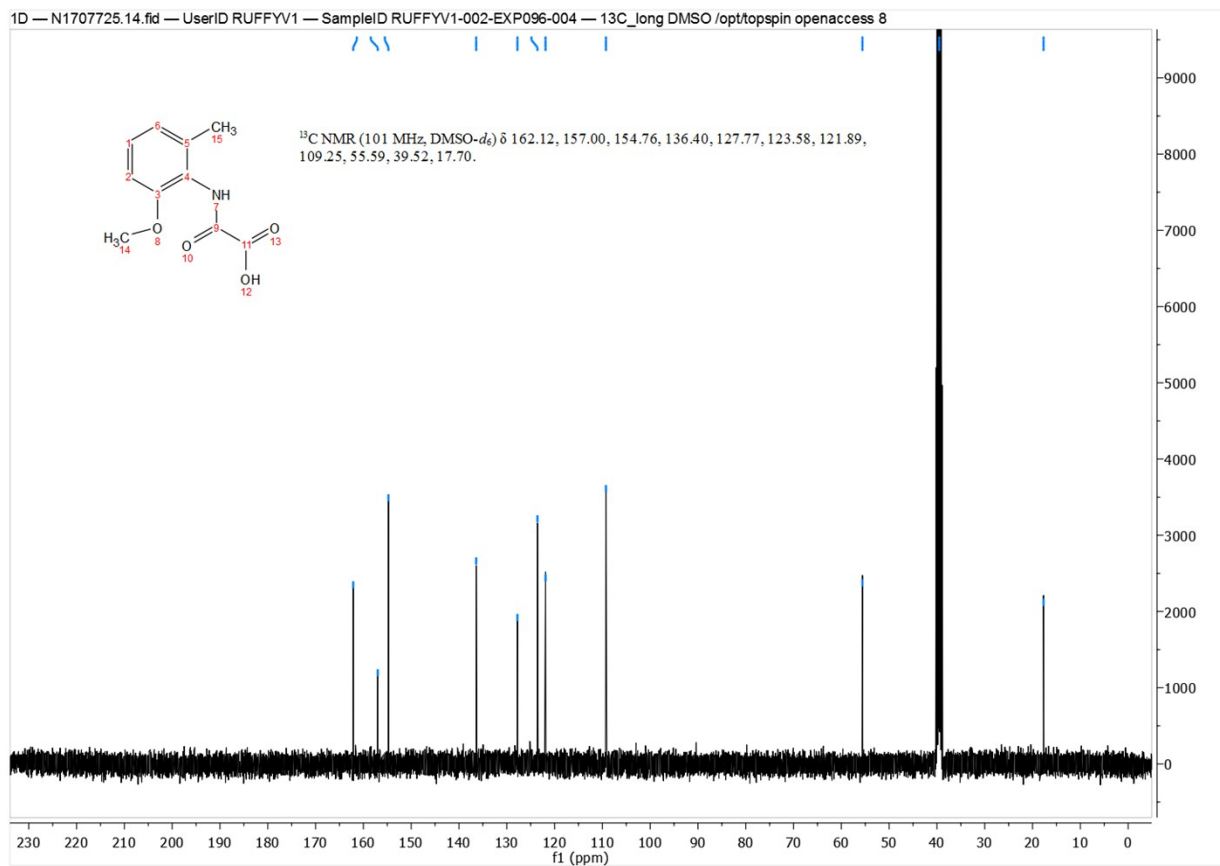


Figure SI III3 : ¹³C NMR analysis of L11.

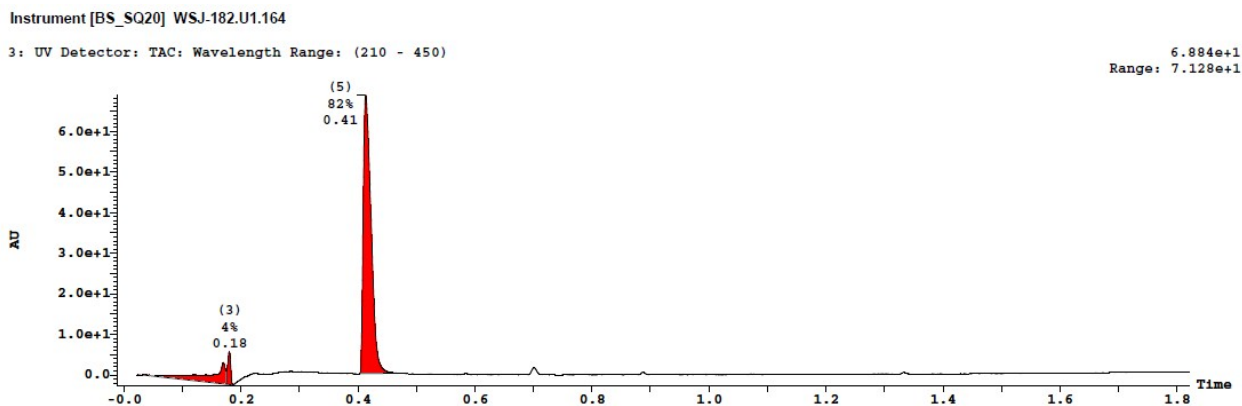


Figure SI III4: UPLC-MS analysis of L11.

UPLC / ESI Method: Column: Acquity UPLC HSS T3 1.8 μ m 2.1x 50mm at 60°C. Gradient: from 5 to 98 % B in 1.4 min - flow 1.0 mL/min Eluent A: water + 0.05 % formic acid + 3.75 mM ammonium acetate, Eluent B: acetonitrile + 0.04 % formic acid. Rt= 0.41 min, m/z =208.1 amu (M-H), interpreted as compatible with the structure of the final compound L11.

Synthesis of L12:

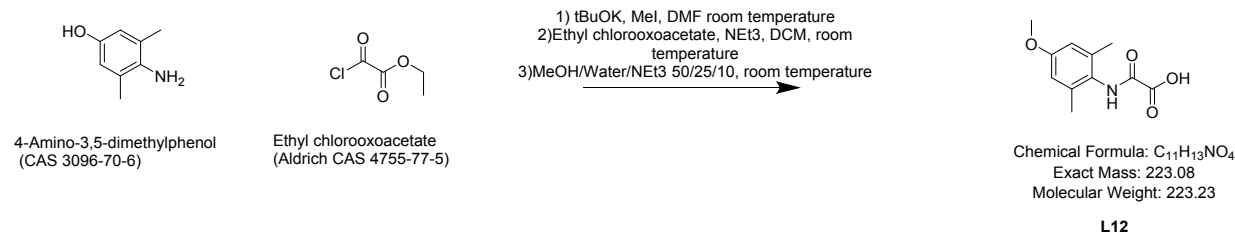


Figure SI III5: Synthesis of **L12**.

4-Amino-3,5-dimethylphenol (CAS 3096-70-6) (3.110 g, 22.67 mmol) was dissolved in DMF (volume: 25 mL) and cooled to 0°C with an ice bath in a 100ml flask. The reaction was placed under a positive pressure of argon and potassium tert-butoxide (3.74 g, 33.3 mmol) added portionwise. The resulting red suspension was treated with methyl iodide (1.2 mL, 19.19 mmol) as a solution in DMF (volume: 25 mL). Addition was performed dropwise using an addition funnel. The reaction was then stirred at room temperature overnight.

Work up

The reaction was diluted with methyl-*t*-butylether and washed three times with 1M KOH in water, then once with brine.

The organic phase was dried with Na₂SO₄, filtered and concentrated in vacuo to give the crude 4-methoxy-2,6-dimethylaniline which was dissolved in 50ml of DCM. Triethylamine (4.74 mL, 34.0 mmol) was added. The reaction was put under argon and cooled at 0°C followed by the addition of ethyl chlorooxoacetate (Aldrich CAS 4755-77-5) (3.80 mL, 34.0 mmol).

The reaction was stirred at room temperature for 2 hours followed by the evaporation of the solvents. The crude product was dissolved in MeOH (50ml) / water (25ml) / trimethylamine (10ml) and hydrolysis of the ethyl ester conducted for 12 hours at room temperature.

Purification

The crude product was recrystallized from cyclohexane/ethyl acetate to yield 1.563g of **L12** as light brown needles.

Analysis of L12

¹H NMR (400 MHz, DMSO-*d*₆) δ 9.64 (s, 1H), 6.65 (s, 2H), 3.72 (s, 3H), 2.09 (s, 6H).

¹³C NMR (101 MHz, DMSO-*d*₆) δ 162.28, 157.73, 157.34, 136.28, 126.68, 112.88, 55.07, 18.13.

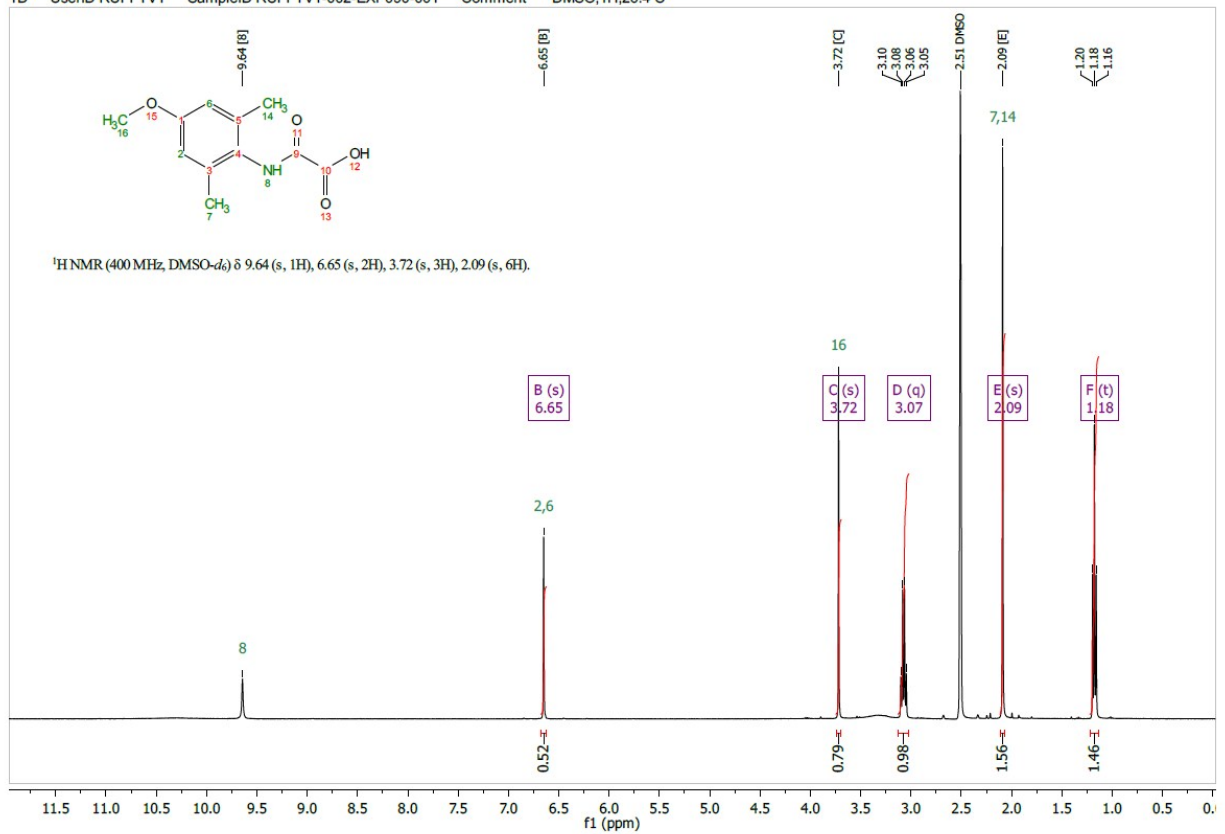


Figure SI III6: ¹H NMR analysis of L12.

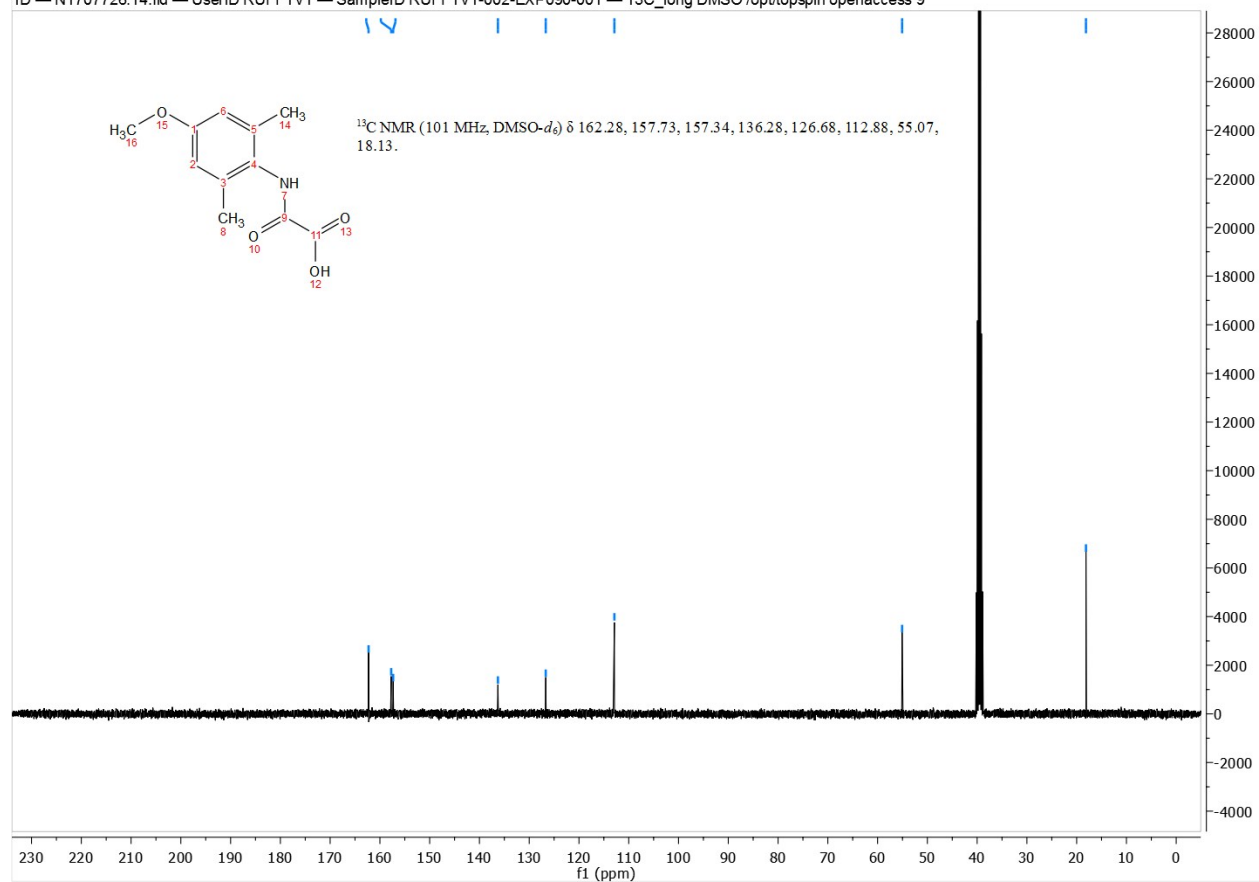


Figure SI III7: ¹³C NMR analysis of L12.

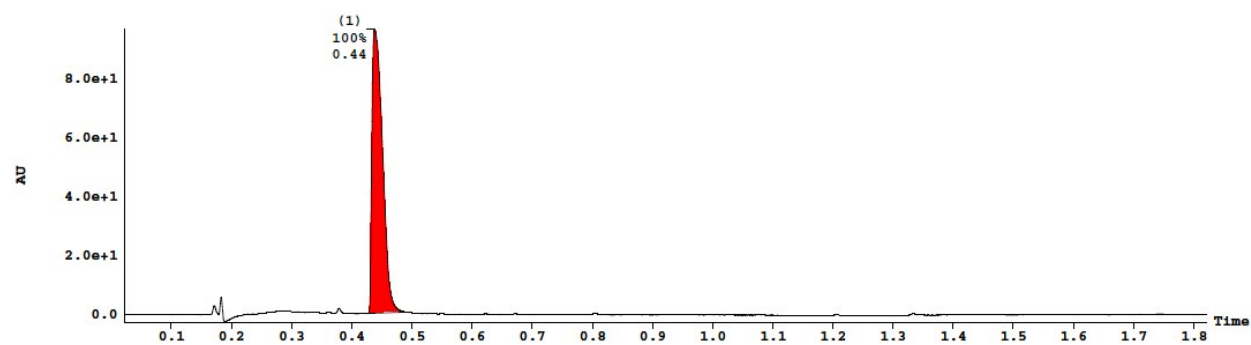


Figure SI III8: UPLC-MS analysis of L12.

UPLC / ESI Method: Column: Acquity UPLC HSS T3 1.8μm 2.1x 50mm at 60°C. Gradient: from 5 to 98 % B in 1.4 min - flow 1.0 mL/min Eluent A: water + 0.05 % formic acid + 3.75 mM ammonium acetate, Eluent B: acetonitrile + 0.04 % formic acid, rt 0.44min, m/z 222.2 (M-H) interpreted as compatible with the structure of L12.

Synthesis of L13

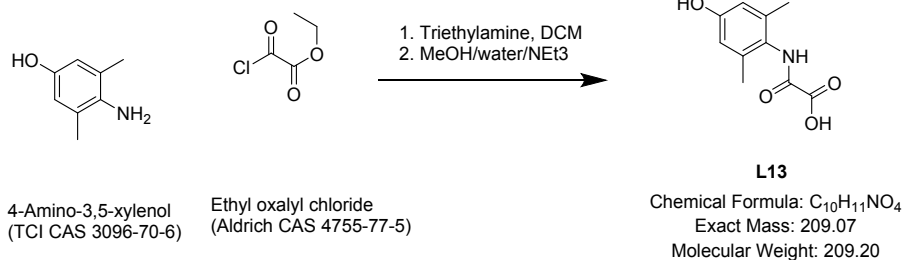


Figure SI III9: Synthesis of L13.

The standard protocol used for the synthesis of L11 was used, starting from 4-amino-3,5-xyleneol (TCI CAS 3096-70-6) (0.896 g, 6.53 mmol) yielding 315mg (23%) of L13.

Analysis of L13

¹H NMR (400 MHz, DMSO-*d*₆) δ 9.90 (s, 1H), 9.24 (s, 1H), 6.48 (s, 2H), 2.03 (s, 6H).

¹³C NMR (101 MHz, DMSO-*d*₆) δ 162.37, 157.39, 155.90, 136.07, 125.13, 114.25, 18.03.

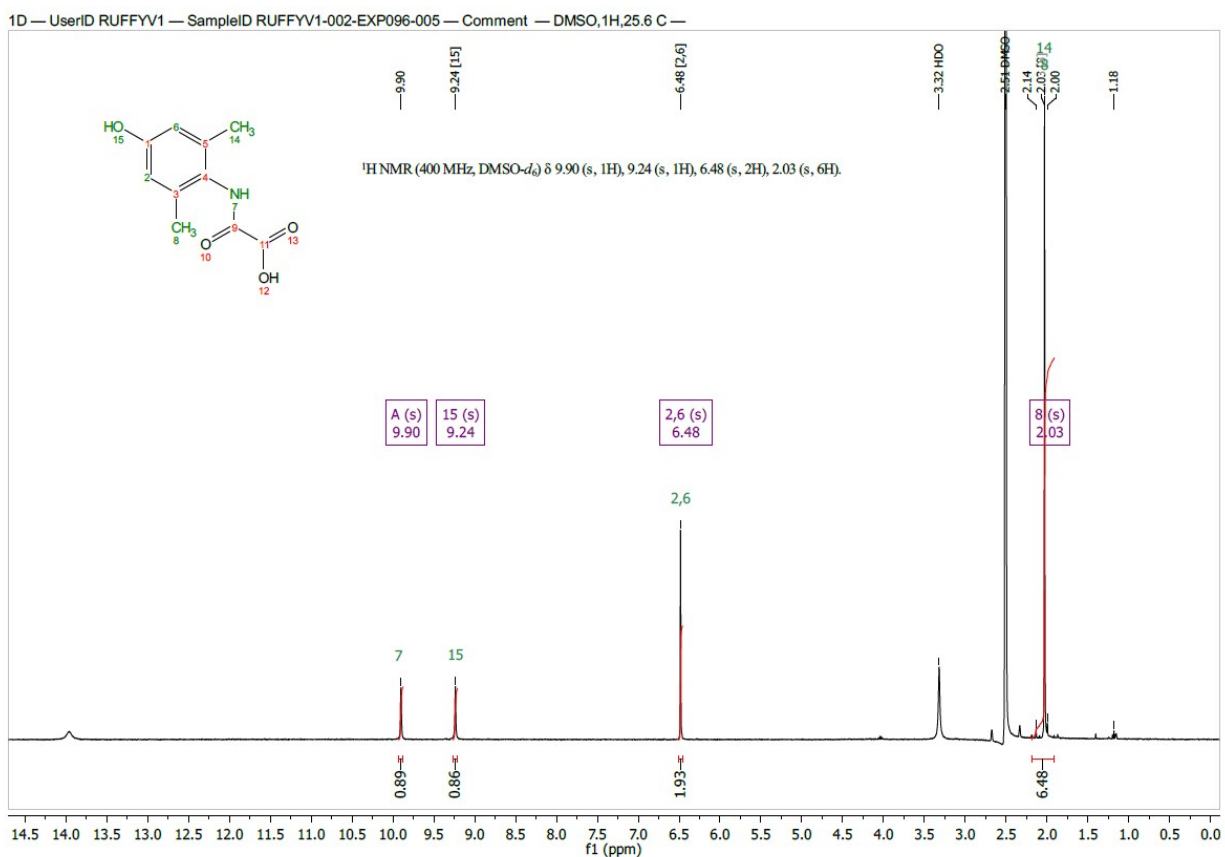


Figure SI III10: ¹H NMR analysis of L13.

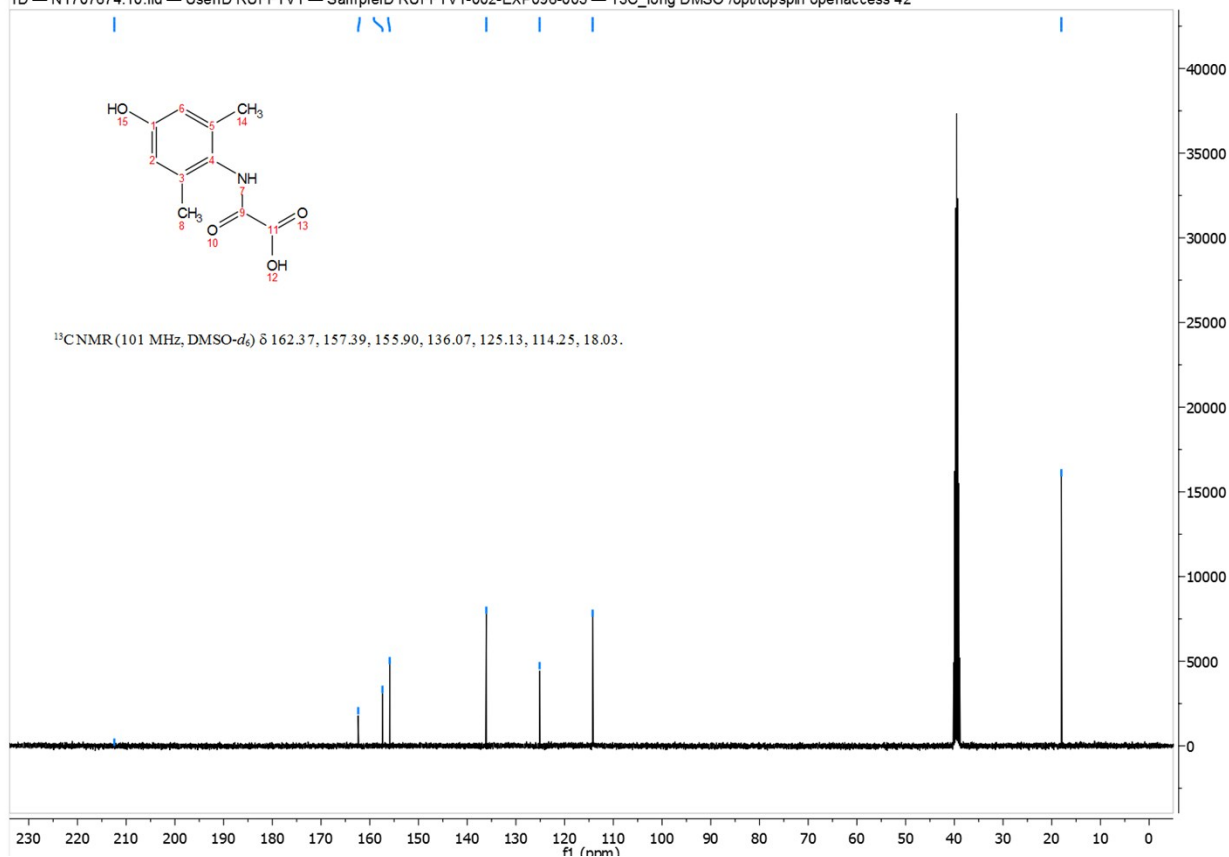


Figure SI III11: ¹³C NMR analysis of L13.

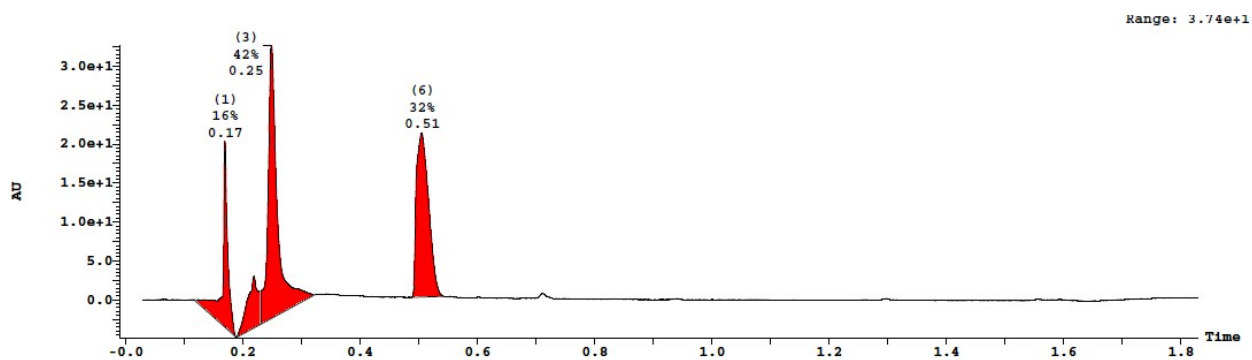


Figure SI III12: UPLC-MS analysis of L13.

UPLC / ESI Instrument: SQ17. **Open_Access_MS/LCMS.** **Method:** Column: Acquity UPLC HSS T3 1.8µm 2.1x 50mm at 60°C **Gradient:** from 5 to 98 % B in 1.4 min - flow 1.0 mL/min **Eluent A:** water + 0.05 % formic acid + 3.75 mM ammonium acetate, **Eluent B:** acetonitrile + 0.04 % formic acid. **Rt= 0.25 min, m/z =208.2 amu (M-H),** interpreted as compatible with the structure of the final compound **L13**. The peak at 0.51min corresponds to ethyl acetate.

Synthesis of L14

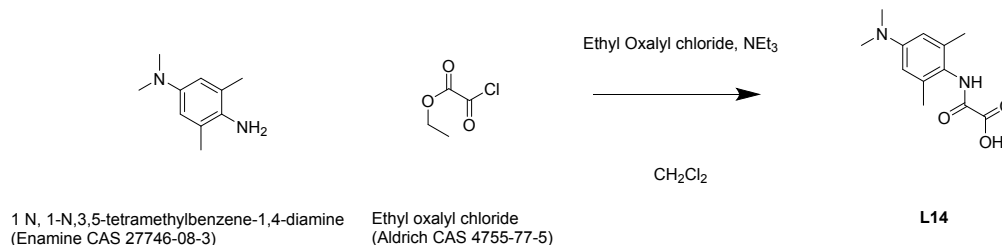


Figure SI III13: Synthesis of L14.

1 N, 1-N,3,5-tetramethylbenzene-1,4-diamine (Enamine CAS 27746-08-3) (0.4 g, 2.435 mmol) was dissolved in DCM (Volume: 24 mL). Triethylamine (0.509 mL, 3.65 mmol) was added followed by the dropwise addition over 10 min of ethyl oxalyl chloride (Aldrich CAS 4755-77-5) (0.300 mL, 2.68 mmol). The reaction was diluted in 100ml of DCM, washed twice with water and the organic phase concentrated in vacuo.

The crude product was saponified by dissolving it in MeOH (50ml)/ water (25ml)/ trimethylamine (10ml) and the hydrolysis of the ethyl ester conducted for 12 hours at room temperature. The solvents were evaporated and the crude product purified by reverse phase HPLC.

Analysis of L14

^1H NMR (400 MHz, $\text{DMSO}-d_6$) δ 9.97 (s, 1H), 6.59 (s, 2H), 2.92 (s, 7H), 2.09 (s, 6H).

^{13}C NMR (101 MHz, $\text{DMSO}-d_6$) δ 162.34, 157.42, 146.96, 135.96, 117.44, 114.53, 41.97, 18.37.

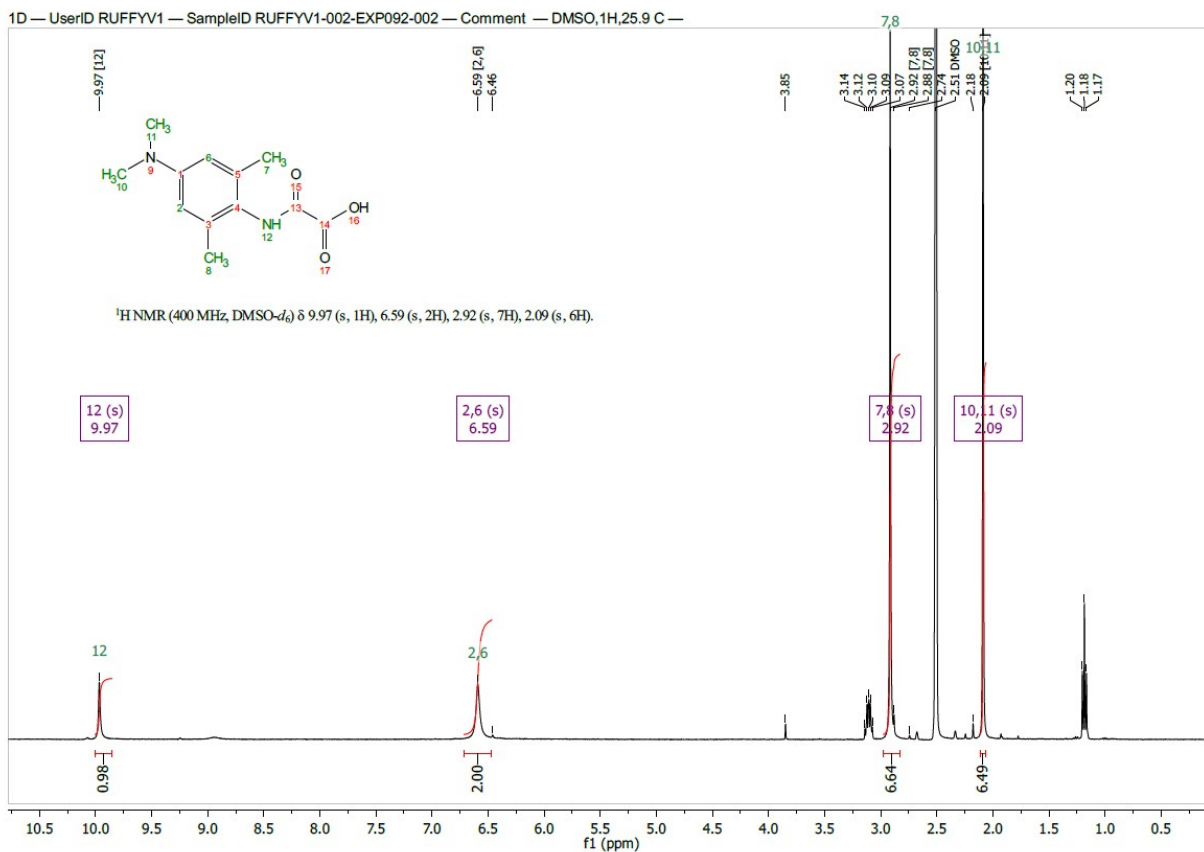


Figure SI III14: ^1H NMR analysis of L14.

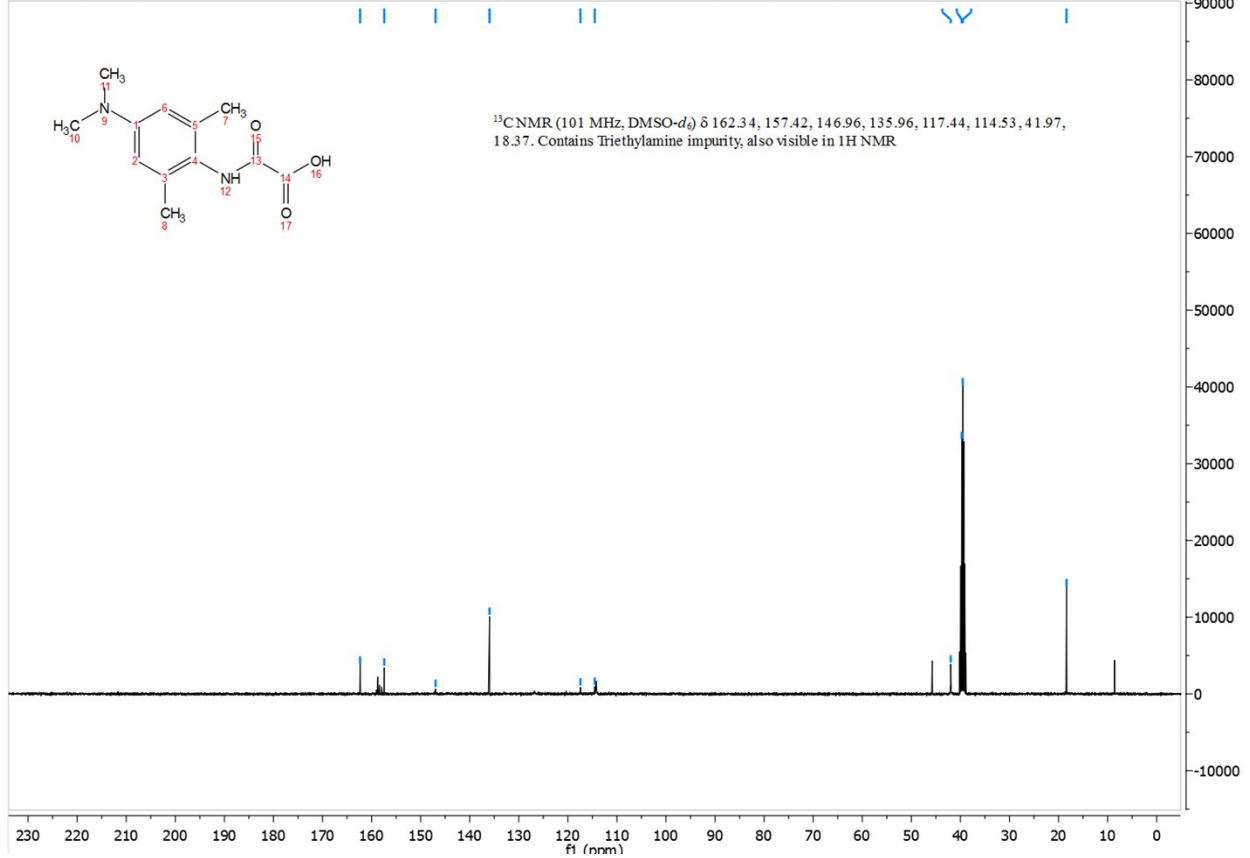


Figure SI III15: ^{13}C NMR analysis of **L14**.

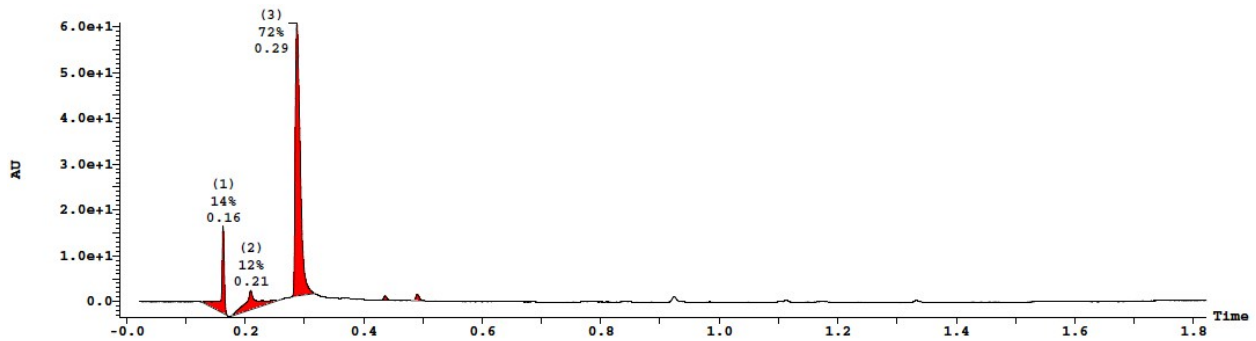
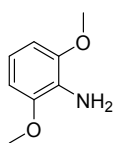


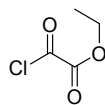
Figure SI III16: UPLC-MS analysis of **L14**.

UPLC / ESI. Method: Column: Acquity UPLC HSS T3 $1.8\mu\text{m}$ $2.1 \times 50\text{mm}$ at 60°C . Gradient: from 5 to 98 % B in 1.4 min - flow 1.0 mL/min Eluent A: water + 0.05 % formic acid + 3.75 mM ammonium acetate, Eluent B: acetonitrile + 0.04 % formic acid, $R_t = 0.29$ min, $m/z = 237.2$ amu ($M+H$), interpreted as compatible with the structure of the final compound **L14**.

Synthesis of L15

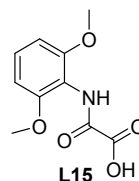


2,6-dimethoxyaniline
(Alfa Aesar CAS -2734-70-5)



Ethyl oxalyl chloride
(Aldrich CAS 4755-77-5)

1. Triethylamine, DCM
2. MeOH/Water/NEt₃



Chemical Formula: C₁₀H₁₁NO₅
Exact Mass: 225.06
Molecular Weight: 225.20

Figure SI III17: Synthesis of L15.

The synthesis of **L15** was performed starting from 2,6-dimethoxyaniline (Alfa Aesar CAS -2734-70-5) (3 g, 19.5mmol) was dissolved in CH₂Cl₂ (Volume: 40 mL) under argon. Triethylamine (3.00 mL, 21.5 mmol) was then added to this solution followed by ethyl oxalyl chloride (Aldrich CAS 4755-77-5) (2.406 mL, 21.5 mmol). The reaction was stirred at room temperature for 2 hr. The reaction was diluted with 100ml of DCM. The organic phase was washed with water and brine. The organic phase was dried on Na₂SO₄, filtered and concentrated in vacuo. The crude ethyl ester intermediate was saponified by dissolving it in MeOH (50ml)/ water (20ml)/ trimethylamine (10ml) and stirring at room temperature until complete hydrolysis (3 hours).

After 3 hours the pH of the solution was adjusted to 2 by addition of concentrated HCl. The reaction was diluted with DCM and washed 3 times with water. The organic phase was dried on Na₂SO₄, filtered and concentrated in vacuo.

Purification

The crude product was recrystallized from water to yield 2.65 g of **L15** as light brown needles.

Analysis

¹H NMR (400 MHz, DMSO-d₆) δ 13.94 (s, 1H), 9.63 (s, 1H), 7.26 (t, J = 8.4 Hz, 1H), 6.71 (d, J = 8.4 Hz, 2H), 3.75 (s, 6H).

¹³C NMR (101 MHz, DMSO-d₆) δ 162.12, 157.00, 155.78, 128.35, 113.20, 104.42, 55.80.

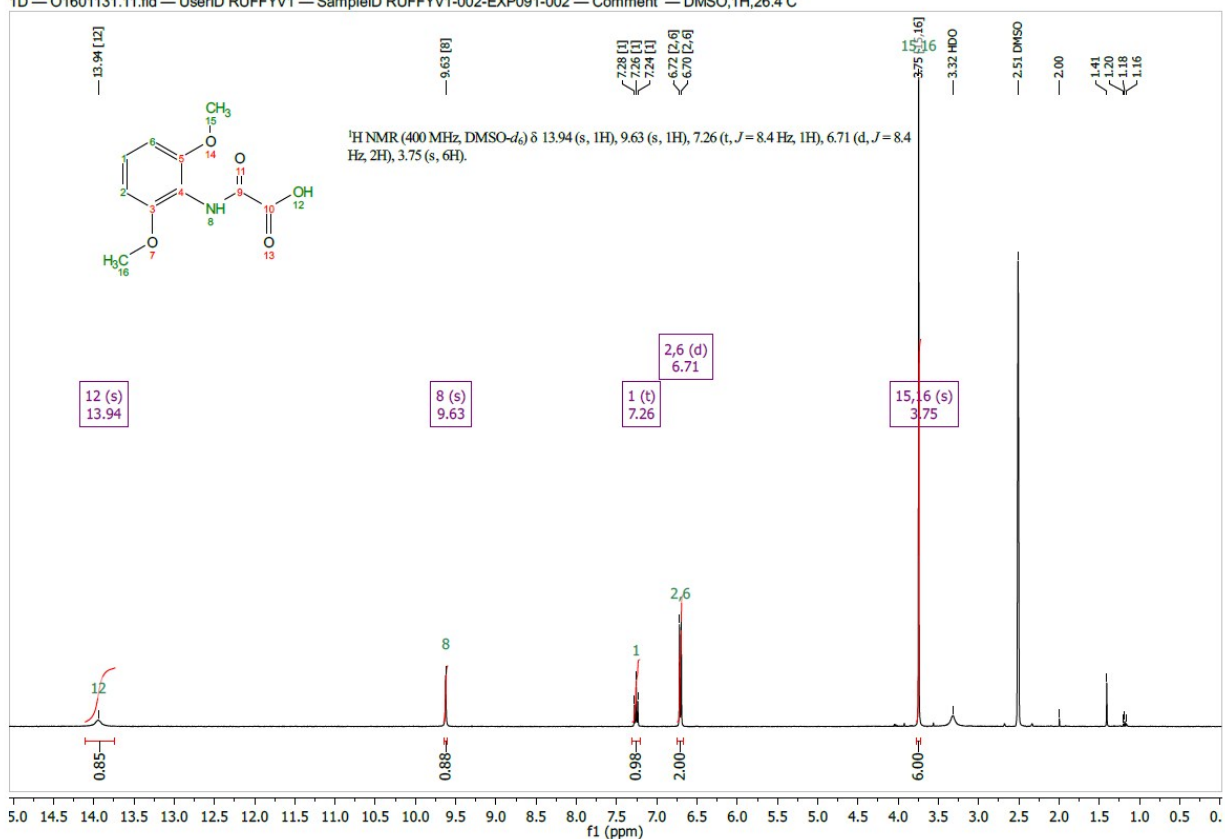


Figure SI III18: ¹H NMR analysis of L15.

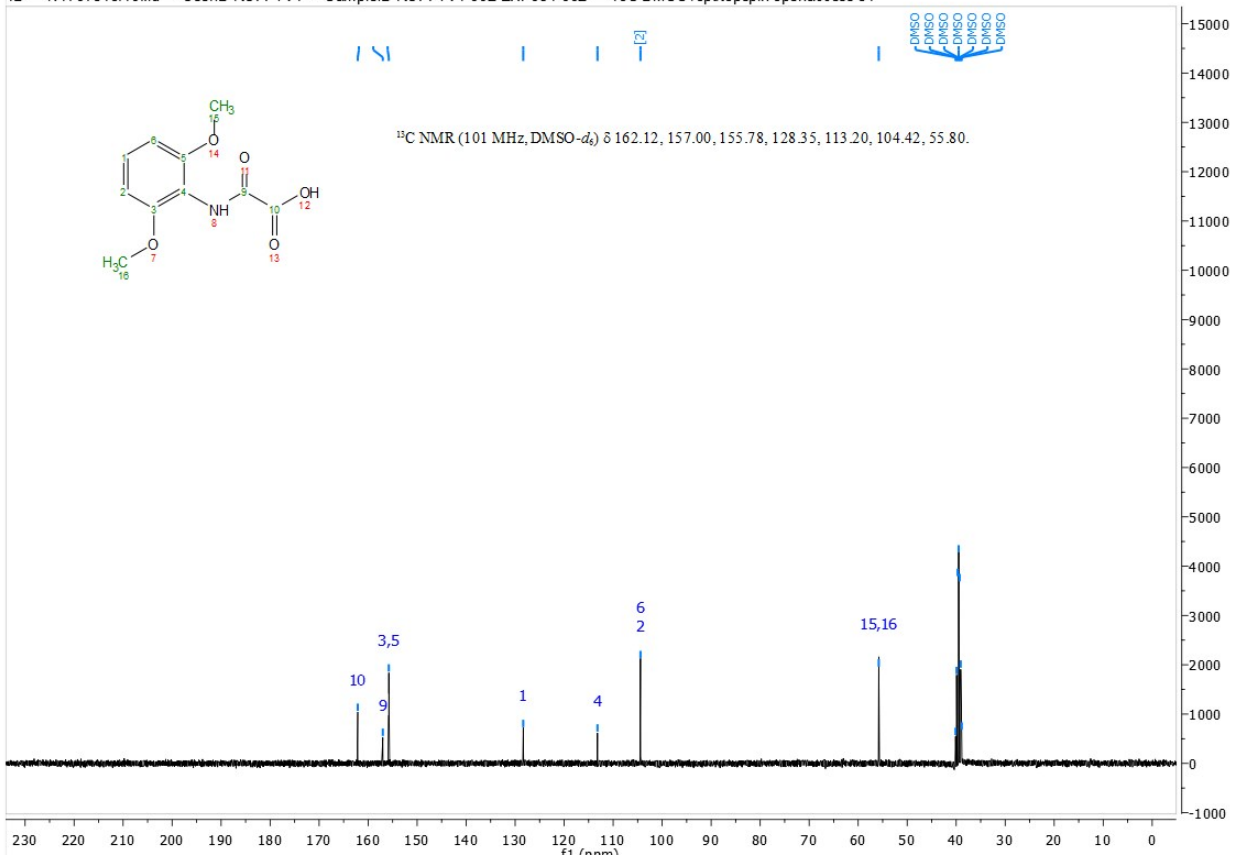


Figure SI III19: ¹³C NMR analysis of L15.

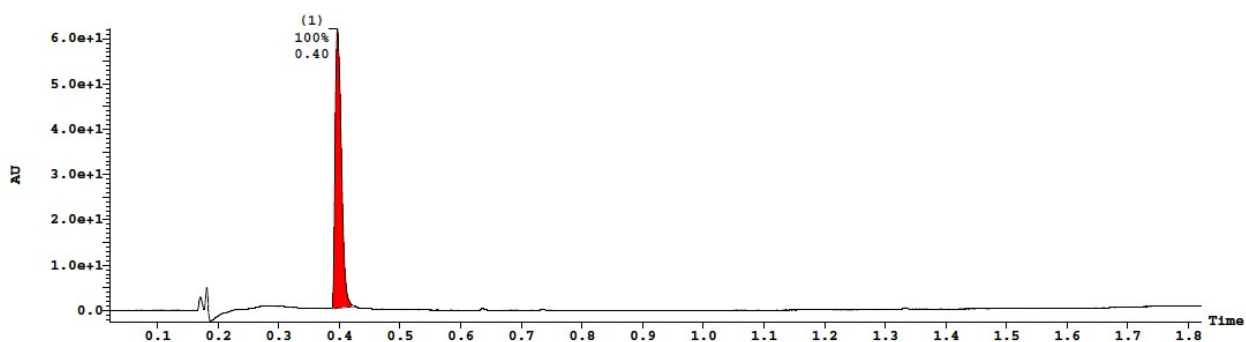
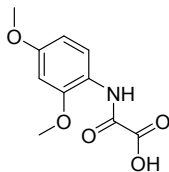


Figure SI III20: UPLC-MS analysis of L15.

UPLC / ESI Method: Column: Acquity UPLC HSS T3 1.8 μ m 2.1x 50mm at 60°C, Gradient: from 5 to 98 % B in 1.4 min - flow 1.0 mL/min Eluent A: water + 0.05 % formic acid + 3.75 mM ammonium acetate, Eluent B: acetonitrile + 0.04 % formic acid, Rt 0.4min, m/z 226.2 (M+H), interpreted as compatible with the structure of the final compound L15.

Synthesis of L16



L16

Chemical Formula: C₁₀H₁₁NO₅

Exact Mass: 225.06

Molecular Weight: 225.20

The standard protocol for the synthesis of **L11** was used starting from 2,4-dimethoxyaniline (Fluka CAS 2735-04-8) (1 g, 6.53 mmol) to yield 1.47gr of **L16** (34%).

Analysis of L16

¹H NMR (400 MHz, DMSO-d₆) δ 9.50 (s, 1H), 7.87 (d, *J* = 8.8 Hz, 1H), 6.68 (d, *J* = 2.6 Hz, 1H), 6.54 (dd, *J* = 8.8, 2.6 Hz, 1H), 3.85 (s, 3H), 3.76 (s, 3H).

¹³C NMR (101 MHz, DMSO-d₆) δ 161.99, 157.44, 155.43, 150.80, 121.74, 118.95, 104.38, 98.92, 56.02, 55.38.

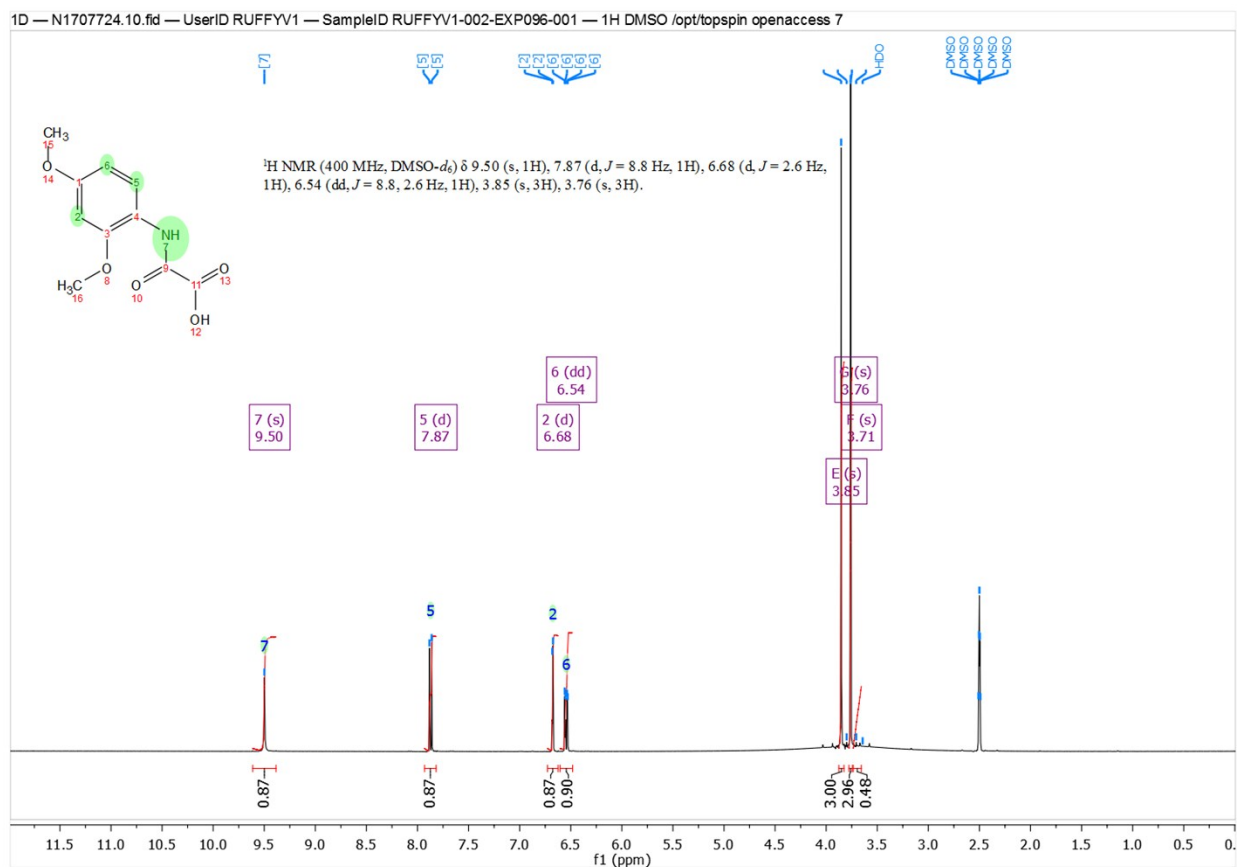


Figure SI III21: ¹H NMR analysis of L16.

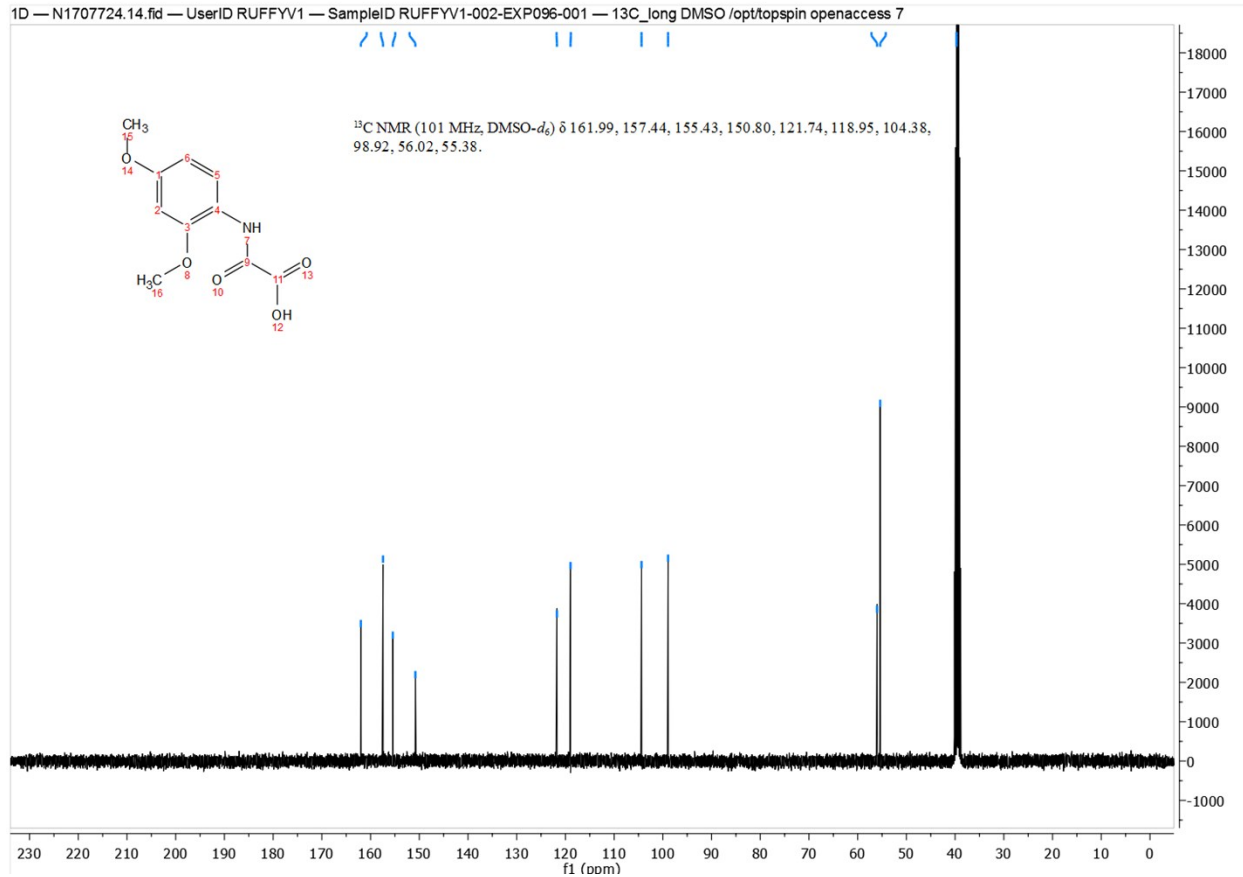


Figure SI III22: ¹³C NMR analysis of L16.

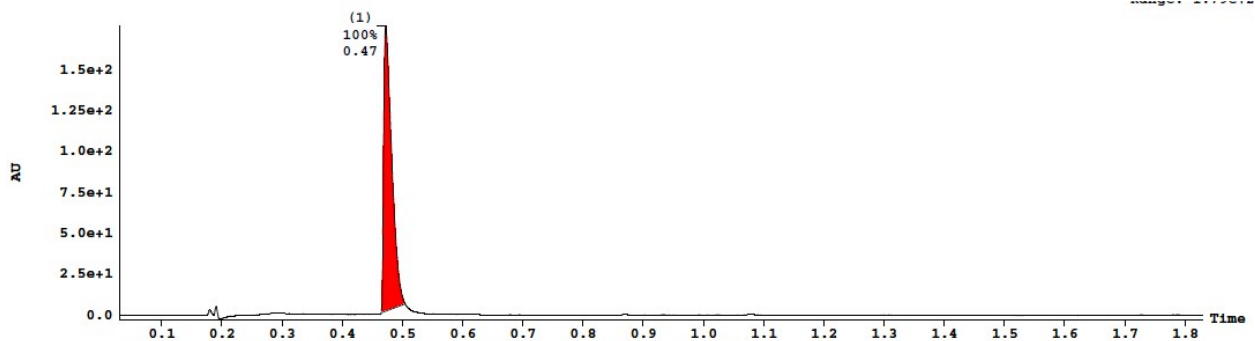
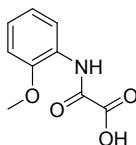


Figure SI III23: UPLC-MS analysis of L16.

UPLC / ESI Method: Column: Acquity UPLC HSS T3 1.8 μ m 2.1x 50mm at 60°C. Gradient: from 5 to 98 % B in 1.4 min - flow 1.0 mL/min Eluent A: water + 0.05 % formic acid + 3.75 mM ammonium acetate, Eluent B: acetonitrile + 0.04 % formic acid. Rt= 0.47 min, m/z =224.0 amu (M-H), interpreted as compatible with the structure of the final compound **L16**

Synthesis of L17



L17

Chemical Formula: C₉H₉NO₄

Exact Mass: 195.05

Molecular Weight: 195.17

The standard protocol for the synthesis of **L11** was used starting from *o*-anisidine (Fluka CAS 90-04-0) (0.804 g, 6.53 mmol) to yield 1.2gr of **L17** (70%).

Analysis of L17

¹H NMR (400 MHz, DMSO-d₆) δ 9.62 (s, 1H), 8.08 (d, J = 7.9 Hz, 1H), 7.22 - 7.08 (m, 2H), 6.99 (t, J = 7.6 Hz, 1H), 3.89 (s, 3H).

¹³C NMR (101 MHz, DMSO-d₆) δ 161.85, 155.59, 149.11, 125.75, 125.52, 120.58, 120.18, 111.25, 55.99.

1D — UserID RUFFYV1 — SampleID RUFFYV1-002-EXP096-002 — Comment — DMSO,1H,25.3 C —

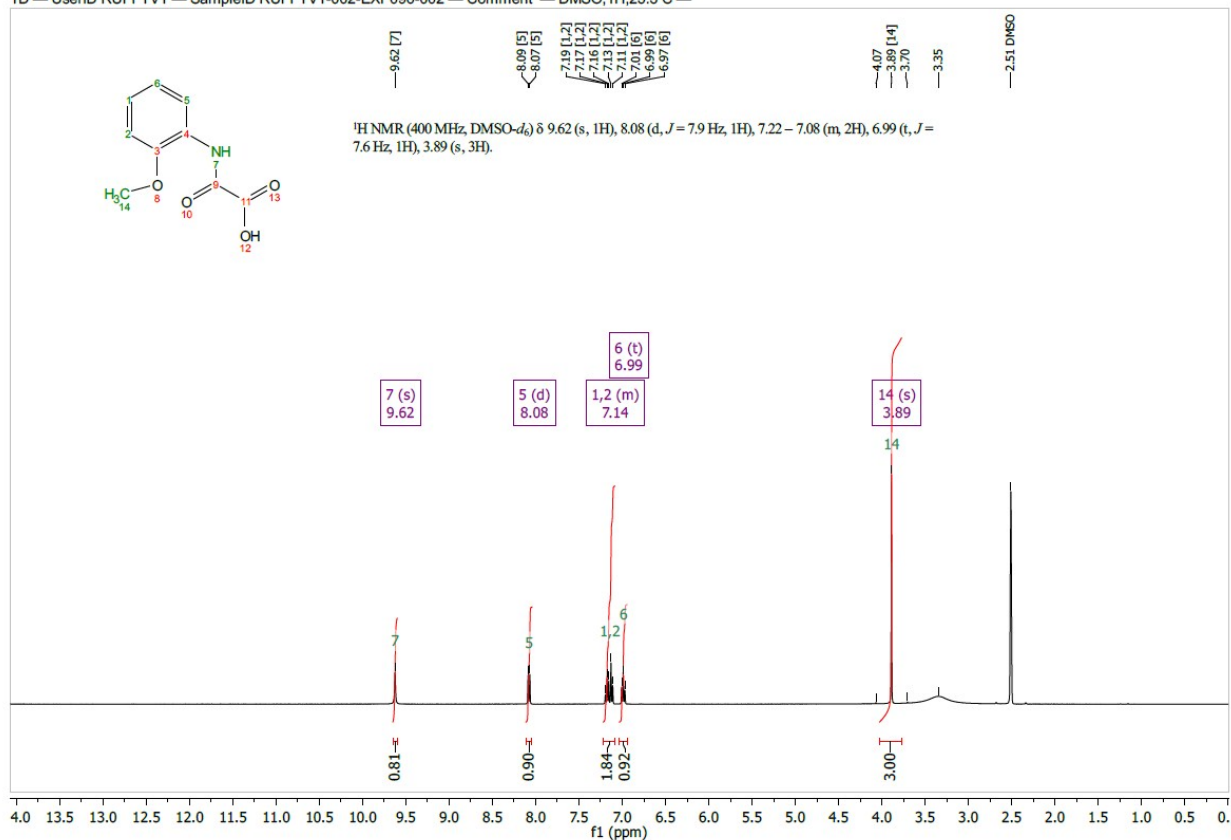


Figure SI III24: ¹H NMR analysis of L17.

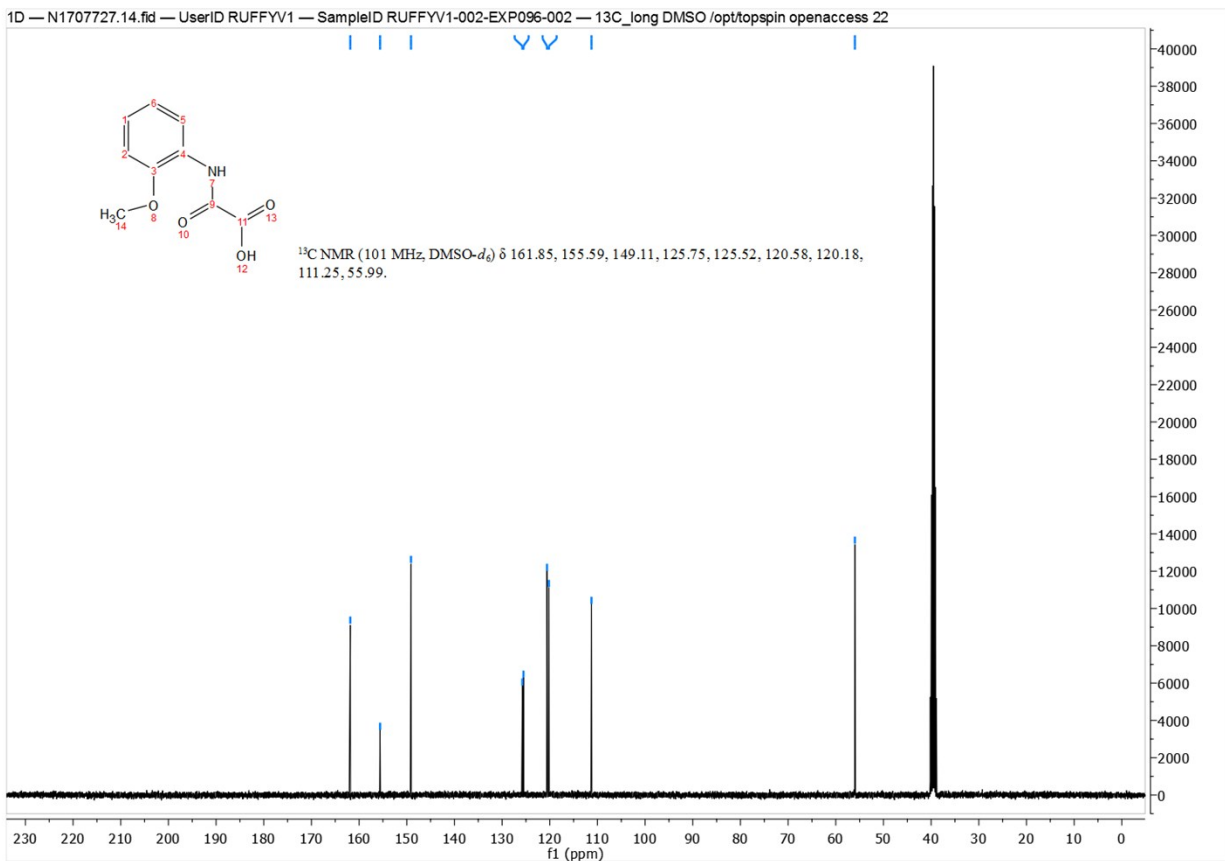


Figure SI III25: ^{13}C NMR analysis of L17.

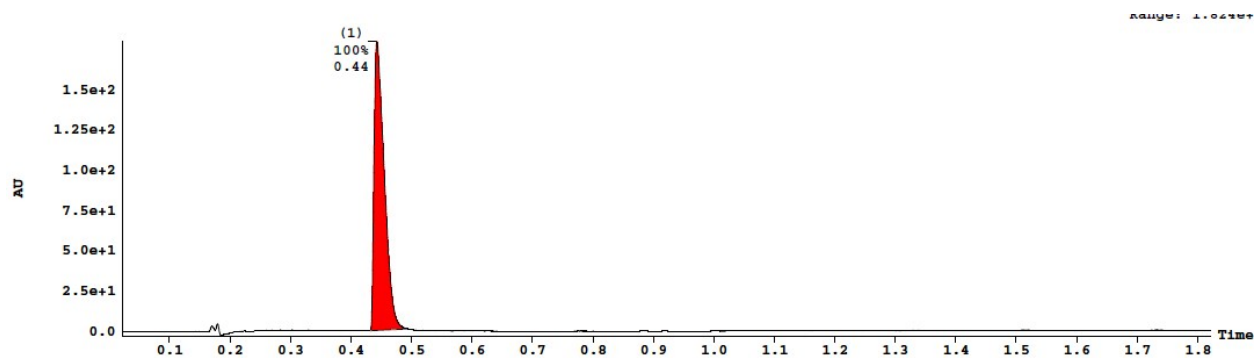
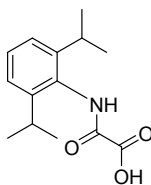


Figure SI III26: UPLC-MS analysis of L17.

UPLC / ESI Method: Column: Acquity UPLC HSS T3 1.8 μm 2.1x 50mm at 60 $^{\circ}\text{C}$ Gradient: from 5 to 98 % B in 1.4 min - flow 1.0 mL/min Eluent A: water + 0.05 % formic acid + 3.75 mM ammonium acetate Eluent B: acetonitrile + 0.04 % formic acid. Rt= 0.44 min, m/z =194.1 amu (M-H), interpreted as compatible with the structure of the final compound L17.

Synthesis of L18



L18

Chemical Formula: C₁₄H₁₉NO₃

Exact Mass: 249.14

Molecular Weight: 249.31

The standard protocol for the synthesis of **L11** was used starting from 2,6-diisopropylaniline (Fluka CAS 24544-04-5) (1 g, 5.64 mmol) to yield 573 mg of **L18** (40%).

Analysis of L18

¹H NMR (400 MHz, DMSO-*d*₆) δ 10.22 (s, 1H), 7.29 (dd, *J* = 8.3, 7.1 Hz, 1H), 7.19 (d, *J* = 7.7 Hz, 2H), 2.98 (hept, *J* = 6.9 Hz, 2H), 1.11 (d, *J* = 6.8 Hz, 12H).

¹³C NMR (101 MHz, DMSO-*d*₆) δ 162.33, 158.42, 145.59, 131.21, 127.93, 122.99, 28.02, 26.31, 23.35.

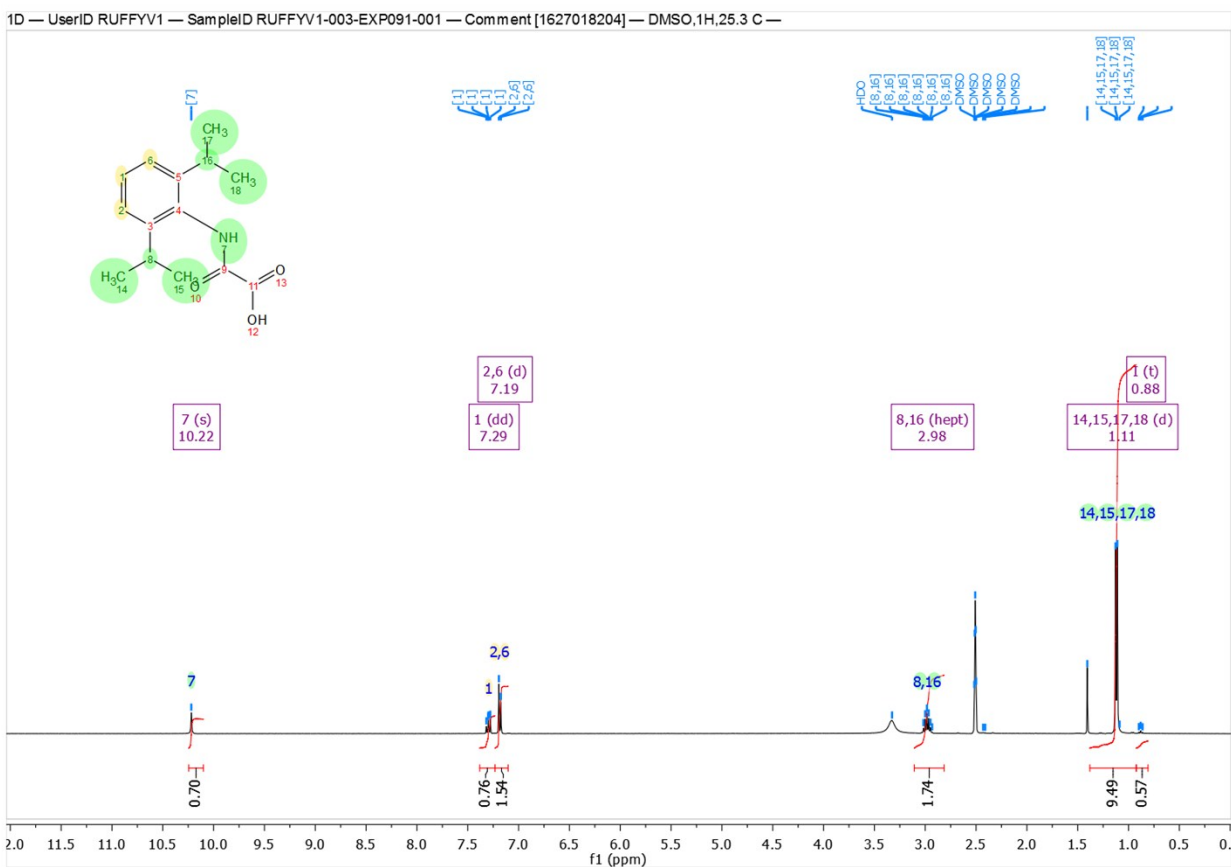


Figure SI III27: ¹H NMR analysis of L18.

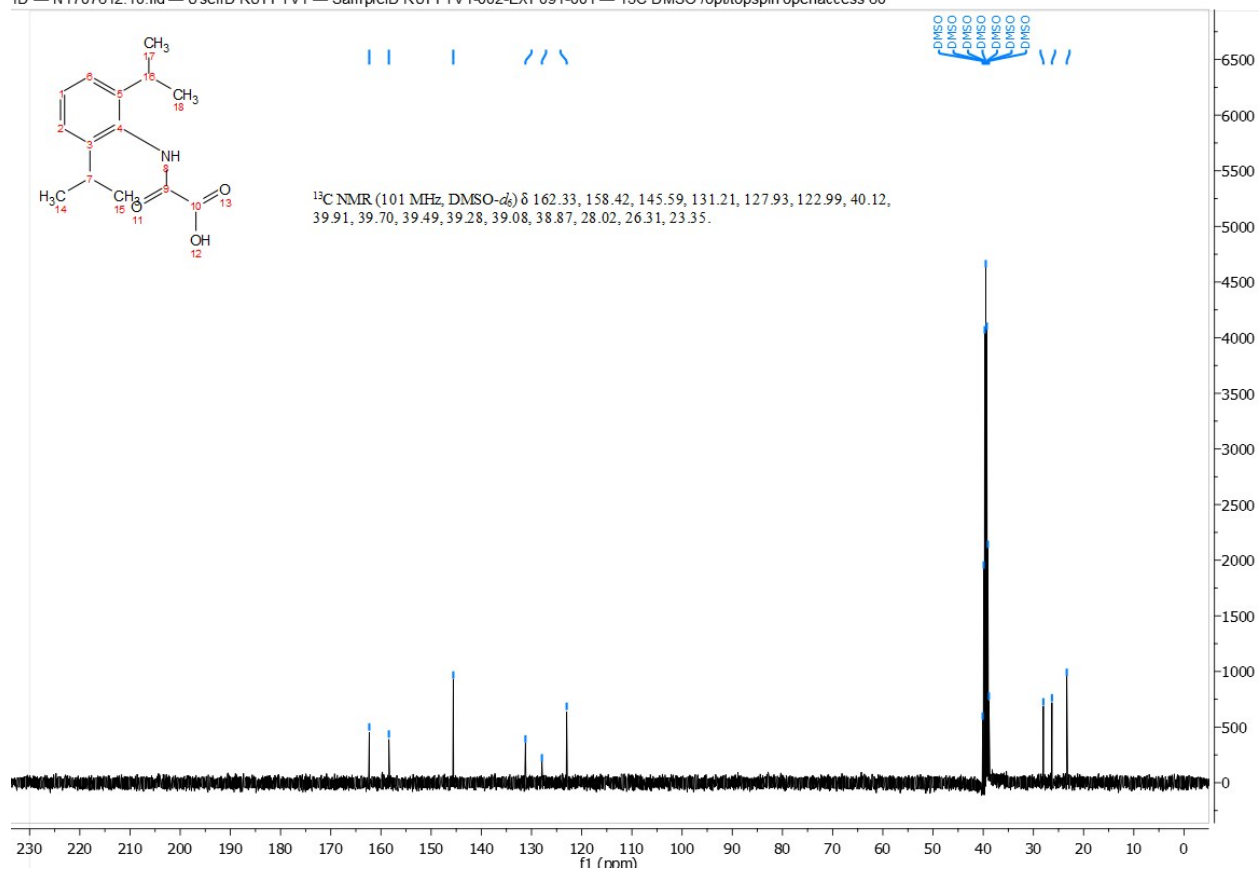


Figure SI III28: ¹³C NMR analysis of L18.

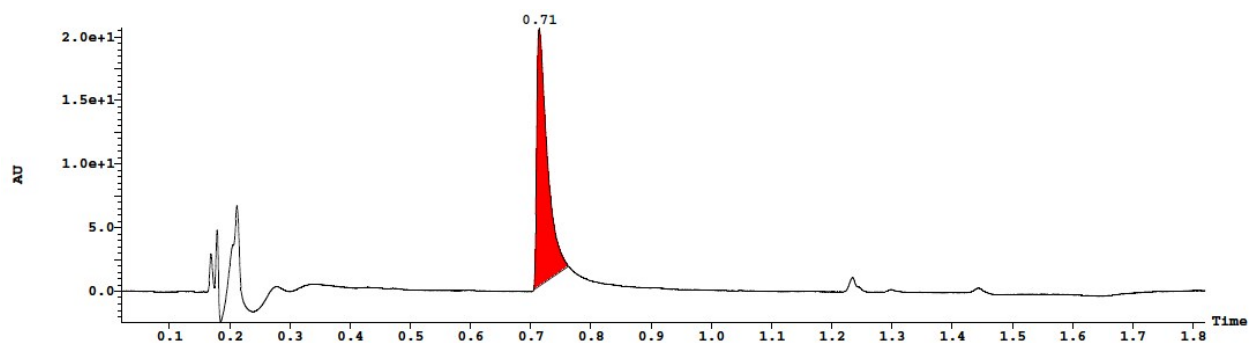
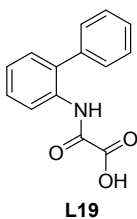


Figure SI III29: UPLC-MS analysis of L18.

UPLC / ESI Method: Column: Acquity UPLC HSS T3 1.8 μ m 2.1x 50mm at 60°C, Gradient: from 5 to 98 % B in 1.4 min - flow 1.0 mL/min Eluent A: water + 0.05 % formic acid + 3.75 mM ammonium acetate, Eluent B: acetonitrile + 0.04 % formic acid. Rt= 0.71 min, m/z =248.3 amu (M-H), interpreted as compatible with the structure of the final compound L18.

Synthesis of L19



Chemical Formula: C₁₄H₁₁NO₃

Exact Mass: 241.07

Molecular Weight: 241.25

The standard protocol for the synthesis of **L11** was used starting from 2-aminobiphenyl (ABCR CAS 90-41-5) (1 g, 5.64 mmol) to yield 71 mg of **L19** (5%).

Analysis of L19

¹H NMR (400 MHz, DMSO-*d*₆) δ 9.95 (s, 1H), 8.10 (d, *J* = 7.4 Hz, 1H), 7.49 (t, *J* = 7.4 Hz, 1H), 7.40 (q, *J* = 9.0, 8.5 Hz, 3H), 7.34 – 7.20 (m, 2H).

¹³C NMR (101 MHz, DMSO-*d*₆) δ 161.89, 159.93, 137.93, 134.20, 133.65, 130.23, 128.84, 128.79, 128.04, 127.66, 125.02, 122.03.

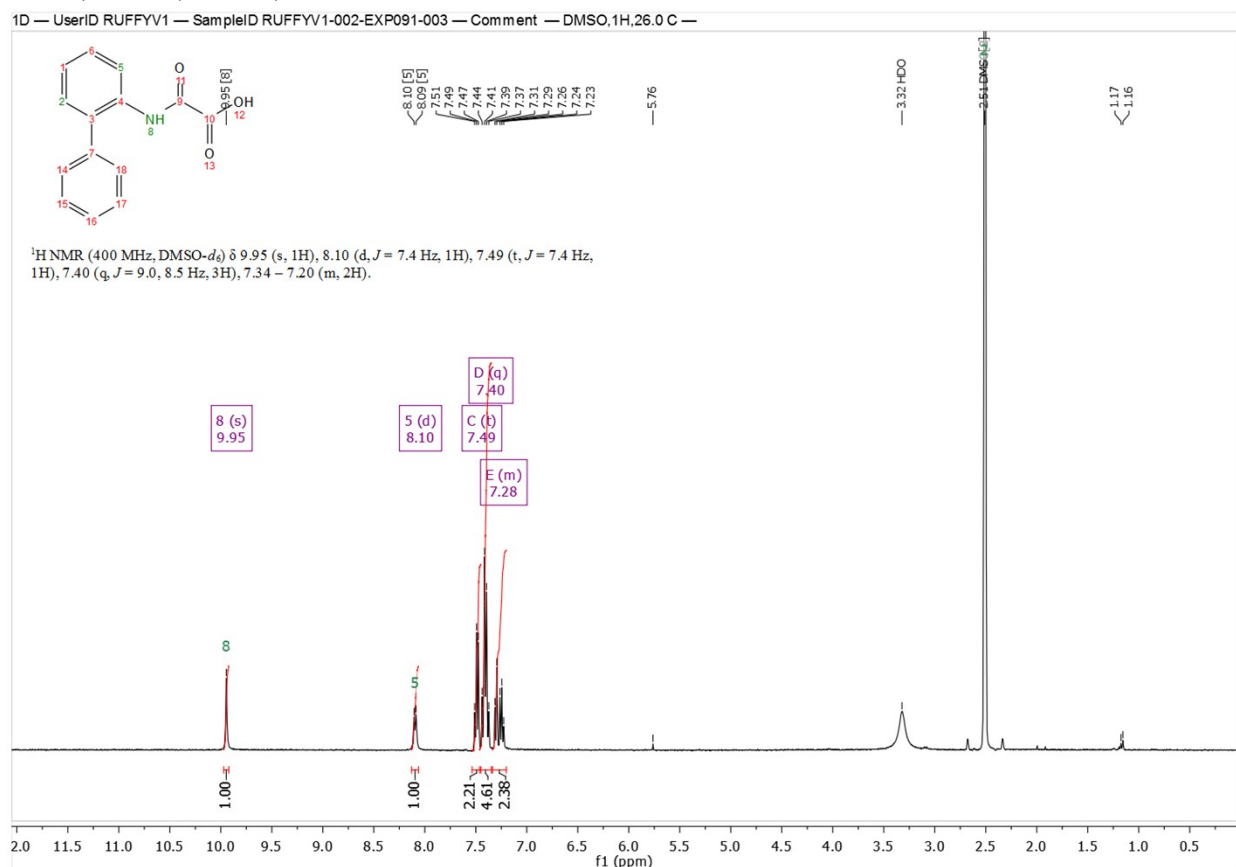


Figure SI III30: ¹H NMR analysis of L19.

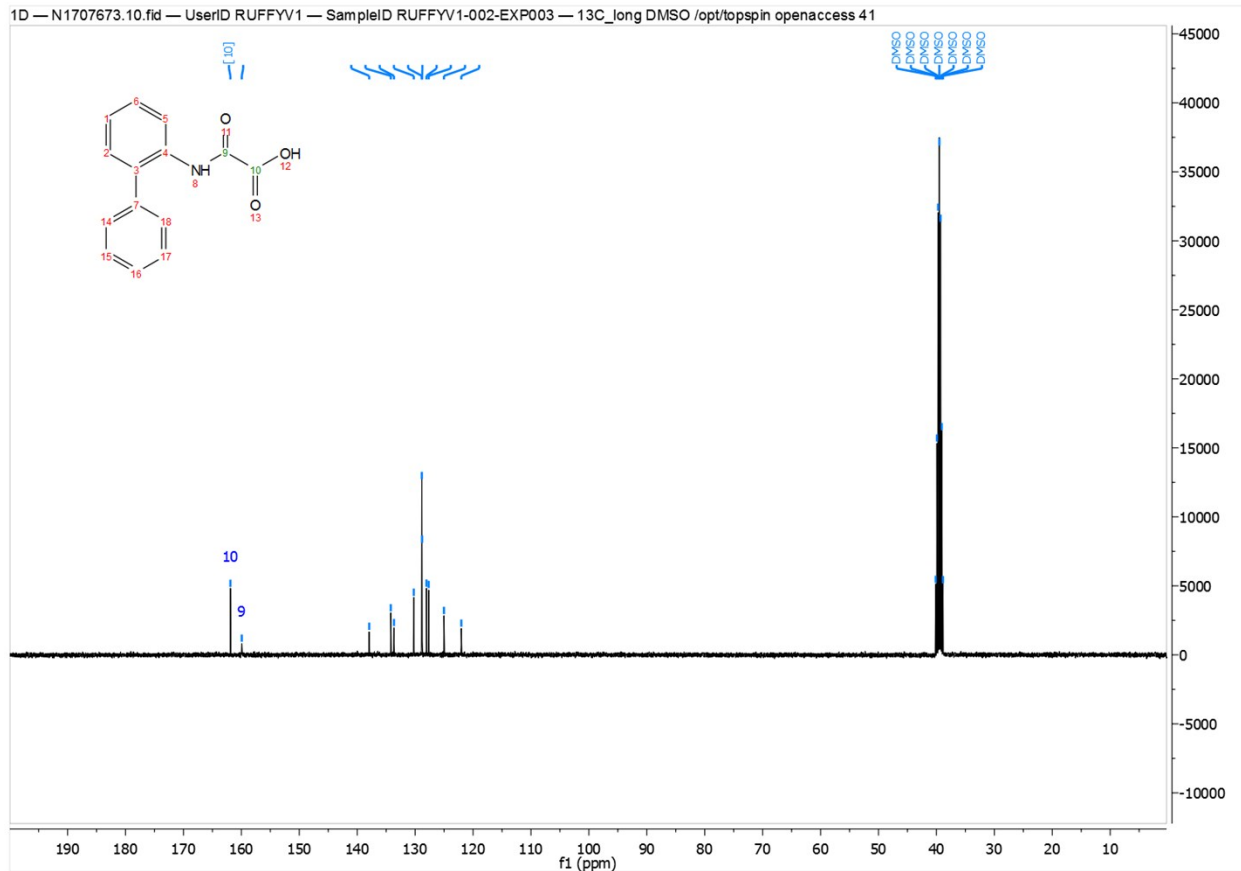


Figure SI III31: ¹³C NMR analysis of L19.

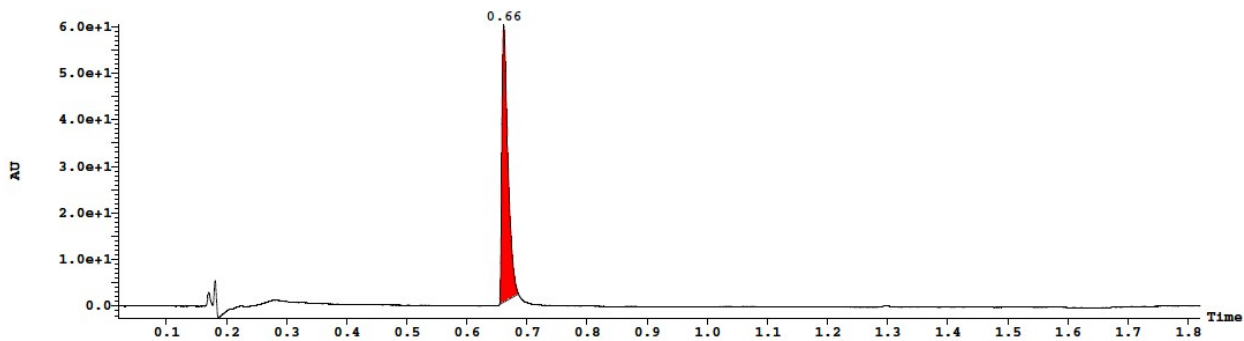
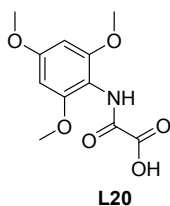


Figure SI III32: UPLC-MS analysis of L19.

UPLC / ESI Method: Column: Acquity UPLC HSS T3 1.8 μ m 2.1x 50mm at 60°C, Gradient: from 5 to 98 % B in 1.4 min - flow 1.0 mL/min Eluent A: water + 0.05 % formic acid + 3.75 mM ammonium acetate, Eluent B: acetonitrile + 0.04 % formic acid, Rt 0.66min, m/z 240.4 (M-H), interpreted as compatible with the structure of the final compound L19.

Synthesis of L20



Chemical Formula: C₁₁H₁₃NO₆

Exact Mass: 255.07

Molecular Weight: 255.23

The standard protocol for the synthesis of **L11** was used starting from 2,4,6-trimethoxyaniline hydrochloride (CAS 14227-17-9) (0.535 g, 2.435 mmol) to yield 341mg of **L20** (55%).

Analysis of L20

¹H NMR (400 MHz, DMSO-d₆) δ 13.90 (s, 1H), 9.44 (s, 1H), 6.27 (s, 2H), 3.80 (s, 3H), 3.73 (s, 6H).

¹³C NMR (101 MHz, DMSO-d₆) δ 162.26, 159.77, 157.54, 156.33, 106.30, 91.02, 55.76, 55.42.

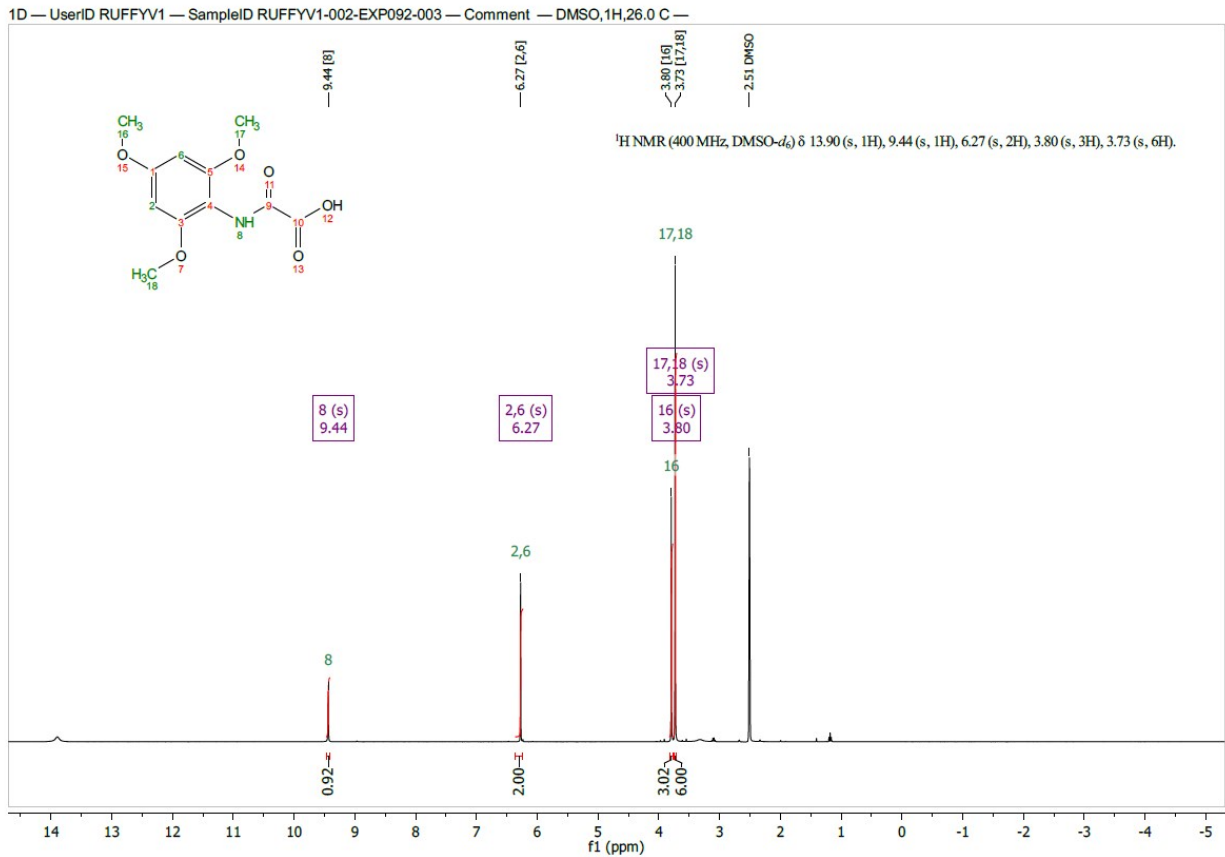


Figure SI III33: ¹H NMR analysis of L20.

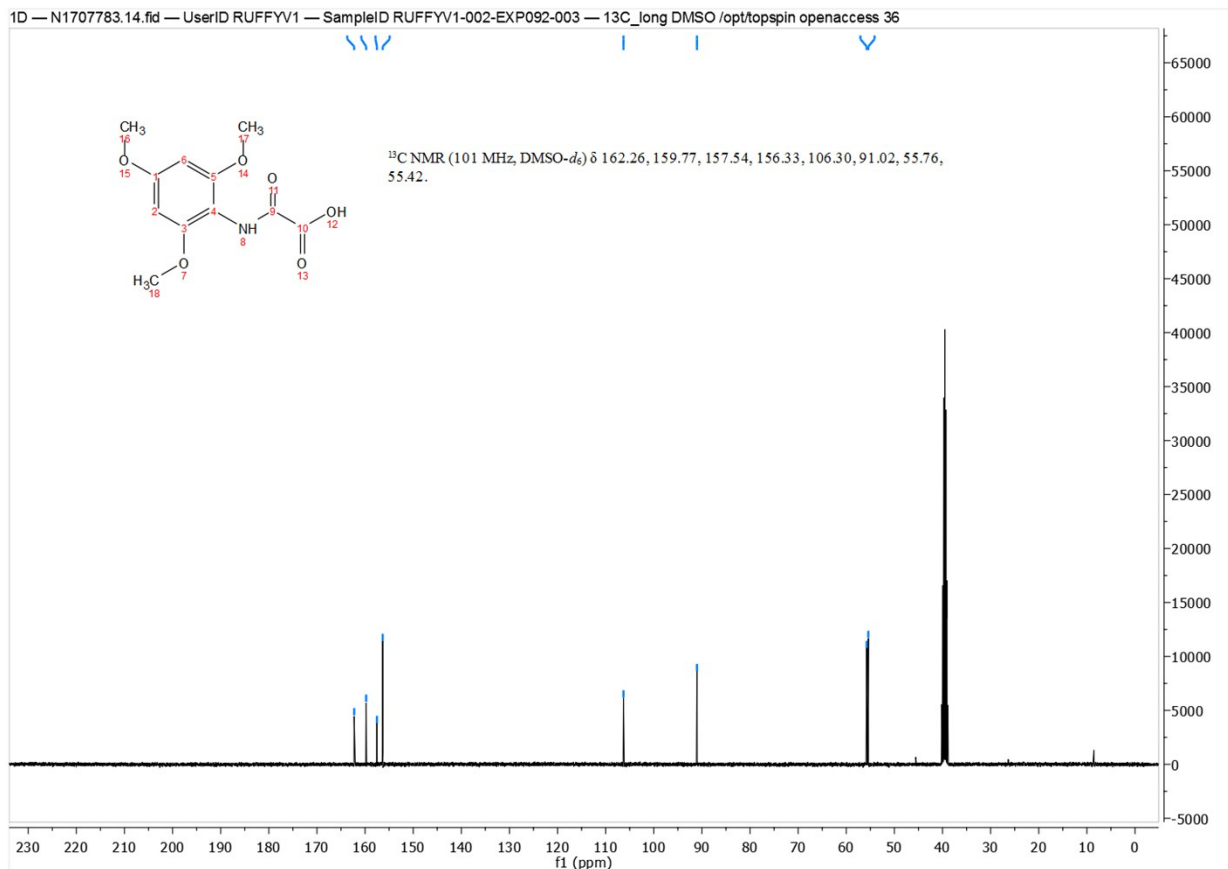


Figure SI III34: ¹³C NMR analysis of L20.

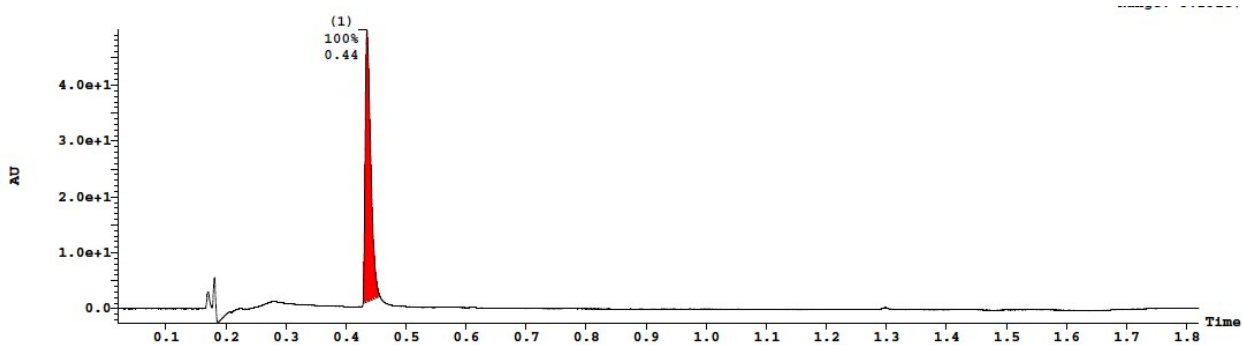
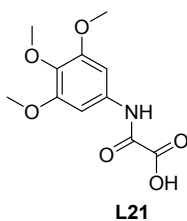


Figure SI III35: UPLC-MS analysis of L20.

UPLC / ESI Method: Column: Acquity UPLC HSS T3 1.8 μ m 2.1x 50mm at 60°C. Gradient: from 5 to 98 % B in 1.4 min - flow 1.0 mL/min Eluent A: water + 0.05 % formic acid + 3.75 mM ammonium acetate. Eluent B: acetonitrile + 0.04 % formic acid Rt 0.44min, m/z 256.2 (M+H), interpreted as compatible with the structure of L20.

Synthesis of L21



Chemical Formula: C₁₁H₁₃NO₆

Exact Mass: 255.07

Molecular Weight: 255.23

The standard protocol for the synthesis of **L11** was used starting from 3,4,5-trimethoxyaniline (Fluka CAS 24313-88-0) (1.196 g, 6.53 mmol) to yield 500 mg of **L21** (30%).

Analysis for L21

¹H NMR (400 MHz, DMSO-d₆) δ 10.58 (s, 1H), 7.23 (s, 2H), 3.75 (s, 6H), 3.64 (s, 3H).

¹³C NMR (101 MHz, DMSO-d₆) δ 162.04, 156.51, 152.66, 134.37, 133.76, 98.18, 60.12, 55.78.

1D — UserID RUFFYV1 — SampleID RUFFYV1-002-EXP096-003 — Comment — DMSO,1H,25.3 C —

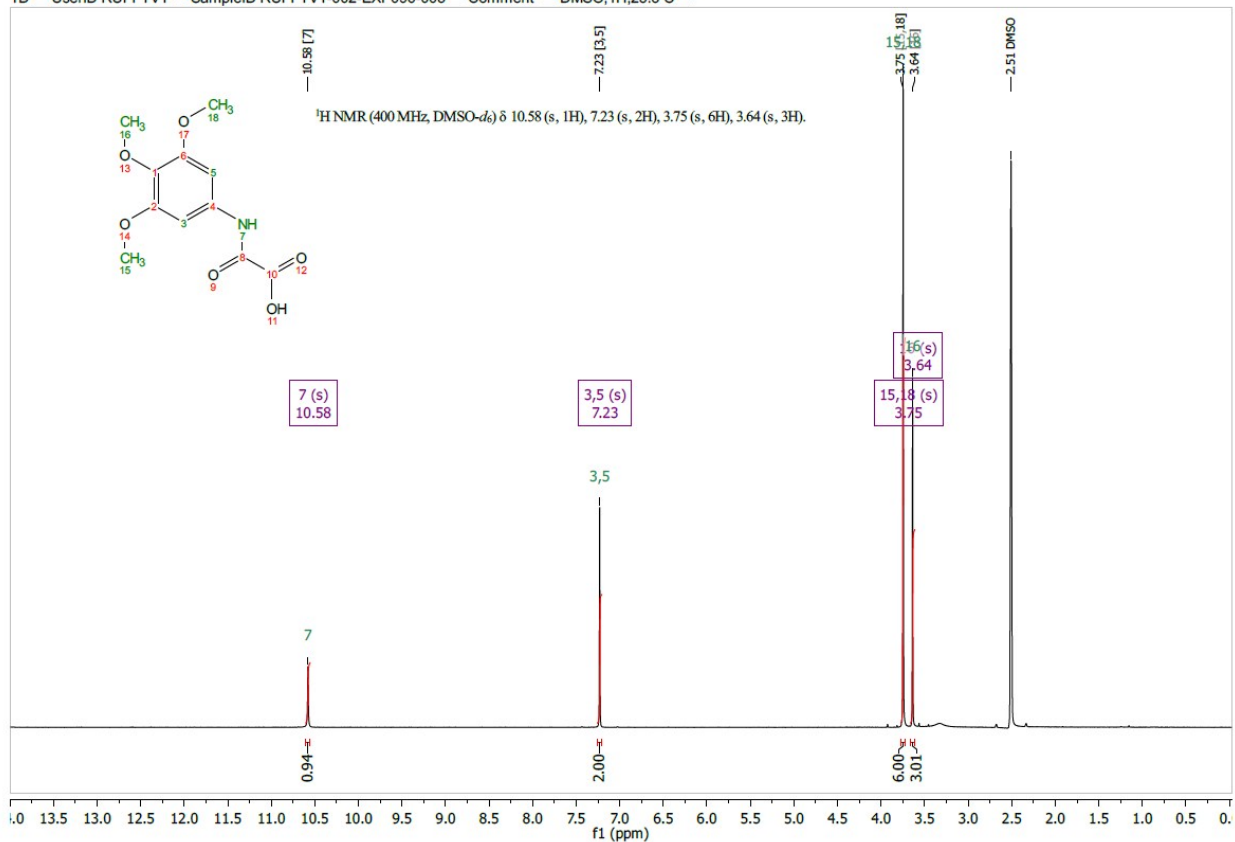


Figure SI III36: ¹H NMR analysis of L21.

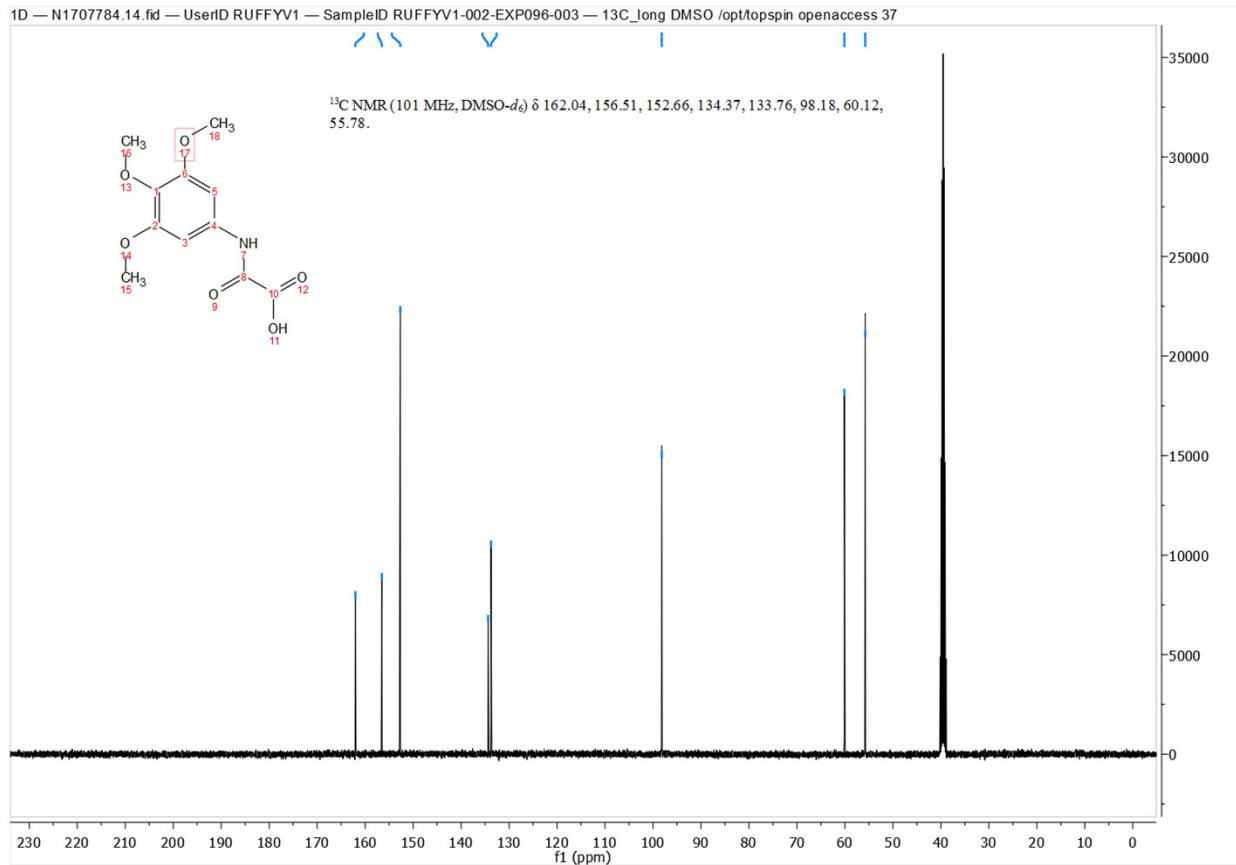


Figure SI III37: ¹³C NMR analysis of L21.

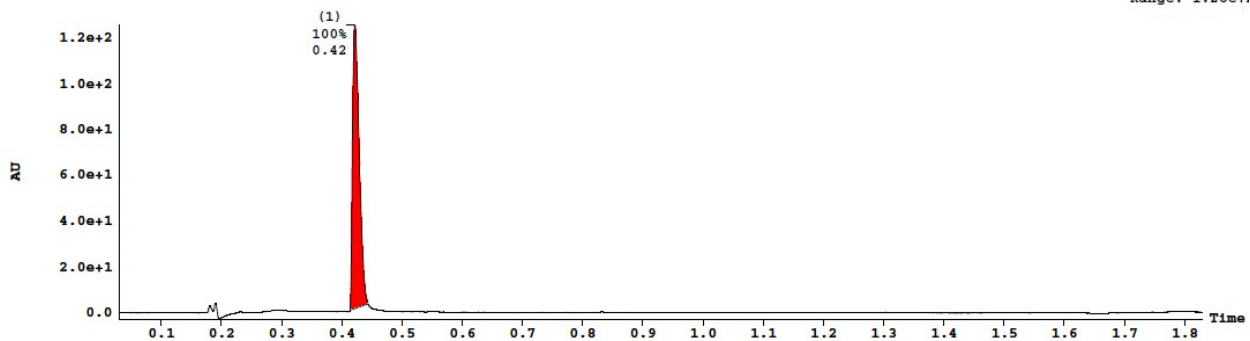


Figure SI III38: UPLC-MS analysis of L21.

UPLC / ESI Method: Column: Acquity UPLC HSS T3 1.8 μ m 2.1x 50mm at 60°C. Gradient: from 5 to 98 % B in 1.4 min - flow 1.0 mL/min Eluent A: water + 0.05 % formic acid + 3.75 mM ammonium acetate. Eluent B: acetonitrile + 0.04 % formic acid. Rt= 0.42 min, m/z =254.1 amu (M-H), interpreted as compatible with the structure of the final compound **L21**.

SI IV Protocol and analysis for the screening of known ligands on **1a** (main text Figure 1)

In a 96 well twin.tec® plate 1 nmole of the DNA-conjugated aryl iodide substrate **1a** was evaporated in each well.

Data set A without ascorbate: The dry pellet of **1a** (6.52 µg, 0.001 µmol) was dissolved in 8 µl of a solution of K₃PO₄ (1M in water) followed by the addition of the amines **2a-f** stock solutions (2M in DMSO) (4 µL, 8.00 µmol). The CuI(100mM)/Ligand **L1-10**(200mM) solutions in DMSO were then added (4 µL) to the reactions. The plate was centrifuged for a minute, to ensure a complete and homogenous mixing, sealed and incubated at 40°C in a PCR thermocycler (Eppendorf) for 6 hours.

Data set B with ascorbate: The same experiment was repeated in the presence of sodium ascorbate (CAS 134-03-2). The dry pellet of **1a** (6.52 µg, 0.001 µmol) was dissolved in 8 µl of a solution of K₃PO₄ (1M in water) and sodium ascorbate (100mM in water), followed by the addition of the amines **2a-f** stock solutions (2M in DMSO) (4 µL, 8.00 µmol). The CuI(100mM)/Ligand **L1-10**(200mM) solutions in DMSO were then added (4 µL) to the reactions. The plate was centrifuged for a minute, to ensure a complete and homogenous mixing, sealed and incubated at 40°C in a PCR thermocycler (Eppendorf) for 6 hours.

Work Up

Each reaction was diluted with 84 ul of a 200mM EDTA disodium salt solution in milliQ water and desalted before UPLC-TOF analysis.

Effect of the addition of sodium ascorbate for the reactions using L8 (DMPAO) as ligand and characterization of the products 3a-f

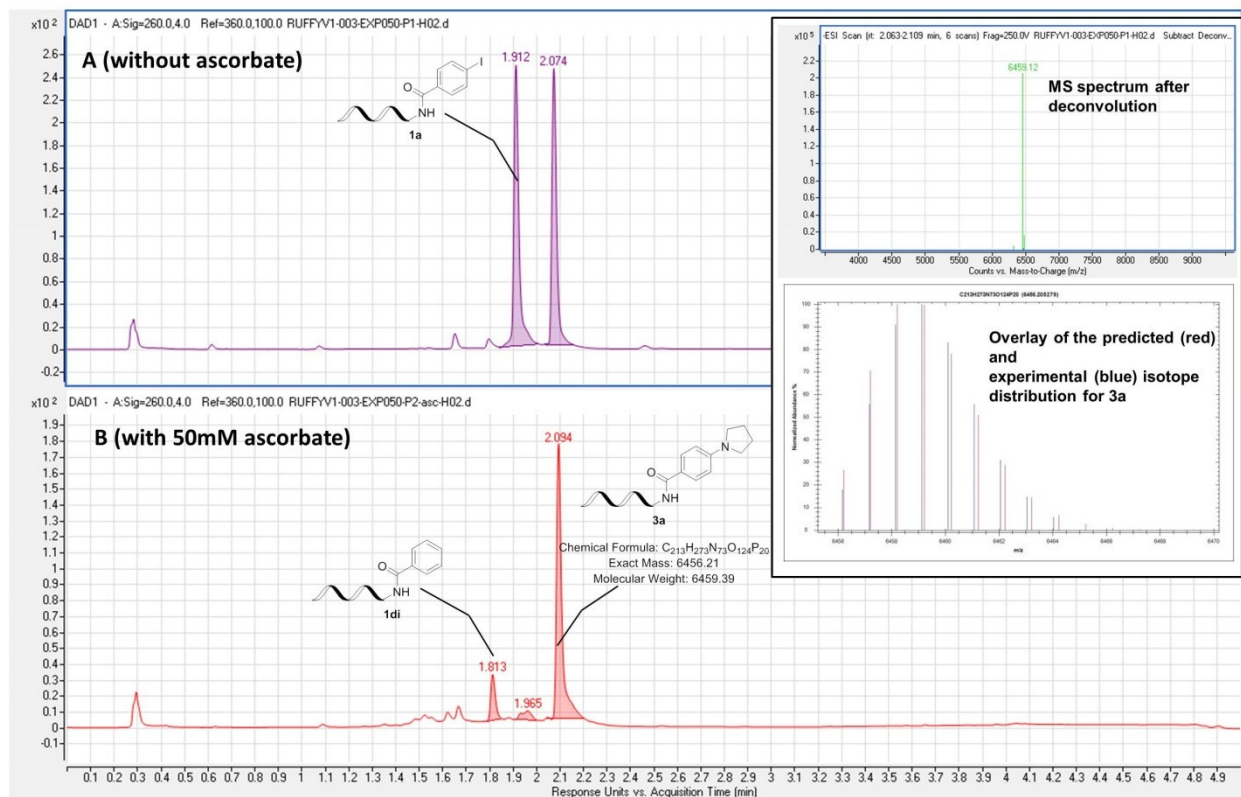


Figure SI IV1: UPLC-TOF analysis of the effect of the addition of sodium ascorbate for the reactions using the amine **2a** and L8 as ligand and characterization of the product **3a**.

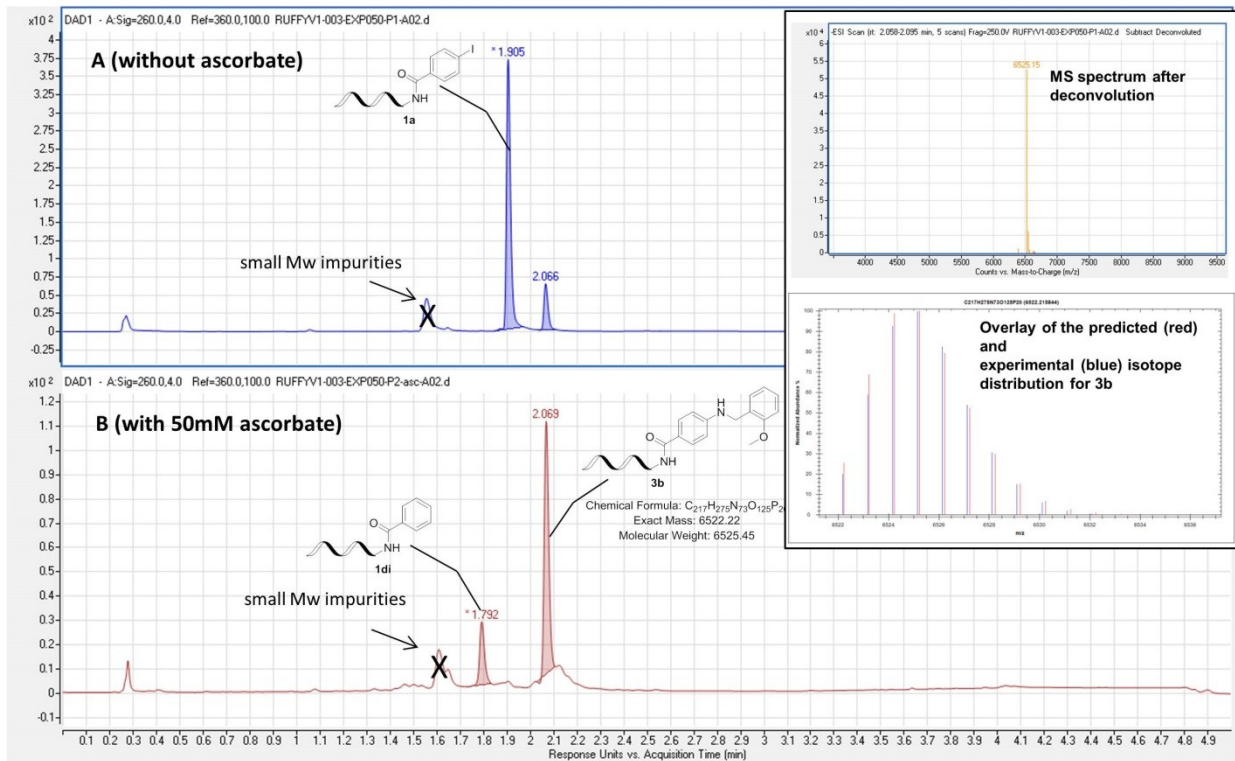


Figure SI IV2: UPLC-TOF analysis of the effect of the addition of sodium ascorbate for the reactions using the amine **2b** and **L8** as ligand and characterization of the product **3b**.

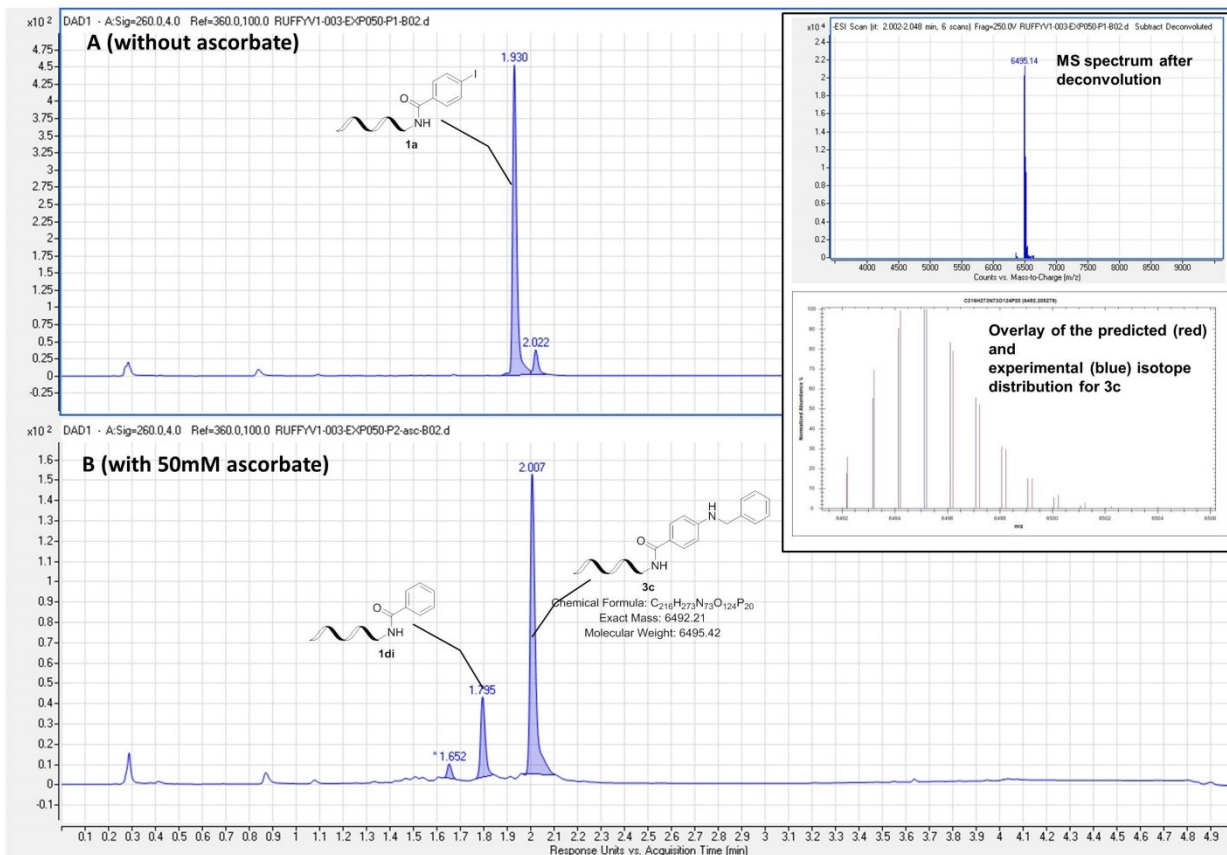


Figure SI IV3: UPLC-TOF analysis of the effect of the addition of sodium ascorbate for the reactions using the amine **2c** and **L8** as ligand and characterization of the product **3c**.

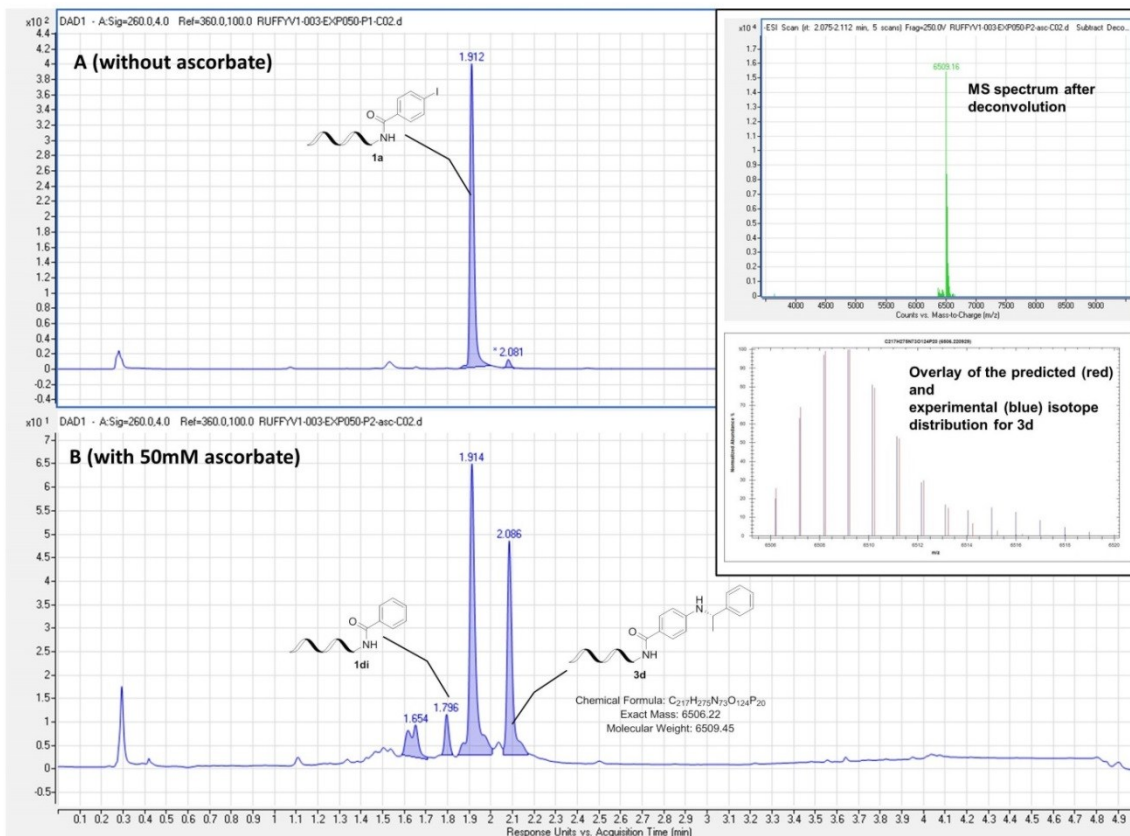


Figure SI IV4: UPLC-TOF analysis of the effect of the addition of sodium ascorbate for the reactions using the amine **2d** and **L8** as ligand and characterization of the product **3d**.

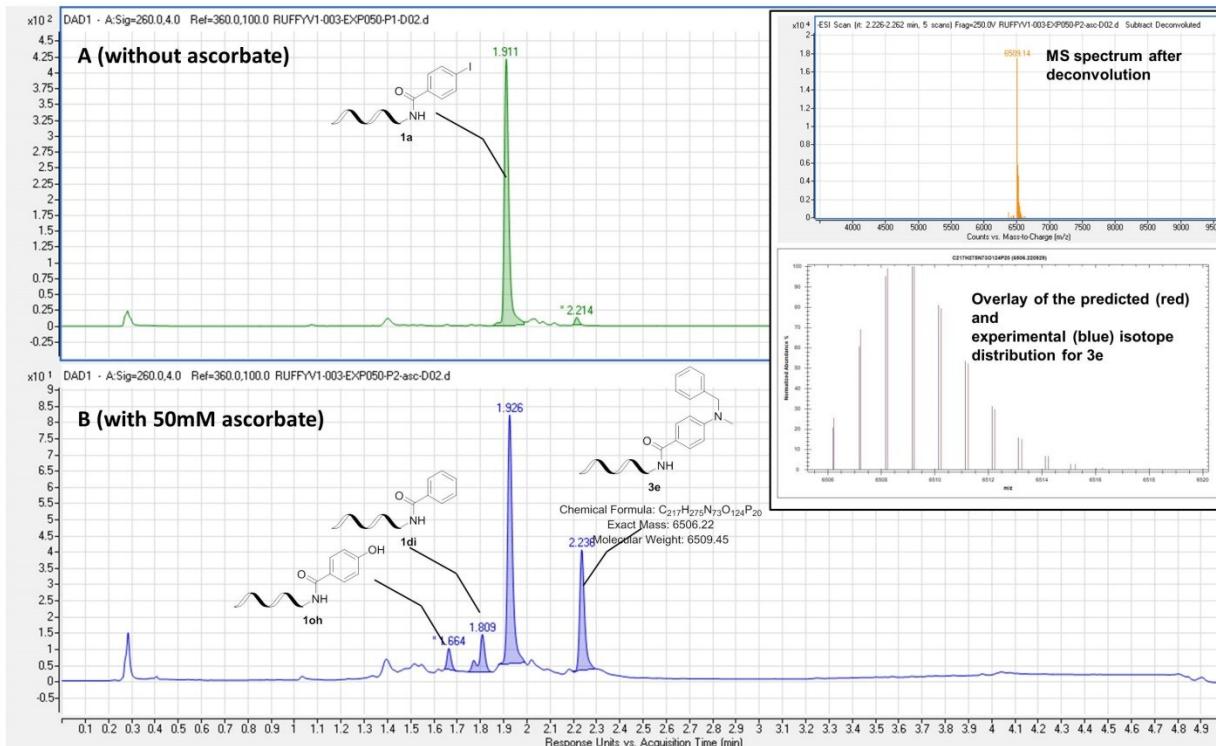


Figure SI IV5: UPLC-TOF analysis of the effect of the addition of sodium ascorbate for the reactions using the amine **2e** and **L8** as ligand and characterization of the product **3e**.

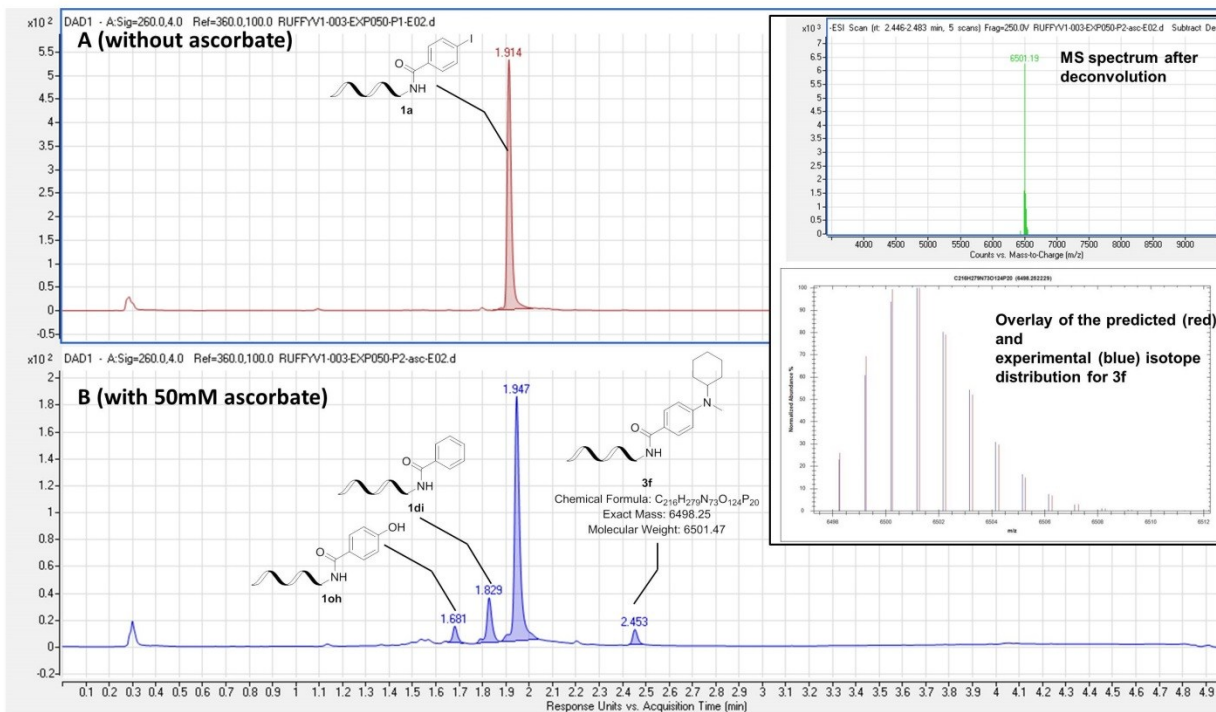


Figure SI IV6: UPLC-TOF analysis of the effect of the addition of sodium ascorbate for the reactions using the amine **2f** and **L8** as ligand and characterization of the product **3f**.

Preliminary evaluation of DNA degradation

During this initial screening we purposefully looked at potential signs of DNA degradation as one of the main criteria for selection of ligands, together with the conversion into the expected *N*-arylated product.

While any obvious signs of DNA cleavage are not observed in the experiment without sodium ascorbate (Figure 1 A main text), in the presence of this reducing agent traces of DNA cleavage product can be observed as illustrated with the figures SI IV 1-6 . Under our screening conditions, these DNA cleavage side products are present in low proportion (2-3%) and can be easily overlooked if one is not looking for them specifically, or if a less sensitive analytical setup is used. Importantly, ligands like **L10** led to higher proportion (10-15%) of DNA cleavage side products (Figure SI IV7). These product likely result from the copper-mediated cleavage of the phosphodiester backbone of the conjugates.³

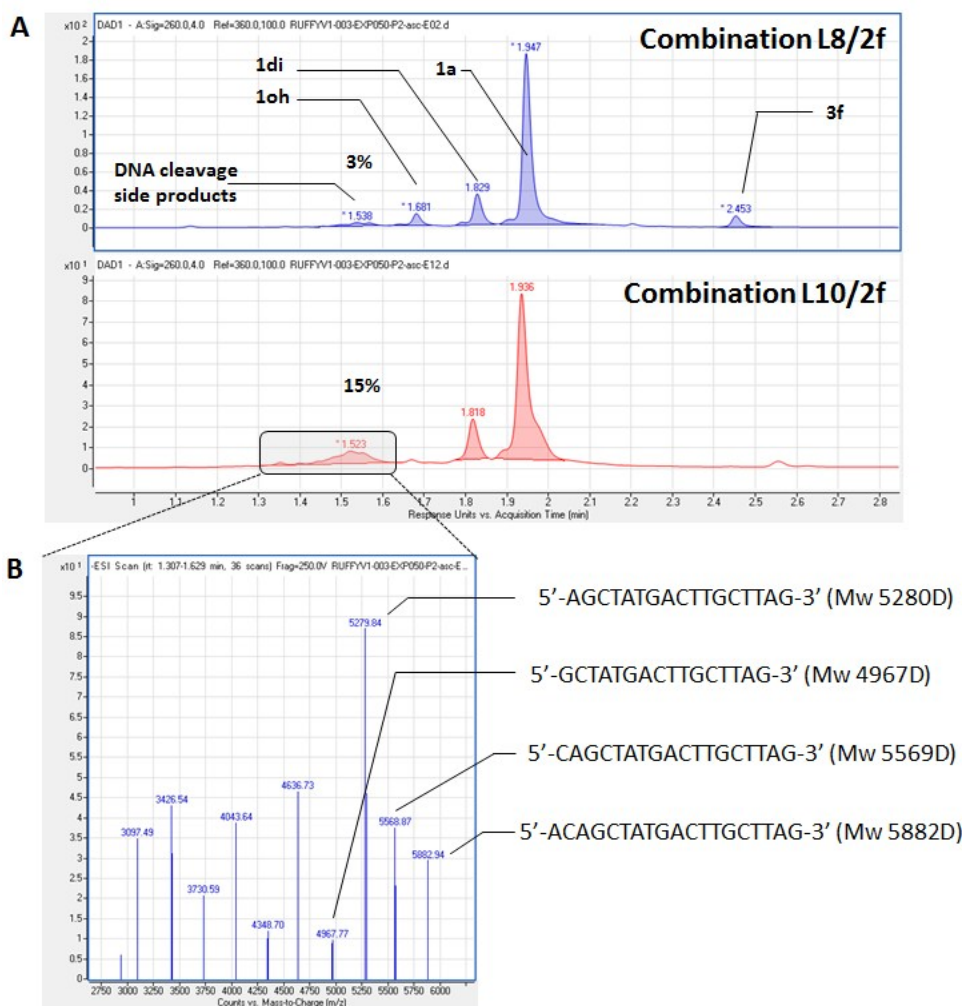


Figure SI IV7: Detection of traces of DNA cleavage side products during the initial ligand screening in **L8/2f** and **L10/2f** combinations in the presence of sodium ascorbate. **B** Integration of the MS counts for the highlighted area for the combination **L10/2f** and identification of DNA cleavage side products. Only 4 cleavage product are highlighted for clarity, but it is clear from the MS distribution that shorter oligonucleotides are also present.

It is important to note that subsequent optimization using **L15** as ligand did reduce even further the small proportion of DNA cleavage products (as illustrated with figures SI V1-8), presumably as increasing the ligand concentration prevents coordination of the copper to the phosphate backbone of the DNA.

SI V Protocol and analysis for the screening of DMPAO analogs on 1a and 1b (main text Figure 2)

2 sets of 96 reactions were set up.

1 nmole of the halogenoaromatic substrate **1a** or **1b** were evaporated in 2 separate twintec® plates (1 nmole per well).

First set of 96 reactions on **1a**:

In each well, **1a** (0.246 µg, 0.001 µmol) was dissolved in 8 µl of a solution of sodium ascorbate (100mM in 1M K₃PO₄ in water) followed by the addition of the amine stock solutions (2M in DMSO) (4 µL, 8.00 µmol) according to the plate map depicted in figure 2 (main text). The Cu(OAc)₂(100mM)/Ligand (200mM) (**L8-L11-21**) solutions in DMSO (4 µl, 0.400 µmol) were then added according to the plate map depicted in figure 2 (main text). The twin.tec® plate was sealed and incubated at 40°C in a PCR thermocycler (Eppendorf) for 3 hours.

Second set of 96 reactions on **1b**:

In each well, **1b** (0.246 µg, 0.001 µmol) was dissolved in 8 µl of a solution of sodium ascorbate (100mM in 1M K₃PO₄ in water) followed by the addition of the amine stock solutions (2M in DMSO) (4 µL, 8.00 µmol) according to the plate map depicted in figure 2 (main text). The Cu(OAc)₂(100mM)/Ligand (200mM) (**L8-L11-21**) solutions in DMSO(4 µl, 0.400 µmol) were then added according to the plate map depicted in figure 2 (main text). The plate was sealed and incubated at 40°C in a PCR thermocycler (Eppendorf) for 3 hours.

Selected screening results and comparison of the efficiency of L8 and L15

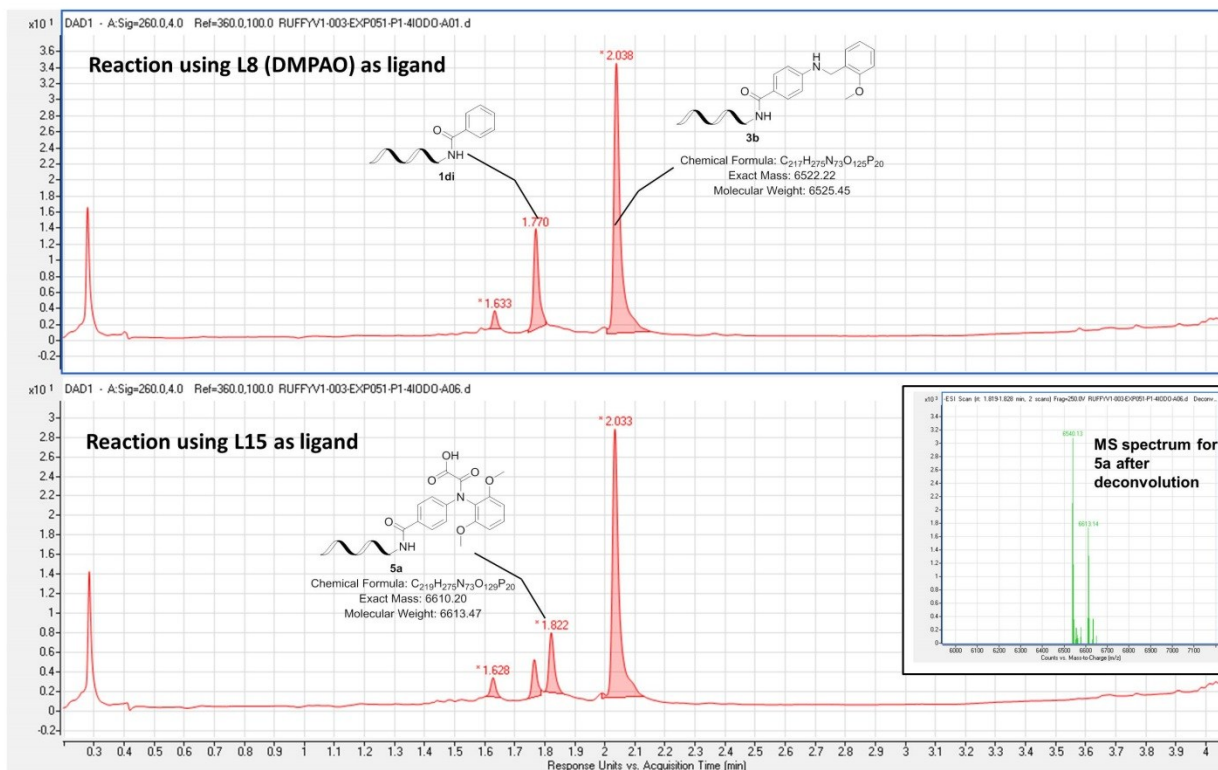


Figure SI V1: UPLC-TOF comparison of the reactions using the amine **2b** and **L8** or **L15** as ligand and characterization of the side product **5a**.

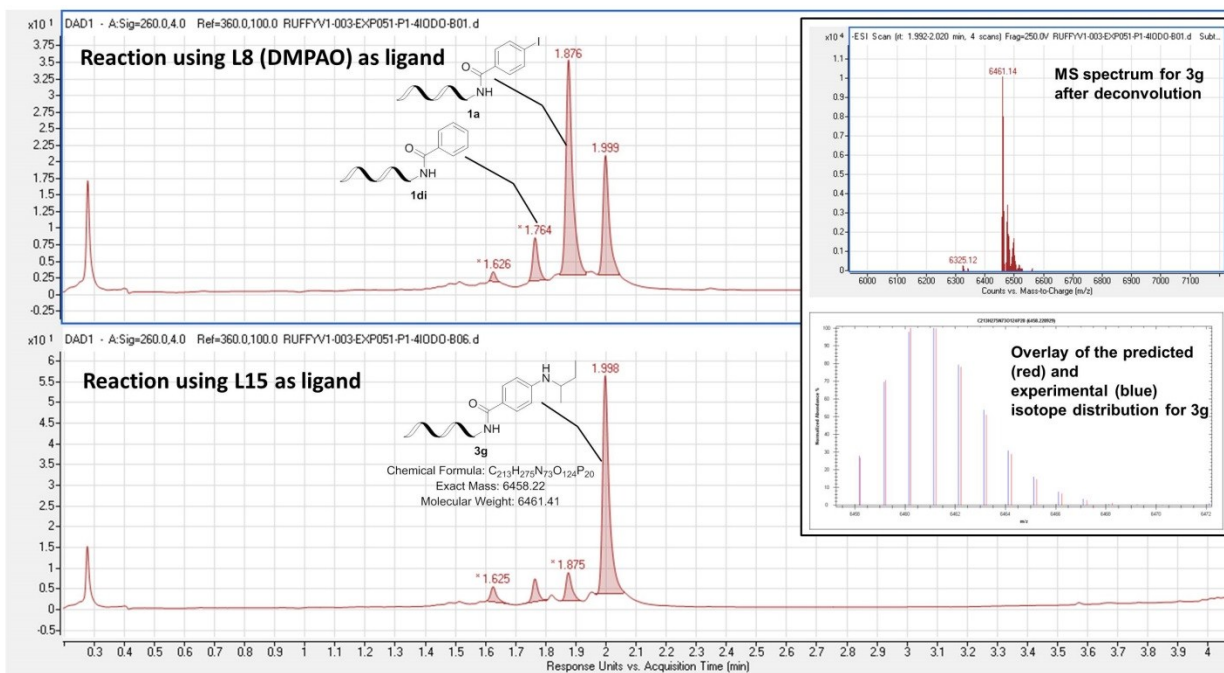


Figure SI V2: UPLC-TOF comparison of the reactions using the amine **2g** and **L8** or **L15** as ligand and characterization of the product **3g**.

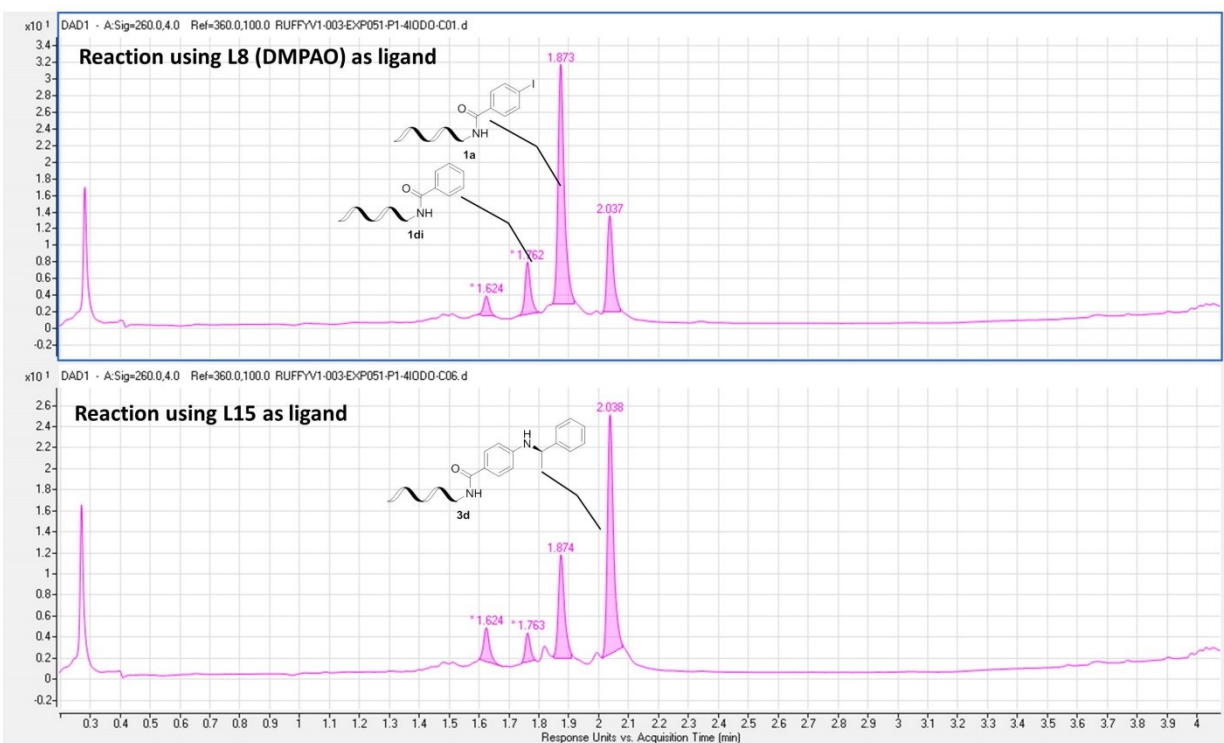


Figure SI V3: UPLC-TOF comparison of the reactions using the amine **2d** and **L8** or **L15** as ligand.

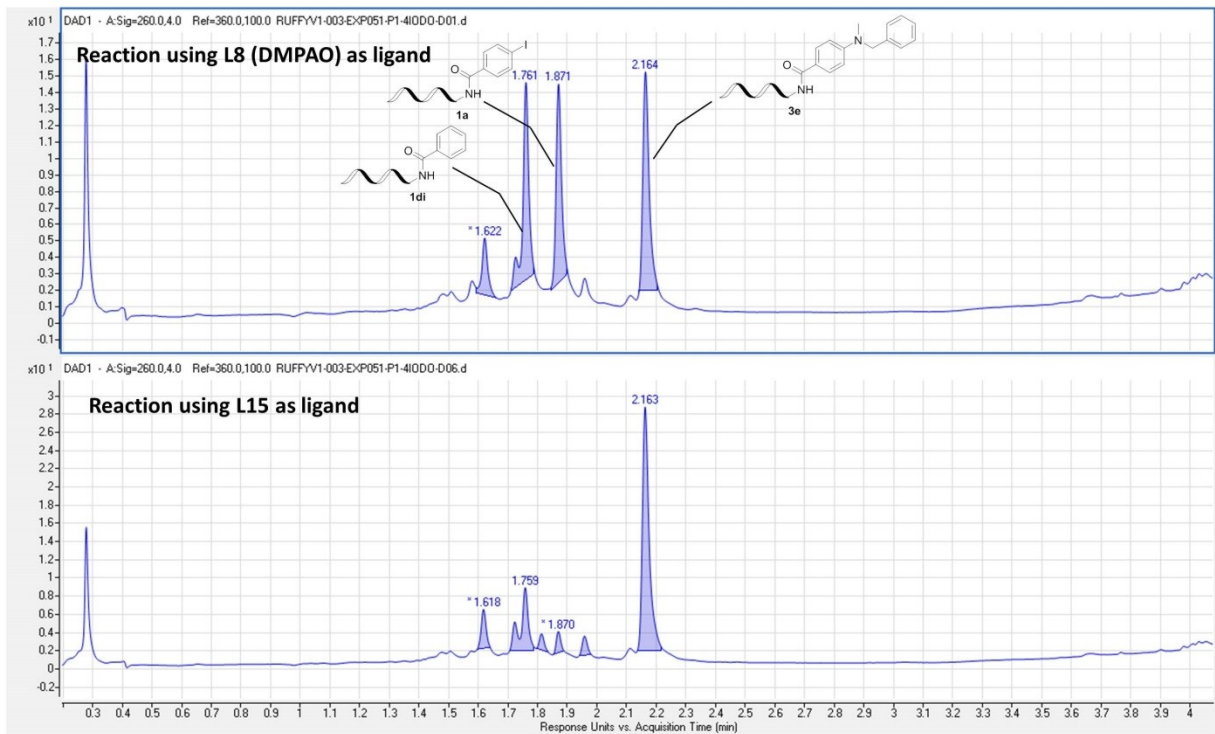


Figure SI V4: UPLC-TOF comparison of the reactions using the amine **2e** and **L8** or **L15** as ligand.

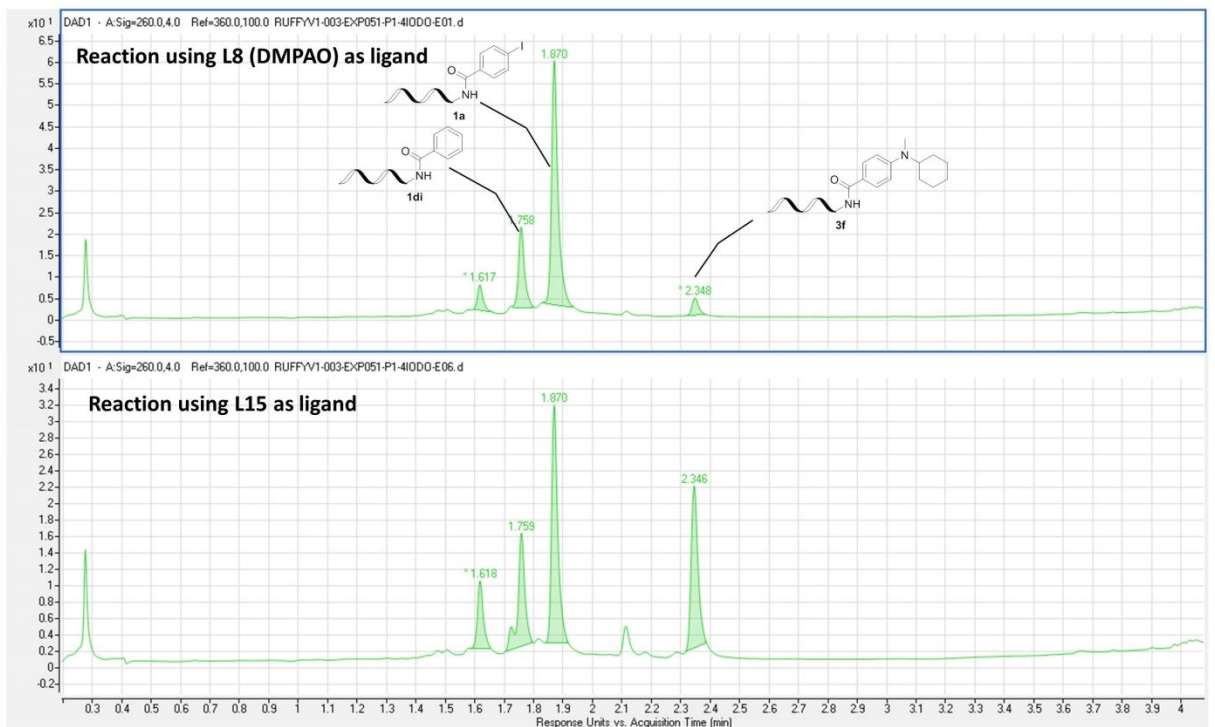


Figure SI V5: UPLC-TOF comparison of the reactions using the amine **2f** and **L8** or **L15** as ligand.

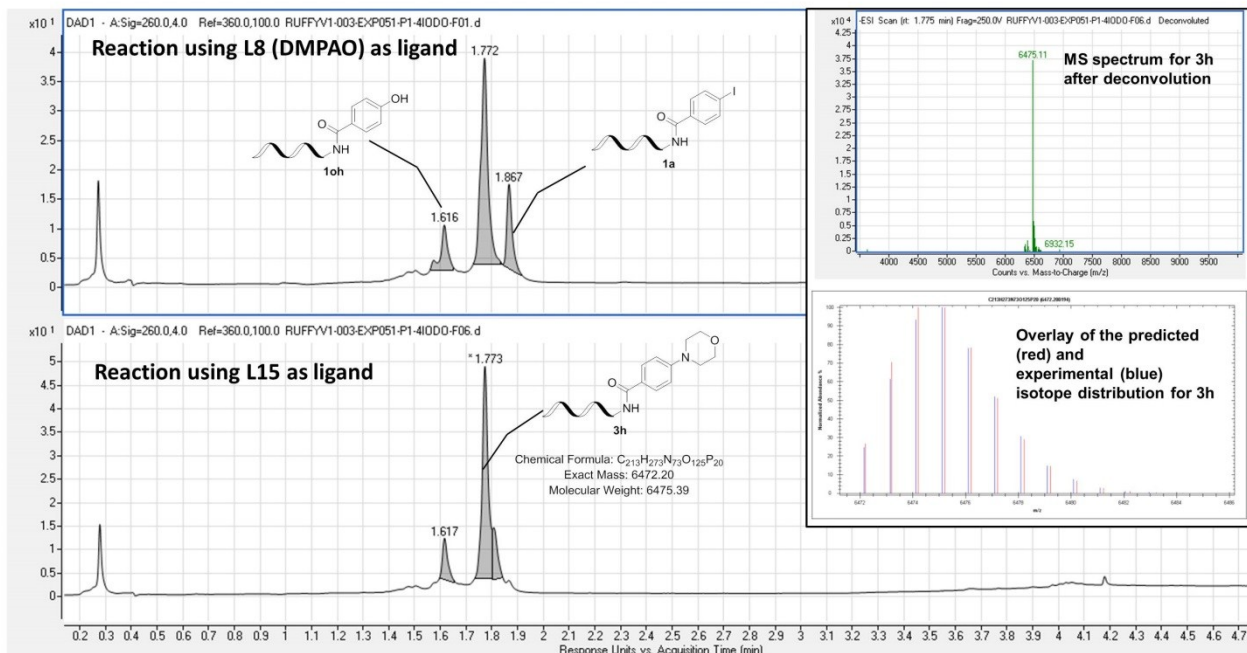


Figure SI V6: UPLC-TOF comparison of the reactions using the amine **2h** and **L8** or **L15** as ligand and characterization of the product **3h**.

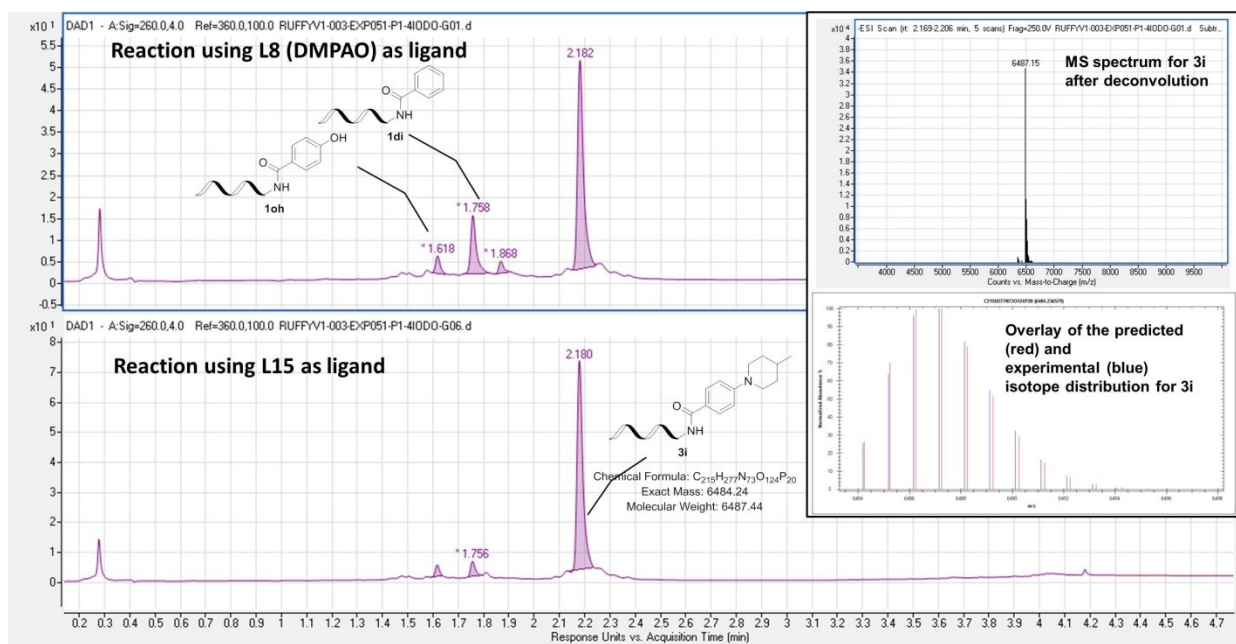


Figure SI V7: UPLC-TOF comparison of the reactions using the amine **2i** and **L8** or **L15** as ligand and characterization of the product **3i**.

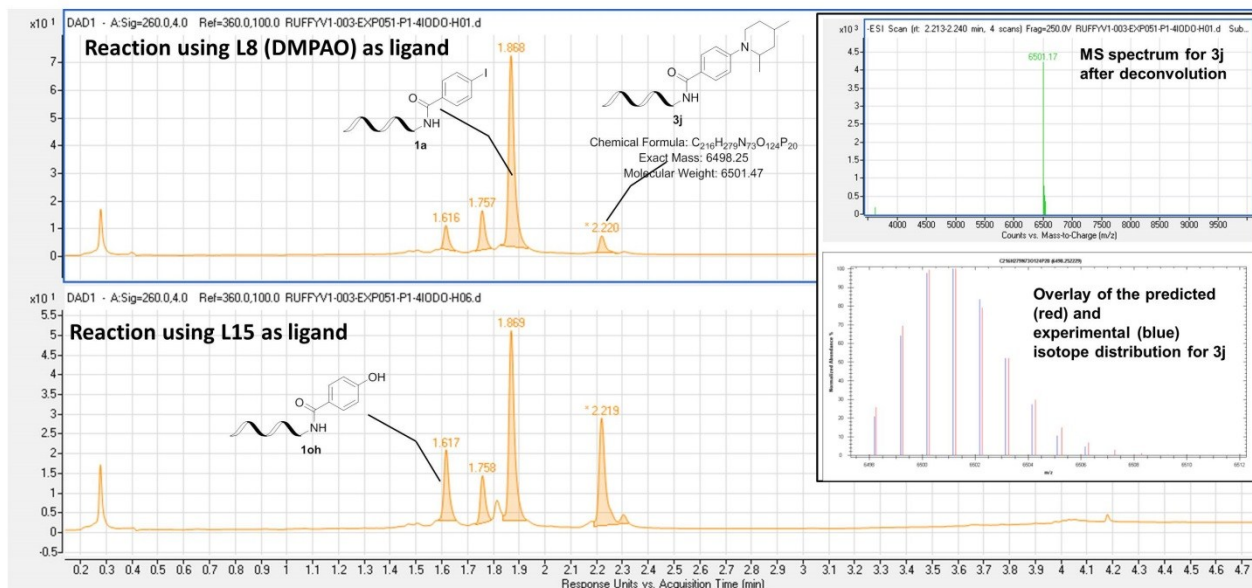


Figure SI V8: UPLC-TOF comparison of the reactions using the amine **2j** and **L8** or **L15** as ligand and characterization of the product **3j**.

SI VI Protocol and analysis for the optimization of the concentration of **L15** and DMSO proportion on the reaction between **1a** and **1b** and amines **2b-j** (main text Figure 3)

1a or **1b** (as a dry pellet in a twin.tec® plate) (6.52 µg, 0.001 µmol) was dissolved in 16 µL of a solution of the base K_3PO_4 (500mM) and ligand **L15** (50-100 or 200mM) solutions. The water was evaporated under reduced pressure.

The dried solids were dissolved by the addition of the aqueous sodium ascorbate solutions (100 or 200 mM, and 4 or 8 µL depending on the final proportion of water). The volume of co-solvent was adjusted by adding water or DMSO when necessary. The stock solutions of amines **2b-j** (2M in DMSO) (4 µL, 8.00 µmol) were added according to the plate map depicted in Figure 3 (main text). The solutions of $Cu(OAc)_2$ in water (for reactions in DMSO/Water 1/3) or DMSO (for reactions in DMSO/Water 1/1-3/1) (100mM, 4 µL) were then added according to the plate map depicted in Figure 3 (main text).

The plate was sealed and incubated at 40°C in a PCR thermocycler (Eppendorf) for 3 hours.

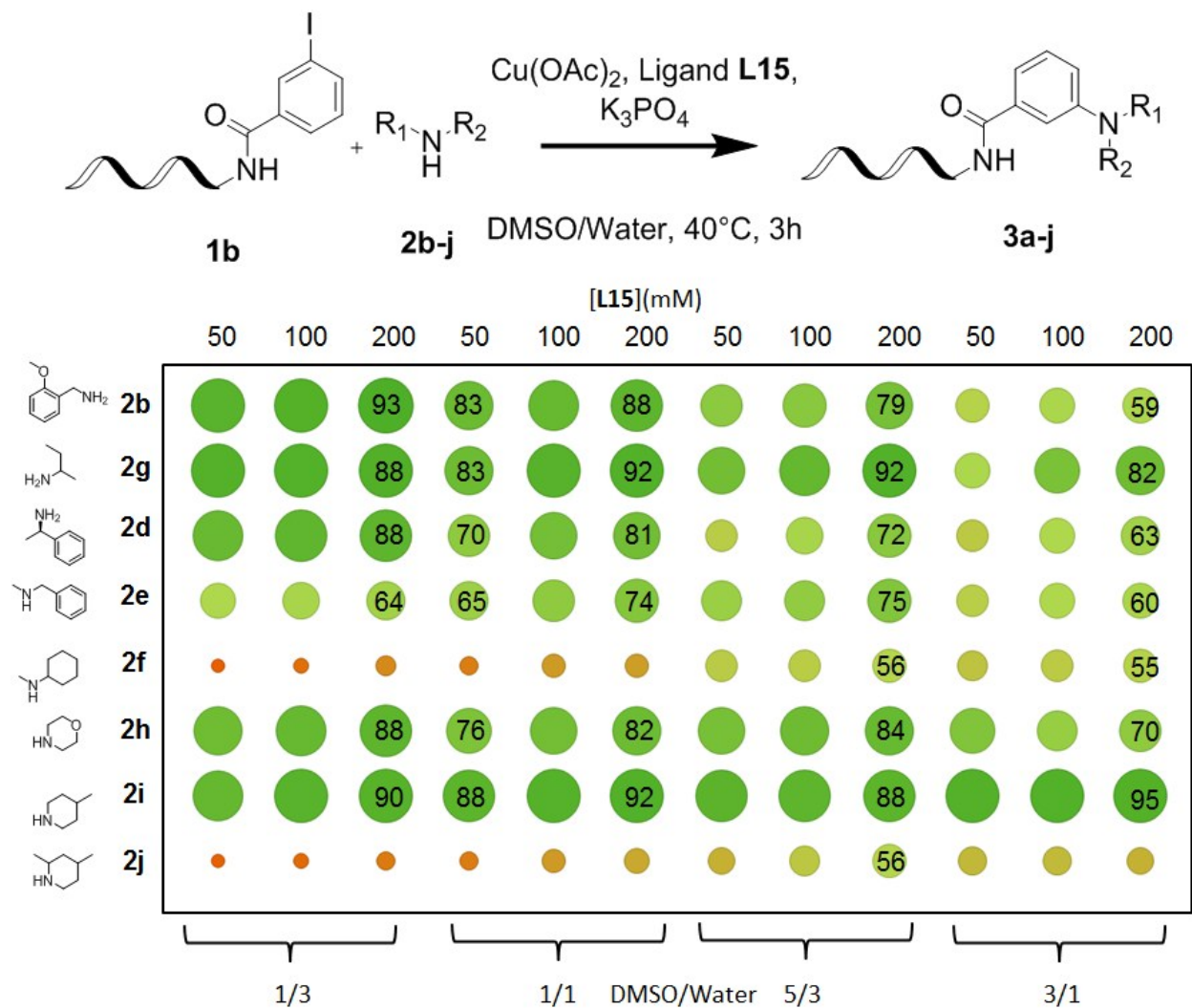


Figure SI VI1: Optimization of the concentration of **L15** and co-solvent proportion in a plate format. **Reaction conditions:** **1b** (1nmol), **2b-j** (500mM), Cu(OAc)_2 (25mM), ligand **L15** (50-200mM), sodium ascorbate (50mM), K_3PO_4 (500mM), DMSO/Water (1/3-3/1, 16ul), 40°C, 3h. The color and diameter of circles correlates with the yield of **3b-j** (A) or **4b-j** (B) determined by UPLC-TOF analysis.

Effect of the concentration of L15 and cosolvent proportion on the proportion of Ligand L15 *N*-arylation side reaction using 1a as substrate.

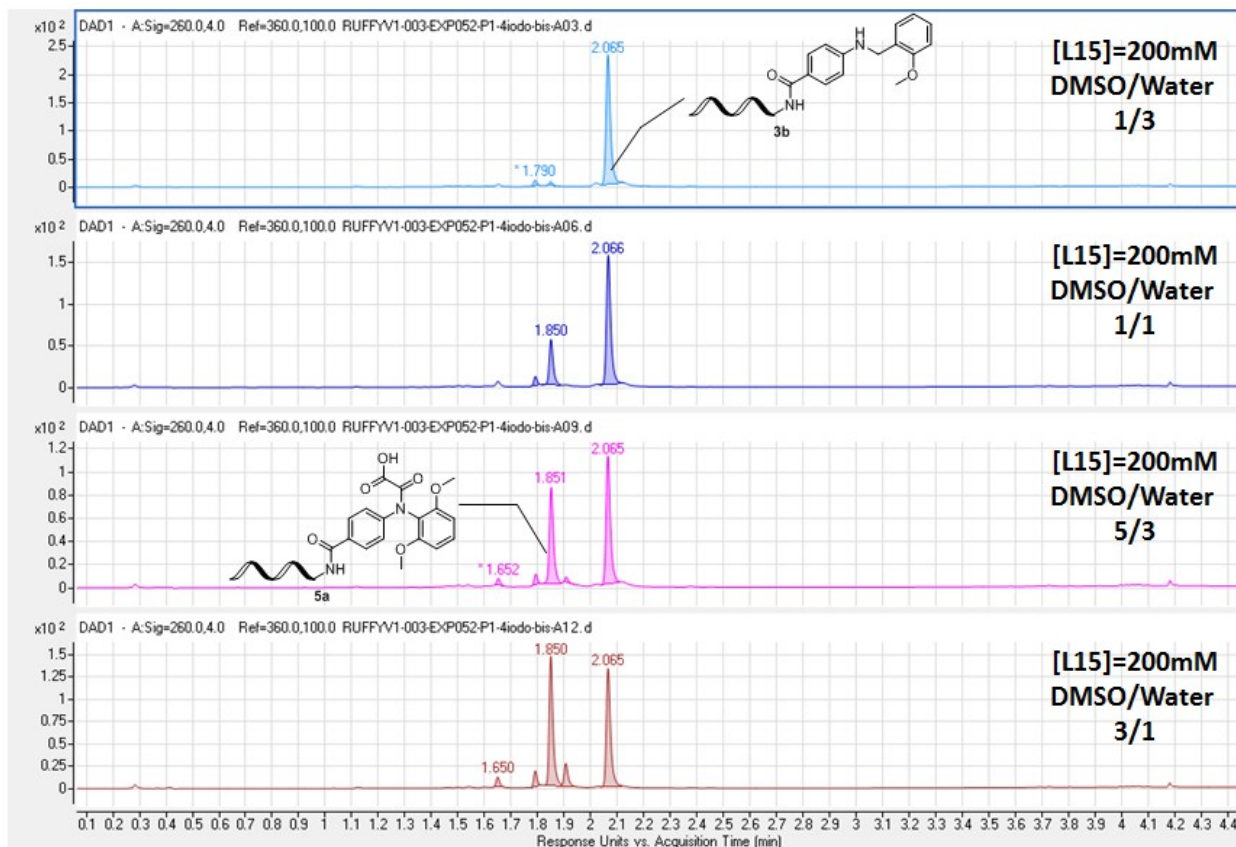


Figure SI VI2: Effect of the concentration of L15 and cosolvent proportion on the proportion of Ligand L15 *N*-arylation side reaction using 1a as substrate.

SI VII Protocol and analysis for the optimization of the DMSO proportion, concentration of L15, amine, and base for the reaction between 1b and amine 2f

In each well of a twin.tec® plate **1b** (as a dry pellet) (3.26 μg , 0.0005 μmol) was dissolved in 4 μL of a solution of the base K_3PO_4 (0, 250, 500, or 1000mM). Sodium ascorbate was then added as a solution in water (4 μL , at 0, 25, 50, 100 or 400mM) The **2f** amine stock solution (400mM, 800mM or 2M in DMSO) (4 μL , 8.00 μmol) was added in each well.

The volume of cosolvent was adjusted by adding water or DMSO to get a final reaction volume of 16 μL . 4 μL of the CuI (100 mM) and ligand **L15** (50,100, 200, 400 or 800mM) solutions in DMSO were then added. After centrifugation, the plate was sealed and incubated at 40°C in a PCR thermocycler (Eppendorf) for 3 hours.

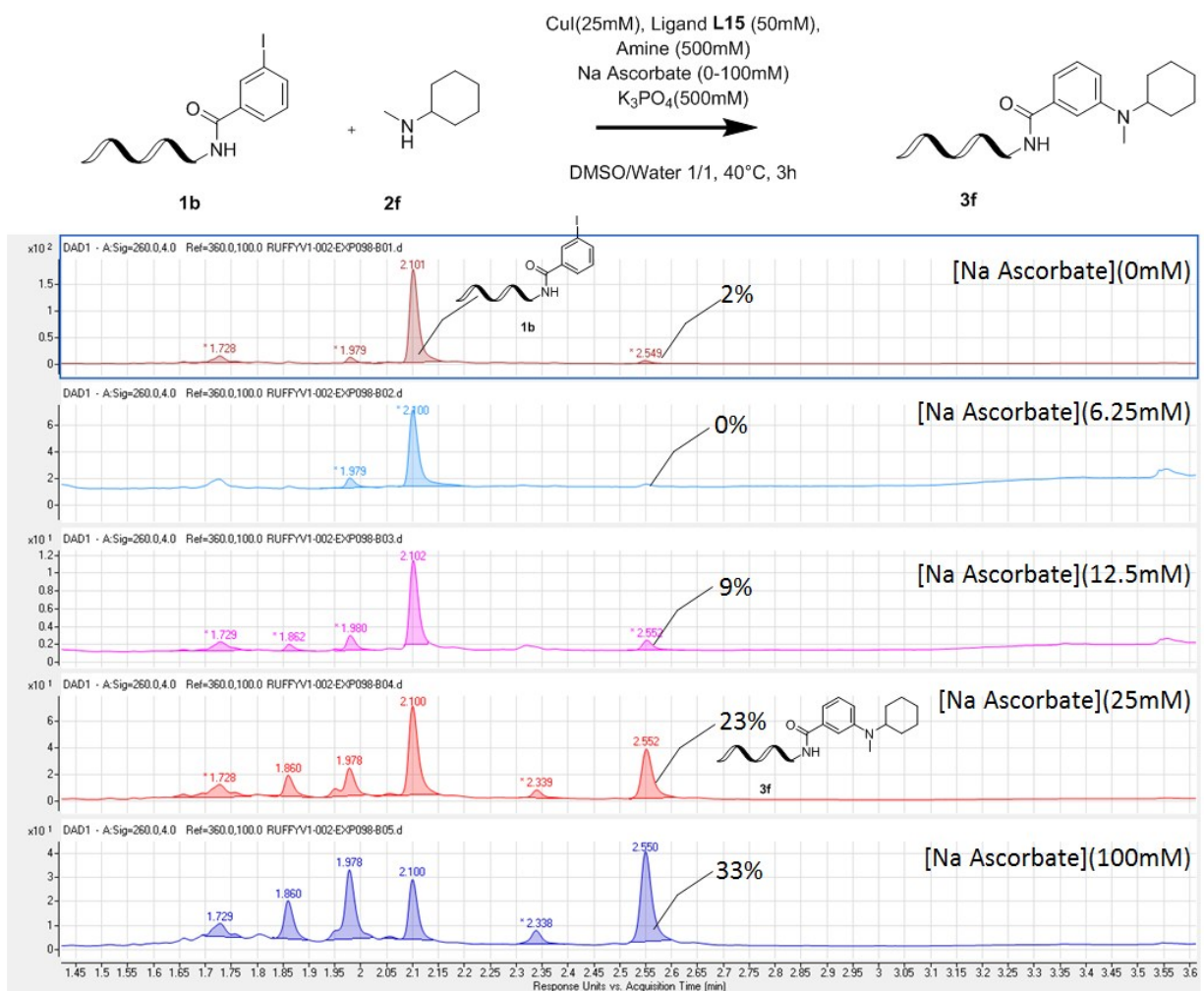


Figure SI VII1: Effect of the sodium ascorbate concentration on the yield of **3f**. Reactions conditions: **1b** (0.5nmol), **2f** (500mM), CuI (25mM), ligand **L15** (50mM), sodium ascorbate (0-100mM), K_3PO_4 (500mM), DMSO/Water (1/1 16 μL), 40°C, 3h.

Increasing the concentration of sodium ascorbate is beneficial for the yield of **3f**, but a 100mM concentration did not increase the yield of **3f** compared to 50 mM (39%).

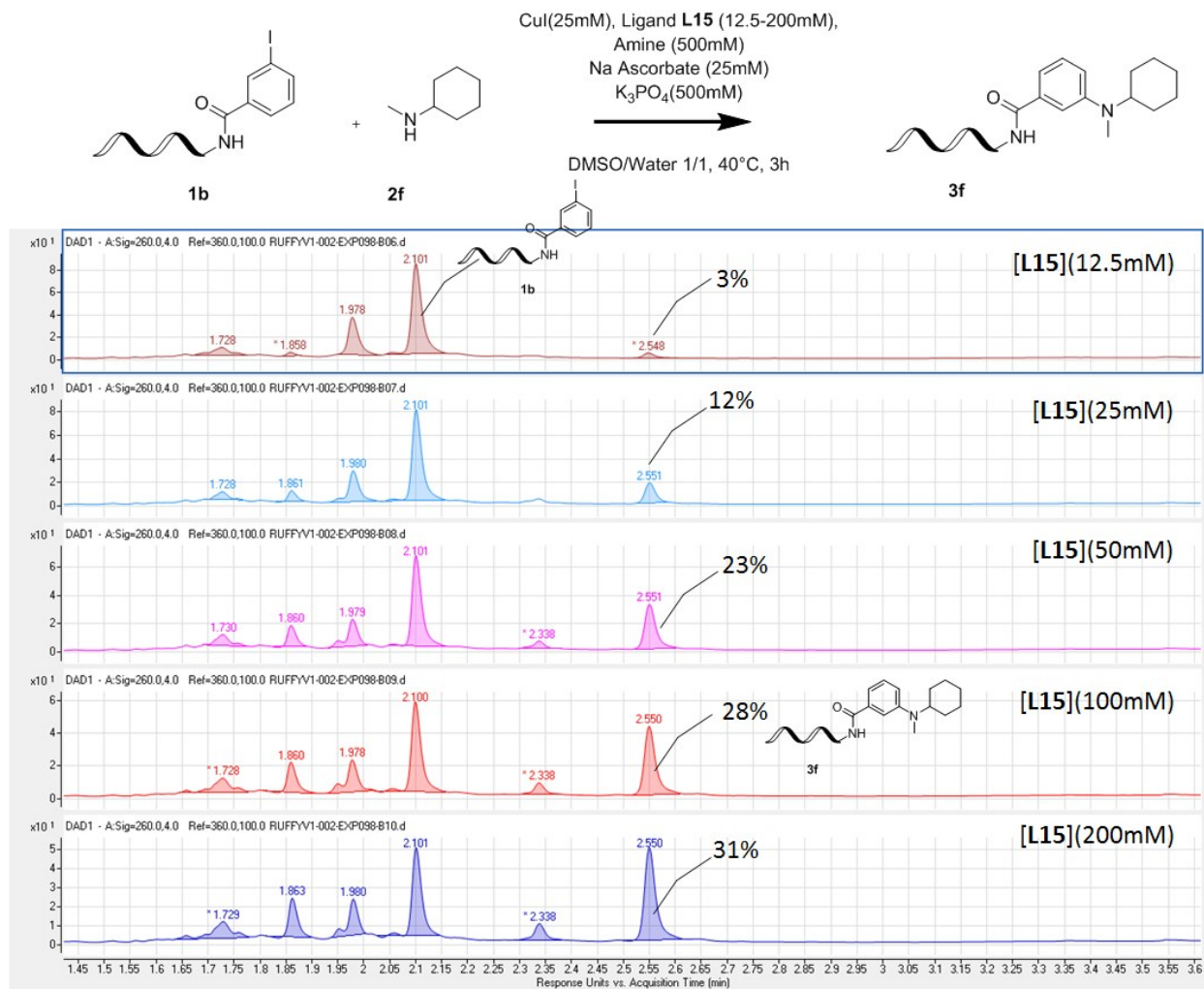


Figure SI VII2: Effect of the ligand **L15** concentration on the yield of **3f**. Reactions conditions: **1b** (0.5nmol), **2f** (500mM), CuI (25mM), ligand **L15** (12.5-200mM), sodium ascorbate (25mM), K_3PO_4 (500mM), DMSO/Water (1/1 16 μ l), 40°C, 3h.

Increasing the concentration of **L15** up to 200mM led to higher yield of **3f**.

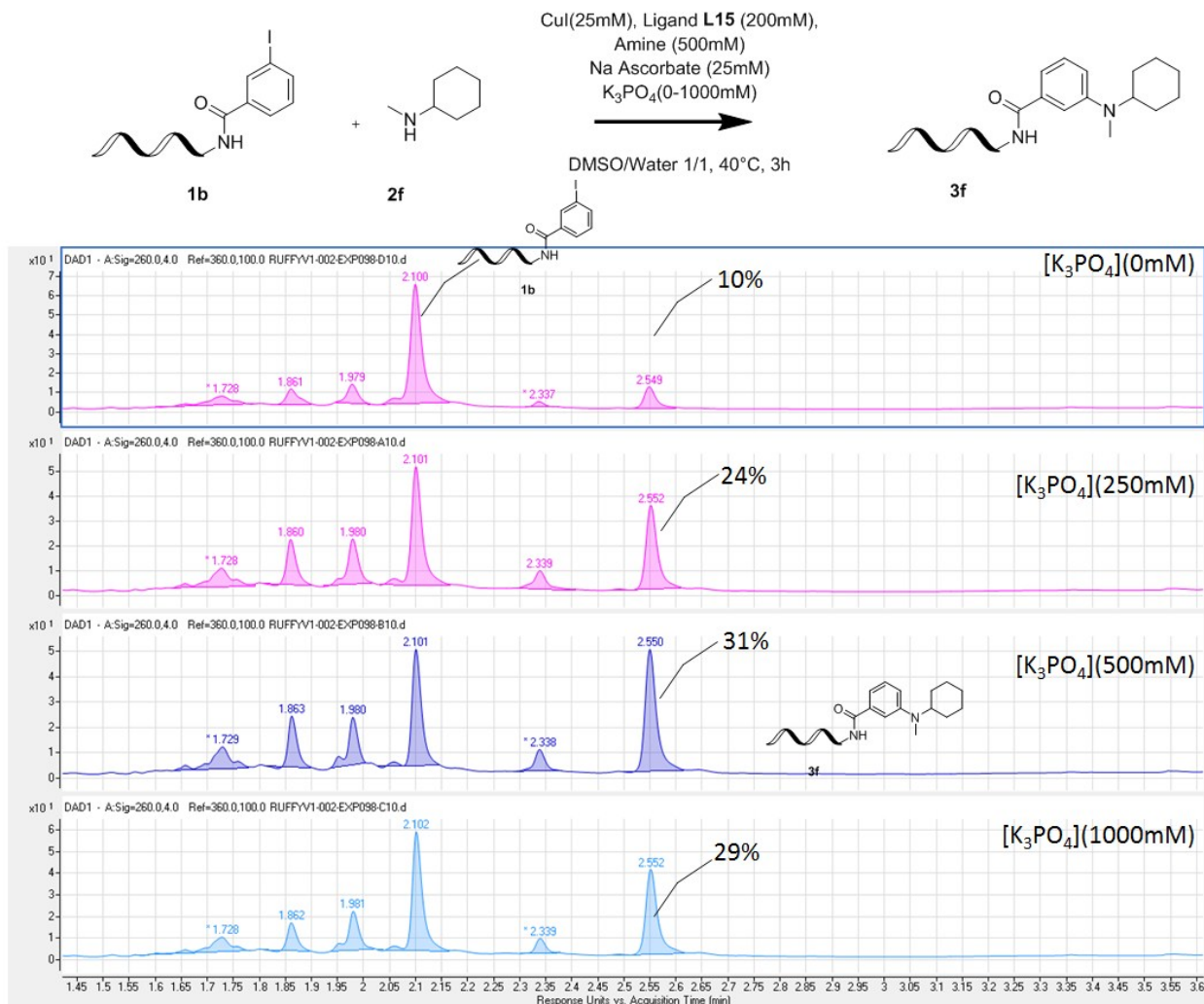


Figure SI VII3: Effect of the concentration of K₃PO₄ on the yield of **3f**. Reaction conditions: **1b** (0.5nmol), **2f** (500mM), CuI (25mM), ligand **L15** (200mM), sodium ascorbate (25mM), K₃PO₄ (0, 250, 500 or 1000mM), DMSO/Water (1/1 16ul), 40°C, 3h.

The initial base concentration appeared optimal as increasing it or lowering it did not have a positive effect on the yield of **3f**.

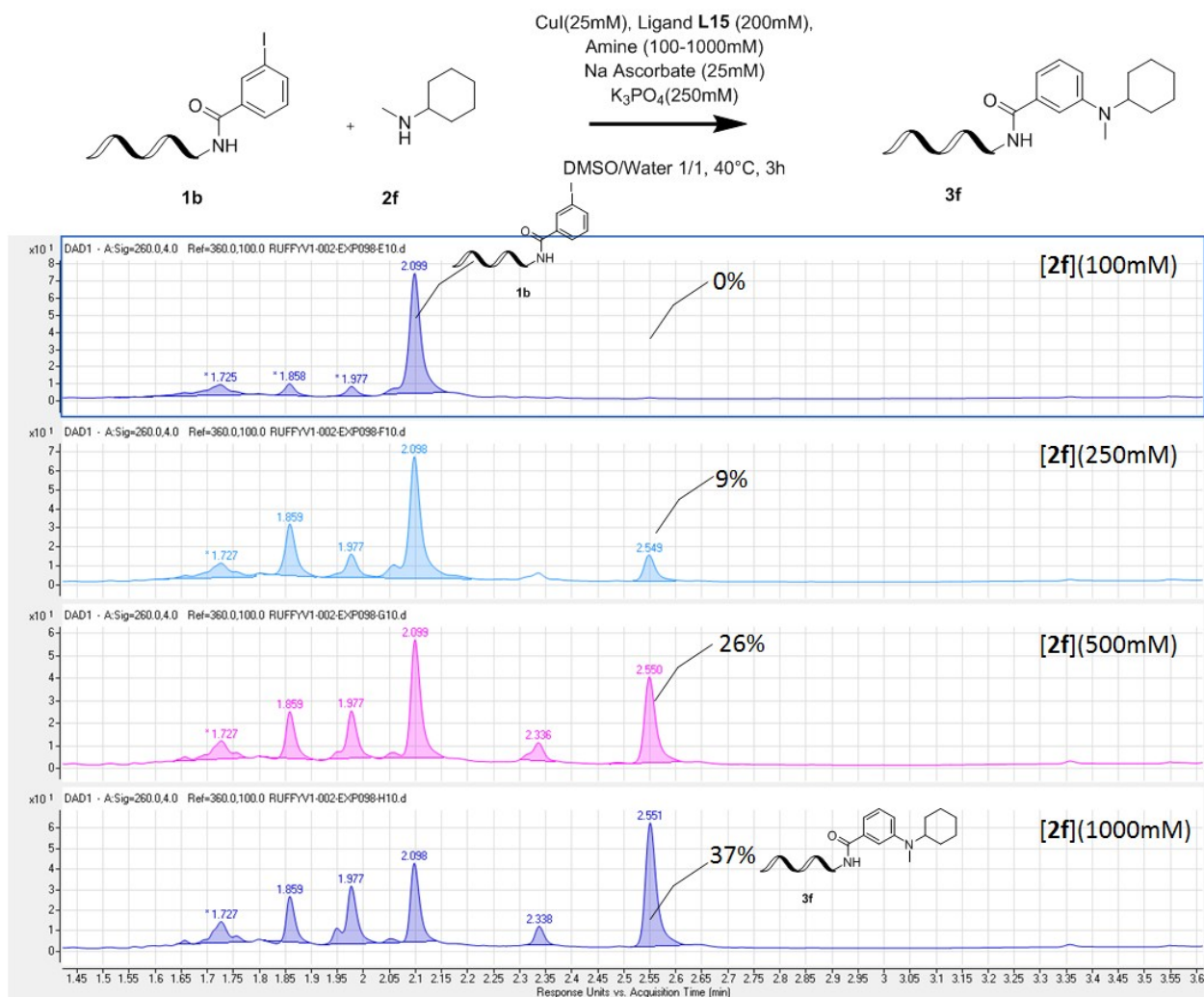
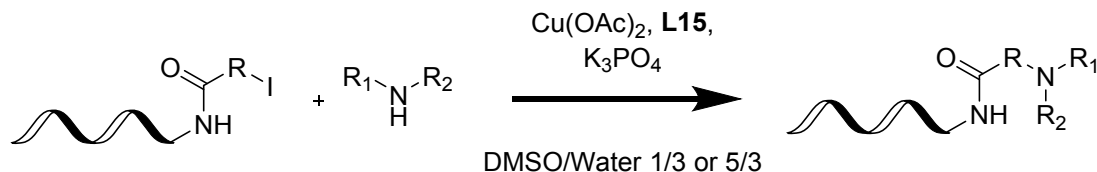


Figure SI VII4: Effect of the concentration of amine **2f** on the yield of **3f**. Reactions conditions: **1b** (0.5nmol), **2f** (100, 250, 500 or 1000mM), CuI (25mM), ligand **L15** (200mM), sodium ascorbate (25mM), K3PO4 (500mM), DMSO/Water (1/1 16ul), 40°C, 3h.

Increasing the concentration of amine up to 1M led to higher yield of **3f**, but for practical reasons, a concentration of 500mM was used in subsequent experiments.

SI VIII Protocol and analysis for the scope determination on 1b-1i (main text Figure 4)



2x96 reactions were set up.

1 nmole of the DNA-conjugated aryl iodide substrate **1b-f** were evaporated in 2 twin.tec® plates according to the plate map depicted in Figure 4 (main text).

First set of 96 reactions in DMSO/Water 1/3 at 40°C (Conditions 1):

The DNA-conjugated aryl iodide (as a dry pellet in a twin.tec® plate) (0.001 μmol) was dissolved in a solution containing **L15** (320 mM), sodium ascorbate (80 mM) and K_3PO_4 (800mM) in water (10 μl) followed by the addition of the amines **2b-2q** stock solutions (2M in DMSO) (4 μL , 8.00 μmol) according to the plate map depicted in Figure 4 (main text). Copper (II) acetate (200 mM in water) (2 μL , 0.400 μmol) was then added (final volume 16 μl) and the twin.tec® plate was sealed and incubated at 40°C in a PCR thermocycler (Eppendorf) for 3 hours.

Second set of 96 reactions in DMSO/Water 5/3 at 40°C (Conditions 2):

The DNA-conjugated aryl iodide (as a dry pellet in a twin.tec® plate) (0.001 μmol) was dissolved in a solution of containing **L15** (533 mM), sodium ascorbate (133 mM) and K_3PO_4 (1333 mM) in water (6 μl) followed by the addition of the amines **2b-2q** stock solutions (2M in DMSO) (4 μL , 8.00 μmol) according to the plate map depicted in Figure 4 (main text). The copper (II) acetate (66.6mM in DMSO) (6 μl , 0.400 μmol) was then added (final volume 16 μl) and the twin.tec® plate was sealed and incubated at 40°C in a PCR thermocycler (Eppendorf) for 3 hours.

SI IX Protocol and analysis for the evaluation of DNA damage under reaction conditions 1

The evaluation of the DNA strands integrity was evaluated using two different techniques.

A first pass method is UPLC-TOF analysis, which is indicative of the yield and integrity of the model oligonucleotides used in reaction optimization. As illustrated in the figure SI IV7, DNA degradation by cleavage (known to be catalyzed by copper under certain reaction conditions³) can be detected by the appearance of new peaks in the UV chromatogram correlating with molecular weights corresponding to shorter oligonucleotides.

Other types of DNA damage include depurination, usually occurring under acidic conditions or at elevated temperature. These can also be detected by the analysis of the MS spectrum.

Under our optimized conditions, no such DNA degradation was detected as illustrated by the clean UPLC-TOF chromatogram and MS spectrum analysis depicted in the figures SI IV1-6 .

However, we wanted to ensure that our conditions would still be compatible with the longer DNA tags which are used during DEL synthesis. We therefore used the DNA-encoded library rehearsal methodology described by Malone *et al.* in which the yield of amplifiable DNA remaining after a chemical transformation is evaluated by quantitative PCR.⁴⁻⁵

In this protocol, a known amount of a reporter oligonucleotide is initially added to the reactions. After work up, the amount of intact/amplifiable DNA in the reaction samples is quantified using qPCR using the Ct method and compared to standard samples. While in the original publication by Malone *et al.* used an oligonucleotide immobilized on magnetic beads,⁴ we added the free reporter oligonucleotides directly in the reactions for a better correlation with the solution-phase library synthesis conditions. We designed a 84-base double stranded reporter oligonucleotide **qPCR-oligo** to mimic the DNA tags. This proprietary oligonucleotide (obtained commercially from IDT) is designed to contain a 20 bases site for hybridization with a qPCR Taqman probe (obtained commercially from Aldrich) in between two 20-base primers. The amount of amplifiable DNA in Ullmann *N*-arylation reaction samples was then determined using standard qPCR protocols after establishing a standard curve for reference samples containing known quantities of this reporter oligonucleotide (Figures SI IX1 and IX2).

1a (as a dry pellet in a twin.tec® plate) (0.001 μmol) and **qPCR-oligo** (1.6 pmol) (as a dry pellet in a twin.tec® plate) was dissolved in a solution containing **L15** (320mM), sodium ascorbate (0-80mM) and K_3PO_4 (800mM) in water (10 μl) followed by the addition of the Amine **2b** stock solutions (2M in DMSO) (4 μL , 8.00 μmol). Copper (II) acetate (0-200mM in water) (2 μL , 0.400 μmol) was then added (final volume 16 μl) and the twin.tec® plate was sealed and incubated at 40°C in a PCR thermocycler (Eppendorf) for 30min, 1, 2 or 3 hours.

Each reaction was diluted with 84 μl of a 100mM sodium diethyldithiocarbamic acid solution in milliQ water and desalted before UPLC-TOF and qPCR analysis.

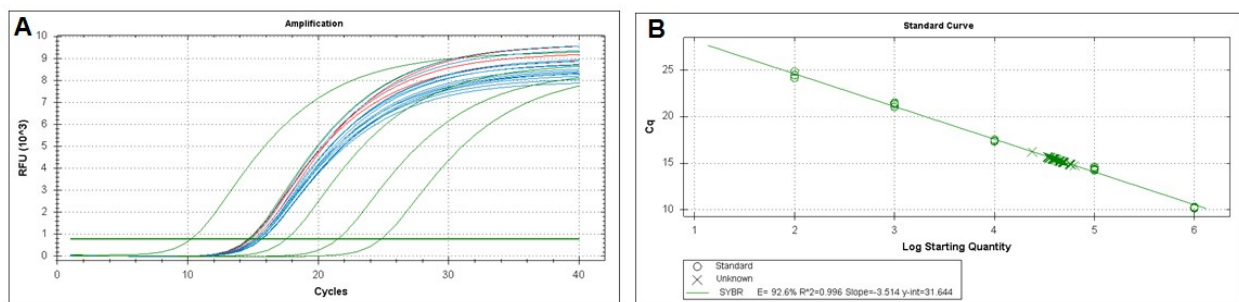


Figure SI IX1: Representative qPCR amplification curves (**A**) and Cq (Calibration cycles) calibration curve (**B**) used for the quantification of remaining amplifiable DNA in Ullmann *N*-arylation reaction samples (Blue) and reference samples (red). Samples (cross symbols) were run in quadruplicates and calibration standards (circles) were run in sextuplicates. For clarity, only one fluorescence trace is represented for each sample.

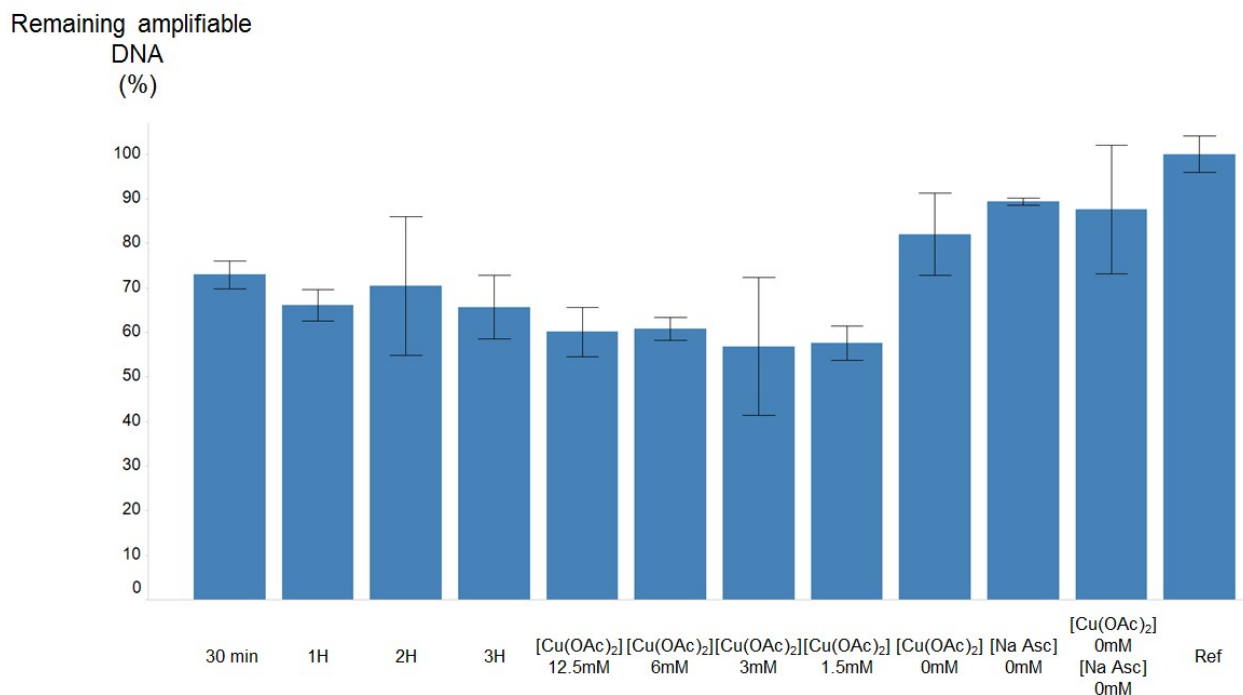


Figure SI IX2: Percentage of remaining amplifiable DNA in Ullmann *N*-arylation reaction samples under reaction conditions 1 after 30min, 2 or 3 hours and different amounts of copper(II) source or sodium ascorbate. The bar labels indicate variations of a single parameter with respect to conditions 1 (**1a** (1nmol), **2b** (500mM), Reporter oligonucleotide (1.6 pmol), Cu(OAc)₂(25mM), ligand **L15** (200mM), sodium ascorbate (50mM), K₃PO₄ (500mM), DMSO/Water(1/3 16ul), 40°C, 3h). The percentage of amplifiable DNA was calculated with respect to a reference sample (Ref) of **qPCR-oligo** which was not submitted to reaction conditions 1 but treated exactly like the reactions samples during the sample preparation for qPCR. Error bars correspond to the standard deviation of the samples.

Gratifyingly, for this particular combination of coupling partners **1a** and **2b**, all reactions led to >95% conversion into **3b** except the 3 reactions that did not contain either Cu(OAc)₂ or sodium ascorbate (0% conversion).

After submitting the **qPCR-oligo** to the standard reaction **conditions 1**, 65% of the reporter dsDNA remains amplifiable (Figure SI IX2). This loss is comparable with other chemical transformations which have successfully been applied to the synthesis of DNA encoded libraries like acylation⁴ or Suzuki cross-couplings^{4, 6}. We conclude that **conditions 1** are compatible with the presence of the DNA tags and with the synthesis of DNA encoded libraries.

SI X Confirmation the structure of DNA conjugates **3a**, **3h** and **4h** by synthesis using commercial reagents

Whenever possible, we confirmed the structure of the *N*-arylated DNA conjugates by acylation of the **Oligo-Amine** starting material using commercial carboxylic acids. The retention time and MS spectrum of the acylation product was directly compared to those of the reaction samples. Co-injection with the reaction mixtures confirmed the identity of the expected *N*-arylated product.

As the acids 4-pyrrolidin-1-ylbenzoic acid, 4-morpholinobenzoic acid and 3-morpholinobenzoic acid are available commercially, we synthesized the conjugates **3a**, **3h** and **4h** by standard acylation protocols adapted from the literature¹.

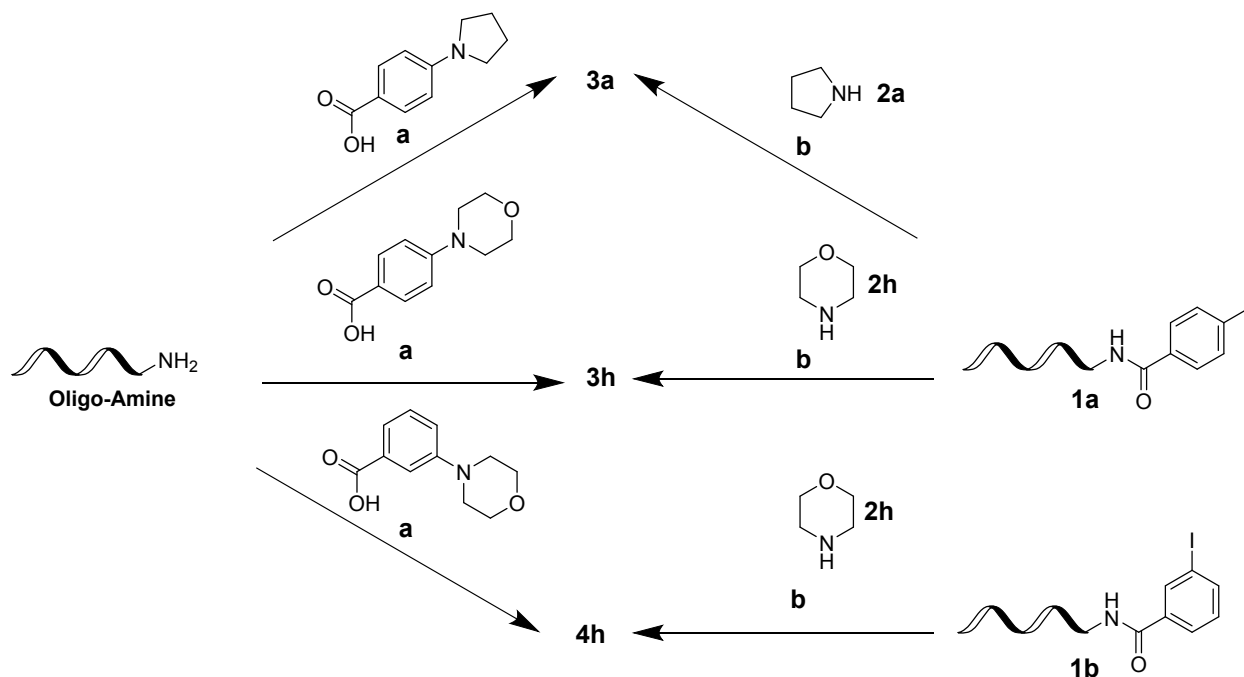


Figure SI X1: Structural confirmation for the DNA conjugates **3a**, **3h** and **4h** by acylation or using Ullmann *N*-arylation protocol 1. Reaction conditions a: standard acylation protocol (SI II) b: Reactions **conditions 1** (SI VIII).

3a, **3h** and **4h** were obtained using the general acylation procedure described in SI II. These compounds were also prepared independently by Ullmann *N*-arylation using our reaction conditions 1 (see section SI VIII), starting from **1a** and **1b**. The perfect match in retention time and isotope distribution for these compounds confirms the structure of the *N*-arylation products (Figures SI X2-4).

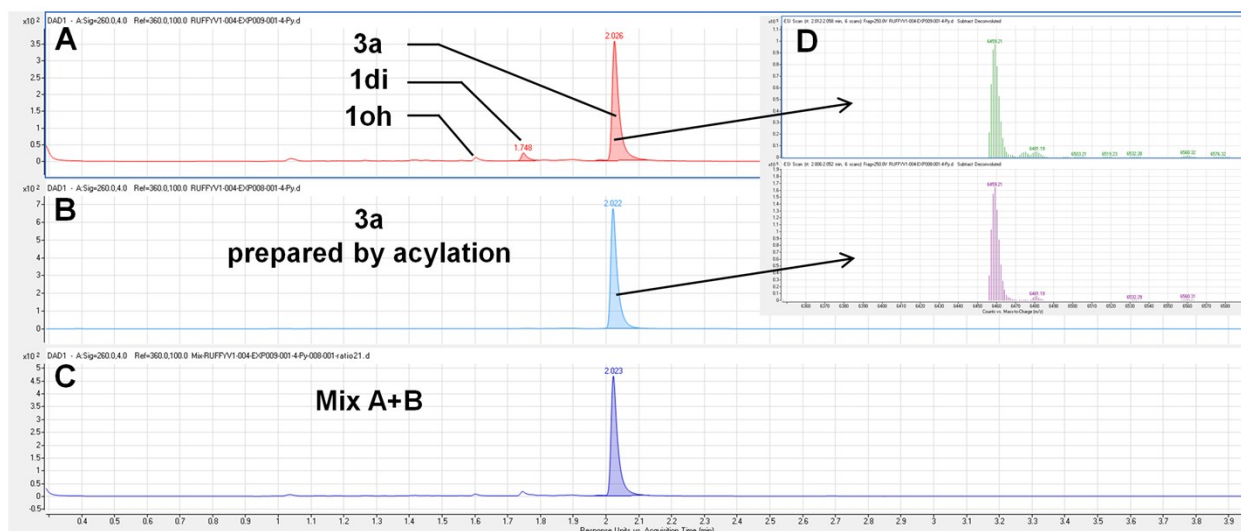


Figure SI X2: Confirmation of the structure of **3a** by two independent synthetic methods. **A** Crude Chromatogram of the *N*-arylation reaction between **1a** and **2a** under reaction **conditions 1** (sample **A**). **B** Chromatogram of a purified sample of **3a** obtained by acylation of **Oligo-Amine** with 4-pyrrolidin-1-ylbenzoic acid (ChemBridge PubChem Substance ID 329781281) under the standard acylation conditions described in SI II and Figure SI X1 (sample **B**). **C** Mixture of the two samples **A** and **B**. **D** Overlay of the MS isotopic distribution of the sample of **3a** prepared by *N*-arylation (top) and acylation (bottom).

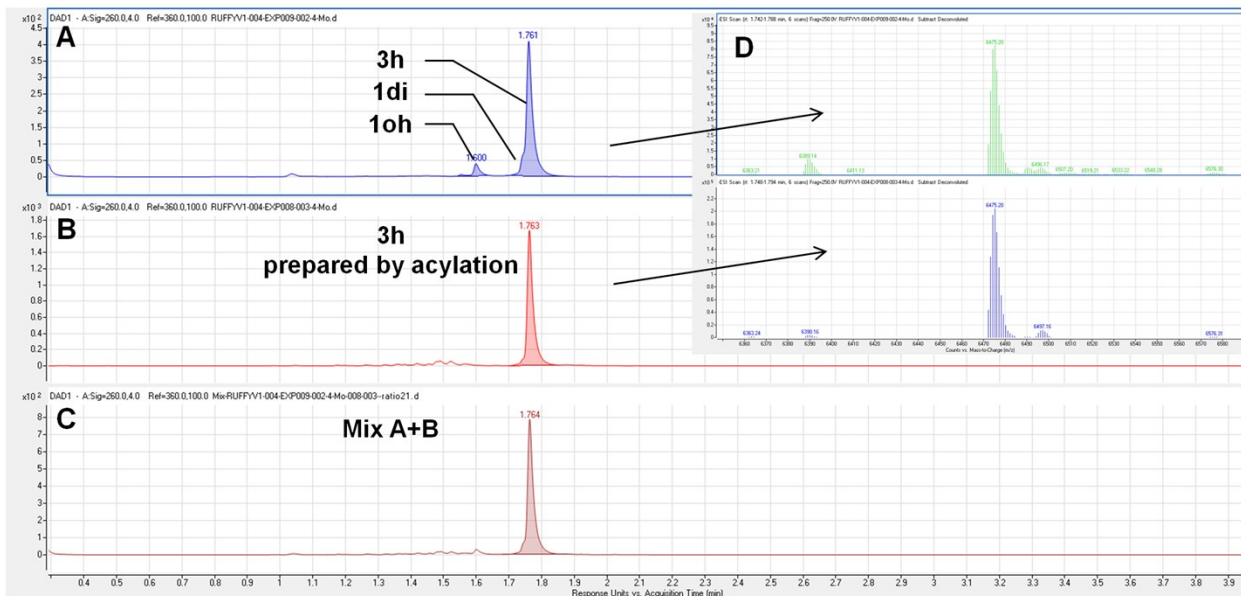


Figure SI X3: Confirmation of the structure of **3h** by two independent synthetic methods. **A** Crude chromatogram of the *N*-arylation reaction between **1a** and **2h** under reaction **conditions 1** (sample **A**). **B** Crude chromatogram of a sample of **3h** obtained by acylation of **Oligo-Amine** with 4-morpholinobenzoic acid (Aldrich CAS 7470-38-4) under the standard acylation conditions described in SI II and Figure SI X1 (sample **B**). **C** Mixture of the two samples **A** and **B**. **D** Overlay of the MS isotopic distribution of the sample of **3h** prepared by *N*-arylation (top) and acylation (bottom).

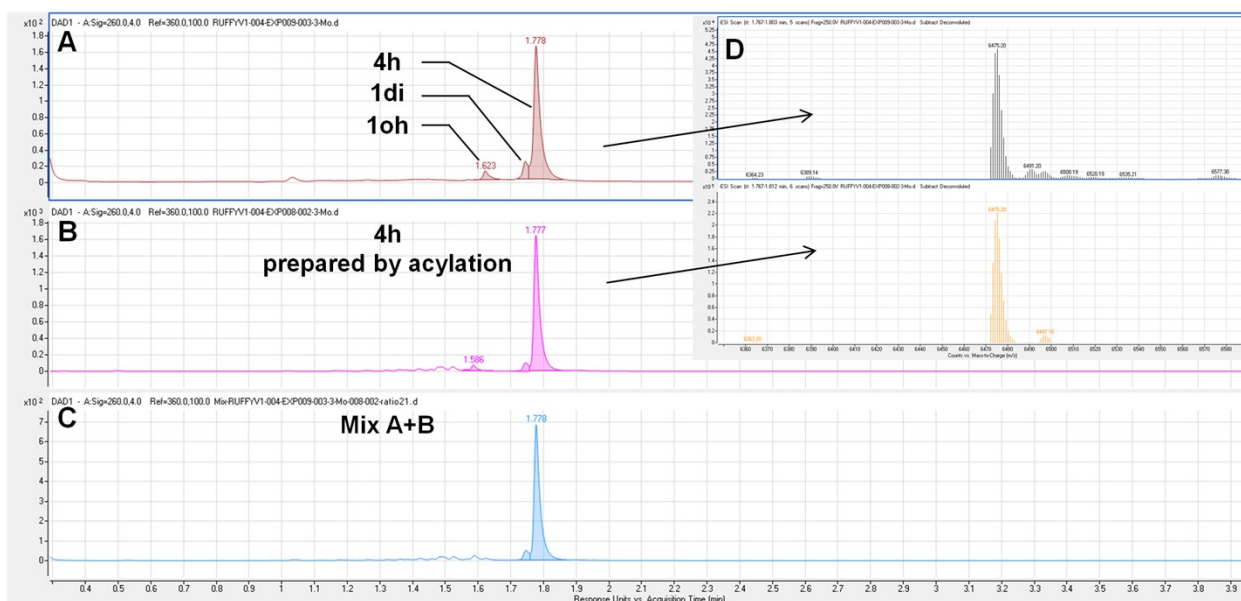


Figure SI X4: Confirmation of the structure of **4h** by two independent synthetic methods. **A** Crude chromatogram of the *N*-arylation reaction between **1b** and **2h** under reaction **conditions 1** (sample **A**). **B** Crude chromatogram of a sample of **3h** obtained by acylation of **Oligo-Amine** with 3-morpholinobenzoic acid (ABCR 215309-00-5) under the standard acylation conditions described in SI II and Figure SI X1 (sample **B**). **C** Mixture of the two samples **A** and **B**. **D** Overlay of the MS isotopic distribution of the sample of **4h** prepared by *N*-arylation (top) and acylation (bottom).

- Halpin, D. R.; Lee, J. A.; Wrenn, S. J.; Harbury, P. B., DNA Display III. Solid-Phase Organic Synthesis on Unprotected DNA. *PLoS Biology* **2004**, *2* (7), e175.
- Zhang, Y.; Yang, X.; Yao, Q.; Ma, D., CuI/DMPAO-Catalyzed *N*-Arylation of Acyclic Secondary Amines. *Org. Lett.* **2012**, *14* (12), 3056-3059.
- Brissos Rosa, F.; Caubet, A.; Gamez, P., Possible DNA - Interacting Pathways for Metal - Based Compounds Exemplified with Copper Coordination Compounds. *Eur. J. Inorg. Chem.* **2015**, *2015* (16), 2633-2645.
- Malone, M. L.; Paegel, B. M., What is a "DNA-Compatible" Reaction? *ACS Combinatorial Science* **2016**, *18* (4), 182-187.
- Schmittgen, T. D.; Livak, K. J., Analyzing real-time PCR data by the comparative CT method. *Nature Protocols* **2008**, *3*, 1101.
- Ding, Y.; Clark, M. A., Robust Suzuki–Miyaura Cross-Coupling on DNA-Linked Substrates. *ACS Combinatorial Science* **2015**, *17* (1), 1-4.