

Supporting Information

Homogeneous and Functional Group Tolerant Ring-Closing Metathesis for DNA-Encoded Chemical Libraries

^{†‡}*Olivier B. C. Monty*, [†]*Pranavanand Nyshadham*, [†]*Kurt M. Bohren*, [†]*Murugesan Palaniappan*, [†]*Martin M. Matzuk*, [†]*Damian W. Young**, [†]*Nicholas Simmons**

[†]Center for Drug Discovery and Department of Pathology and Immunology, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, United States, [‡]Rice University Department of Chemistry, 6100 Main Street, Houston, TX 77005

*Corresponding Authors: Damian.young@bcm.edu (D. W. Y.); nsimmons@its.jnj.com (N. S.)

Contents

General information	3
Materials.....	3
DNA Headpiece (S1) Synthesis.....	4
General analytical procedure for the analysis of DNA oligonucleotide compositions.	5
General DNA Precipitation Procedure (ethanol precipitation).....	5
General procedure for the ligation of DNA oligonucleotides.	6
General HPLC Purification Procedure	7
General Reaction Conditions for the preparation of substrates	8
Syntheses.....	11
Synthesis of Catalyst A.....	11
.....	11
Synthesis of Catalyst B.....	11
.....	11
Synthesis of Building Blocks.....	12
Synthesis of On-DNA Scaffolds	14
Synthesis of On-DNA Substrates.....	15
Synthesis of Single-Substrate Library	20
Experimental Information	21
Buffering the RCM Reaction Mixture.....	21
General Protocol for RCM Reaction under Our Main Reaction Conditions (see below for special protocol for substrate 24a).....	21

General Protocol for RCM Reaction under our Alternative #1 Conditions	21
General Protocol for RCM Reaction under our Alternative #2 Conditions	22
RCM, Precipitation Procedure and Post-RCM Yield of Substrate 24a	22
RCM and Sequencing Prep of Single-Substrate DECL	24
Quenching the RCM Reaction and Preparation for LC/MS Analysis	24
Protocol for Replication of Reaction Conditions Reported by Lu, X. et al. ⁹	25
Method for the Determination of the Percent Conversion from MS Deconvolution Results	25
Comparison of Percent Conversions Determined from MS Deconvolution Spectra v. UV spectra	26
Confirmation of Cyclization via Chemical Modification	31
Confirmation of Cyclization via HTRF Assay of Cyclized v. Uncyclized Substrate 26a against target protein, ER alpha	33
Sequencing of the Single-Substrate DECL: a Pre- and Post-RCM Comparison	34
Illumina sequencing and analysis	34
Screening Tables	36
Characterization Information	37
Post-Metathesis (RCM and CM) MS Deconvolution Spectra	69
Post-Metathesis (RCM and CM) LCMS Traces for Reactions Ran under Scheme 1(iii) Conditions	97

General information

Many of the materials, equipment and general procedures within this work that were used to ligate, chemically modify and analyze the oligonucleotide–chemical conjugates are adapted from related works we have published previously.¹⁻⁴

Materials

DTSU (“DEC-Tec Starting Unit”) (Figure S1) from LGC Biosearch Technologies and 5’-phosphorylated oligonucleotides were obtained from LGC Biosearch Technologies or Sigma-Aldrich; all were assessed for purity through the general analytical procedure for DNA oligonucleotides; T4 DNA ligase was obtained from Enzymatics (Qiagen) and its activity was experimentally determined through test ligations on various DNA substrates. Chemical building blocks and reagents were sourced from a variety of suppliers. Barcoded tubes used to store DNA oligomers were read using a SampleScan 96 scanner (BiomicroLab) and decoded using Vortex software (Dotmatics). All buffers, including HEPES 10X ligation buffer (300 mM 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid, 100 mM MgCl₂, 100 mM dithiothreitol, 10 mM adenosine triphosphate, pH 7.8) and basic borate buffer (250 mM sodium borate/boric acid, pH 9.5), were prepared in-house. Various DNA working solutions were prepared using DNase free ultra-pure water (Invitrogen), HPLC-grade acetonitrile (Fisher) or high-purity absolute ethanol (Koptec). LC/MS running solvents were made from Optima LC/MS grade water (Fisher), Optima LC/MS grade methanol (Fisher), 99+% purity hexafluoroisopropanol (Sigma) and HPLC-grade triethylamine (Fisher). Solutions were generally transferred or pooled utilizing Biotix or Fisher brand pipette tips and reservoirs (various sizes), reactions were generally performed in polypropylene PCR plates or Eppendorf tubes. Plates were sealed for incubation with AlumaSeal II foil seals (Excel Scientific). Large volume DNA precipitations were performed in polypropylene 250 mL screw-cap bottles or 50 mL Falcon tubes. Heated reactions were either performed in ep384 Mastercyclers (Eppendorf), benchtop heat blocks (Fisher) or in laboratory ovens (Fisher). Solutions were centrifuged in either Avanti J-30I or Allegra X-15R centrifuges (Beckman-Coulter). Optical density measurements were made using a Biophotometer (Eppendorf).

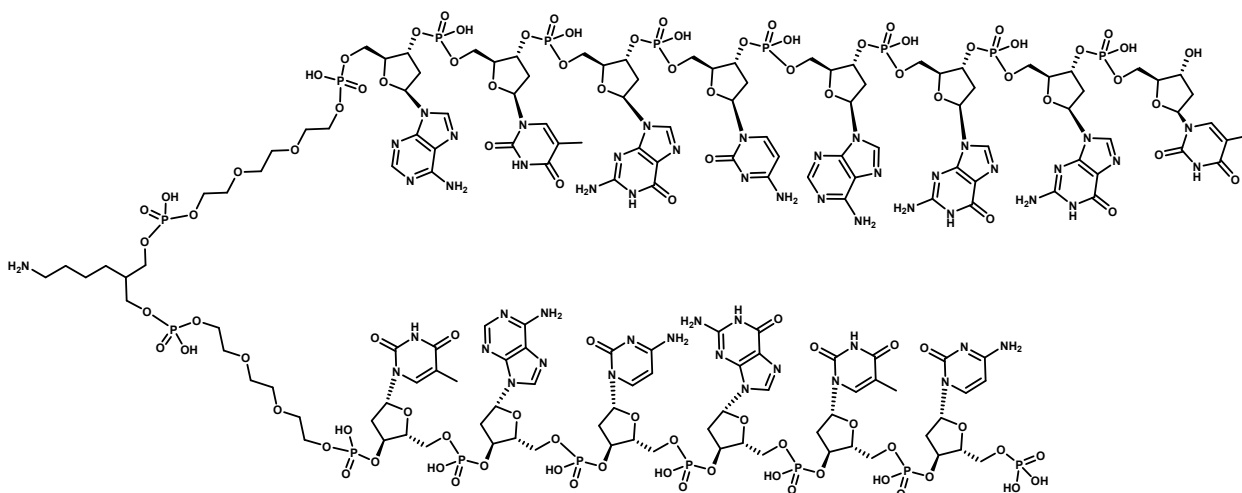


Figure S1. Structure of “DTSU” (5’-Phos-CTGCAT-Spacer 9-Amino C7-Spacer 9 ATGCAGGT 3’).

DNA Headpiece (S1) Synthesis



S1

To eight 250 mL, fluorinated ethylene propylene (FEP) centrifuge bottles each charged with DTSU (**Figure S1**) (1282 μL , 15 μmol , 1 equiv, 11.7 mM in H_2O), “FPU_upper” (1394 μL , 15.75 μmol , 1.05 equiv, 11.3 mM in H_2O , 11 bp DNA oligomer with 5’-Phos, ATTCACCTCAGG), “FPU_lower” (1302 μL , 15.75 μmol , 1.05 equiv, 12.1 mM in H_2O , 11 bp DNA oligomer with 5’-Phos, TGAGTGAATAC), water (41 mL), and 10X HEPES ligation buffer (4.5 mL) were added, followed by T4 DNA ligase (225 μL). The solns were mixed and incubated at 25 $^\circ\text{C}$ overnight. After completion of ligation was confirmed by LCMS by the general procedure, each soln was precipitated by the general procedure (limiting centrifugal speed to 4000 \times G due to FEP). After the pellets were each reconstituted in H_2O (10 mL), pH 9.5 borate buffer (20 mL, 5000 μmol , 333 equiv, 250 mM aq. stock) and CH_3CN (12 mL) were added to all bottles, followed by a soln of Fmoc-15-amino-4,7,10,13-tetraoxapentadecanoic acid (3 mL, 1200 μmol , 80 equiv, 400 mM stock in CH_3CN) and 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (“DMTMM”, 3 mL, 1200 μmol , 80 equiv, 400 mM fresh stock in H_2O). After mixing, the solns were incubated at 25 $^\circ\text{C}$ for 2h. After completion of acylation was confirmed by LCMS by the general procedure, each soln was precipitated by the general procedure (limiting centrifugal speed to 4000 \times G due to FEP). After the pellets were each reconstituted in H_2O (15 mL), an aq. piperidine soln (5 mL, 10% aq. piperidine v/v, fresh) was added and the solns were incubated at 25 $^\circ\text{C}$ for 4h. After completion of Fmoc deprotection was confirmed by LCMS by the general procedure, additional H_2O (20 mL) was added to each bottle and then all were precipitated by the general procedure (limiting centrifugal speed to 4000 \times G due to FEP). After the pellets were reconstituted in H_2O (5 mL) and combined (with additional washes), a stock of HP for further experiments was prepared (75.7 mL, 93.11 μmol , 78% yield, 1.23 mM aq. stock). This stock may be purified by HPLC or used directly in other reactions. In this case, it was purified before use.

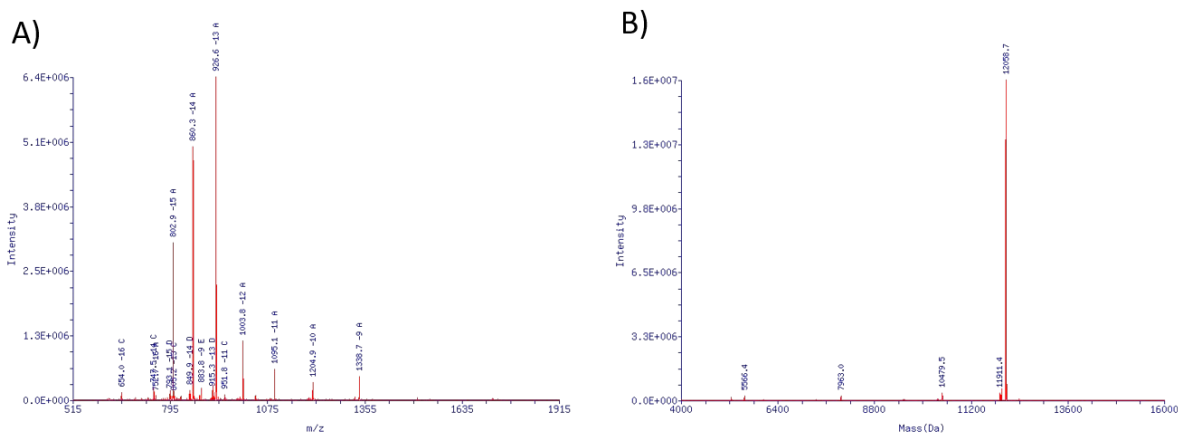


Figure S2. Representative deconvolution analysis of oligonucleotide MS data on the HP (“Headpiece”) S1. A) The crude MS data showing the various m/z ions observed in the 500–2000 mass region; B) The deconvoluted spectrum showing the parent ion mass (12059, the expected molecular weight of S1).

General analytical procedure for the analysis of DNA oligonucleotide compositions.

A Vanquish UHPLC system was integrated with LTQ XL ion trap mass spectrometer (ThermoFisher Scientific) for LC/MS analysis of oligonucleotides. Injection amounts were typically 5–10 μ L containing 50–200 pmol DNA analyte.

LC/MS Parameters for Thermo Vanquish UHPLC with LTQ Ion Trap MS Instrument

(i) LC settings

Column: Thermo DNAPac RP (2.1 x 50 mm, 4 μ m)

Solvent A: 15mM triethylamine (TEA)/100mM hexafluoroisopropanol (HFIP) in water

Solvent B: 15mM TEA/100mM HFIP in 50% methanol

Solvent C: Methanol

Flow rate: 0.65 mL/min

Run time: 2 mins (gradient)

Column temperature: 100 °C (post column cooler at 40 °C)

(ii) MS settings

Source: ESI in negative mode

Spray voltage: 4100 V

Source heater temperature: 390 °C

Sheath Gas: 28 (instrument units)

Auxiliary Gas: 8 (instrument units)

Sweep Gas: 2 (instrument units)

Capillary temperature: 350 °C

Capillary voltage: -33.0 V

Tube lens: -92.0 V

MS Scan: 500 – 2000 *m/z*

Samples were analyzed on a Thermo Vanquish UHPLC system coupled to an electrospray LTQ ion trap mass spectrometer. An ion-pairing mobile phase comprising of 15mM TEA/100mM HFIP in a water/methanol solvent system was used in conjunction with an oligonucleotide column Thermo DNAPac RP (2.1 x 50 mm, 4 μ m) for all the separations. All mass spectra were acquired in the full scan negative-ion mode over the mass range 500–2000*m/z*. The data analysis was performed by exporting the raw instrument data (.RAW) to an automated biomolecule deconvolution and reporting software (ProMass) which uses a novel algorithm known as ZNova to produce artifact-free mass spectra. The following deconvolution parameters were applied: peak width 3.0, merge width 0.2, minimum and normalize scores of 2.0 and 1.0 respectively. The noise threshold was set at S/N 2.0. The processed data was directly exported to Microsoft Excel worksheets for further data comparisons. A sample MS analysis using ProMass software is presented in Figure S2.

General DNA Precipitation Procedure (ethanol precipitation).

Based on the theoretical solution volume *n* (ignoring any loss from heating, etc.), *n*/20–*n*/10 volume of a 5 M NaCl stock solution was added and the solution was mixed. Then absolute ethanol (3*n* volume, 75% *v/v* final ethanol concentration) was added, the solution was thoroughly mixed, and then stored at -20 °C overnight to precipitate the DNA. The resulting slurry was centrifuged (10,000 \times G for 1 h), the supernatant decanted, an addition *n*–2*n* volume of chilled 75% ethanol (*v/v*) was added, and the pellet was centrifuged

again ($10,000 \times G$ for 30 min). After decantation of the supernatant, the pellet was dried (in open air or under gentle vacuum) and reconstituted in neutral water or buffer (to a concentration of ~ 1 mM; assessed by optical density measurements). The solution was then centrifuged ($10,000 \times G$ for 10 min) to pellet any left-over solids (unremoved chemical building blocks or byproducts, denatured ligase, etc.), and the solution was transferred to leave these solids behind. The DNA may undergo a second round of precipitation if the purity is insufficient (as assessed by the general analytical procedure). In addition, if the initial solution contains high amounts of organic co-solvent or chaotropic reagents (e.g., piperidine), the solution may be diluted with neutral water to enhance the overall precipitation yield. Typically, precipitations were conducted in polypropylene 96-well plates or polypropylene bottles which can withstand high centrifugal speeds. However, polypropylene is incompatible with piperidine—reactions with this reagent were run in fluorinated ethylene propylene (FEP) bottles and spun with a maximum speed of $4,000 \times G$.

*In the case of piperidine deprotection procedures, the reaction mixture was first diluted 4X in order to limit the solubilizing effects of piperidine, which would reduce the DNA recovery. Sodium chloride was then added and the rest of the procedure was carried out as described in the previous paragraph.

^Δ In the case of Alloc removal, the quenched reaction was diluted 10X before the addition of sodium chloride and the rest of the precipitation procedure.

General procedure for the ligation of DNA oligonucleotides.

To a ~ 1 mM solution of the HP-containing library intermediate (1 equiv), a premixed solution of the pre-duplexed oligonucleotide (“codon”) with the appropriate 2-bp overhang was added (1 mM stock soln in neutral water, 1.05-1.1 equiv). Separately, a master mix consisting of additional water, HEPES 10X ligation buffer, and T4 DNA ligase was prepared and added to the wells or container with mixing and incubated at room temperature overnight. The concentration of the HP-containing library intermediate in the final solution was 0.24 mM (thus the amount of HEPES 10X ligation buffer was $1/10^{\text{th}}$ of this final volume). The amount of T4 DNA ligase stock added depended on the assayed activity of the ligase batch—however we routinely included 100–200X (i.e., full ligation observed with the addition of ligase stock $1/100^{\text{th}}$ – $1/200^{\text{th}}$ overall volume). After the overnight incubation, the ligation progress was assessed by LC/MS with the general analytical procedure (due to the large MW increase, the ligation is obvious even on complex post-pool samples) as well by gel electrophoresis. If incomplete, additional buffer, ligase or codon may be added. Typically, ligation samples were run on a denaturing 6% TBE-Urea gel (Invitrogen), in TBE buffer at 150–180 V for 30–40 min. Gels were stained with ethidium bromide, visualized with a Gel Doc (Bio rad) or equivalent imager, and assessed for transformation into a new, higher-MW band. A typical gel result is shown in Figure S3.

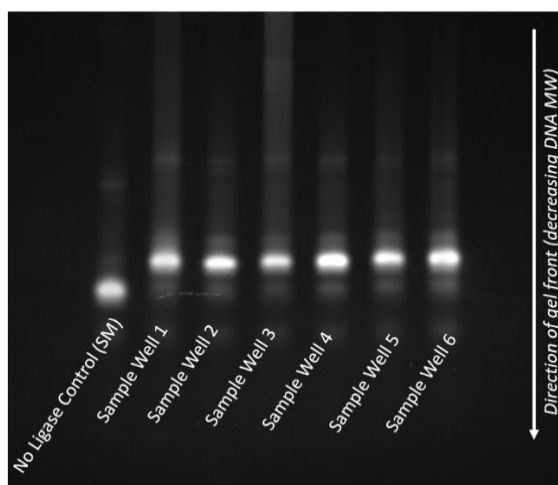


Figure S3. Representative 6% TBE-Urea gel for the analysis of ligation of DNA codons. As shown here, the disappearance of the no ligase control well's band (the starting material) to a higher-MW band signifies a finished codon ligation.

General HPLC Purification Procedure

All the RCM substrates were purified using Agilent 1100 series HPLC system consisting of an autosampler, degasser, quaternary pump and a diode array detector coupled to an analytical scale fraction collector. The mobile phase system comprising of triethylammonium acetate (TEAA) was prepared by titrating glacial acetic acid with triethylamine (TEA) in water. For example, 1L of 0.1 M TEAA was prepared by adding 5.6 ml of glacial acetic acid into 950 ml of water and slowly adding 13.86 ml of TEA. The final pH was adjusted to 7.0 by careful addition of acetic acid, and the final volume adjusted to 1L. HPLC separations were performed under gradient conditions at a flow rate of 1.0 ml/min using Waters WBridge C18 column (2.1 x 50 mm, 3.5 μ m). A linear gradient of 5 to 95% B was used with a run time of 15 minutes to collect the fractions. The collected fractions were analyzed by the general procedure, combined, and later purified by ethanol precipitation.

HPLC system: Agilent 1100 series

Column: Waters XBridge C18 (2.1 x 50 mm, 3.5 μ m)

Solvent A: 0.1 M TEAA (pH 7.0)

Solvent B: Acetonitrile / 0.1M TEAA, 40/60, v/v

Flow rate: 1.0 mL/min

Run time: 15 mins (gradient)

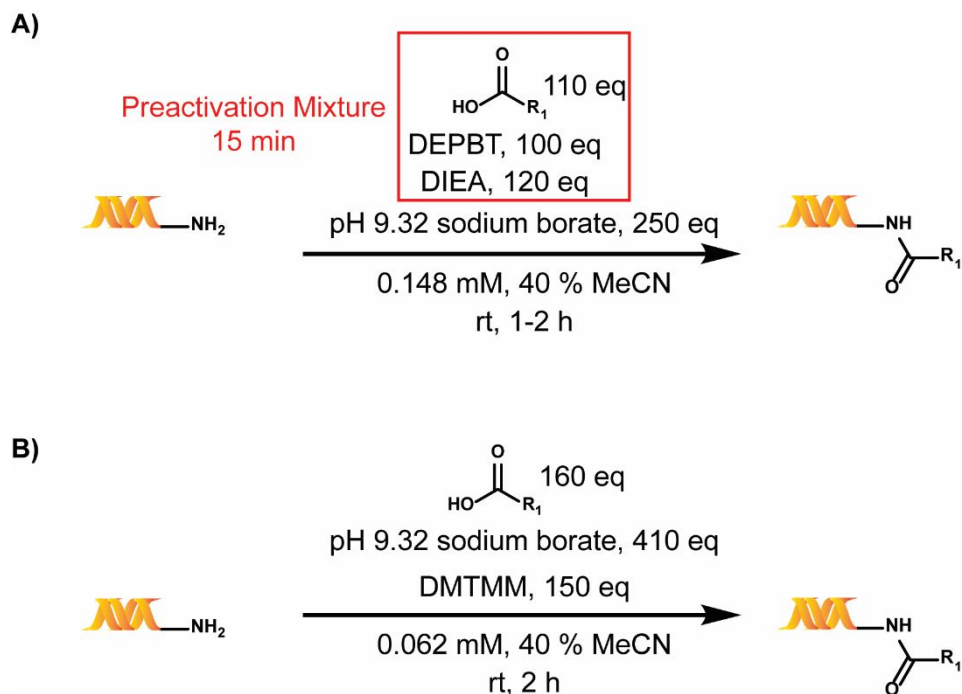
Gradient: 5 to 95% B in 15 minutes

Column temperature: 60 °C

Detection: 260nm UV

General Reaction Conditions for the preparation of substrates

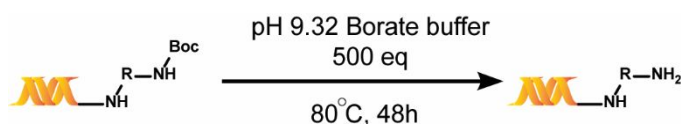
General Acylation Reaction Conditions



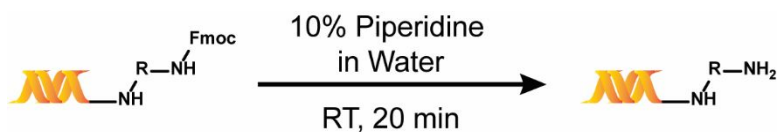
In a few cases, the carboxylic acid was insoluble in acetonitrile (MeCN) and was dissolved in DMSO instead. In such cases, the organic portion of the solvent mixture consisted of MeCN:DMSO 3:1.

In a few cases, the DMTMM coupling did not go to completion after 2 h. In such cases, a supplemental 150 eq of DMTMM were added at the 2h time point and the reaction was allowed to run for another hour. This was then quenched through DNA precipitation. This procedure generally increased the conversion to 80% and above.

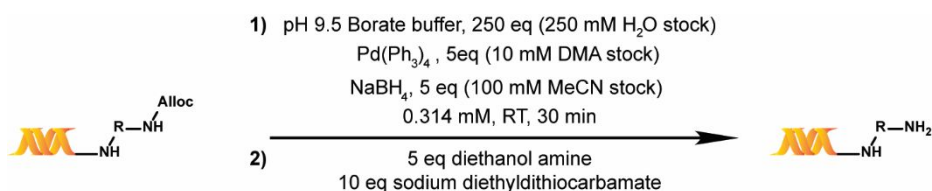
General N-Boc Removal Reaction Conditions



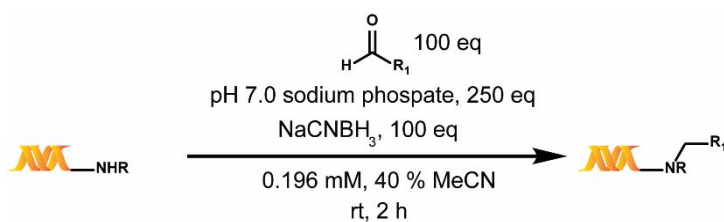
General *N*-Fmoc Removal Reaction Conditions



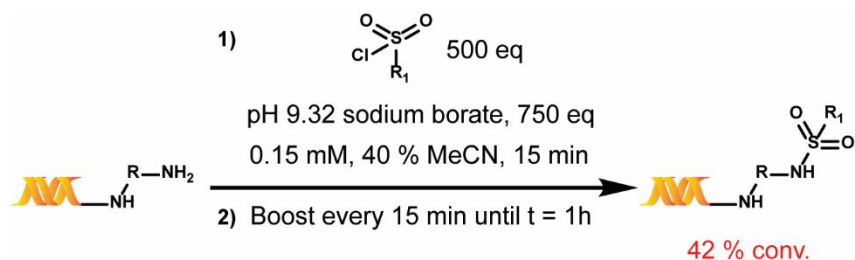
General *N*-Alloc Removal Reaction Conditions



General Reductive Amination Reaction Conditions



General Sulfonamide Formation Reaction Conditions

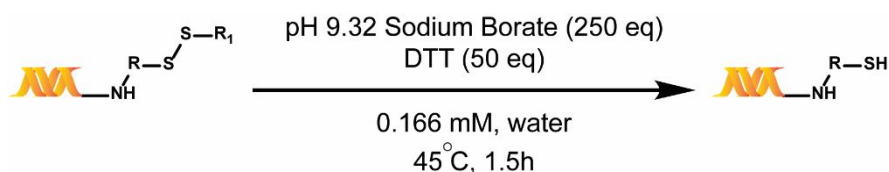


Boost procedure: 500 eq of sulfonyl chloride building block (in MeCN), 750 eq borate buffer and water to maintain the solvent composition at 40% (v/v) MeCN.

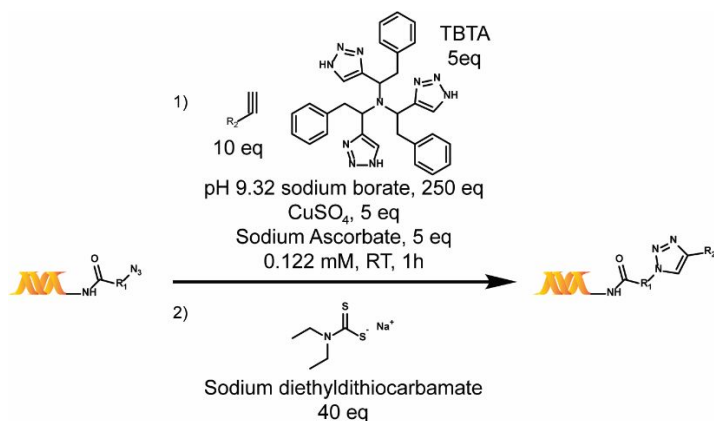
The percent conversion was determined by LC/MS.

These are not optimized reaction conditions. They were sufficient to provide enough purified material for carrying out the RCM reaction. It is to be noted that stock solutions of the sulfonyl chloride appear to undergo hydrolysis (most likely from the water present in our MeCN solvent containers). Stock solutions should therefore be made right before the reaction is run. Moreover, the sulfonyl chloride is most certainly hydrolyzed once added to the aqueous reaction mixture, which would explain the need for a high equivalence. We tried 1000 eq and it made no difference in conversion. Multiple additions did prove to be useful (15.6% conversion with only one addition for 1h of reaction time).

General Disulfide Removal Reaction Conditions

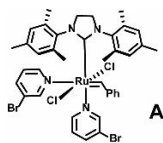


General Conditions for the Cu(I)-Catalyzed Azide–Alkyne Cycloaddition (CuAAC)



Syntheses

Synthesis of Catalyst A



Grubbs 2nd Generation catalyst (Sigma-Aldrich, 30 mg, 0.034 mmol) was vortexed with 3-bromopyridine (200 μ L, 2.08 mmol) in a 1.5 mL Eppendorf tube and allowed to incubate for 10 min. The reaction mixture was then spun at 12,000 x G for 10 min. The supernatant was transferred to another tube, leaving a small amount of black residue behind (impurities in commercial Grubbs G2). Hexanes (1 mL) was added to the reaction mixture to precipitate the Grubbs III catalyst. This was vortexed thoroughly and spun for 2 min at 20,000 x G to precipitate the catalyst and then the supernatant was removed and discarded. This hexanes wash step was repeated four times to remove the excess 3-bromopyridine. After the residue was dried under a brief stream of nitrogen it was placed under high vacuum for 2h to provide a bright green solid. This catalyst slowly decomposes upon storage (at 4 $^{\circ}$ C), so catalyst batches should be used within a week for optimal results. Commercial sources of Grubbs III should not be used.

*Note that this procedure was carried out in open air.

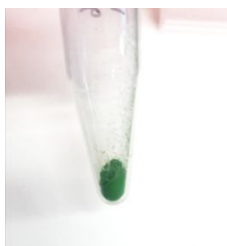
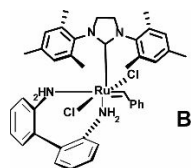


Figure S 4. GIII post synthesis; picture taken by Olivier Monty during the investigation.

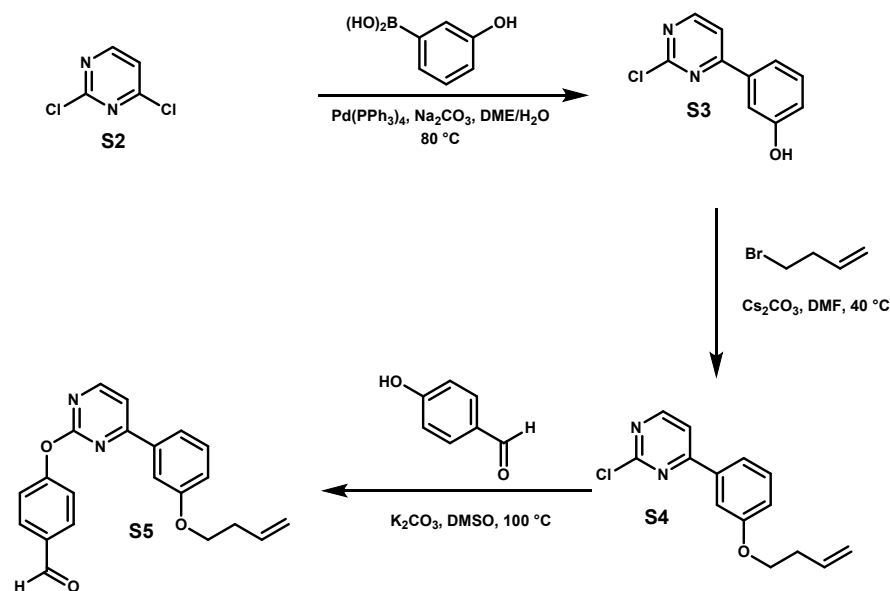
Synthesis of Catalyst B



Catalyst B may be prepared through use of a glove box from Grubbs III type catalysts and 2,2'-bipyridyl diamine.⁵ However, the catalyst stock used for the work that fueled this communication was graciously provided by Dr. Deryn E. Fogg's research group.

Synthesis of Building Blocks

S5



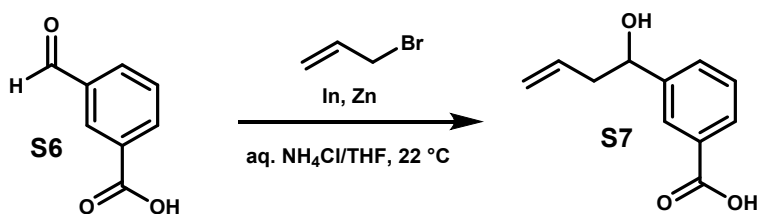
Pyrimidine S3. Preparation of **S3** was previously reported by William et al.⁶ To flask charged with 2,4-dichloropyrimidine, **S2** (0.075 g, 0.503 mmol, 1 equiv), Pd(PPh₃)₄ (0.029 g, 0.0251 mmol, 0.05 equiv), Na₂CO₃ (0.080 g, 0.750 mmol, 1.5 equiv), and (3-hydroxyphenyl)boronic acid (0.083 g, 0.604 mmol, 1.2 equiv), 1,2-dimethoxyethane (1.2 mL) and H₂O (120 μL) were added, and the resultant slurry was heated at 80 °C overnight. Upon cooling, the reaction was quenched with sat. aq. NH₄Cl soln. (1 mL), diluted with H₂O (5 mL) and CH₂Cl₂ (5 mL) and the aqueous layer extracted with CH₂Cl₂ (3 × 5 mL). The combined organics were then dried (MgSO₄), filtered and concentrated. Purification by flash column chromatography (silica, 9:1 → 7:3 hexanes:EtOAc) provided **S3** (0.027 g, 0.131 mmol, 26% yield) as a solid, which was consistent with the reported characterization data. Significant formation of a side product from phenolic arylation of **S3** with dichloropyrimidine was also observed. **S3**: R_f = 0.10 (silica, 8:2 hexanes:EtOAc); ¹H NMR (800 MHz, CDCl₃) δ = 8.65 (d, *J* = 5.25 Hz, 1H), 7.64–7.62 (m, 2H), 7.60 (d, *J* = 7.66 Hz, 1H), 7.39 (t, *J* = 7.89 Hz, 1H), 7.04 (dd, *J* = 8.36 Hz, 2.10, 1H).

S4. Preparation of **S4** was previously reported by William et al.⁶ To a soln of **S3** (0.027 g, 0.131 mmol, 1 equiv) in DMF (0.5 mL) was added Cs₂CO₃ (0.192 g, 0.590 mmol, 4.5 equiv) and 4-bromo-1-butene (80 μL, 0.784 mmol, 6 equiv), and the soln was heated to 40 °C overnight. Upon cooling, the soln was diluted with H₂O (10 mL) and EtOAc (5 mL), and the organic layer washed with H₂O (3 × 10 mL) and brine (5 mL) and then dried (MgSO₄), filtered and concentrated. Purification by flash column chromatography (silica, 9.5:0.5 → 8:2 hexanes:EtOAc) provided **S4** (0.022 g, 0.084 mmol, 64% yield) as a white solid that matched the reported characterization data. **S4**: R_f = 0.15 (silica, 9:1 hexanes:EtOAc); ¹H NMR (800 MHz, CDCl₃) δ = 8.63 (d, *J* = 5.24 Hz, 1H), 7.66 (t, *J* = 2.04 Hz, 1H), 7.63 (d, *J* = 5.23 Hz, 1H), 7.62 (d, *J* = 7.79 Hz, 1H), 7.41 (t, *J* = 7.90 Hz, 1H), 7.08 (dd, *J* = 8.19, 2.42 Hz, 1H), 5.96–5.89 (m, 1H), 5.20 (dd, *J* = 17.23, 1.38 Hz, 1H), 5.13 (dd, *J* = 10.35, 1.38 Hz, 1H), 4.11 (t, *J* = 6.59 Hz, 2H), 2.60–2.57 (m, 2H) ppm; ¹³C NMR (200 MHz, CDCl₃) δ = 167.10, 161.84, 159.94, 159.63, 136.47, 134.39, 130.26, 119.84, 118.35, 117.42, 115.47, 113.29, 67.49, 33.73 ppm.

S5: To a flask charged with **S4** (0.055 g, 0.211 mmol, 1 equiv), 4-hydroxybenzaldehyde (0.051 g, 0.422 mmol, 2 equiv) and K₂CO₃ (0.058 g, 0.422 mmol, 2 equiv), DMSO (1 mL) was added and the soln was

heated to 100 °C for 2 h. The soln was then poured into a mixture of H₂O (20 mL), brine (5 mL) and EtOAc (10 mL), and organic layer washed with H₂O (2 × 20 mL) and brine (5 mL). The organics were then dried (MgSO₄), filtered and concentrated. Purification by flash column chromatography (silica, 9:1 → 0:10 hexanes:CH₂Cl₂) provided **S5** (0.020 g, 0.058 mmol, 27% yield). **S5**: R_f = 0.24 (silica, CH₂Cl₂); IR (film): ν_{\max} = 1696, 1575, 1545, 1370, 1273, 1221, 1156, 834, 781 cm⁻¹; ¹H NMR (800 MHz, CDCl₃) δ = 10.03 (s, 1H), 8.59 (d, *J* = 4.99 Hz, 1H), 7.98 (d, *J* = 8.36 Hz, 1H), 7.61–7.58 (m, 2H), 7.50 (d, *J* = 5.19 Hz, 1H), 7.44 (d, *J* = 8.36 Hz, 1H), 7.38 (t, *J* = 7.85 Hz, 1H), 7.06 (dd, *J* = 8.24, 2.05 Hz, 1H), 5.93–5.87 (m, 1H), 5.18 (d, *J* = 16.8 Hz, 1H), 5.13 (d, *J* = 10.27 Hz, 1H), 4.05 (t, *J* = 6.68 Hz, 2H), 2.56 (q, *J* = 6.62 Hz, 2H) ppm; ¹³C NMR (200 MHz, CDCl₃) δ = 191.11, 167.14, 164.94, 160.25, 159.64, 158.13, 137.12, 134.35, 133.63, 131.54, 130.18, 122.43, 119.79, 118.39, 117.40, 113.15, 112.64, 67.54, 33.72 ppm; HRMS: calcd for C₂₁H₁₉N₂O₃ [M+H]⁺: 347.1390, found 347.1381.

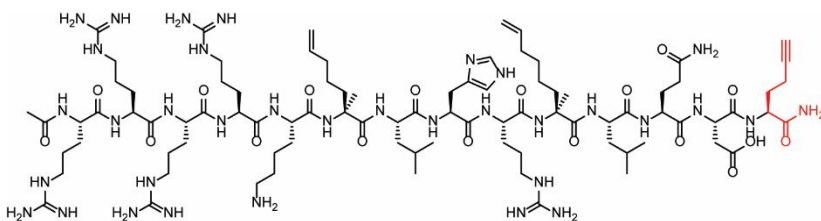
S7



S7. To a vigorously stirred slurry of 3-formylbenzoic acid, **S6** (0.100 g, .667 mmol, 1 equiv), indium (0.076 g, 0.667 mmol, 1 equiv) and zinc (0.080 g, 1.23 mmol, 1.84 equiv) in a mixture of THF (2 mL) and aq. NH₄Cl (4M, 200 μL), allyl bromide (115 μL, 1.33 mmol, 2 equiv) was added slowly at room temperature. After 1 h, the soln was poured into a mixture of EtOAc (5 mL), aq. HCl (2M, 2 mL), H₂O (10 mL) and brine (10 mL), and extracted with EtOAc (2 × 10 mL). The combined organics were dried (MgSO₄), filtered and concentrated. Purification by flash column chromatography (silica, 9:1 → 6:4 hexanes:EtOAc) provided **S7** (0.112 g, 0.583 mmol, 87% yield). **S7**: R_f = 0.3 (silica, 1:1 hexanes:EtOAc); IR (film): ν_{\max} = 1690, 1271, 1193, 695 cm⁻¹; ¹H NMR (800 MHz, CDCl₃) δ = 8.11 (s, 1H), 8.03 (d, *J* = 7.60 Hz, 1H), 7.64 (d, *J* = 7.60 Hz, 1H), 7.48 (t, *J* = 7.60 Hz, 1H), 5.85–5.78 (m, 1H), 5.22–5.16 (m, 2H), 4.86–4.81 (m, 1H), 2.60–2.55 (m, 1H), 2.54–2.49 (m, 1H) ppm; ¹³C NMR (200 MHz, CDCl₃) δ = 171.39, 144.54, 133.99, 131.41, 129.53, 129.49, 128.83, 127.77, 119.27, 72.83, 44.05 ppm; HRMS: calcd for C₁₁H₁₁O₃ [M-H]⁻: 191.0714, found 191.0692.

X

Peptide X was ordered from New England Peptide. The QC information is provided in the Characterization section below.

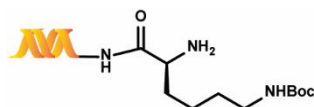


X

Synthesis of On-DNA Scaffolds

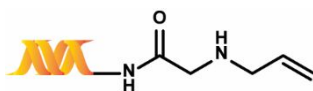
Scaffold 1

DNA headpiece **S1** (2 μmol , 1.03 mM stock solution) was acylated (General Acylation Reaction Conditions B) using Fmoc-Lys(Boc)-OH (CAS# 71989-26-9). LC/MS analysis revealed 90% conversion after 2h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Fmoc Removal Reaction Conditions and was precipitated using the piperidine-specific precipitation procedure (General DNA Precipitation Procedure*). The synthesized scaffold is shown below.



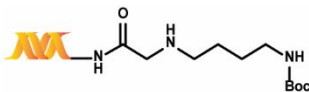
Scaffold 2

DNA headpiece **S1** (2 μmol , 1.03 mM stock solution) was acylated (General Acylation Reaction Conditions A) using Fmoc-N-(allyl)-glycine (CAS# 222725-35-1). LC/MS analysis revealed ~100% conversion after 1h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Fmoc Removal Reaction Conditions and was precipitated using the piperidine-specific precipitation procedure (General DNA Precipitation Procedure*). The synthesized scaffold is shown below.



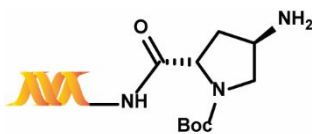
Scaffold 3

DNA headpiece **S1** (800 nmol, 1.03 mM stock solution) was acylated (General Acylation Reaction Conditions A) using Fmoc-N-(4-Boc-aminobutyl)-Gly-OH (CAS# 171856-09-0). LC/MS analysis revealed ~100% conversion after 1h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Fmoc Removal Reaction Conditions and was precipitated using the piperidine-specific precipitation procedure (General DNA Precipitation Procedure*). The synthesized scaffold is shown below.



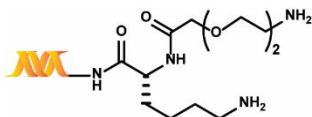
Scaffold 4

DNA headpiece **S1** (200 nmol, 1.03 mM stock solution) was acylated (General Acylation Reaction Conditions A) using N-Boc-*trans*-4-N-Fmoc-amino-L-proline (CAS# 176486-63-8). LC/MS analysis revealed ~100% conversion after 1h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Fmoc Removal Reaction Conditions and was precipitated using the piperidine-specific precipitation procedure (General DNA Precipitation Procedure*). The synthesized scaffold is shown below.



Scaffold 5

DNA headpiece **S1** (2 μmol , 1.03 mM stock solution) was acylated (General Acylation Reaction Conditions B) using Fmoc-D-Lys(Alloc)-OH (CAS# 71989-26-9). LC/MS analysis revealed 90% conversion after 2h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Fmoc Removal Reaction Conditions and was precipitated using the piperidine-specific precipitation procedure (General DNA Precipitation Procedure*). The recovered material was acylated (General Acylation Reaction Conditions B) using {2-[2-(Fmoc-amino)ethoxy]ethoxy}acetic acid (CAS# 166108-71-0). After 2h, a boost of acid/DMTMM (160eq/150 eq) was given. The reaction was allowed to run for another hour. LC/MS analysis revealed ~100% conversion. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Alloc Removal Reaction Conditions. LC/MS analysis revealed completion of the reaction within 30 min. The material was precipitated using the N-Alloc-specific precipitation procedure (General DNA Precipitation Procedure^A). The synthesized scaffold is shown below.



Synthesis of On-DNA Substrates

It is to be noted that all substrates were purified by HPLC before the RCM reaction was carried out. Impure material may exhibit low to no conversion under the developed reaction conditions. It is therefore likely that DNA-encoded libraries should be purified by HPLC before being subjected to RCM reaction.

Substrate 1a

Scaffold 1 (300 nmol, 1.32 mM stock) was subjected to the General Boc Removal Procedure. The material was precipitated using the General DNA Precipitation Procedure. The recovered material was acylated (General Acylation Reaction Conditions A) using 5-Hexenoic acid (CAS# 1577-22-6). LC/MS analysis revealed complete conversion after 1h. This was precipitated using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified material is provided below.

Substrate 2a

Scaffold 2 (40 nmol, 1.58 mM stock) was acylated (General Acylation Reaction Conditions A) using 5-Hexenoic acid (CAS# 1577-22-6). LC/MS analysis revealed complete conversion after 1h. This was precipitated using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified material is provided below.

Substrate 3a

Scaffold 5 (40 nmol, 0.92 mM stock) was acylated (General Acylation Reaction Conditions A) using 5-Hexenoic acid (CAS# 1577-22-6). LC/MS analysis revealed complete conversion after 1h. This was

precipitated using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified material is provided below.

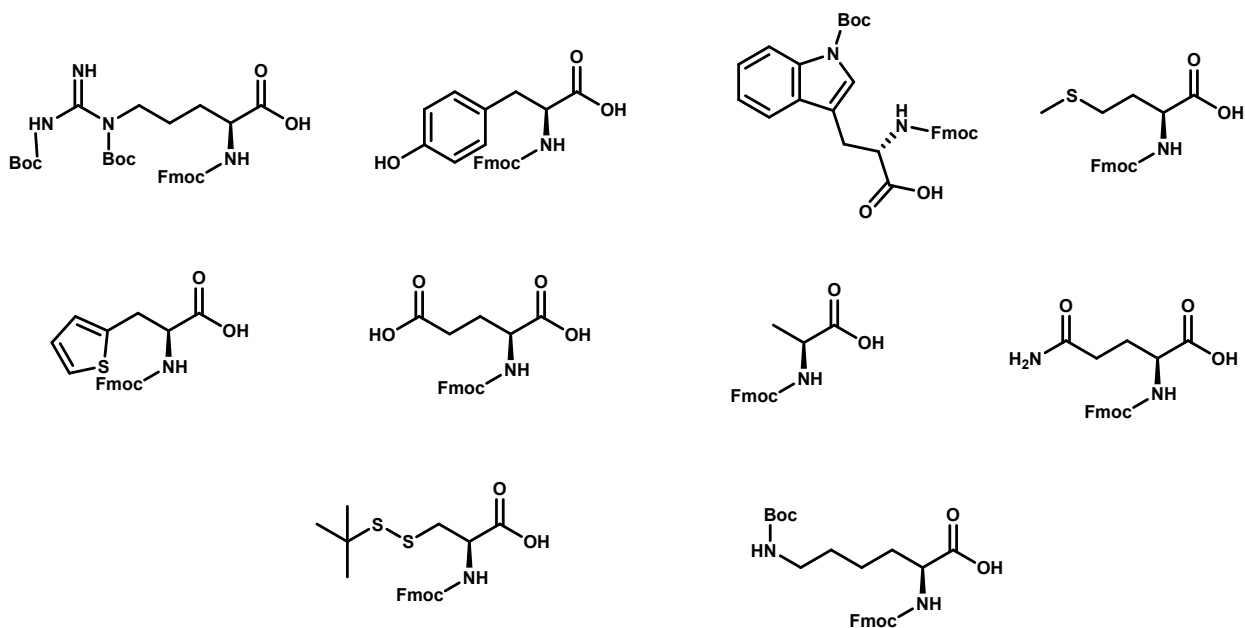
Substrates 4a–13a

For all substrates, Scaffold 2 was acylated (General Acylation Reaction Conditions B) using the corresponding amino acids (shown below).

[In the case of Gln, Met and Tyr, a DMTMM boost (See General Acylation Reaction Conditions) was given to drive the reaction forward. In the case of Gln, DMSO was used to make a stock solution of the amino acid. The reaction therefore contained 40% DMSO/MeCN and 60% water. In the case of Arg, the General Acylation Reaction Conditions B gave the best conversion, although of only 1%. A large amount of Scaffold 2 was therefore used in order to recover enough for the steps that followed. We have since then developed new reaction conditions for the generation of Arg on DNA. The latter work will soon be reported in another publication and therefore cannot be described here.]

Each reaction was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Fmoc Removal Reaction Conditions. This was precipitated using the piperidine-specific precipitation procedure (General DNA Precipitation Procedure*). The recovered material was acylated (General Acylation Reaction Conditions A) using 5-hexenoic acid (CAS# 1577-22-6). LC/MS analysis revealed complete conversion after 1h. Each reaction was quenched using the General DNA Precipitation Procedure. Substrates bearing a protecting group (Cys(StBu), Arg(DiBoc), Lys(Boc) and Trp (Boc)) were subjected to the Disulfide and Boc Removal Reaction Conditions. The recovered material was precipitated using the General DNA Precipitation Procedure. All substrates were finally purified using the General HPLC Purification Procedure. The LC/MS traces of the purified substrates are provided below.

Please note that disulfide deprotection led to dimerization of the released substrates as they formed intermolecular disulfides. The protecting group was therefore kept on the Cys(StBu) substrate. The RCM results are reported in Table 3.



Substrate 14a

Scaffold 1 (60 nmol, 1.32 mM stock) was acylated (General Acylation Reaction Conditions A) using 5-Hexenoic acid (CAS# 1577-22-6). LC/MS analysis revealed complete conversion after 1h. This was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Boc Removal Procedure. This was precipitated using the General DNA Precipitation Procedure. The recovered material was acylated (General Acylation Reaction Conditions A) using 3-methylpent-4-enoic acid (CAS# 1879-03-4). LC/MS analysis revealed complete conversion after 2h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified material is provided below.

Substrate 15a

Scaffold 1 (60 nmol, 1.32 mM stock) was acylated (General Acylation Reaction Conditions A) using 5-Hexenoic acid (CAS# 1577-22-6). LC/MS analysis revealed complete conversion after 1h. This was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Boc Removal Procedure. This was precipitated using the General DNA Precipitation Procedure. The recovered material was acylated (General Acylation Reaction Conditions A) using 4-methylpent-4-enoic acid (CAS# 1001-75-8). LC/MS analysis revealed complete conversion after 2h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified material is provided below.

Substrate 16a

Scaffold 3 (150 nmol, 1.13 mM stock) was subjected to the General Reductive Amination Reaction Conditions using 5-hexenal (CAS# 764-59-0). LC/MS analysis revealed complete conversion after 2h. The material was precipitated using the General DNA Precipitation Procedure. The recovered material was acylated (General Acylation Reaction Conditions A) using 6-heptenoic acid (CAS# 1119-60-4). LC/MS analysis revealed complete conversion after 1h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified material is provided below.

Substrate 17a

Scaffold 3 (150 nmol, 1.13 mM stock) was subjected to the General Reductive Amination Reaction Conditions using 5-hexenal (CAS# 764-59-0). LC/MS analysis revealed complete conversion after 2h. The material was precipitated using the General DNA Precipitation Procedure. The recovered material was acylated (General Acylation Reaction Conditions A) using 3-(allyloxy)propanoic acid (CAS# 22577-15-7). LC/MS analysis revealed complete conversion after 1h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified material is provided below.

Substrate 18a

Scaffold 1 (150 nmol, 1.32 mM stock) was acylated (General Acylation Reaction Conditions B) using Fmoc-(S)-3-amino-3-(3-pyridyl)propionic acid (CAS# 507472-06-2). LC/MS analysis revealed complete conversion after 1 h 45 min. This was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Fmoc Removal Reaction Conditions and was precipitated using the piperidine-specific precipitation procedure (General DNA Precipitation Procedure*). The recovered material was acylated (General Acylation Reaction Conditions A) using 3-butenic acid (CAS# 625-38-7). LC/MS analysis revealed complete conversion after 2 h 45 min. This was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Boc Removal Reaction Conditions. The material was precipitated using the General DNA Precipitation Procedure. The recovered material was acylated (General Acylation Reaction Conditions B) using 5-hexenoic acid (CAS#

1577-22-6). This was quenched using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified material is provided below.

Substrate 19a

Scaffold 4 (160 nmol, 1.03 mM) was subjected to the General Sulfonamide Formation Reaction Conditions. LC/MS analysis revealed a low conversion of 42%. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Reductive Amination Reaction Conditions using 5-hexenal (CAS# 764-59-0). LC/MS analysis revealed complete conversion after 1h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified material is provided below.

Substrate 20a

DNA headpiece (200 nmol, 1.03 mM stock solution) was acylated (General Acylation Reaction Conditions A) using N-Boc-*cis*-4-N-Fmoc-amino-L-proline (CAS# 174148-03-9). LC/MS analysis revealed ~100% conversion after 1h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Fmoc Removal Reaction Conditions and was precipitated using the piperidine-specific precipitation procedure (General DNA Precipitation Procedure*). This was acylated (General Acylation Reaction Conditions A) using 4-Pentenoic acid (CAS# 591-80-0). LC/MS analysis revealed complete conversion after 1h. This was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Boc Removal Procedure. This was precipitated using the General DNA Precipitation Procedure. The recovered material was acylated (General Acylation Reaction Conditions B) using **S7**. LC/MS analysis revealed complete conversion after 2h. This was quenched using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified material is provided below.

Substrate 21a

DNA headpiece (200 nmol, 1.03 mM stock solution) was acylated (General Acylation Reaction Conditions A) using N-Boc-*cis*-4-N-Fmoc-amino-L-proline (CAS# 174148-03-9). LC/MS analysis revealed ~100% conversion after 1h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Fmoc Removal Reaction Conditions and was precipitated using the piperidine-specific precipitation procedure (General DNA Precipitation Procedure*). This was acylated (General Acylation Reaction Conditions A) using 4-Pentenoic acid (CAS# 591-80-0). LC/MS analysis revealed complete conversion after 1h. This was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Boc Removal Procedure. This was precipitated using the General DNA Precipitation Procedure. This was subjected to the General Reductive Amination Reaction Conditions using **S5**. LC/MS analysis revealed complete conversion after 2h. The reaction was precipitated using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified material is provided below.

Substrate 22a

DNA headpiece (200 nmol, 1.03 mM stock solution) was acylated (General Acylation Reaction Conditions A) using Boc-4-(Fmoc-aminomethyl)-D-phenylalanine (CAS# 215302-77-5). LC/MS analysis revealed ~100% conversion after 1h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Fmoc Removal Reaction Conditions and was precipitated using the piperidine-specific precipitation procedure (General DNA Precipitation Procedure*). This was acylated (General Acylation Reaction Conditions A) using 4-Pentenoic acid (CAS#

591-80-0). LC/MS analysis revealed complete conversion after 1h. This was quenched using the General DNA Precipitation Procedure. The recovered material was subjected to the General Boc Removal Procedure. This was precipitated using the General DNA Precipitation Procedure. The recovered material was acylated (General Acylation Reaction Conditions B) using **S7**. LC/MS analysis revealed complete conversion after 2h. This was quenched using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified material is provided below.

Substrate 23a

DNA headpiece (100 nmol, 1.03 mM stock solution) was acylated (General Acylation Reaction Conditions A) using 2-(allylthio)acetic acid (CAS# 20600-63-9). LC/MS analysis revealed complete conversion after 1h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified substrate is provided below.

Synthesis of 24a

Substrate **1a**, two 39-bp DNA oligomers with 5'-Phos and duplexed with 2-bp 3' overhangs were ligated and precipitated by the general procedures. The design of the two oligomers mapped onto the codon 1-3 DNA regions used in the Single-Substrate library (**Figure S6**). The ligated substrate was purified via HPLC and was subjected to the RCM reaction, as described in the main text.

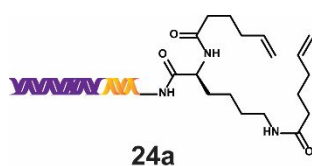


Figure S5. Substrate **24a** after ligation of substrate **1a** to 39 bp DNA oligomers

Synthesis of 25a

DNA headpiece (50 nmol, 1.03 mM stock solution) was acylated (General Acylation Reaction Conditions B) using 2-azidoacetic acid (CAS# 18523-48-3). LC/MS analysis revealed ~90% conversion after 2h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was then clicked with peptide X using the General CuAAC conditions. LC/MS analysis revealed 84% conversion after 1h. This was quenched using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified material is provided below.

Synthesis of 26a

DNA headpiece (50 nmol, 1.03 mM stock solution) was acylated (General Acylation Reaction Conditions B) using 15-Azido-4, 7, 10, 13-tetraoxapentadecanoic acid (CAS# 1257063-35-6). LC/MS analysis revealed ~90% conversion after 2h. The reaction was quenched using the General DNA Precipitation Procedure. The recovered material was then clicked with peptide X (General CuAAC conditions). LC/MS analysis revealed 84% conversion after 1h. This was quenched using the General DNA Precipitation Procedure. The recovered material was purified using the General HPLC Purification Procedure. The LC/MS trace of the purified material is provided below.

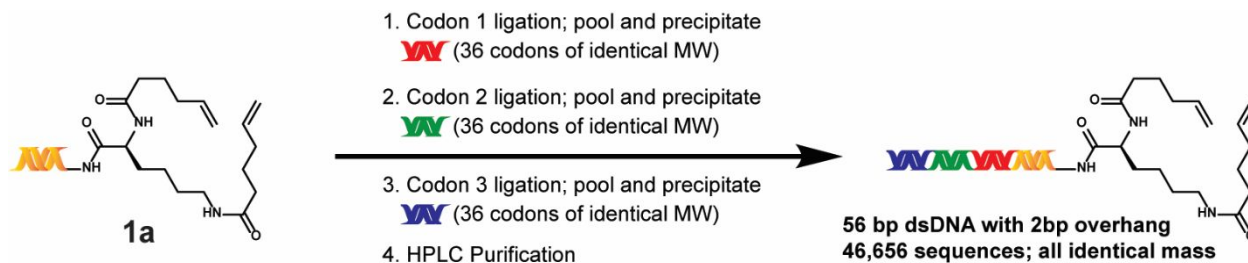
Synthesis of Single-Substrate Library



Figure S6. The architecture of the DNA-only RCM library.

The DNA-only RCM library is a three-cycle library (three split-and-pool cycles, **Scheme S1**). However, the DNA oligomers are used solely to create DNA sequence diversity rather than encode chemical transformations or building blocks. The library is created from substrate **1a** (HPLC purified), which already contains the DTSU, first overhang, forward primer unit “FPU” and upper second overhang regions shown in gold and blue (**Figure S6**). Three sets (codons 1–3) of 36 duplexed complimentary pairs of 13-bp dsDNA oligomers with 5’-Phos and 2-bp overhangs were iteratively ligated by the general procedure, pooled and precipitated by the general procedure. All codons within each set had unique sequences but equivalent molecular weights to provide a single mass upon deconvolution of the pool. Codon ligations were conducted pool amounts of 2 nmol, 1.78 nmol and 1.77 nmol for codon 1, codon 2 and codon 3, respectively. After ligation of the final codon set, the library pool was purified by HPLC by the general procedure to provide a stock soln of the DNA-only RCM library (20 nmol, 0.1 mM stock in H₂O) for further experiments.

Scheme S1. Synthesis of DNA-Only Library from Substrate **1a**



Experimental Information

Buffering the RCM Reaction Mixture

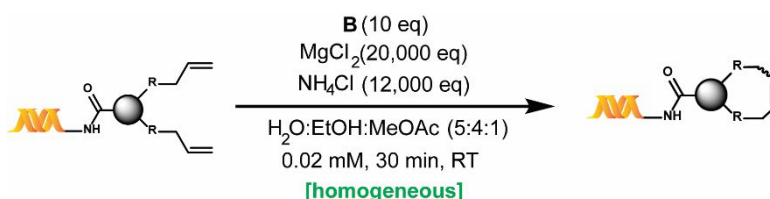
Only non-coordinating buffers can be used in metal catalyzed reactions. Given that DNA is present and that MgCl_2 is present to quench coordination of the phosphate backbone – as well as of the DNA bases – typical phosphate buffers used in DNA-encoded chemistry were tested, among others. The collection of tested buffers is as follows.

pH 3.54, 5.5 and 7.0 sodium phosphate, pH 5.54 MES, pH 7.0 PIPES, and pH 7.5 MOPS buffers

The buffers were added at the typical equivalence of 250 relative to the DNA conjugate and all of them quenched the reaction.

Only high equivalence of ammonium gave a non-quenching and acidic reaction mixture (pH ~5), and was adopted as a part of the reaction conditions.

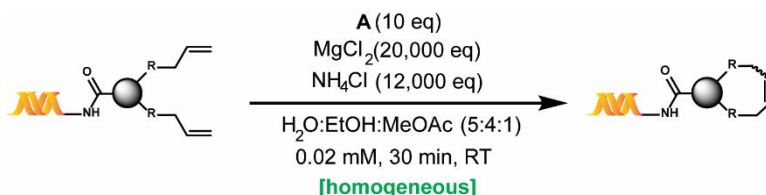
General Protocol for RCM Reaction under Our Main Reaction Conditions (see below for special protocol for substrate 24a)



All reactions were run at 0.02 mM. This concentration was required due to the limited solubility of GIII in MeOAc (2.5 mM gave a clear solution while 5 mM formed a suspension), which drove the reaction volume higher in order to maintain the ideal solvent percentages.

For each reaction, the required volume of DNA stock was added to the reaction vessel followed by the calculated amount of water required by the final solvent composition. MgCl_2 and NH_4Cl were then added from 2M and 4M aqueous stock solutions, respectively, followed by the required volume of EtOH. A 2.5 mM stock solution of **B** was then prepared in MeOAc and the calculated volume was added to and thoroughly mixed with the rest of the reaction mixture. The reaction was allowed to run for 30 min and was quenched as follows.

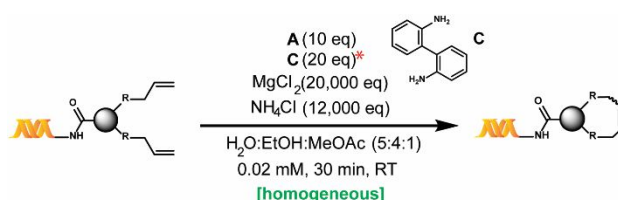
General Protocol for RCM Reaction under our Alternative #1 Conditions



All reactions were run at 0.02 mM. This concentration was required due to the limited solubility of GIII in MeOAc (2.5 mM gave a clear solution while 5 mM formed a suspension), which drove the reaction volume higher in order to maintain the ideal solvent percentages.

For each reaction, the required volume of DNA stock was added to the reaction vessel followed by the calculated amount of water required by the final solvent composition. MgCl₂ and NH₄Cl were then added from 2M and 4M aqueous stock solutions, respectively, followed by the required volume of EtOH. A 2.5 mM stock solution of **A** was then prepared in MeOAc, and the required volume was added to and thoroughly mixed with the rest of the reaction mixture. The reaction was allowed to run for 30 min and was quenched as follows.

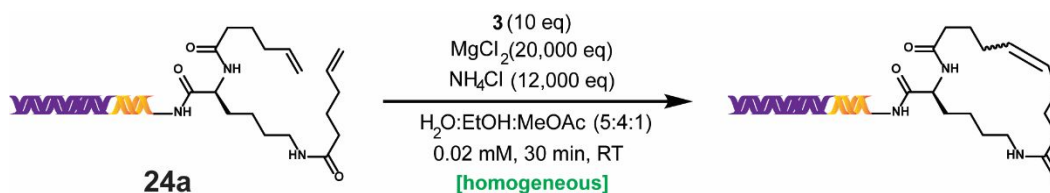
General Protocol for RCM Reaction under our Alternative #2 Conditions



All reactions were run at 0.02 mM. This concentration was required due to the limited solubility of GIII in MeOAc (2.5 mM gave a clear solution while 5 mM formed a suspension), which drove the reaction volume higher in order to maintain the ideal solvent percentages.

For each reaction, the required volume of DNA stock was added to the reaction vessel followed by the calculated amount of water required by the final solvent composition. MgCl₂ and NH₄Cl were then added from 2M and 4M aqueous stock solutions, respectively. The required volume from a 25 mM stock solution of **C** in EtOH was then added, followed by a complementary volume of EtOH. A 2.5 mM stock solution of **A** was then prepared in MeOAc, and the required volume was added to and thoroughly mixed with the rest of the reaction mixture. The reaction was allowed to run for 30 min and was quenched as follows.

RCM, Precipitation Procedure and Post-RCM Yield of Substrate 24a



The distinguishing feature of this substrate (relative to **1a–23a**) is the length of the DNA tag. In a DNA encoded library, the RCM reaction is likely to be carried out at the end of the build, or at least with a longer DNA tag than the DNA headpiece.

The reaction setup was the same as described in the General RCM protocol. It is to be noted that the reaction mixture turns slightly cloudy after addition of EtOH – not the case when the DNA tag is only the DNA headpiece. The reaction mixture was therefore shaken for 30 min as an additional measure.

The reaction was then quenched, as described below.

Before precipitation, however, 12,000 eq of NaOH were added in order to neutralize the ammonium ions in solution. It was found that, without this step, DNA recovery is <5%.

The material was then precipitated using the general precipitation procedure. The reconstituted material was then washed in a 3k Amicon filter to remove the large amount of salt left after precipitation. Quantification of the yield of the reaction was then performed via Bioanalyzer, as described below.

To visualize the composition of the elongated 56-bp dsDNA substrate **24a** and post-RCM product **24b**, diluted samples of both were ran on a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA) using an Agilent DNA 1000 kit. Representative electropherogram results of this analysis are shown for the starting material **24a** in Figure S7 and post-RCM mixture of **24a** in Figure S8. Within these electropherograms, peaks at bp = 15 and bp = 1500 are DNA standards included in the loading matrix. Due to the modified nature of the DNA, DECL samples do not migrate at retention times that directly correspond to DNA standards—however retention times corresponding to bp = 79–81 have been consistently observed for other 56-bp dsDNA samples that contain identical overall DNA–chemical conjugate architecture within our DECL pipeline. Based on sample volumes and the integration of these peaks on several diluted samples, 20–25% of the expected bp = 80 DNA material was recovered after the RCM reaction, quench and precipitation procedures. However as shown in Figure S7, additional small peaks of larger DNA length were observed, which may correspond to low-level intermolecular metathesis products or other intermolecular complexes. It is important to note that these impurities were not readily detected using LC-MS or polyacrylamide gel electrophoresis. It is likely that during a large-scale library production, these undesired impurities may be removed by HPLC purification of the post-RCM, pooled library material before use in protein-binding experiments.

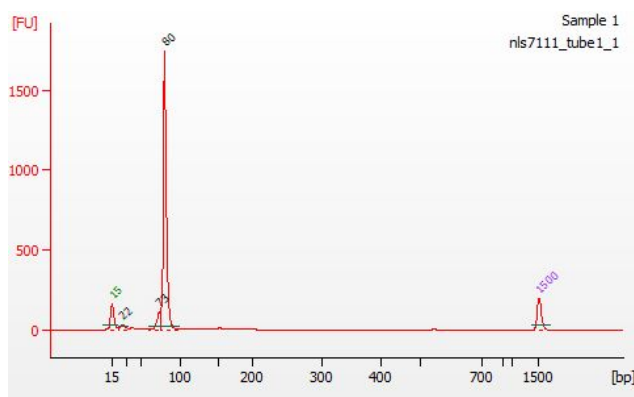


Figure S7. Bioanalyzer electropherogram of substrate **24a**.

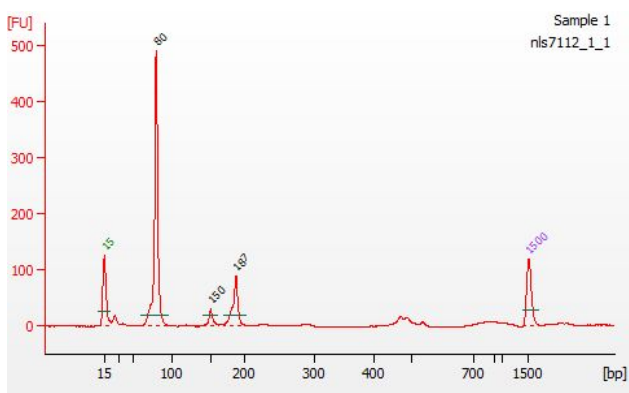
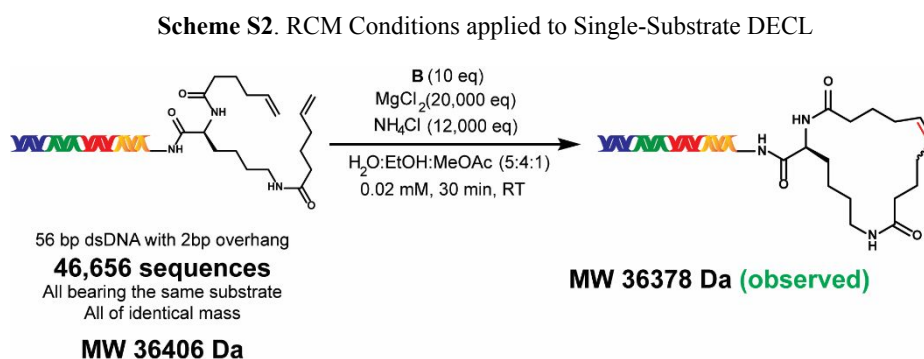


Figure S8. Bioanalyzer electropherogram of the post-RCM mixture from **24a**.

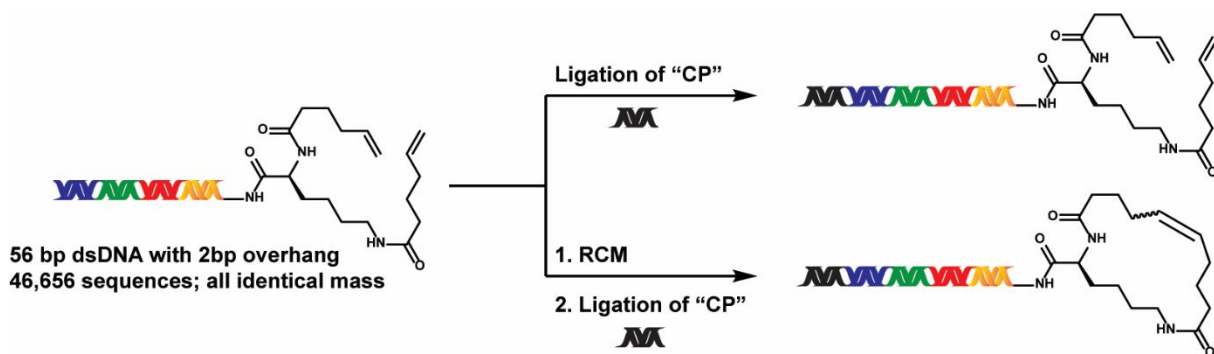
RCM and Sequencing Prep of Single-Substrate DECL

A sample of the DNA-only library then was subjected to the RCM conditions described in Scheme 1(iii), as shown in **Scheme S2**



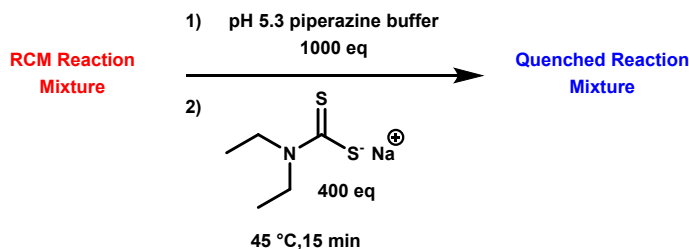
After precipitation, the library sample was prepared for PCR amplification by ligation with a pool of three DNA oligomers “CPs” with 5’-Phos (Scheme S3) as well as a smaller complimentary lower strand oligomer to allow annealing. Concomitantly a sample of the post-HPLC library that did not undergo the RCM reaction was ligated with a unique set of “CPs” and an annealing lower-strand DNA oligomer to serve as a control. “CPs” contain additional segments to enable amplification and bioinformatics analysis. After quantification by qPCR, portions of the two samples (3×10^7 copies) were amplified by PCR with primer adaptors compatible with Illumina sequencing flowcells, quantified (Agilent Bioanalyzer 2100) and sequenced (Illumina NextSeq 500). After sequencing, the samples were compared for sequence content and distribution. The principles used for codon and CP design, as well as the methods used to analyze the naïve sequencing results, were discussed in our recent publication of Faver and coworkers⁷

Scheme S3. Treatment of DNA-Only Library in Preparation for Sequencing



Quenching the RCM Reaction and Preparation for LC/MS Analysis

Sodium diethyldithiocarbamate proved to be the best metal capture agent for our purposes. However, its basic nature represents the threat of pH swings towards the basic end of the spectrum during quenching, as it neutralizes the surrogate buffer, NH_4Cl . Ru-based catalysts are sensitive to basic pH.⁸ Indeed, when sodium diethyldithiocarbamate was simply added after 30 min of reaction, a brown precipitate formed, which contained most of the DNA conjugate. This led to negligible DNA recovery despite precipitation. To avoid this, an actual buffer (unlike the surrogate NH_4Cl), pH 5.3 piperazine, was first added in high equivalence. Due to the protonation state of piperazine at pH 5.3 it does not create large amounts of undesired precipitates. The metal chelator was then added, which caused a color change from colorless/very light green to light yellow. The mixture was heated at 45°C for 15 min to enhance metal capture and separation from the DNA. The quenched reaction mixture was then sampled for injection onto the LC/MS instrument for analysis. The quenching protocol is illustrated below.



Protocol for Replication of Reaction Conditions Reported by Lu, X. et al.

The only difference between the reported protocol⁹ and ours concerned reaction mixing method. While Lu, X. et al.⁹ used stir bars and larger scale reactions (100 nmol compared to 1 nmol in our case), we placed our reactions on a vortexer (VWR Analog Vortexer), which allowed for the dispersion of the catalyst suspension. Additionally, the quenched reaction (supernatant left after dithiocarbamate and centrifugation) was not purified by HPLC before LCMS analysis.

Figure S9 provides a visual of the homogeneity of our reaction conditions.

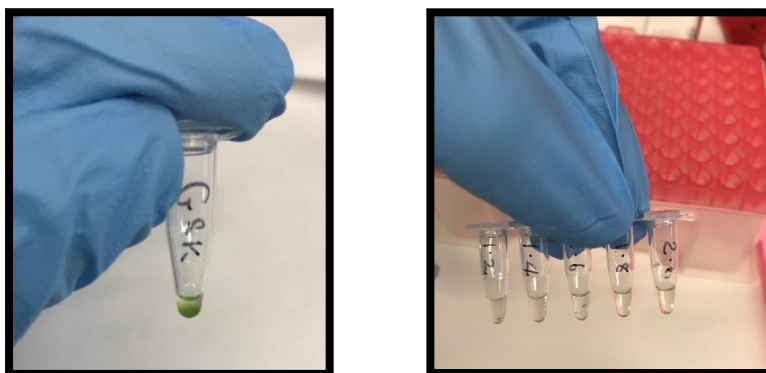
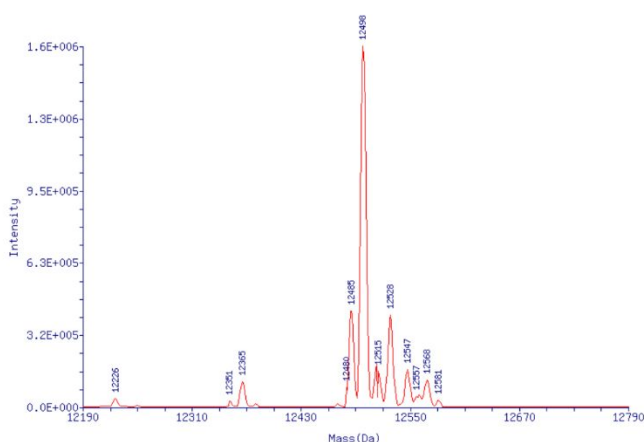


Figure S9 Pictures depicting the heterogeneity of the conditions reported by Lu, X. et al. (left) and the homogeneity of the conditions reported in this work (right); both pictures were taken by Olivier Monty during the investigation.

Method for the Determination of the Percent Conversion from MS Deconvolution Results

Below is a data table obtained after the processing of the MS trace of a post-RCM reaction sample of substrate **1a**. The processing was done by the MS deconvolution software, ProMass. The table is representative of that obtained for all post-RCM samples. The table lists the percentage areas under the graph attributable to the various species peaks. A percentage calculation of the relevant percentages gives the percent conversion. An example is here provided and is representative of the procedure used for the determination of all % conv.

MS Deconvolution Data post-RCM of Substrate **3a** (deconvolution spectrum of starting material available in Characterization Information section)



[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [View Data] [Log File]
Deconvolution Peak Report:

Mass (Da)	Intensity	Std. Dev.	Score	Delta Mass	%Relative	%Total
12497.5	1.58E+006	0.3	15.29	0.0	100.00	49.02
12482.9	4.19E+005	0.4	11.03	-14.6	26.52	13.00
12527.8	4.02E+005	0.4	10.97	30.3	25.45	12.48
12510.0	1.67E+005	1.1	8.40	12.5	10.57	5.18
12546.4	1.64E+005	0.5	7.55	48.9	10.37	5.08

The desired product has $M+2 = 12526$ g/mol. The desired post RCM product should therefore have $M-28+2 = 12498$ g/mol. The latter mass is indeed observed, as shown above.

Three side products are also formed, 12482.9, 12527.8 and 12510.0 (image depicts peak averages).

The % conv. was calculated as follows.

$$\% \text{ conv.} = \left(\frac{49.02}{49.02 + 13.00 + 12.48 + 5.18} \right) * 100 = 61.5$$

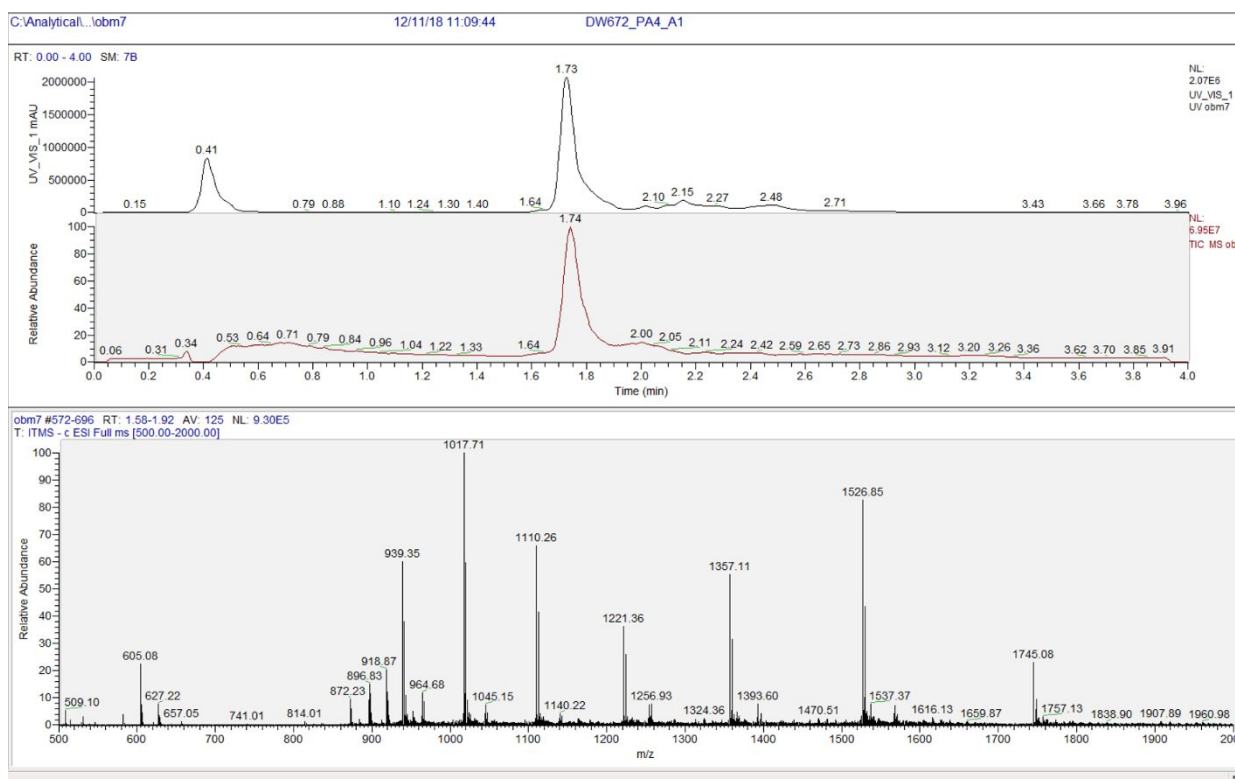
Comparison of Percent Conversions Determined from MS Deconvolution Spectra v. UV spectra

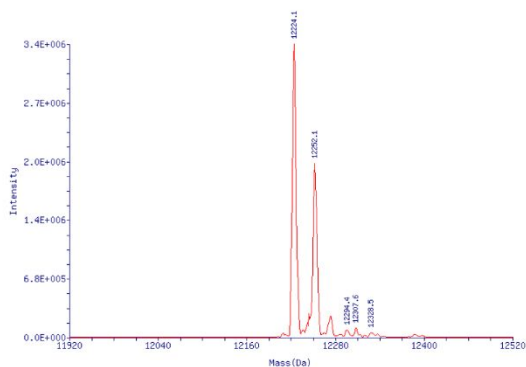
Reaction conversions in DNA-encoded chemistry are typically calculated from the deconvolution spectra obtained from LC-MS runs, as shown above. Unlike small molecules, DNA-chemical conjugates possess numerous charges, thus resulting in m/z values many folds smaller than the molecular weight. Deconvolution software is therefore essential to process the relatively large amounts of data obtained from LC-MS spectra.

However, it is possible to extract the UV spectra of specific DNA-chemical conjugates whose mass can be calculated. Assuming the contribution to the UV signal of the chemical is drowned by the high UV activity of DNA in a DNA-chemical conjugate, percent conversions could be calculated from the UV trace as well. A comparison was thus made to help satisfy the curiosity of the reader, as described below.

Compound 2a (Calculated mass 12252.12) post RCM to give Compound 2b (calculated mass 12224.09)

The LCMS and deconvolution spectra are shown below.





[<<] [Top] [ESI Mass Spectrum] [Deconvolution] [View Data] [Log File]

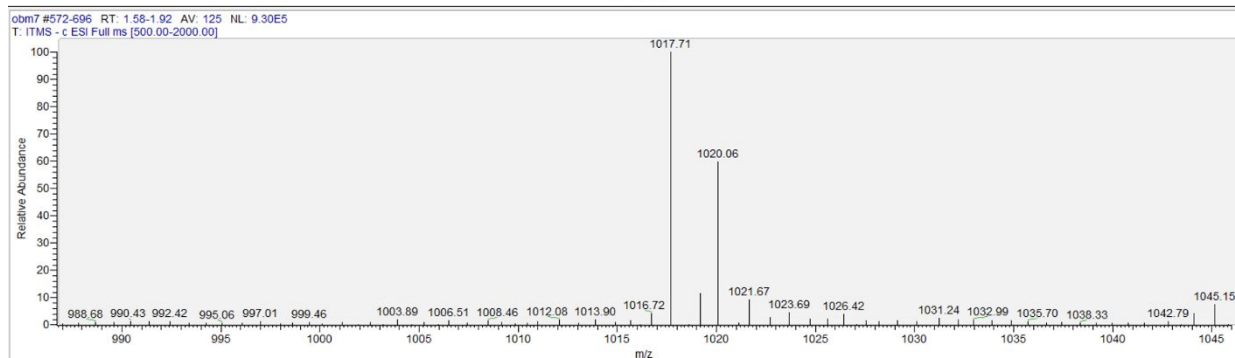
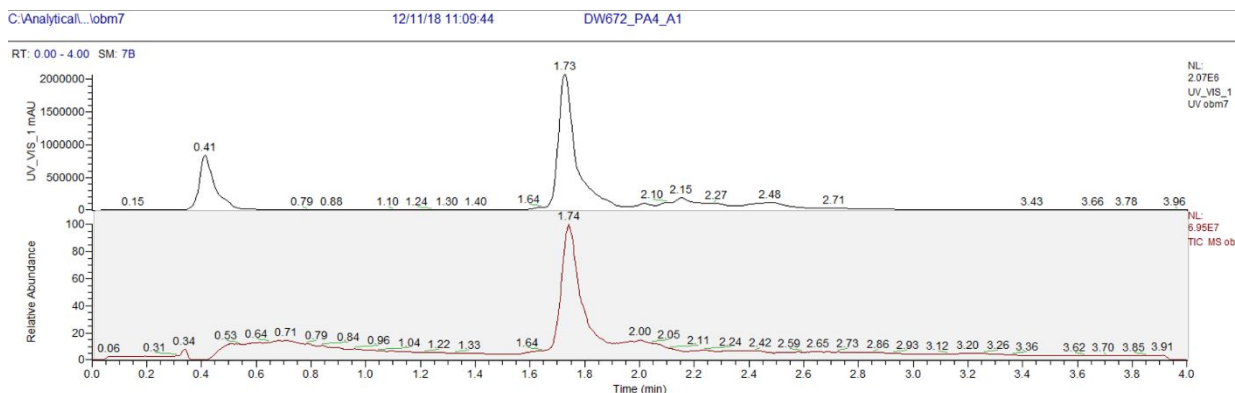
Deconvolution Peak Report:

Mass (Da)	Intensity	Score	Delta Mass	%Relative	%Total
12224.1	3.38E+006	19.31	0.0	100.00	53.48
12252.1	2.00E+006	16.67	26.0	59.26	31.69
12552.7	2.90E+005	11.13	328.6	8.58	4.59
12580.4	1.83E+005	9.06	356.3	5.43	2.90
12307.6	1.18E+005	4.16	-5939.5	3.49	1.86
12541.4	1.10E+005	3.28	83.5	3.26	1.75
12541.4	1.01E+005	5.77	317.3	2.98	1.59
12294.4	8.13E+004	5.74	70.3	2.40	1.29
12328.5	5.39E+004	4.55	104.4	1.60	0.85

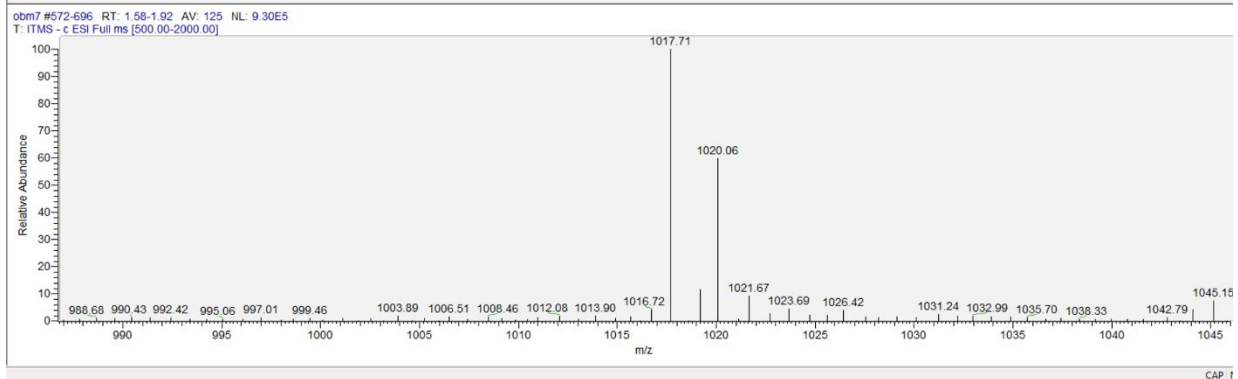
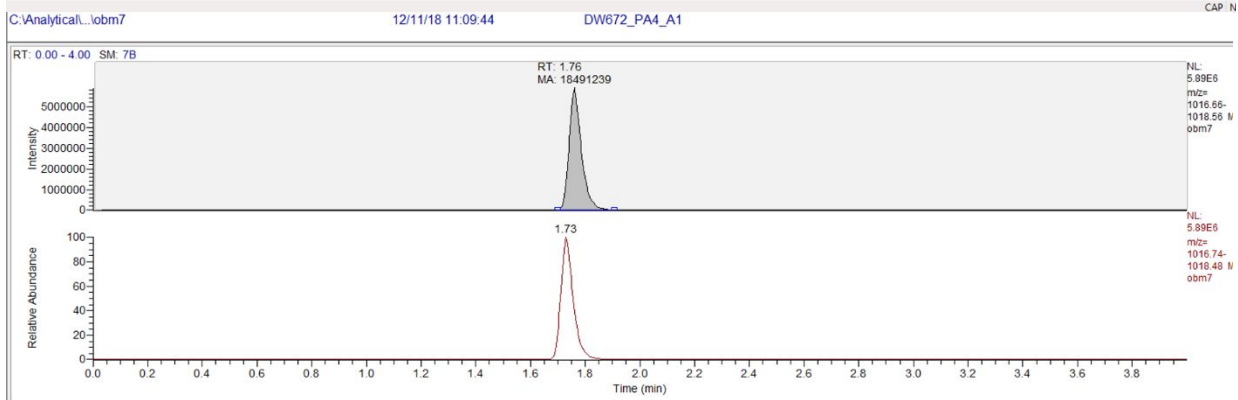
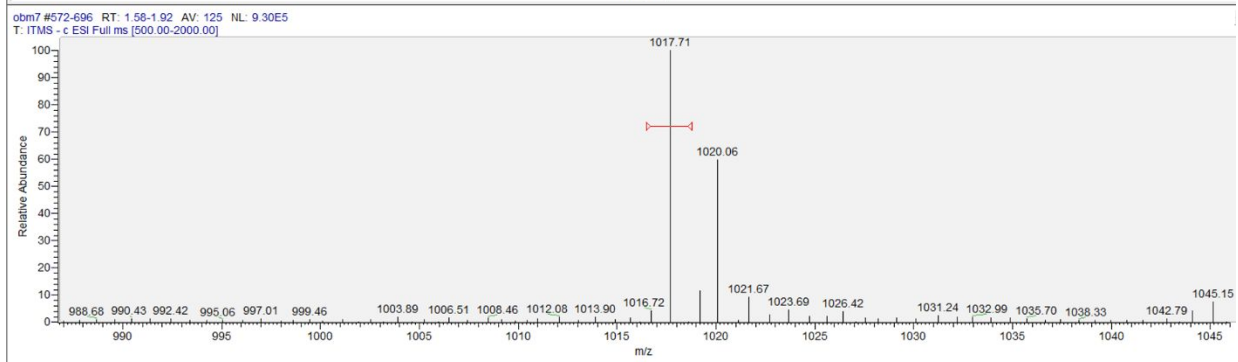
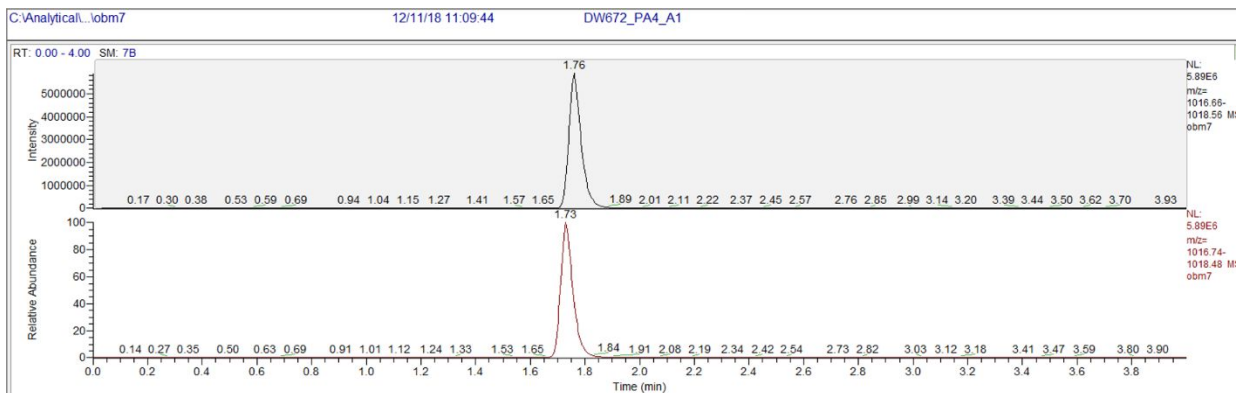
According to the deconvolution results, there are two peaks of interest, the starting material (SM)-like (SM-like refers to the fact that the material left undergoes no further cyclization. We hypothesize it has been modified to an unreacted form, as discussed in the next section) peak (31.69% by area) and the product peak (53.48% by area). Using the formula above, a ~62% conversion can be calculated, as reported in the main paper.

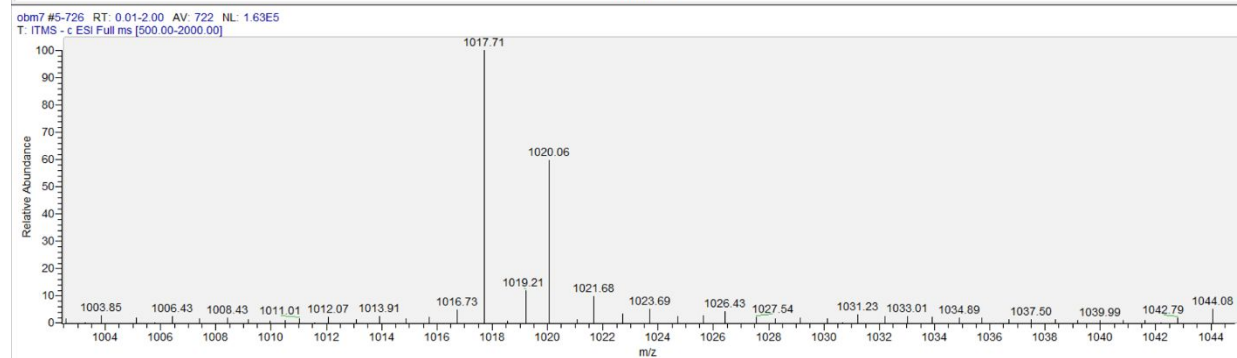
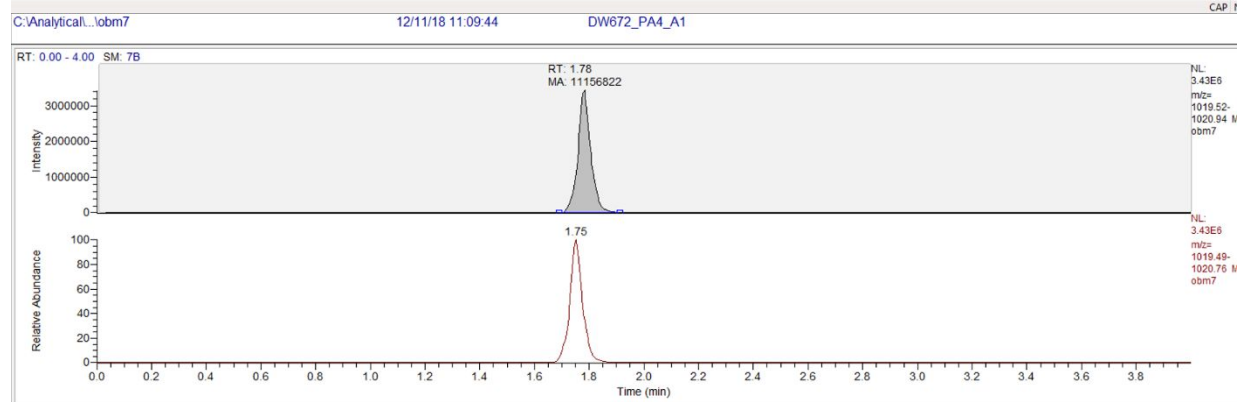
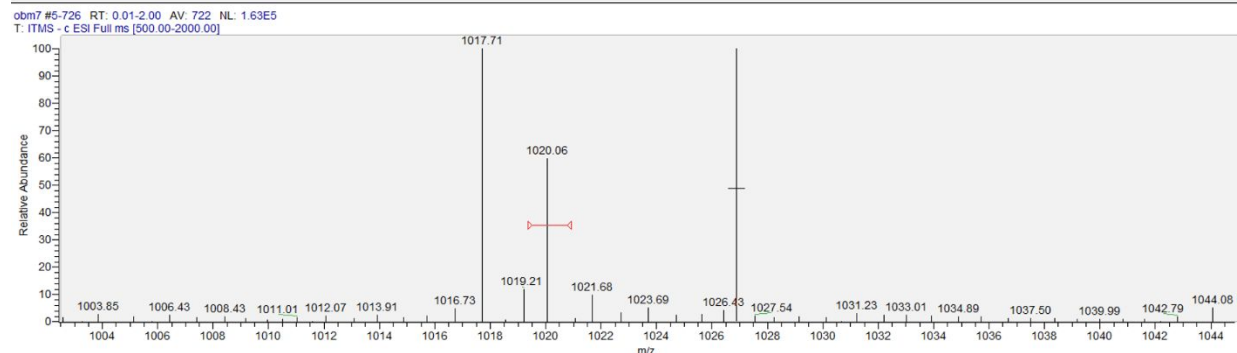
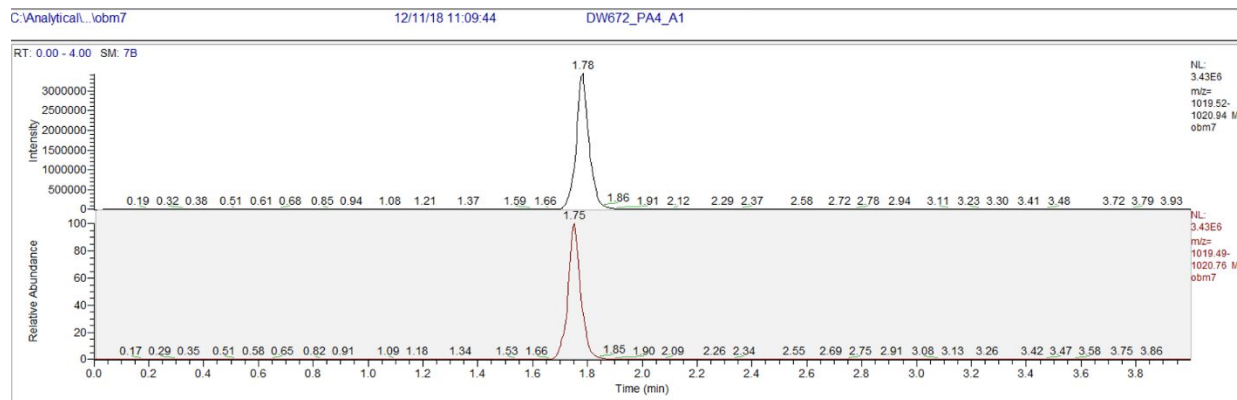
As visible from the LCMS spectrum, there is no LC resolution between the cyclized and un-cyclized material and the UV peaks have to be extracted. The extracted peaks are as follows.

The DNA headpiece (**S1**) typically has a charge (z) of +12 or +13. The m/z base peak therefore corresponds to one or the other. In the case of **2a**, +12 gives the base peak and this region is shown zoomed in below.



The SM-like peak should be ~1021.01 (12252.12/12) and the product peak should be ~1018.7 (12224.12/12). Those are 1020.06 and 1017.71 in the spectrum above. Each of those correspond to a UV peak with an area under the peak, as shown below.





The UV spectrum area corresponding to m/z 1020.06 is 11156822 (A1) and that corresponding to m/z 1017.71 is 18491239 (A2). The percent conversion is therefore ~62% $[(A2/(A1+A2))*100]$, in agreement with the result obtained from the deconvolution results.

Confirmation of Cyclization via Chemical Modification

As described in the main text, chemical modification was used to ascertain the occurrence of ring closure during the reaction, as well as to try and determine the identity of the side products. The test was designed as follows.

The non-conjugated DNA headpiece, the DNA conjugate before RCM, the crude RCM reaction (after imposition of the conditions described in Scheme 1(iii), post quench) and the HPLC-purified RCM product were subjected to the conditions illustrated in Scheme 3 (reproduced from main paper). The changes in mass were followed, and conclusions drawn, as elaborated below.

Scheme 3. Confirmation of Ring Formation via Chemical Modification

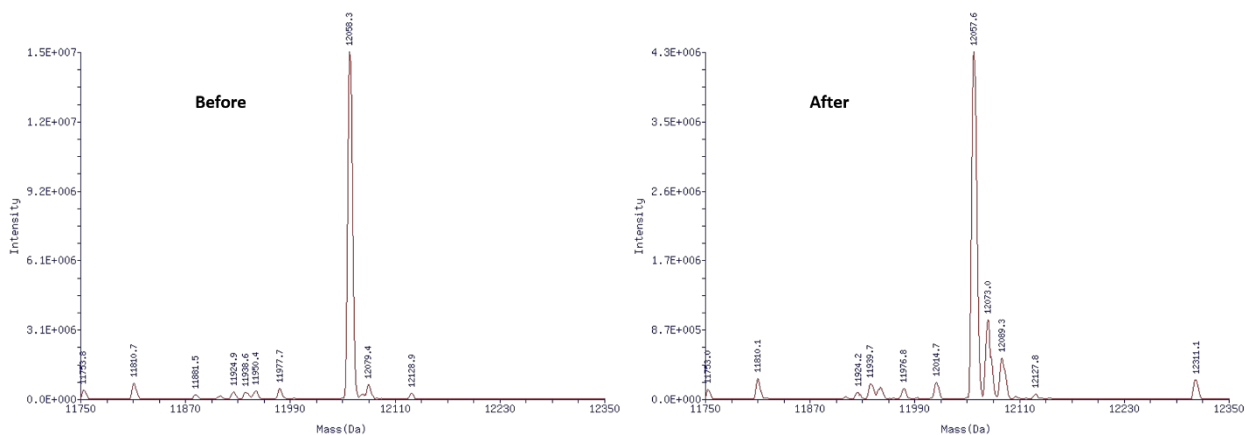
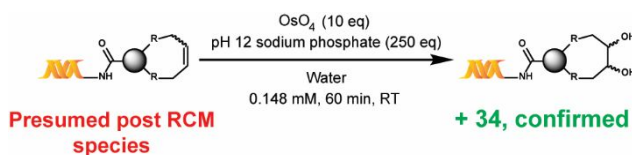


Figure S 10. DNA Headpiece before and after conditions described in Scheme 3 (as labeled in main paper).

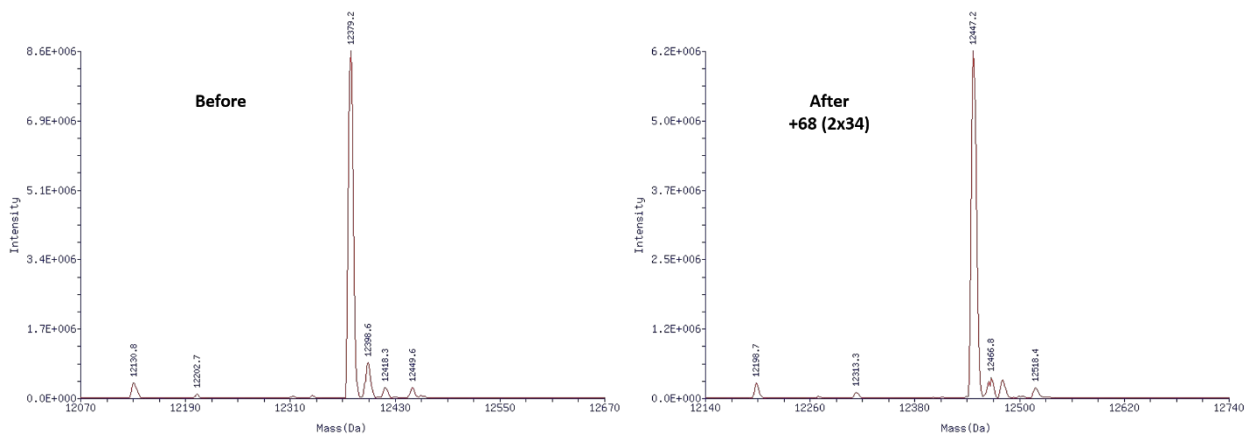


Figure S 11. Substrate 1a before and after conditions described in Scheme 3 (as labeled in main paper).

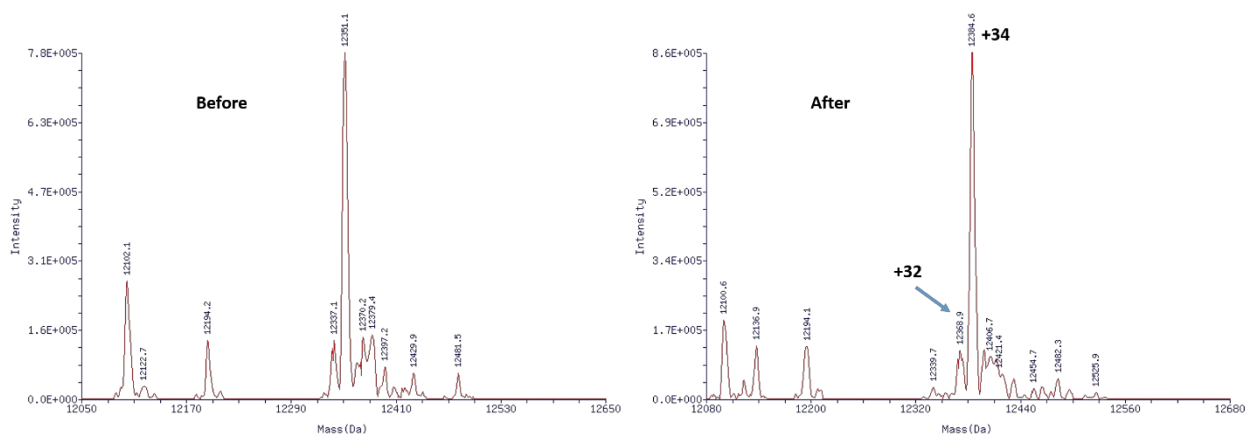


Figure S 12. Crude RCM of substrate **1a** before and after conditions described in Scheme 3 (as labeled in main paper).

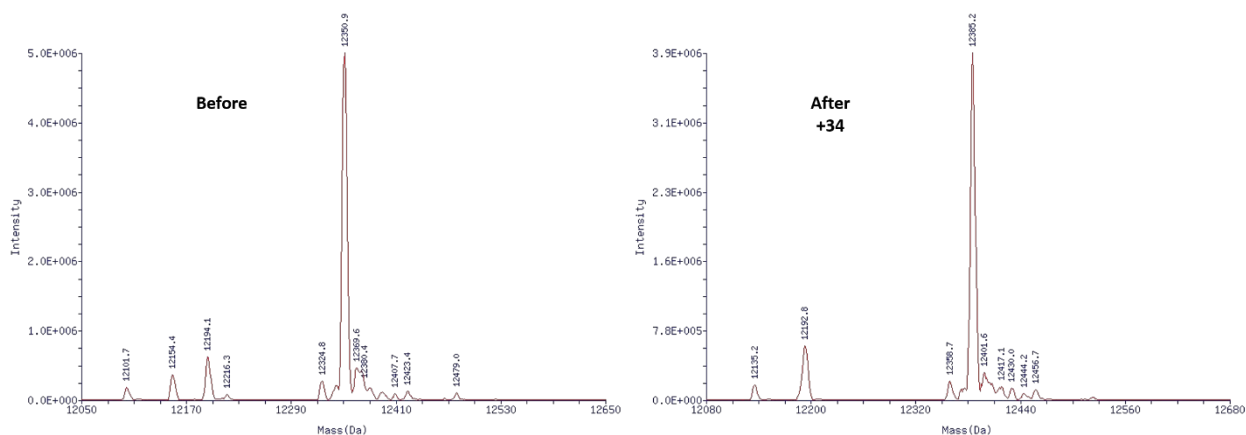


Figure S 13. HPLC-purified RCM product of substrate **1a** before and after conditions described in Scheme 3 (as labeled in main paper). Reaction complete after 10 min.

As shown in Figure S10, the mass of the DNA headpiece is unaffected by the reaction conditions. Changes in mass observed in the case of the DNA conjugates therefore originated in the reactivity of the attached substrate.

As shown in Figure S11, the diene **1a** behaved according to prediction, with a change in mass of +68 (2×34) that corresponds to the addition of two hydroxyl groups to each olefin.

Both Figure S12 and S13 show that the target product mass (12351 Da) indeed corresponds to the desired cyclized product, as a shift of +34 (dihydroxylation of only one olefin) is observed. The absence of a +68 peak provides strong evidence of cyclization.

Figure S12 provides insight into the identity of the side products. The side product at 12337 Da showed a +32 change in mass. This is suggestive of a cyclized side product, although +34 would have been diagnostic. We believe that mono-olefin isomerization (from terminal to 1,2-disubstituted) and eventual cyclization (loss of methylene first and then cyclization, or RCM of the internal olefin) gave the side product. The peak at 12379 (mass of **1a**) seems to suggest the reaction was incomplete and could have benefitted from running longer than 30 min. However, allowing the RCM reaction to run longer does not lead to a significant change in conversion to the desired product (**Table 1** in main paper). Additionally, as shown in Figure S12, there is no signal at 12447, thus indicating that the starting material has been converted to a different chemical entity.

Confirmation of Cyclization via HTRF Assay of Cyclized v. Uncyclized Substrate 26a against target protein, ER alpha

To have phenotypic evidence of successful cyclization of substrate **26a**, the post-RCM reaction mixture containing **26b** (quenched and precipitated) was contrasted with the pure starting material, **26a**, with respect to their level of binding to the co-activator region of the estrogen receptor alpha. Three samples were therefore prepared from the no-compound headpiece, **S1**, as control, the unstapled peptide substrate **26a**, and the post-RCM reaction mixture containing a majority of the stapled peptide **26b** for use within an homogenous, time-resolved fluorescence protein interaction assay (HTRF). Samples P1 and P2 were prepared by ligation of a “T1 tag” and sample P3 was prepared by ligation of a “T6 tag”.

P1: A DNA-only control of plain headpiece, **S1**, elongated with a 60/62-mer DNA T1 tag

P2: Substrate **26a** elongated with a 60/62-mer DNA T1 tag

P3: Post-RCM **26b** elongated with a 60/62-mer DNA T6 tag

T1 upper: CTCACCTAGTCACCTCCATTTCGTCGCCAGTGGCACATCCGGCACGCTCATTTCGCAATACT

T1 lower: AGTATTGCGAATGAGCGTGCCGGATGTGCCACTGCGGACGAATGGAGGTGACTAAGTGAGCC

T6 upper: TAGGAGGAGTTGTCGTGGAACGTCCGCAGTGGCACATCCGGCACGCTCATTTCGCAATACT

T6 lower: AGTATTGCGAATGAGCGTGCCGGATGTGCCACTGCGGACGTTTCGACGACAACCTCCTCCTACC

The concentration of each of these samples was quantified by qPCR before use within the HTRF assay. A homogenous time-resolved fluorescence protein interaction assay (Cisbio) was developed in which the estrogen receptor (6His ESR1-LBD) AA298-554 was indirectly labeled with MAb Anti-6HIS-Tb cryptate Gold (Cisbio # 61HI2TLF) to make the donor, and fluorescein-SRC3-1 coactivator peptide (Life Technologies # PV4590) was used as the acceptor. Excitation at 340 nm of the donor results in a signal that is measured ratiometrically as a quotient of two emission wavelength 520nm/620 nm (acceptor/donor), and is proportional to the binding of the 2 proteins. The assay was done in a total volume of 50 μ L, containing 25 nM estrogen receptor, 200 nM estradiol (E2), 100 nM SRC3-1, 0.7 nM Tb-cryptate, and the signal was measured for compounds P1–P3 at 0.28, 0.56, and 1.12 μ M concentration for the effect on coactivator peptide binding. As shown in Figure S14, a dose dependent effect was observed for stapled peptides P2 and P3, with enhanced blockage of coactivator peptide binding for P3.

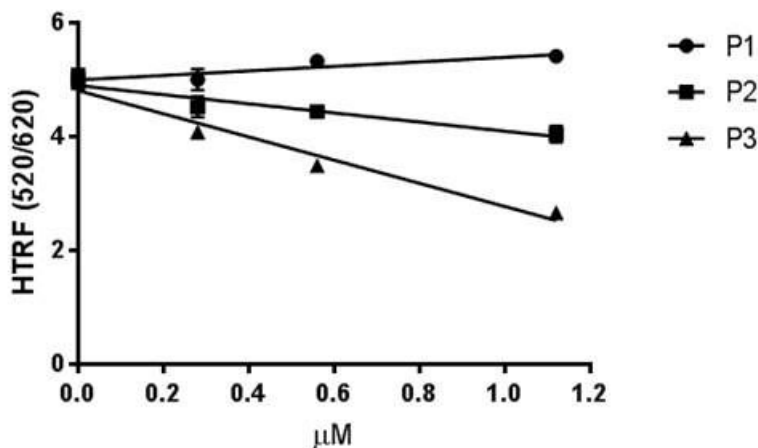


Figure S14. Plot of the HTRF signal for P1–P3 at 0.28, 0.56, and 1.12 μ M concentrations.

Sequencing of the Single-Substrate DECL: a Pre- and Post-RCM Comparison

Amplifiable samples of the DECL before and after the RCM condition were quantified by quantitative Real-Time PCR (qPCR) and then a total of 3×10^7 DNA copies were amplified by PCR with primers adaptors to add the sequences compatible with Illumina sequencing flowcells. Platinum Taq DNA Polymerase High Fidelity (Thermo Fisher Scientific) PCR reagent was used for PCR amplification. A total of 15 PCR cycles were used for amplification and the following PCR conditions were used (Initial denaturation at 95°C for 2.5min, denaturation at 95°C for 30s, annealing at 58°C for 30s, extension at 72°C for 1min and final extension at 72°C for 10min). PCR library template was purified using Agencount AMPure XP SPRI beads according to the manufacturer's instructions. The purified library was analyzed in Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA) by using Agilent high sensitivity DNA kit to verify library size and concentration before clustering. Illumina NextSeq 500 was used for sequencing.

Illumina sequencing and analysis

Raw DNA sequence reads (in the form of FASTQ files), quality metrics, and sequencing index-to-sample attribute value pairs were obtained from Illumina BaseSpace at the conclusion of sequencing. Samples were linked to their respective FASTQ files based on their sequencing index (DTSU) and were expanded into individual experiments if they were part of a larger pool. Individual samples were then decoded by perfectly matching individual oligonucleotide sub-structures without gaps and in the order defined by the known DNA encoding structure (Main Library Build). Valid DNA barcodes were annotated with the corresponding oligonucleotide sequence-lookup for each of the three codon cycles. The degenerate UMI (unique molecular identifier) portions of the DNA barcodes were accumulated into a list of UMIs for each unique codon tuple as a method to distinguish experimental vs. amplification events. Unique molecule counts were then evaluated using a directed-graph counting model as described previously.⁷ The set of unique codon tuples with unique molecule counts was then aggregated across all possible combinations of codons (all *n-synthons*), and enrichment for each *n-synthon* was evaluated independently. The populations of each *n-synthon* in the DECL samples with and without the RCM condition were compared by plotting observed-to-expected *n-synthon* population ratios (Figure S15). Observed populations were evaluated by using total counts for the library sample, counts for a specific *n-synthon*, and the Agresti-Coull estimation interval for proportions. The expected populations were evaluated by using the codon diversity of the library and an assumption of uniform yields (i.e., equal probability of observation for each codon within a cycle). In Figure S15, the observed-to-expected population ratios are plotted for the DECL with RCM against the DECL without RCM. The comparisons are separated by "axis" which correspond to all *n-synthons* in the 3 cycle library: axis 0 represents cycle 1 *mono-synthons*, axis 1 represents cycle 2 *mono-synthons*, axis 3 represents cycle 1 x cycle 2 *di-synthons*, etc. Along each combinatorial axis, most *n-synthons* are near the expected population (i.e., 100%) in each DECL sample. We observed that many *n-synthons* which were under- or over-populated in the RCM DECL sample were similarly under- and over-populated in the non-RCM DECL sample. This corresponds to points which follow the $y=x$ line. Importantly, deviations from the $y=x$ line of equal populations were symmetrically distributed about the $y=x$ line, and increased with increasing dimension of *n-synthon*, which is consistent with random sampling effects. We therefore conclude that we observe no significant differences in codon populations between the two DECL samples with and without RCM.

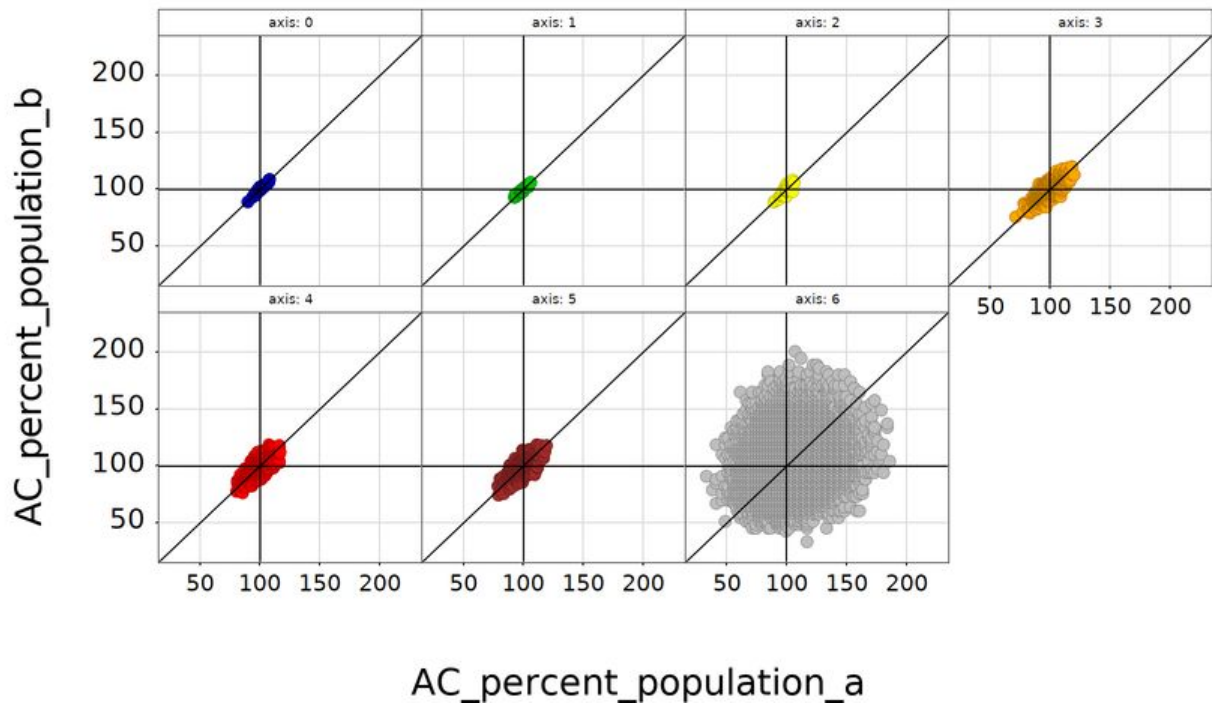
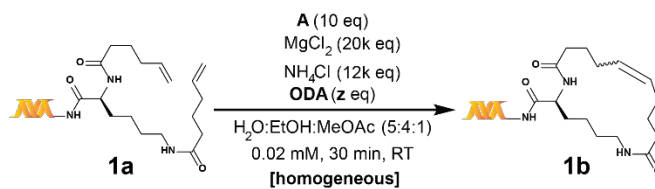


Figure S 15. Comparison of the codon populations of the pre- (population a) and post-RCM (population b) DECLs.

Screening Tables

Table S1. Optimization Screen for the equivalence of the ODA ligand, **C**, relative to the DNA conjugate

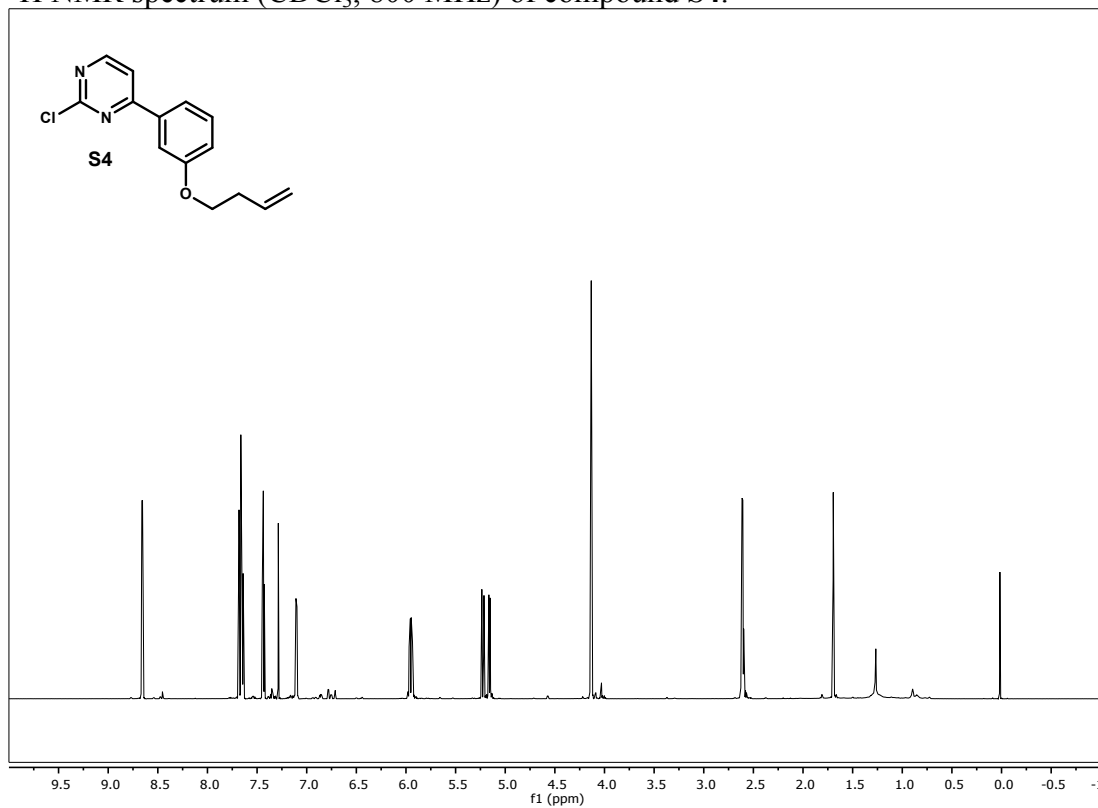


Trial# ^a	z	% conv. 1b ^b
1	0	66
2	0.5	68
3	1.0	72
4	1.5	74
5	2.0	75
6	3.0	74

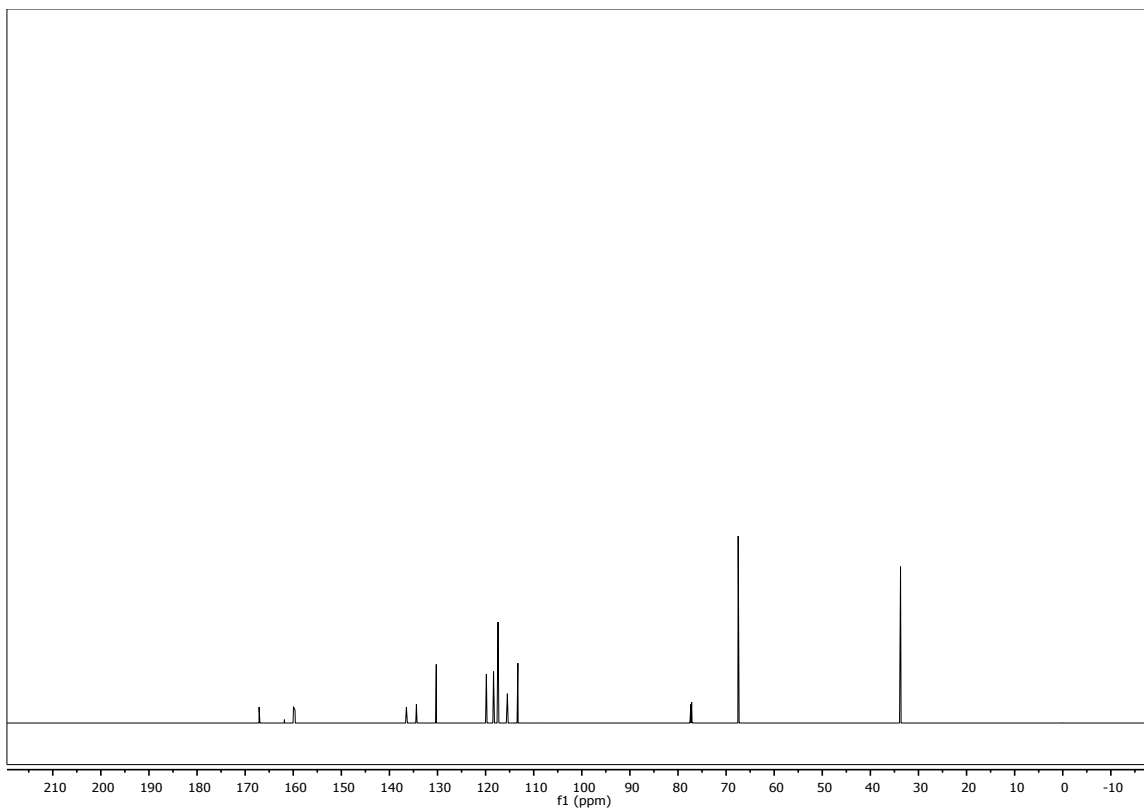
^aAll reactions were run with 1 nmol of **1a**. ^bThe percent conversions (% conv.) were determined by LC/MS after the quenching procedure, as described above.

Characterization Information

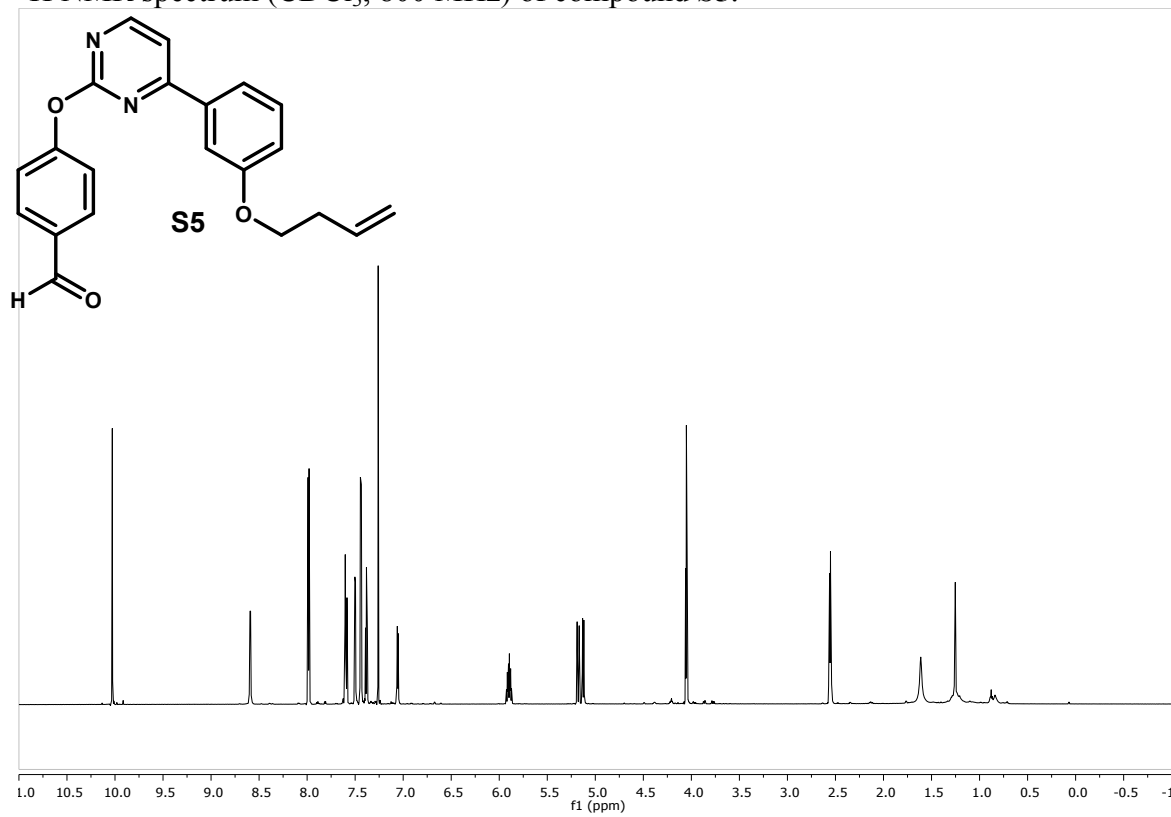
^1H NMR spectrum (CDCl_3 , 800 MHz) of compound **S4**.



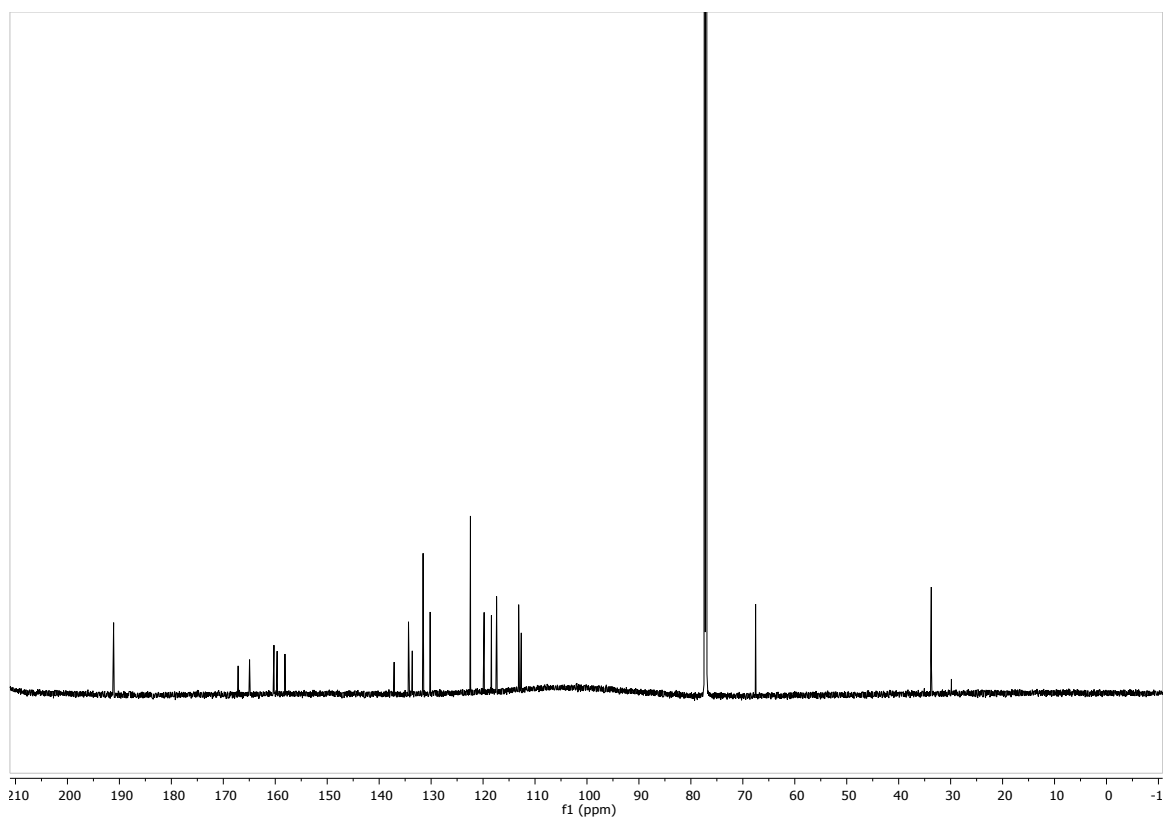
^{13}C NMR spectrum (CDCl_3 , 200 MHz) of compound **S4**.



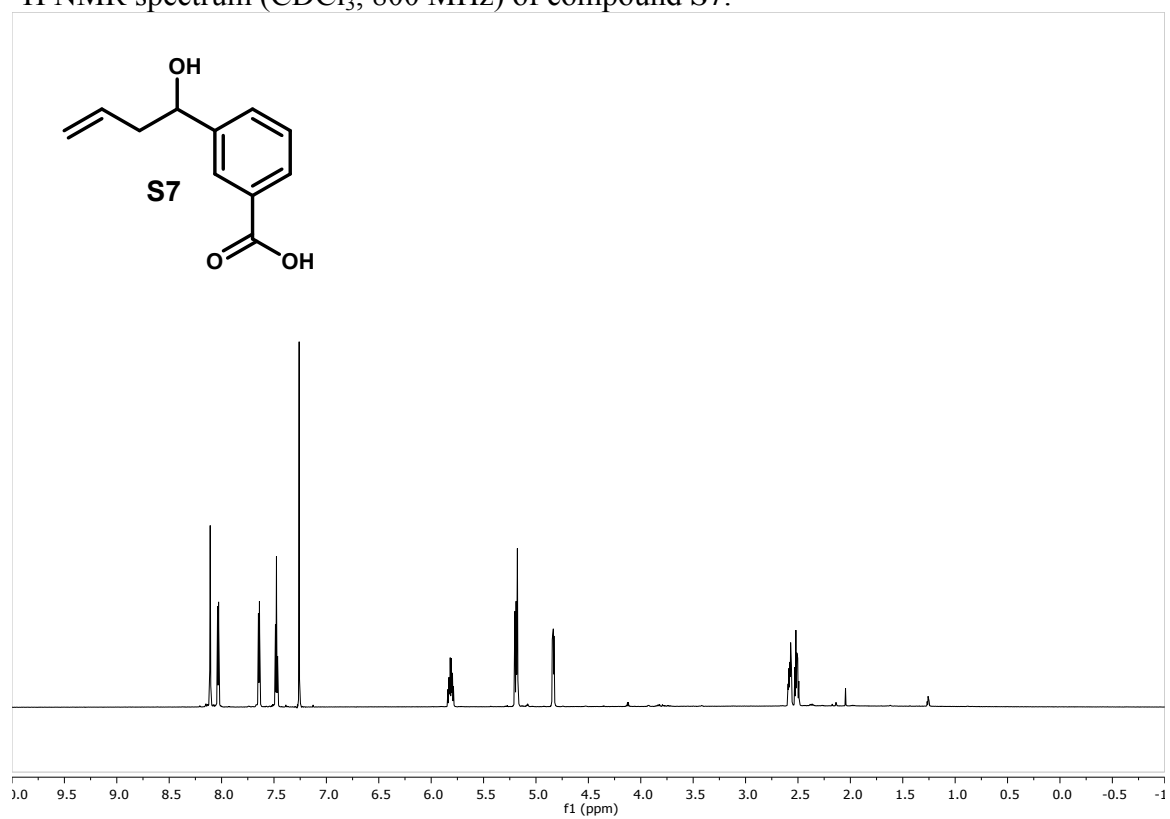
^1H NMR spectrum (CDCl_3 , 800 MHz) of compound **S5**.



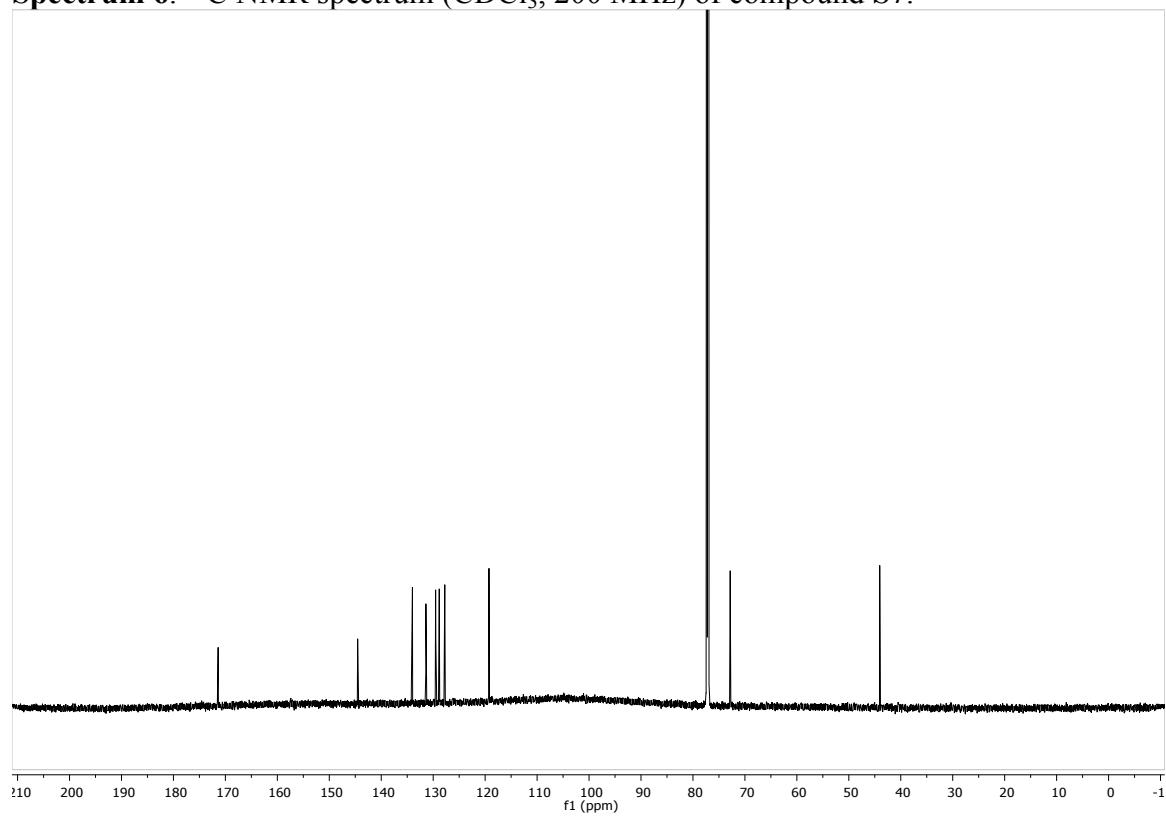
^{13}C NMR spectrum (CDCl_3 , 200 MHz) of compound **S5**.

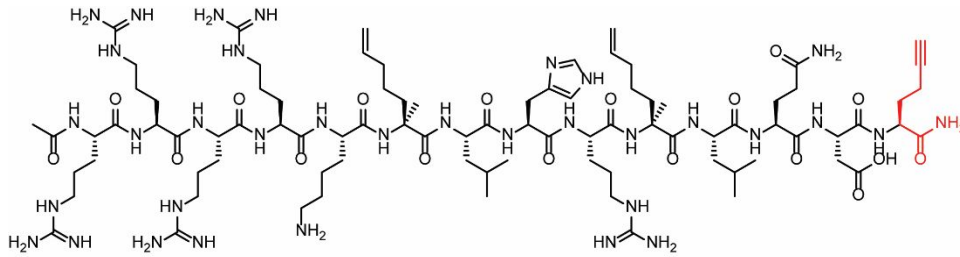


^1H NMR spectrum (CDCl_3 , 800 MHz) of compound S7.



Spectrum 6. ^{13}C NMR spectrum (CDCl_3 , 200 MHz) of compound **S7**.





X

Certificate of Analysis

Sequence: Ac-RRRRK(S5)LHR(S5)LQD(hPra)-amide		
Peptide Name:	Date: January 29, 2019	
Order#: P620616	Lot#: LB7098	Amount: 10.0mg

Quality Control Specifications:

QC Test	QC Specifications	Results
Purity by HPLC	>90% by percent area	Pass
Mass Identification by Mass Spectral Analysis	Calculated Mass within 0.1% of Molecular Weight: 1963	Pass
Concentration/ Net Peptide	Amino Acid Analysis (AAA) determining original concentration/net peptide content.	N/A

Product: Research Grade Custom Peptide containing traces of Trifluoroacetate (TFA) salts.

Formulation:

Final concentration: N/A

Final form: Dry

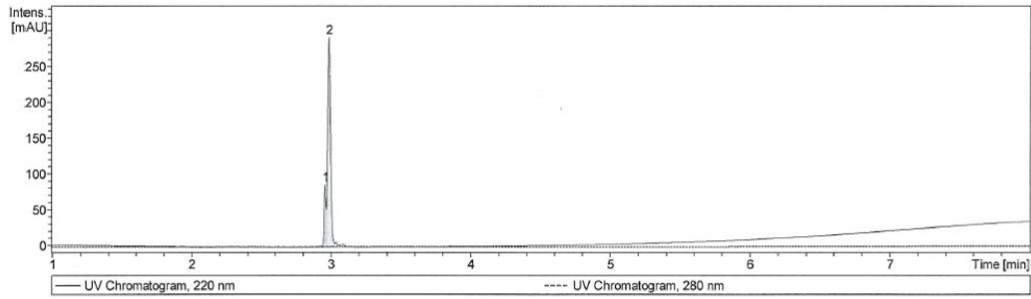
Stability and Conditions: Refer to the Quality Control Detail Information on our website at www.newenglandpeptide.com/support/quality-control-information. As always, NEP has individual batch records stored electronically for each peptide that includes traceable lot numbers of raw materials used during synthesis. Should you require this information, email sales@newenglandpeptide.com with your peptide lot number.

Notes (if applicable): Split peaks on HPLC trace. Both peaks to be counted toward final purity.

BS

Approval/Initials

Analysis Name D:\Data\LB709832-60_212773_P1-C-4_01_11740.D
 Sample Name LB7098 32-60
 Method APRIL20171.2mLperMIN_NEPOAHIGH_11740.m
 Instrument amaZon SL



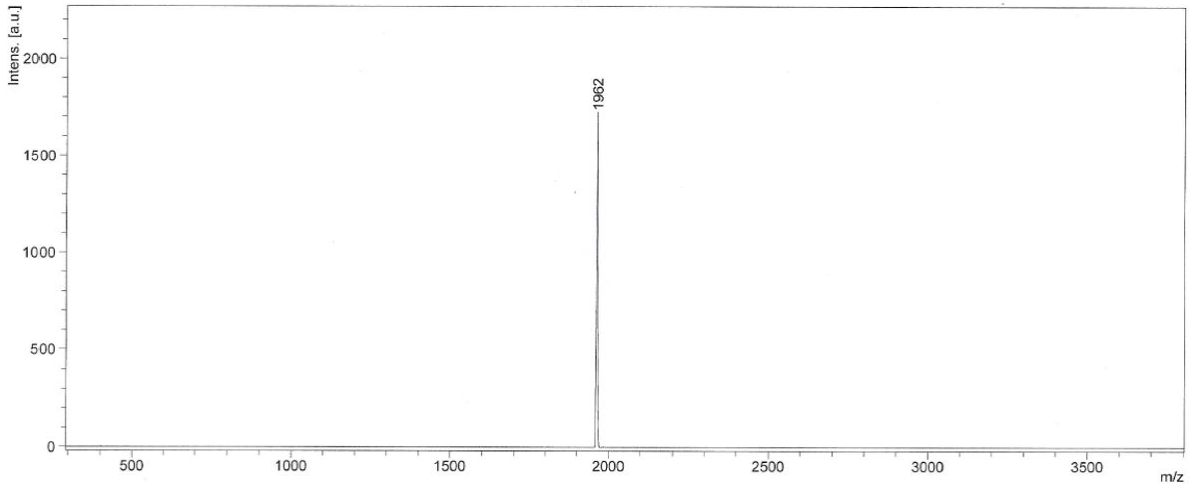
Target Mass	Meas. Mass	Expec. Mass	Delt. Mr [Da]	Intensity	Area	Area Fraction [%]
Compd 2; 2.99 min; Pep Mr: 1962.32	1962.32	1963.00	-0.68	289	409	83.9

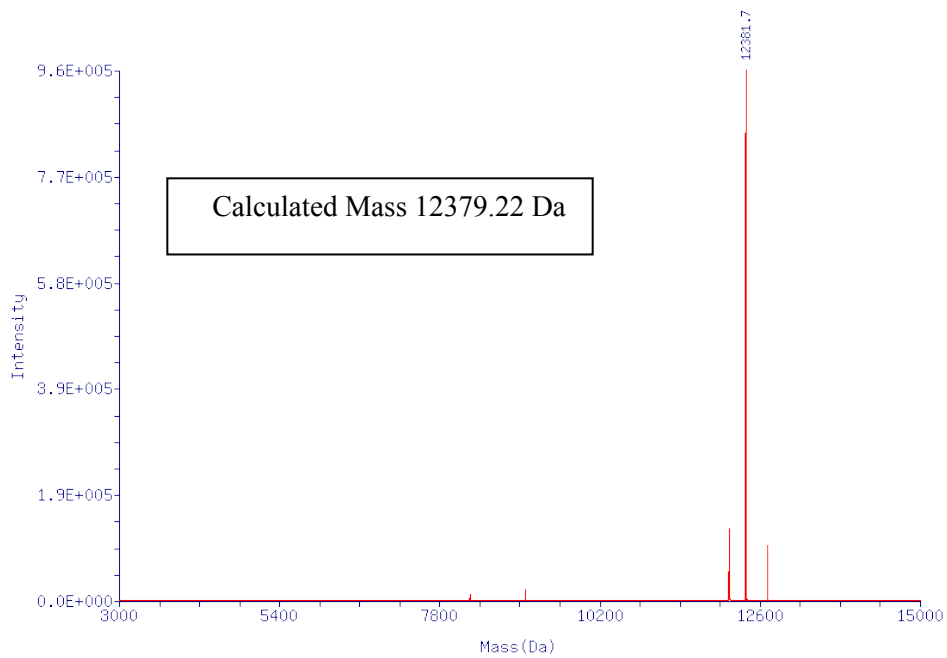
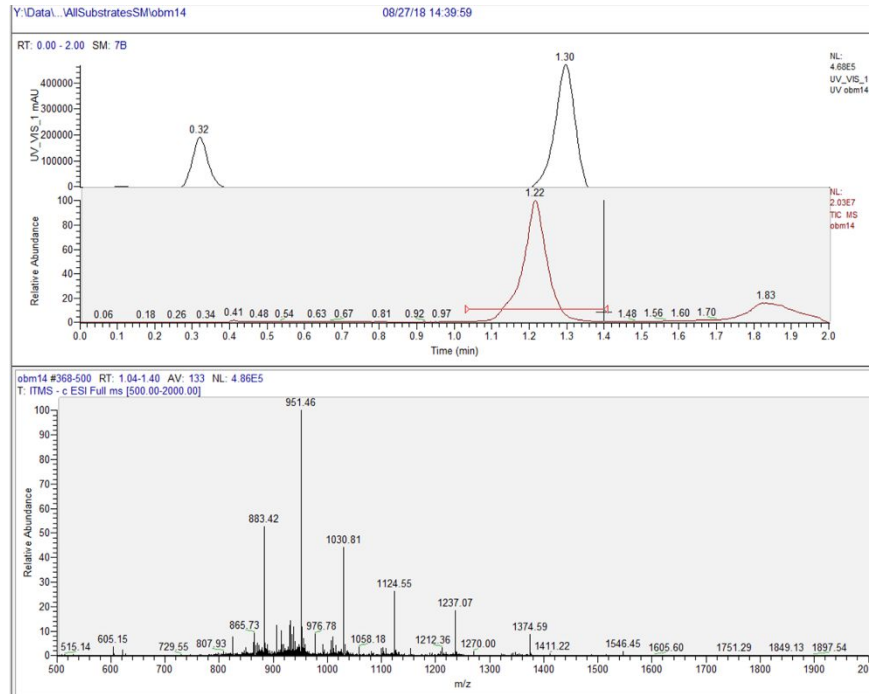
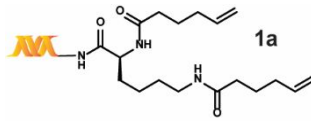
#	RT [min]	Area	Area Frac. %
1	2.96	78.536	16.11
2	2.99	408.850	83.89

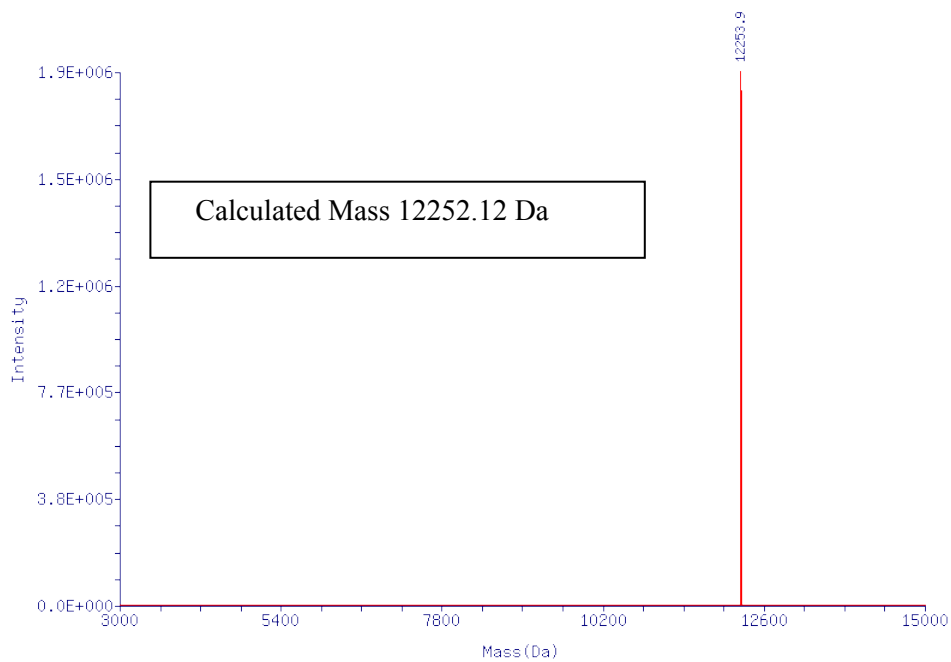
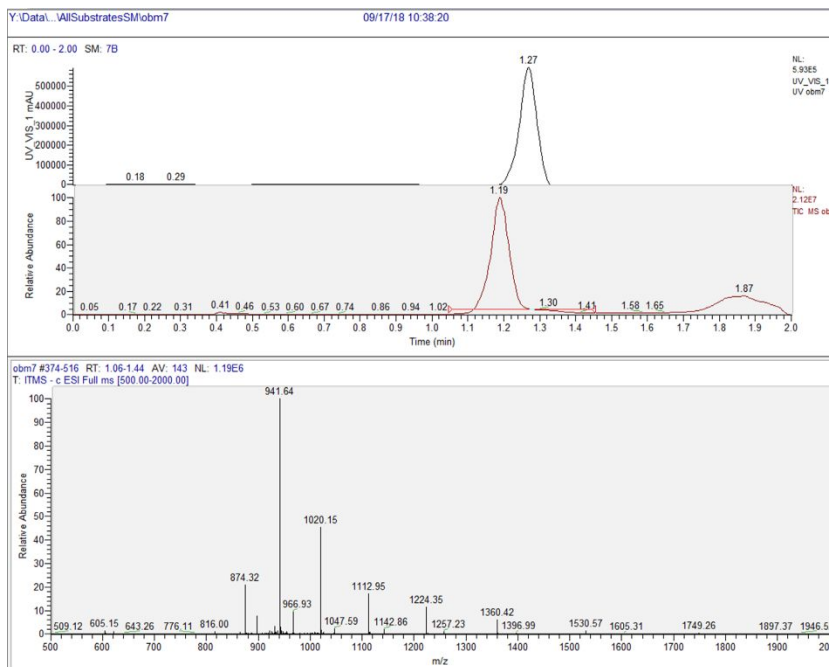
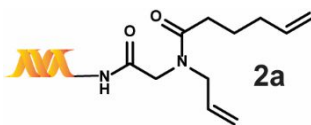
Acquisition method name D:\Methods\flexControlMethods\Negative
 Peptides 500-3500_Sep_2018.par
 Date of Acquisition 2019-01-25T13:00:53.921-05:00
 Comment 2 LB7098 32-60
 LaserPower = 100% - Laser beam attenuation 62.778
 Number of shots 181
 Current absolute file path of the spectrum Z:\Classical\LB7098 32-60\0_D11\111SLin
 Sample name (file name prefix) LB7098 32-60_0_D11\1
 Position D11
 positive voltage polarity NEG

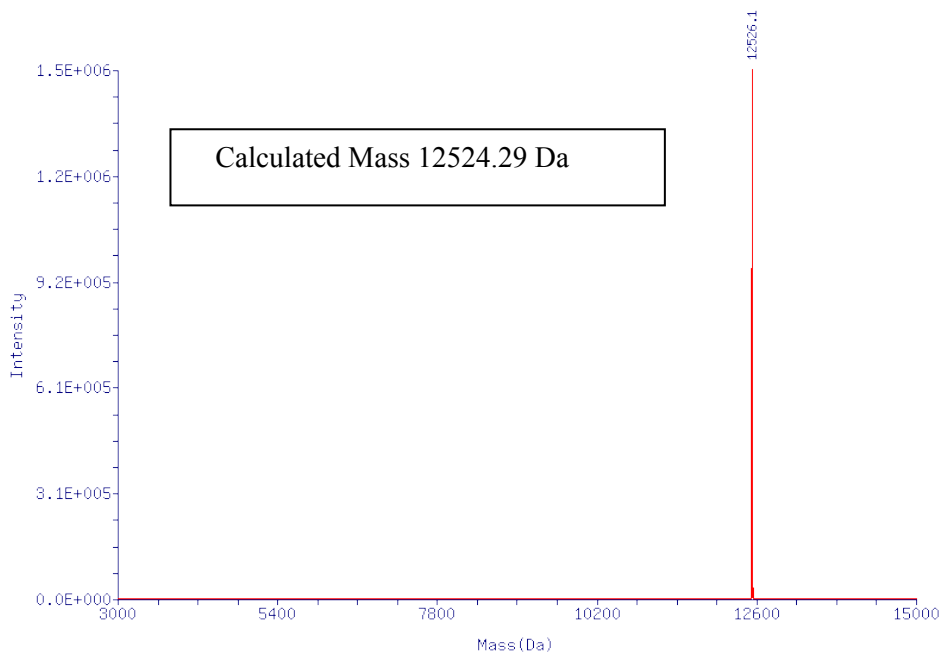
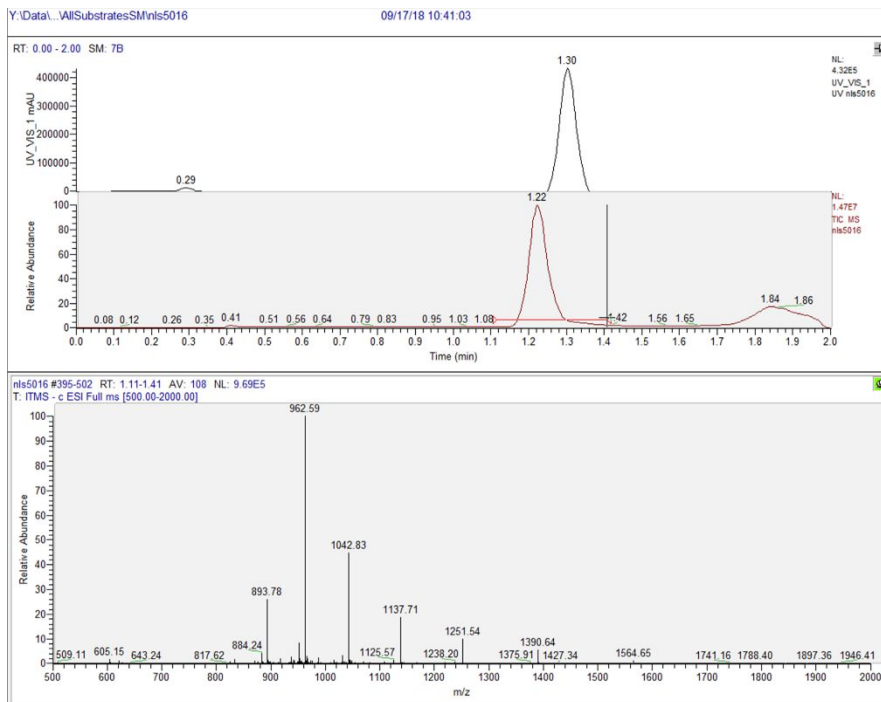
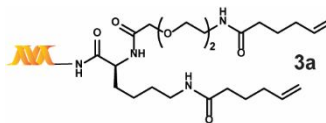
Sample ID: LB7098 32-60

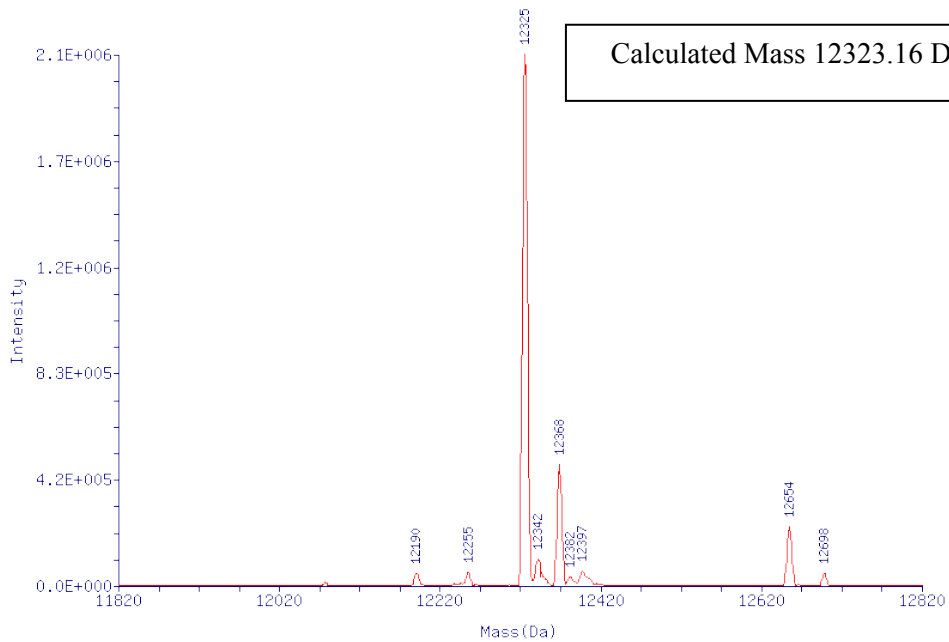
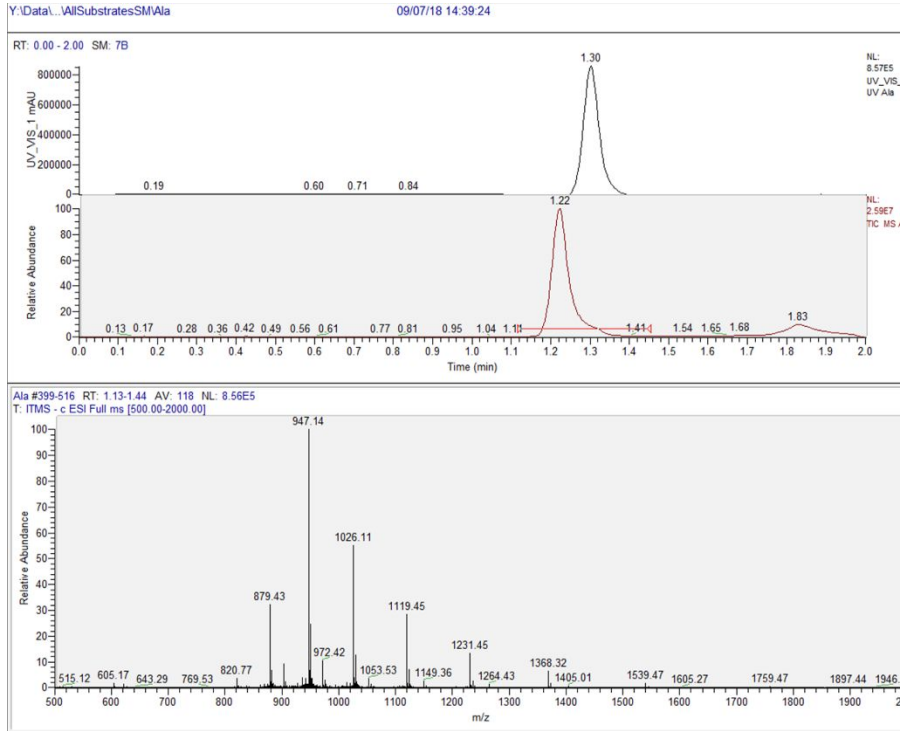
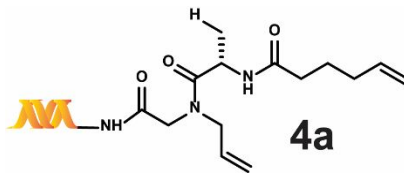
Molecular Weight: 1963

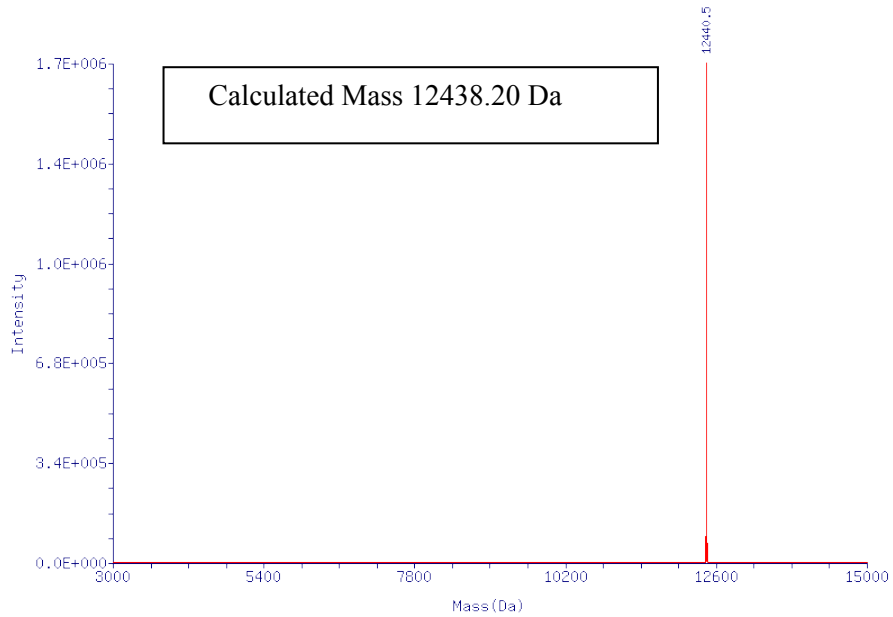
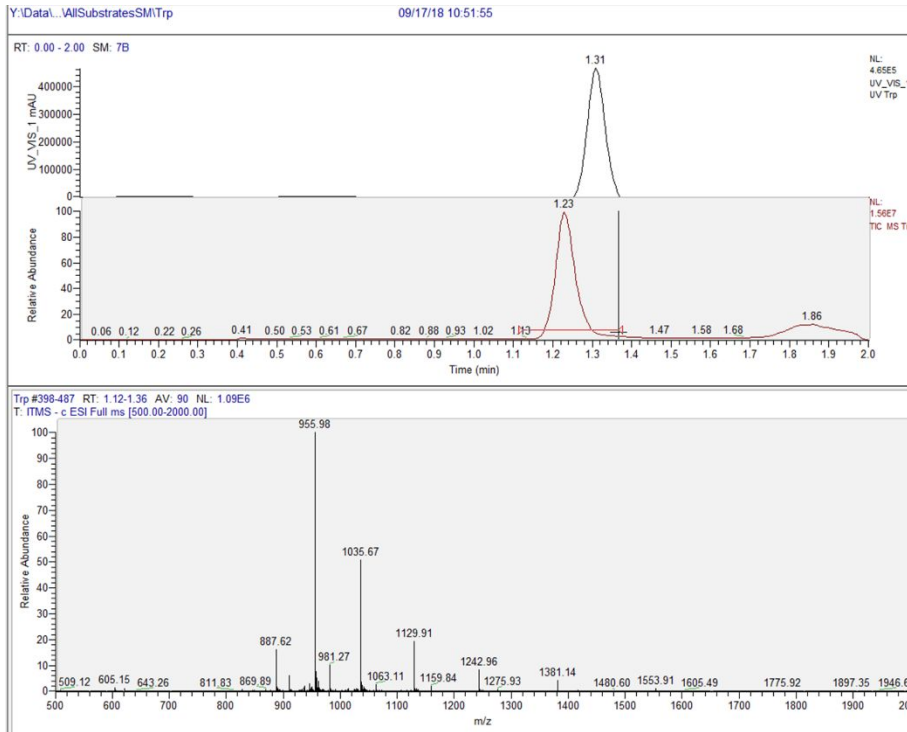
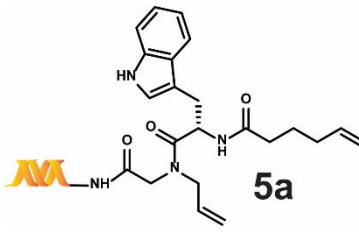


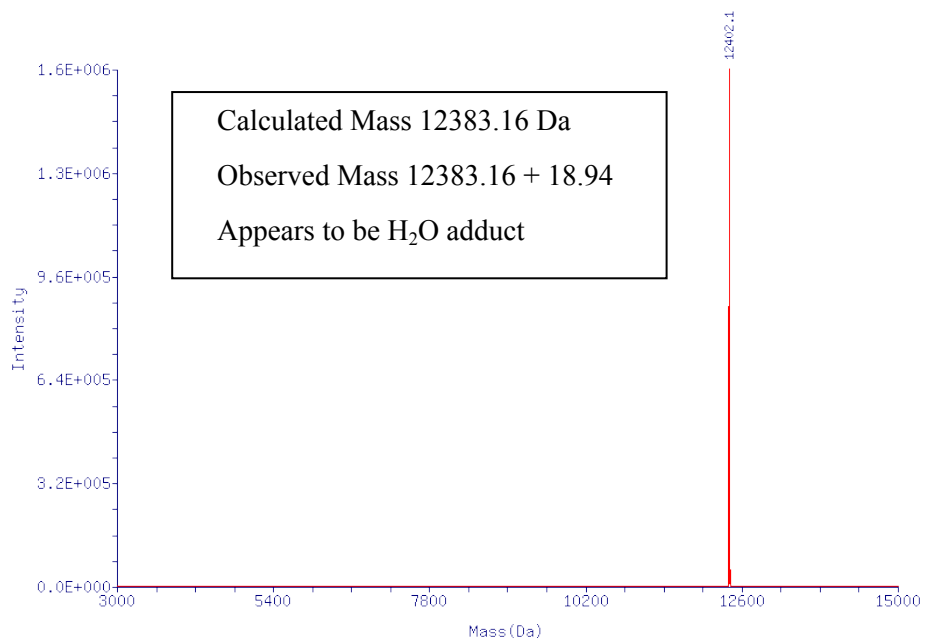
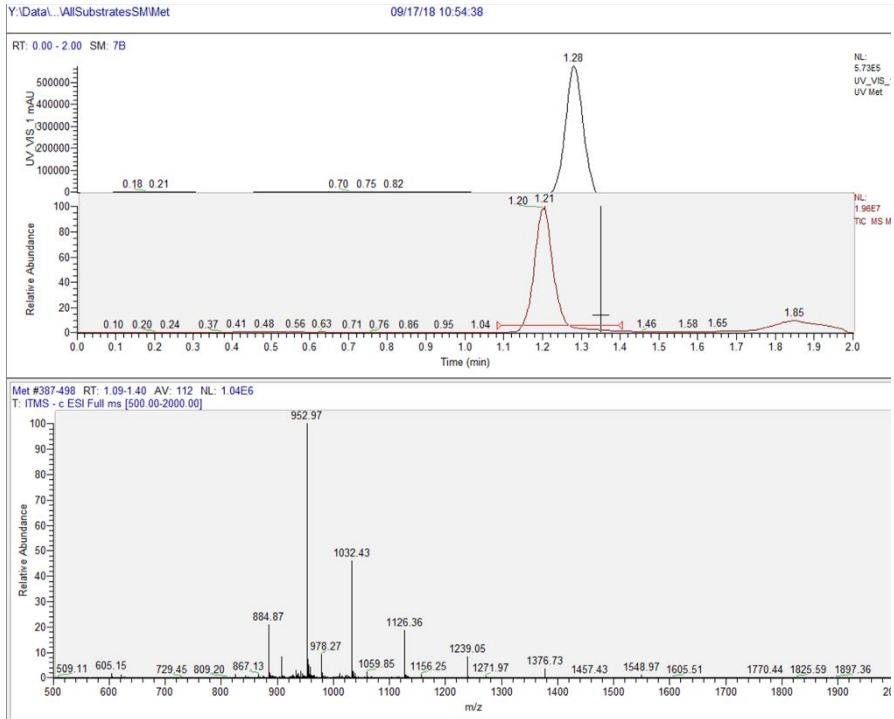
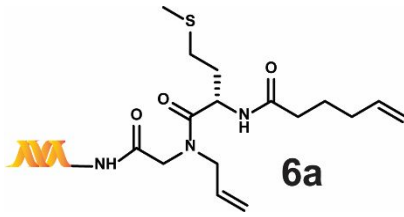


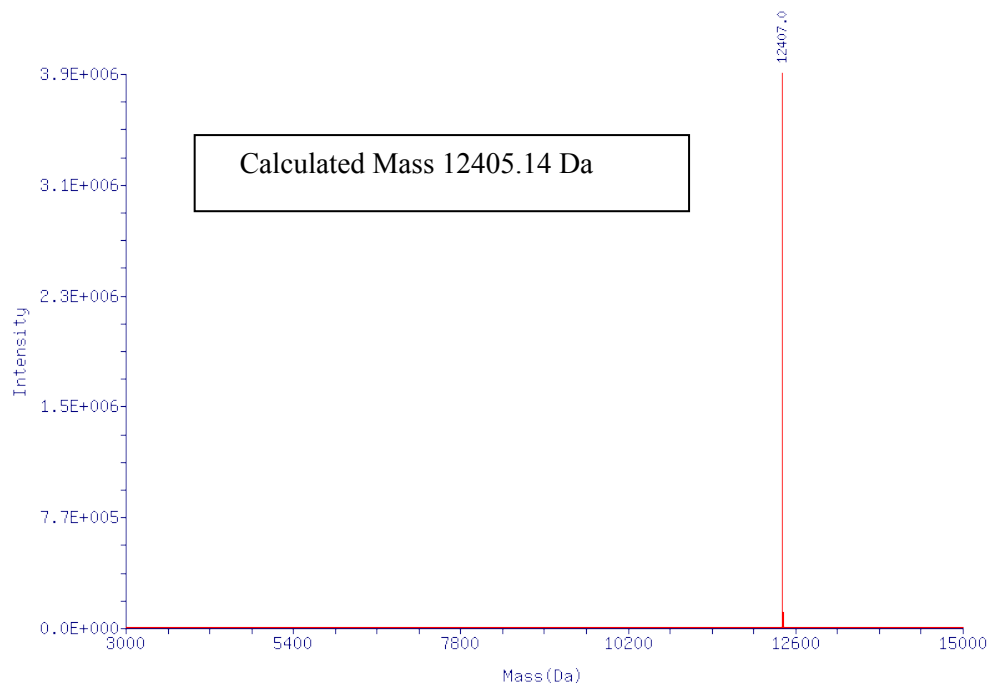
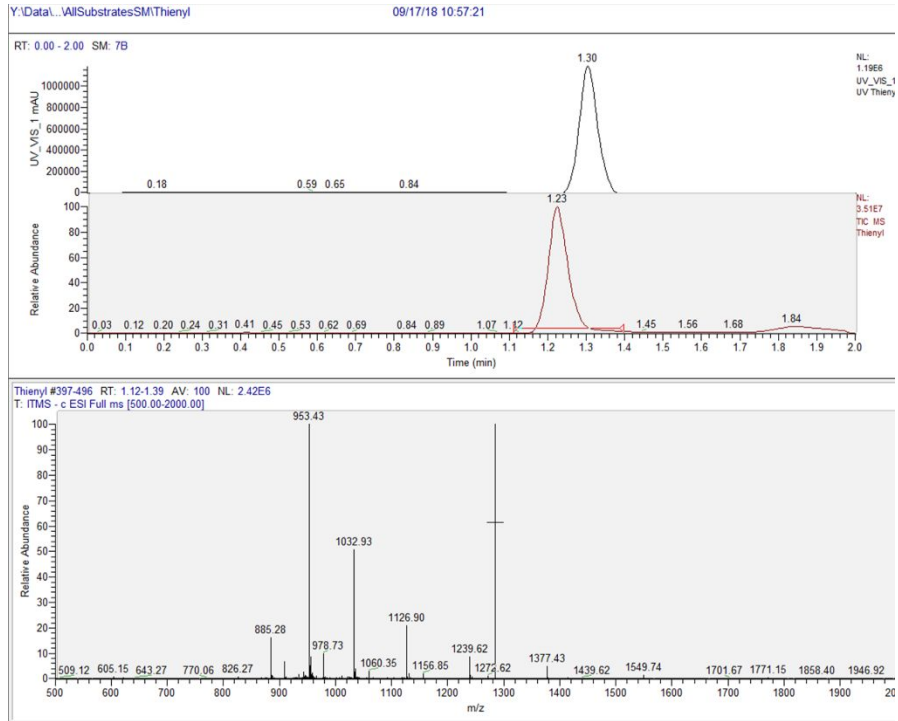
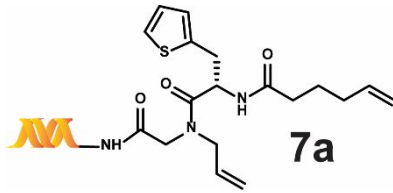


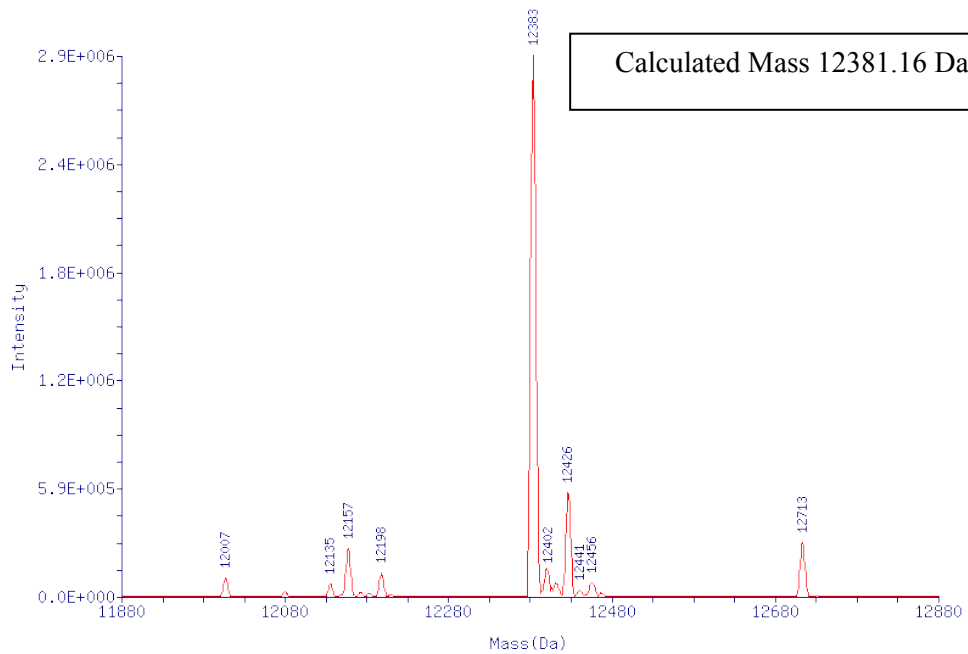
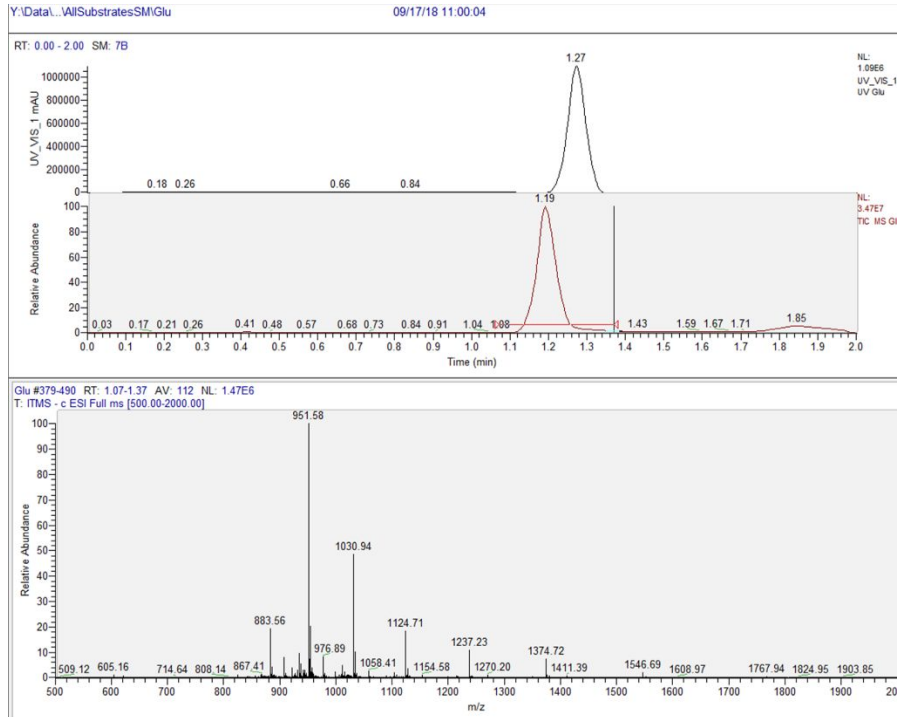
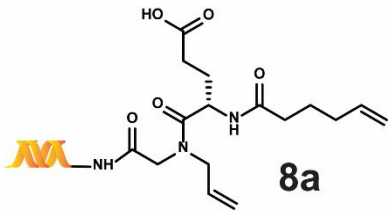


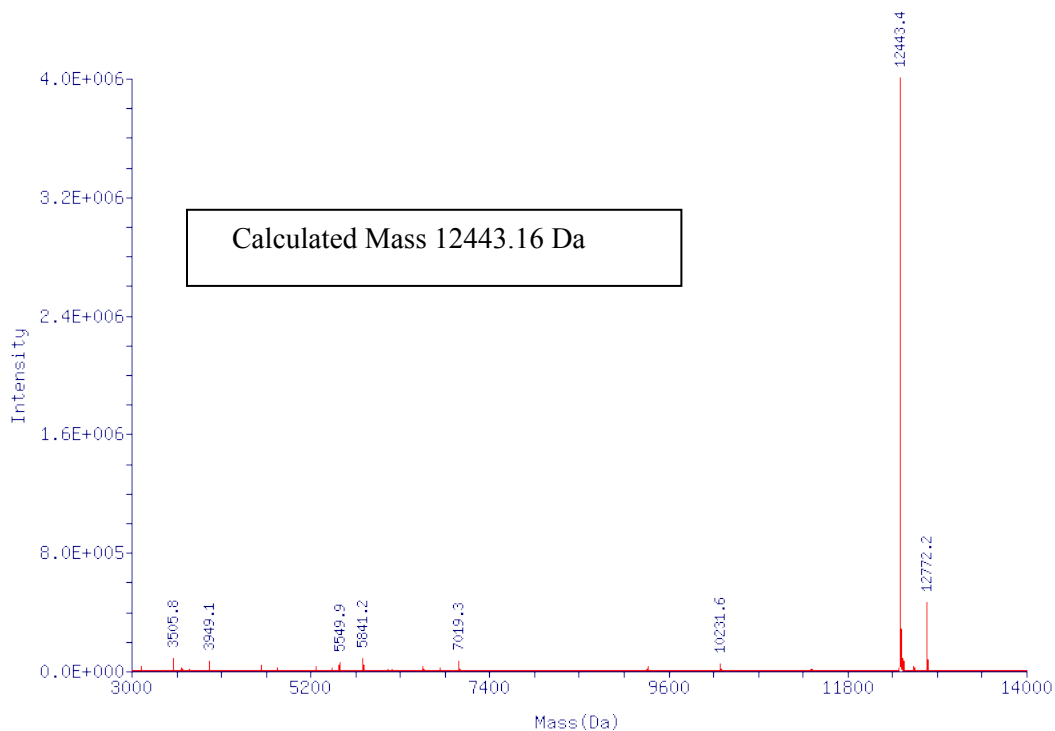
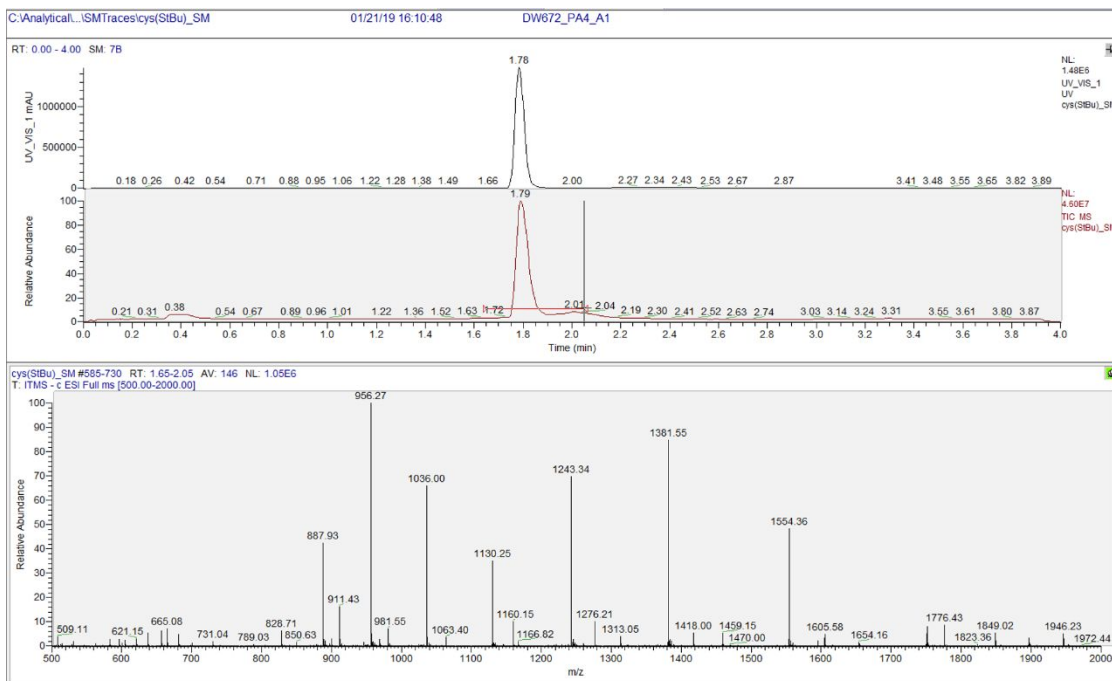
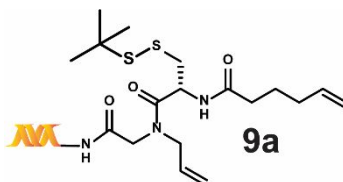


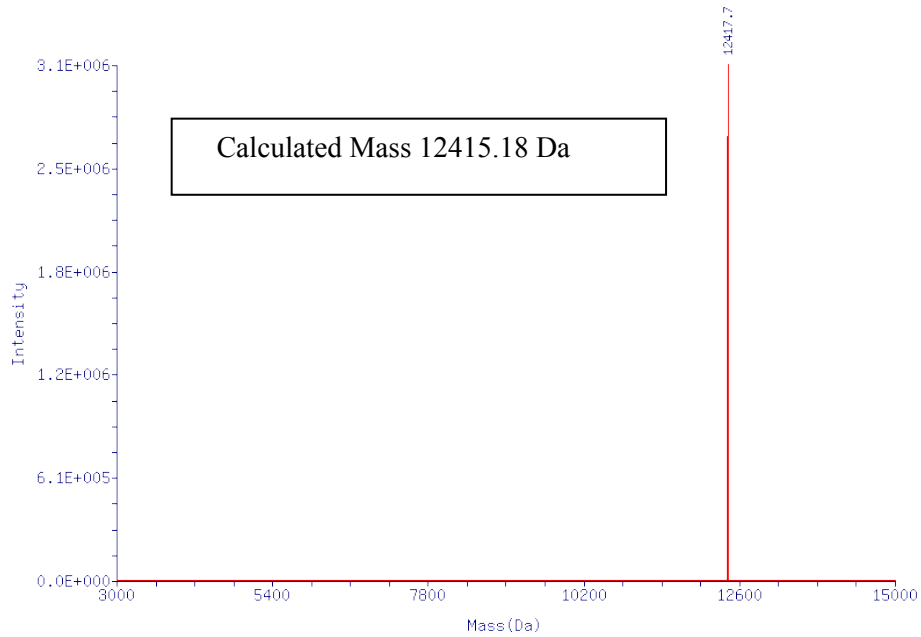
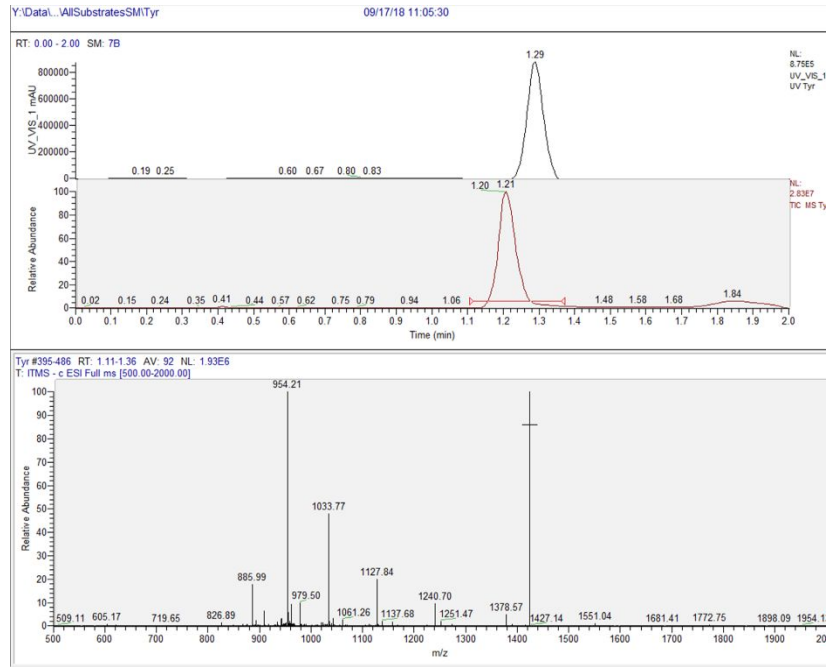
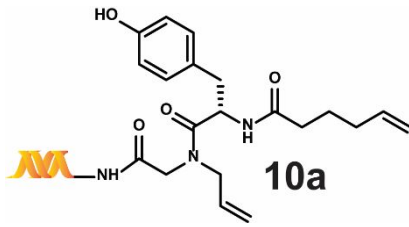


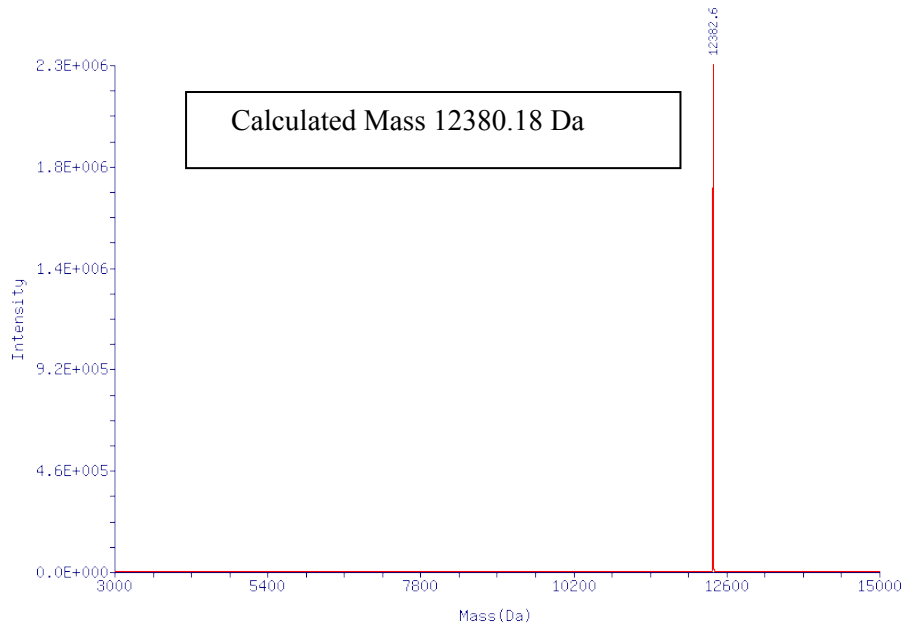
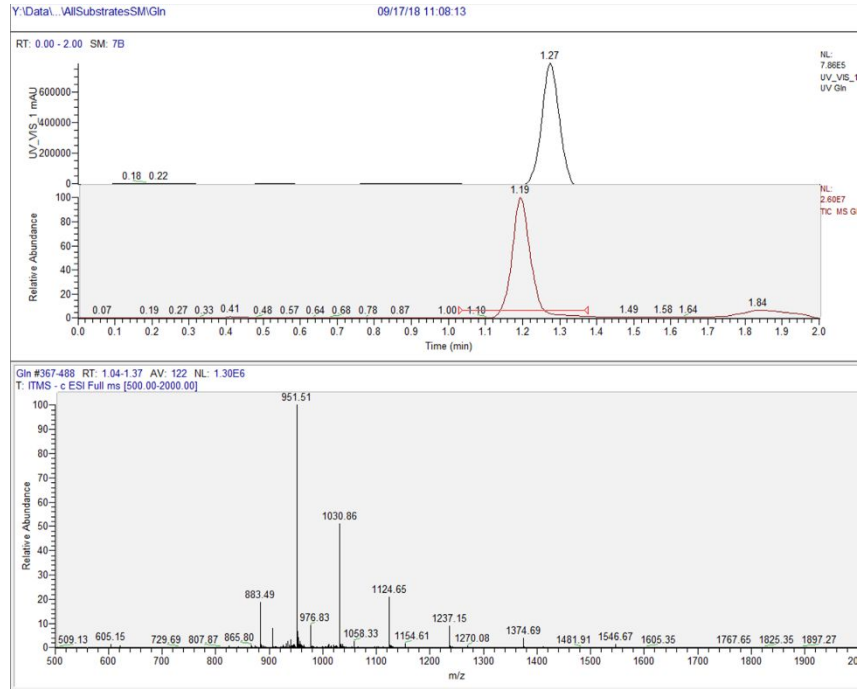
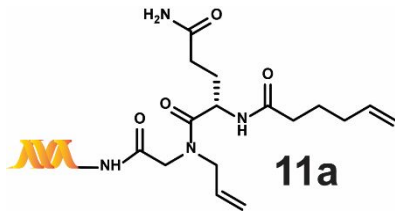


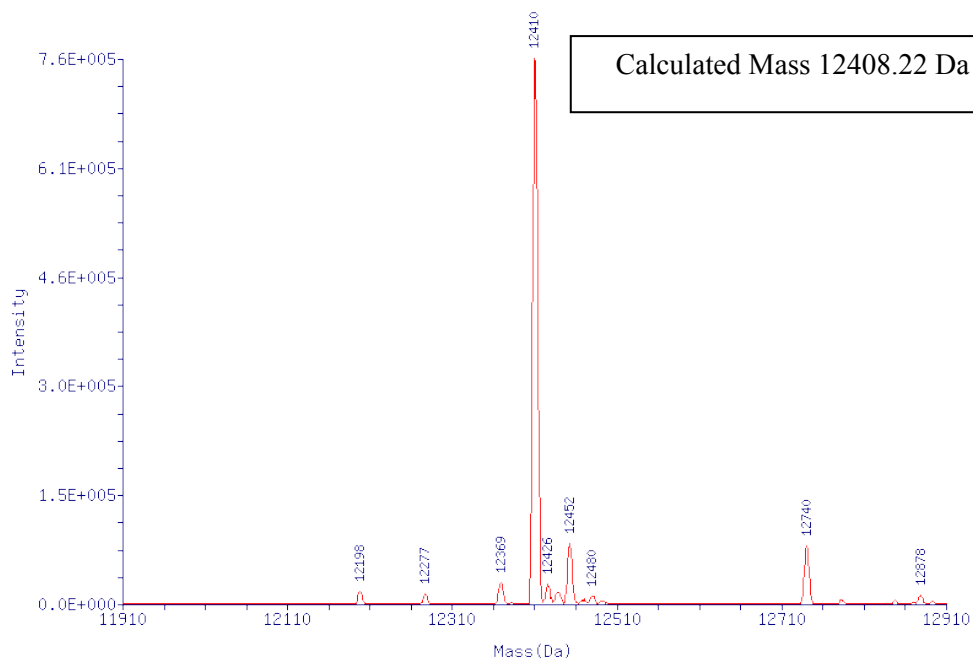
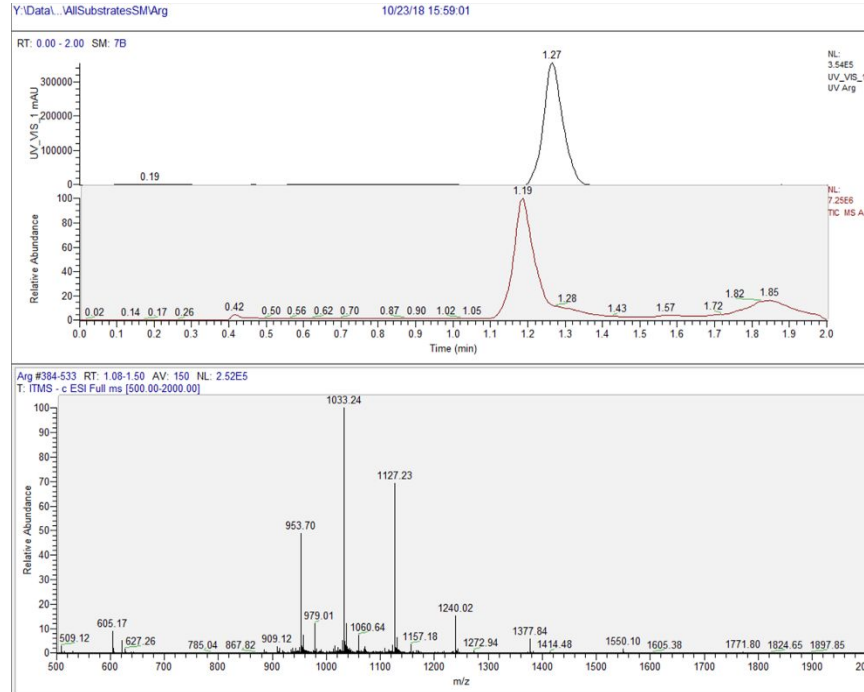
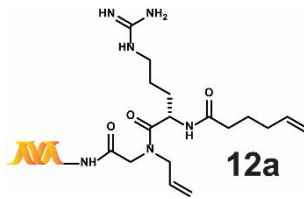


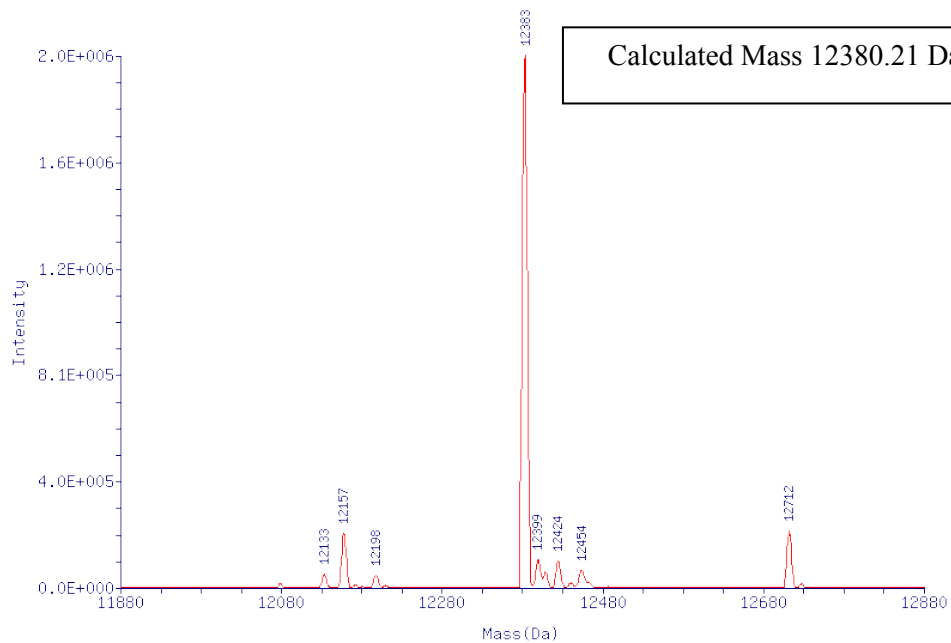
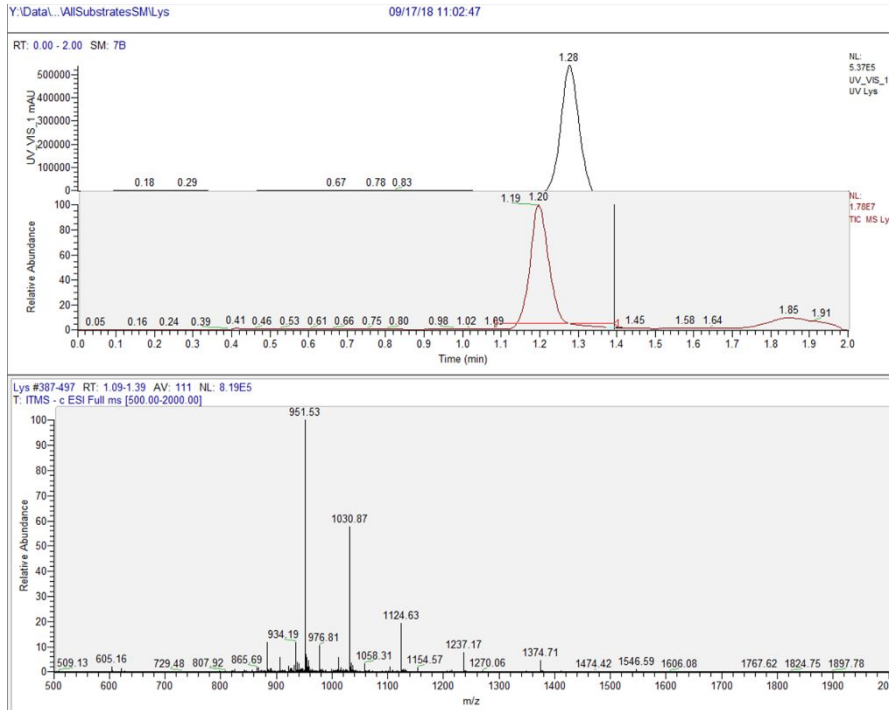
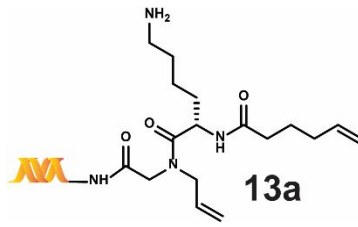


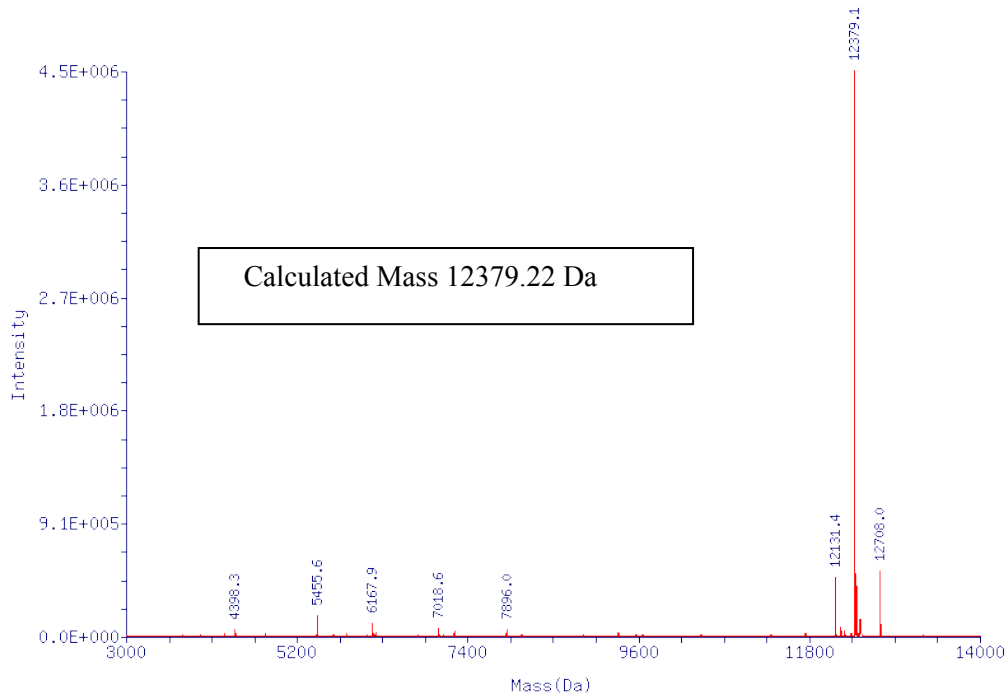
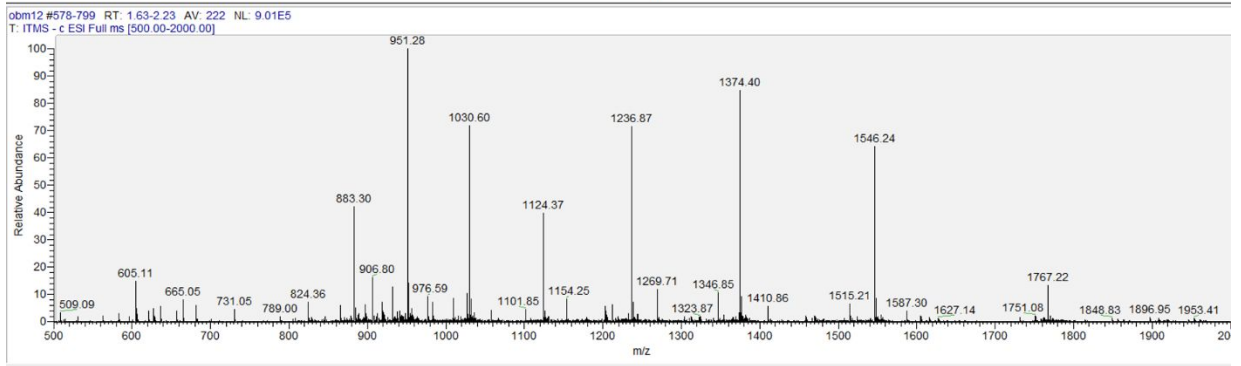
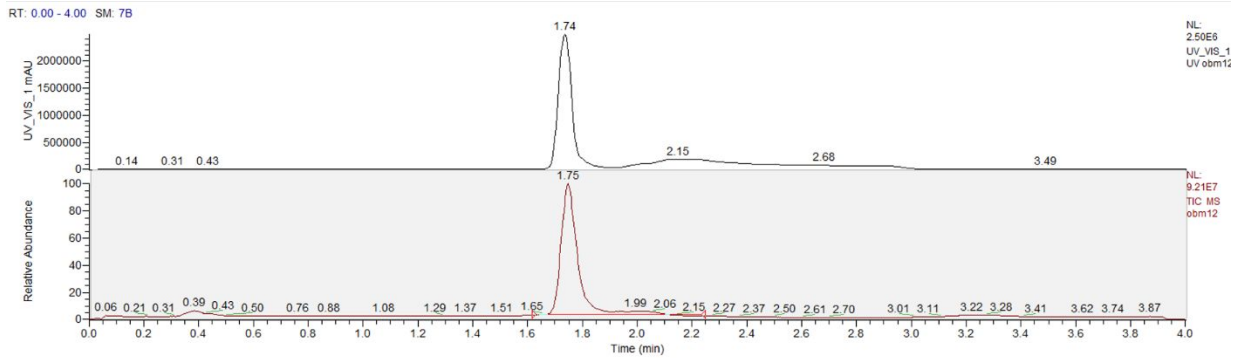
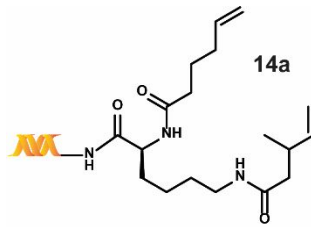


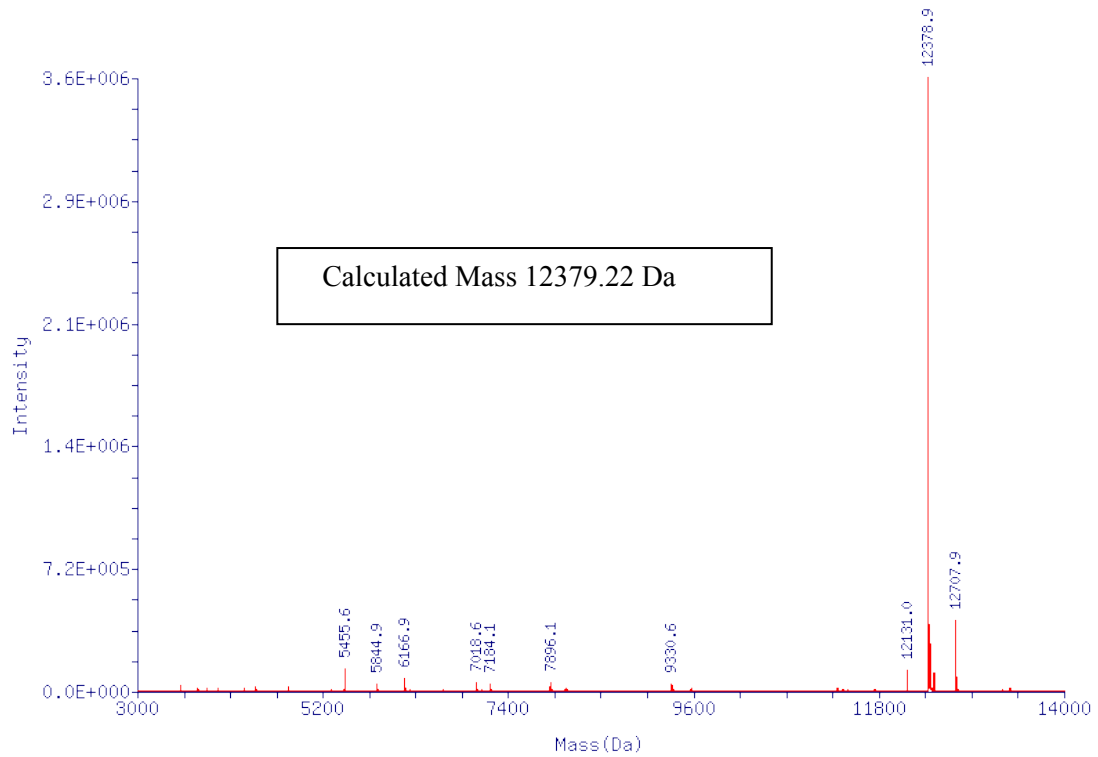
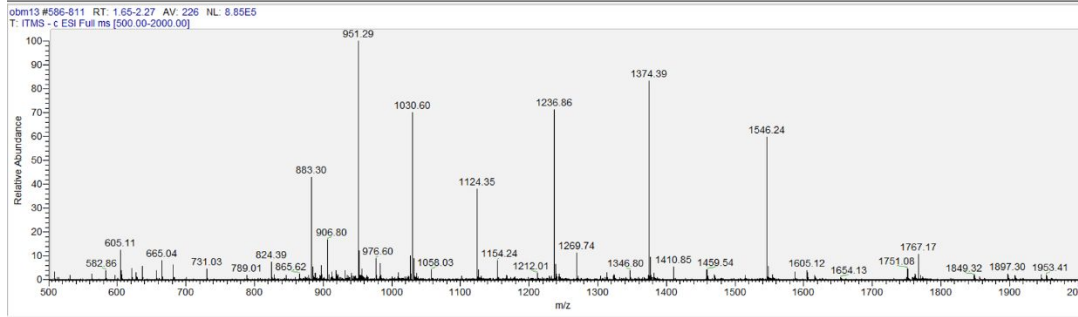
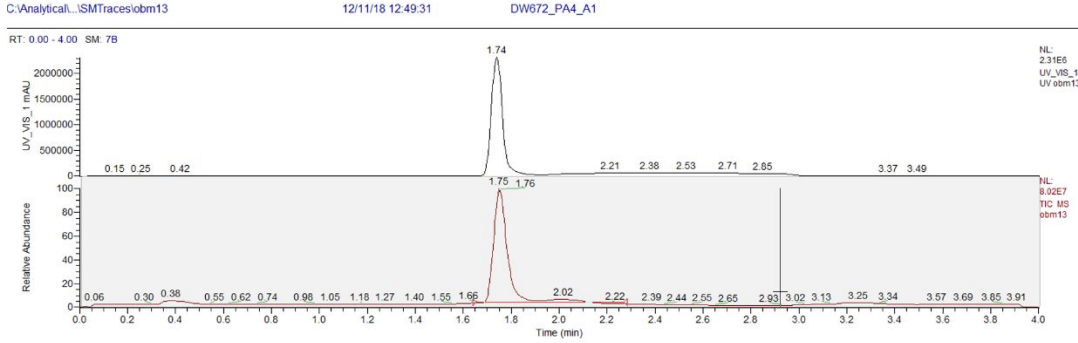
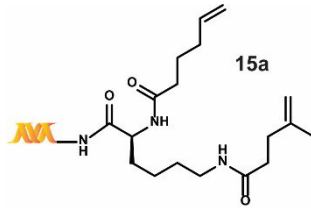


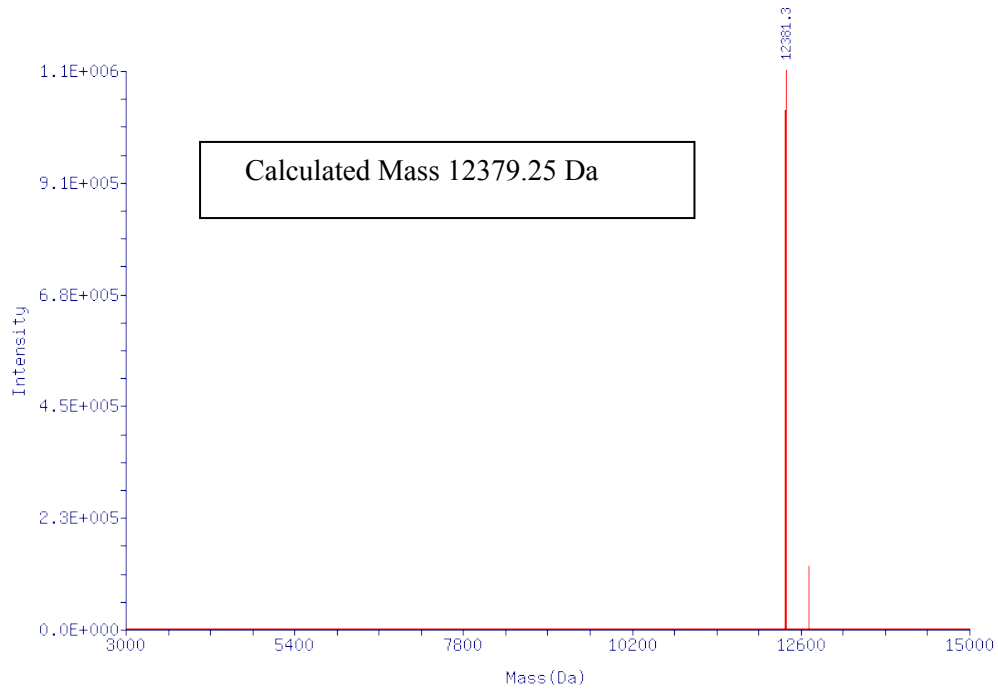
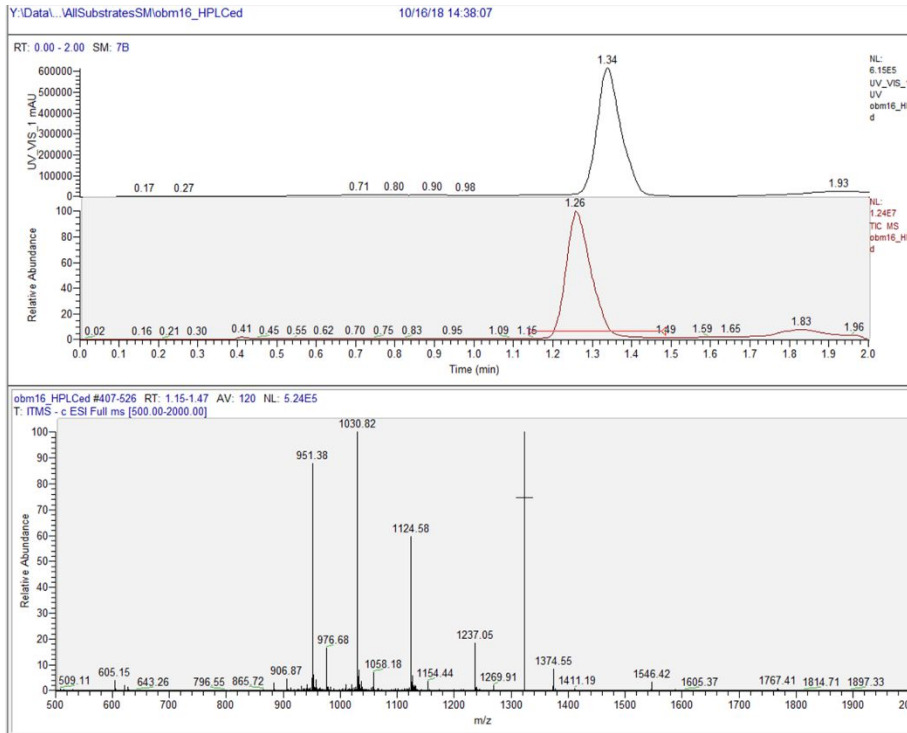
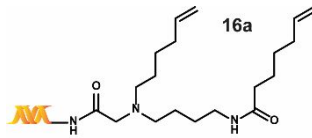


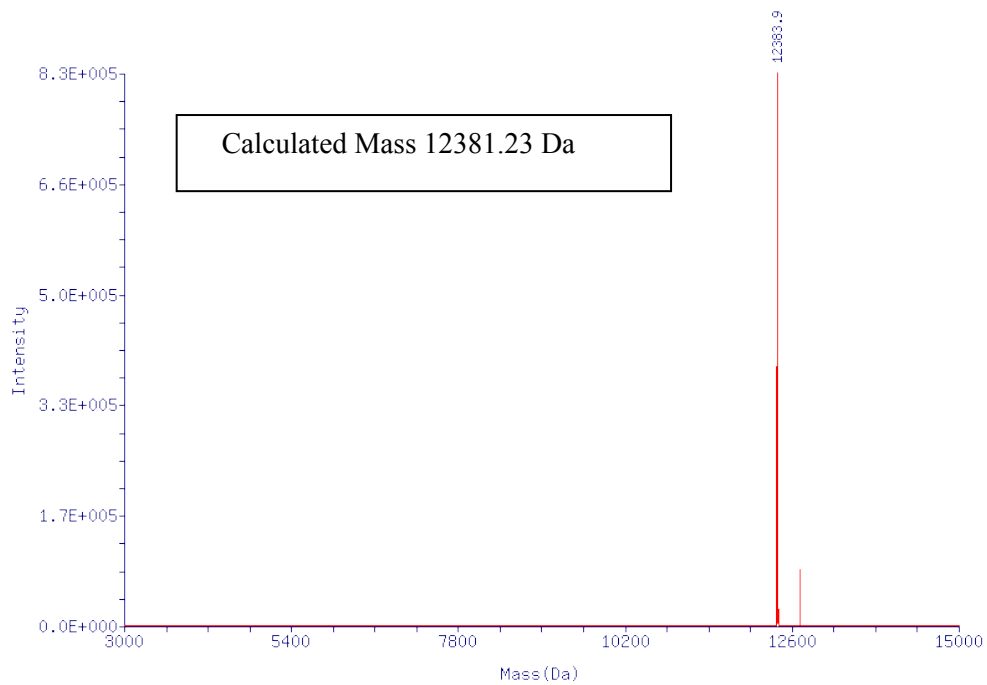
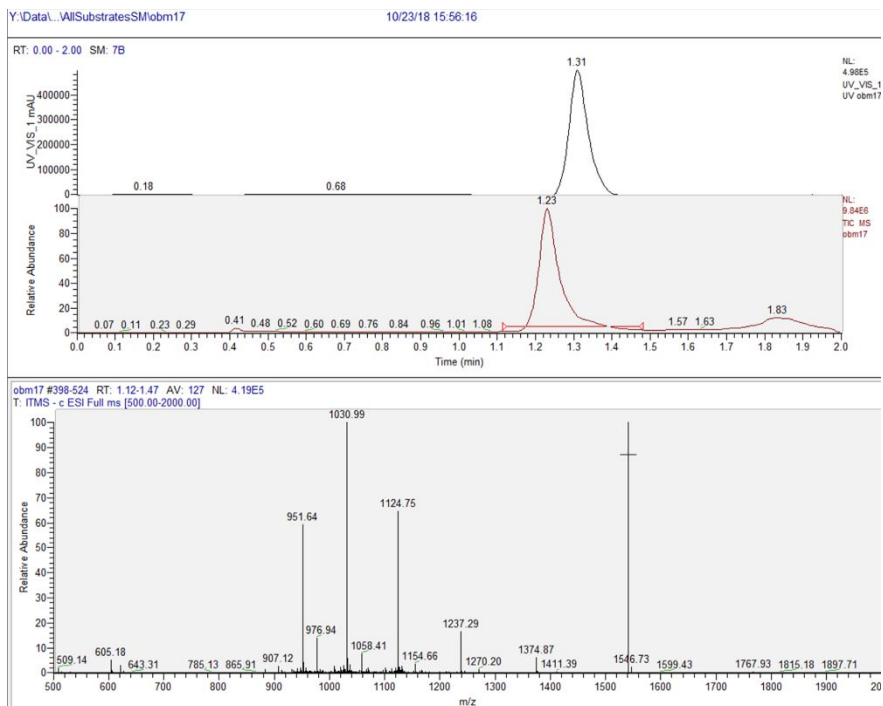
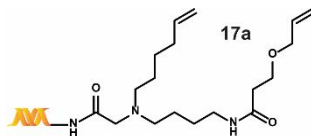


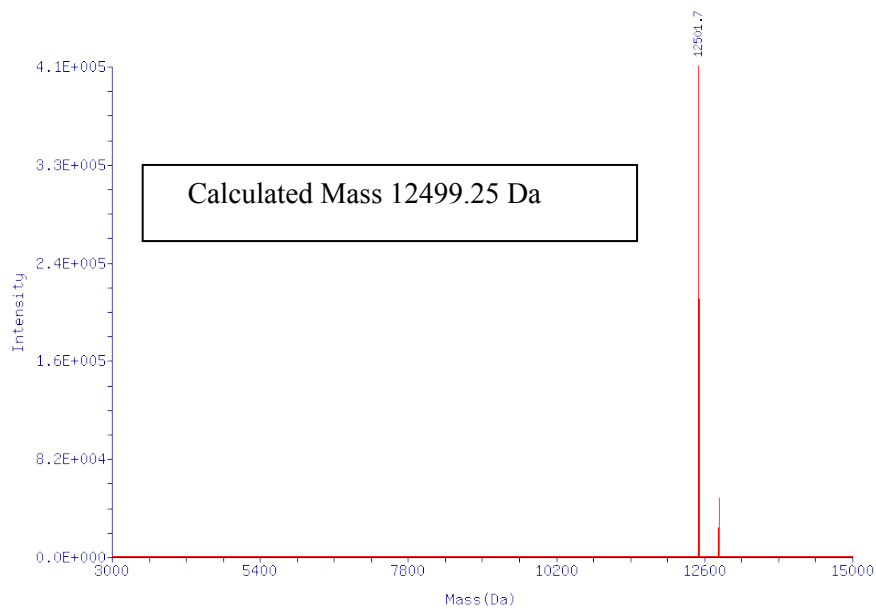
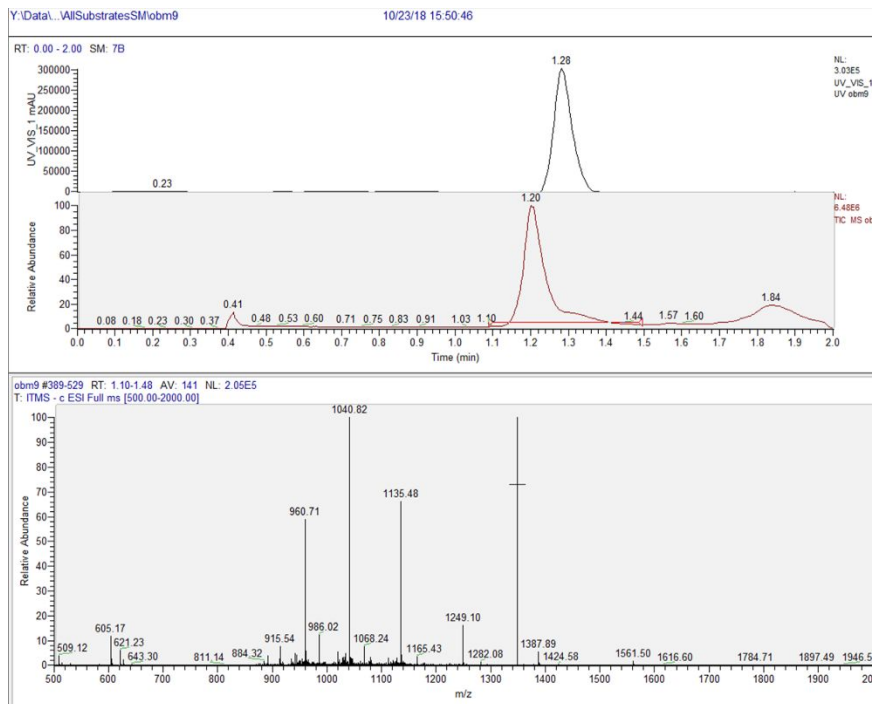
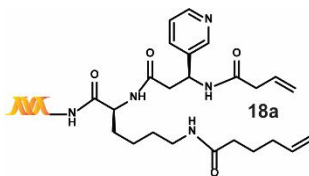


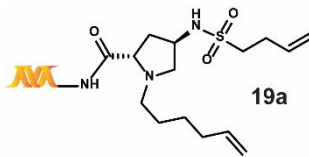








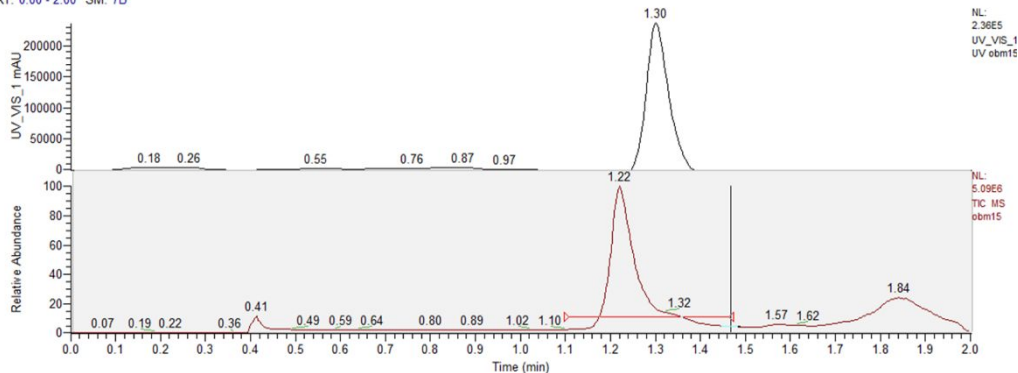




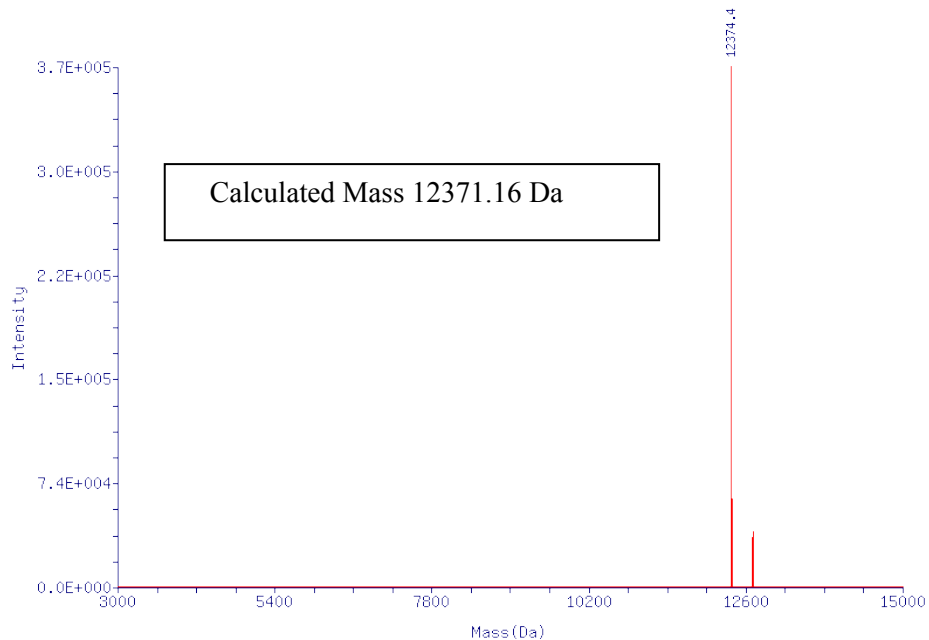
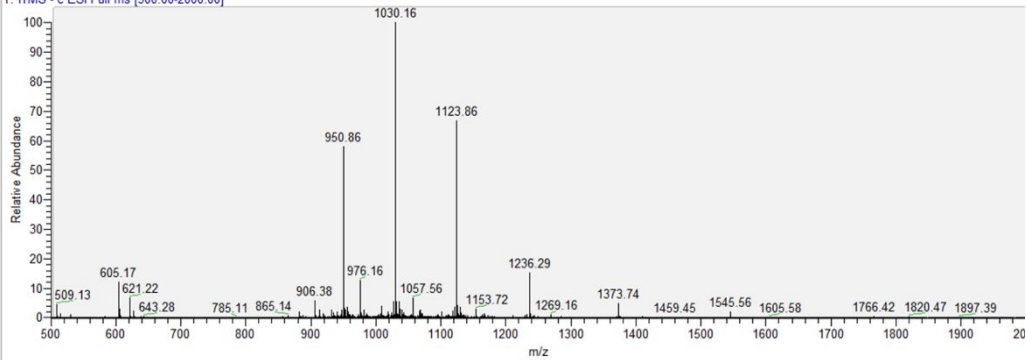
Y:\Data\...\AllSubstratesSM\obm15

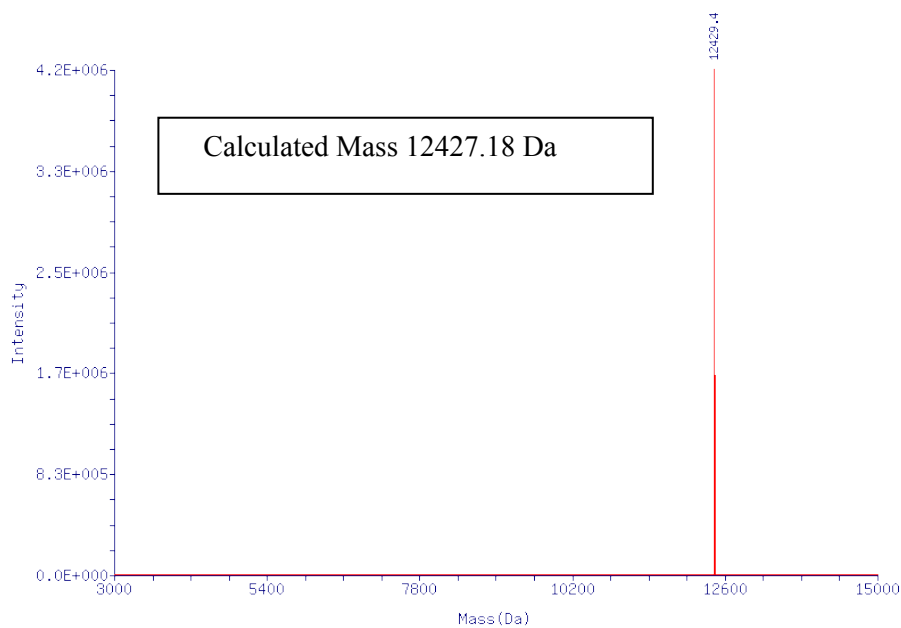
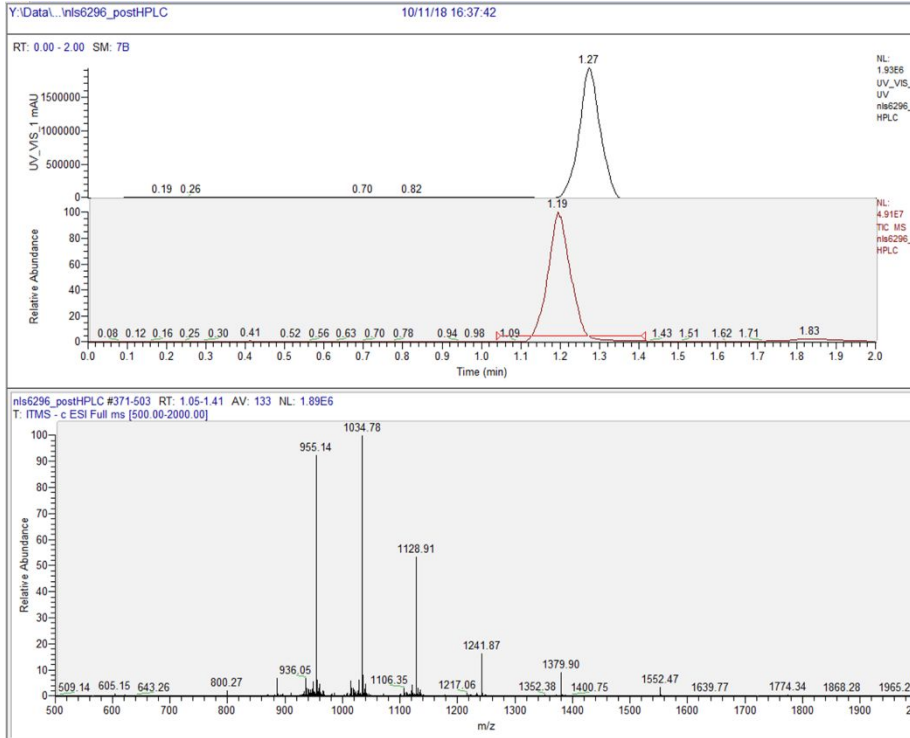
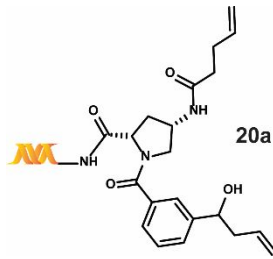
10/23/18 15:53:31

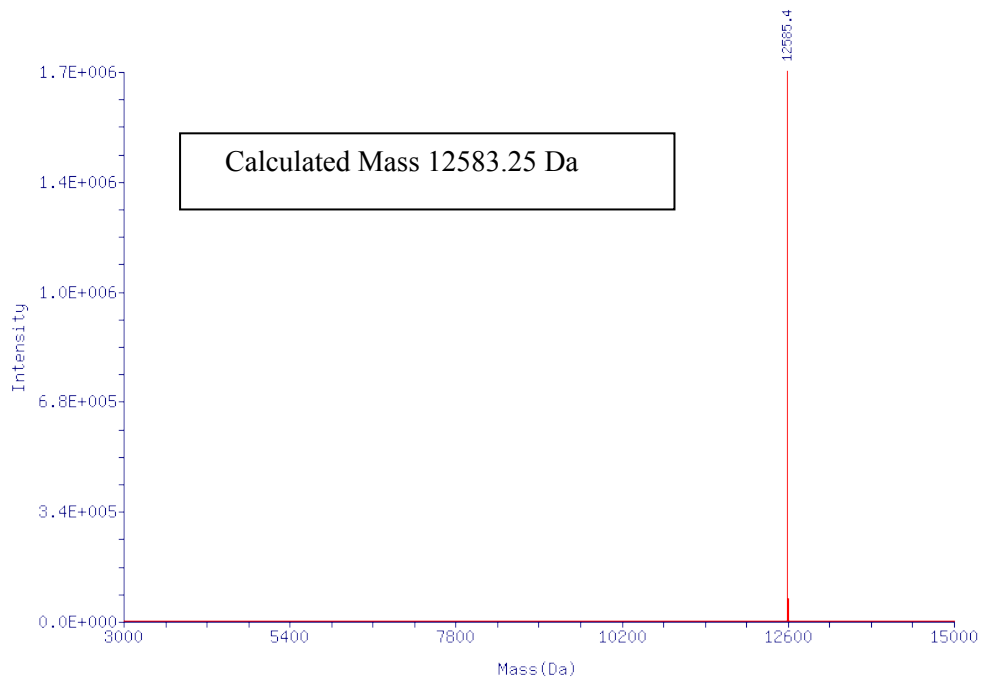
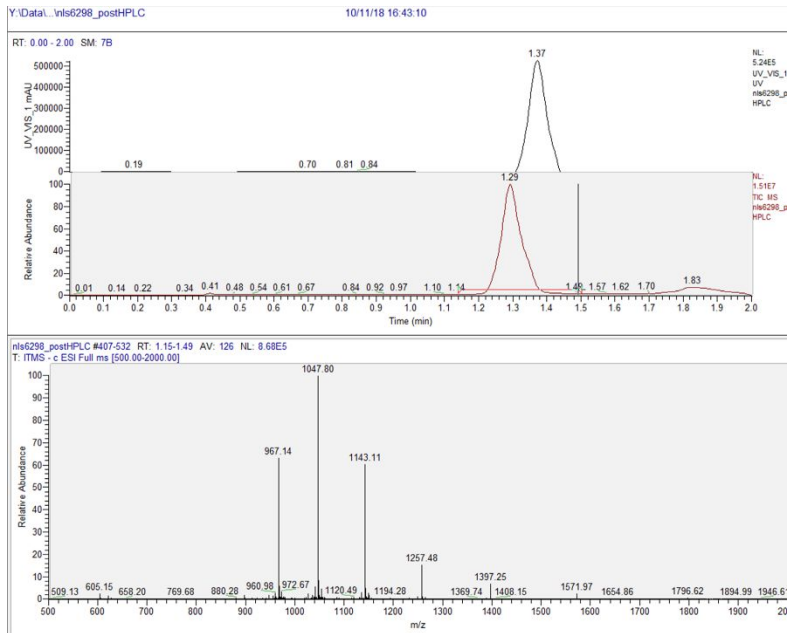
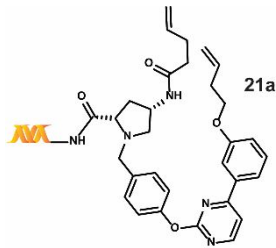
RT: 0.00 - 2.00 SM: 7B

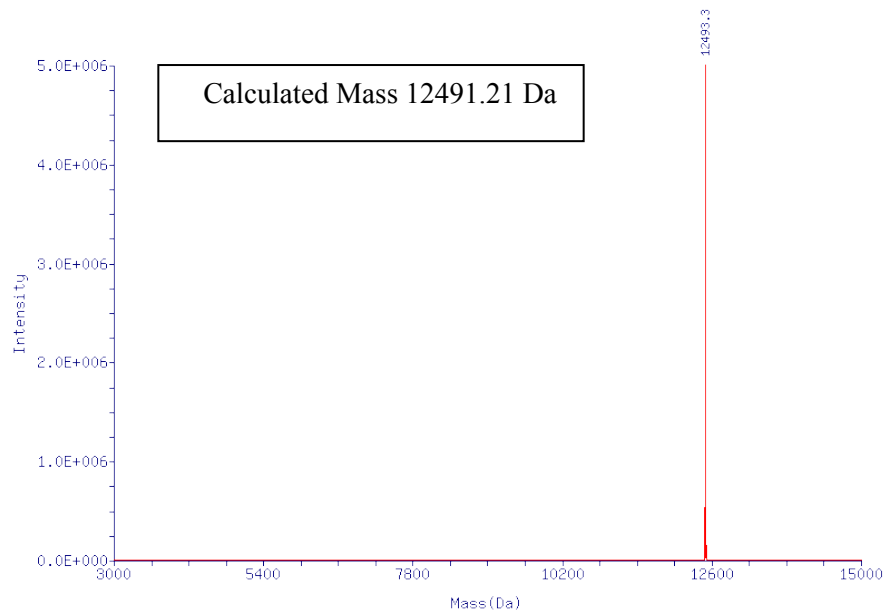
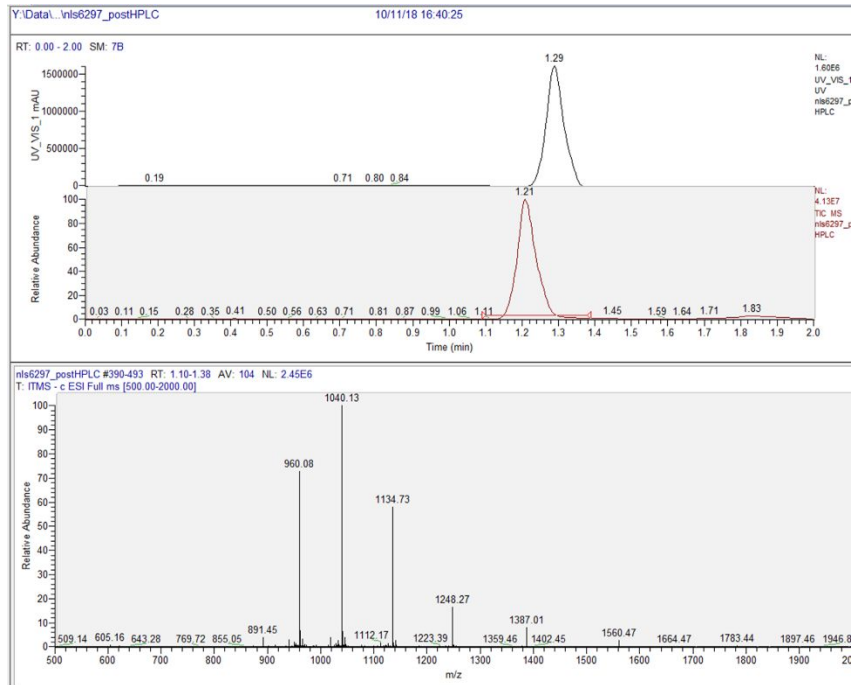
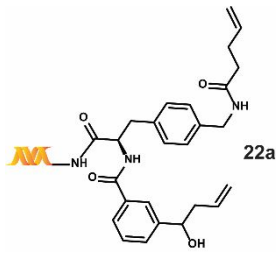


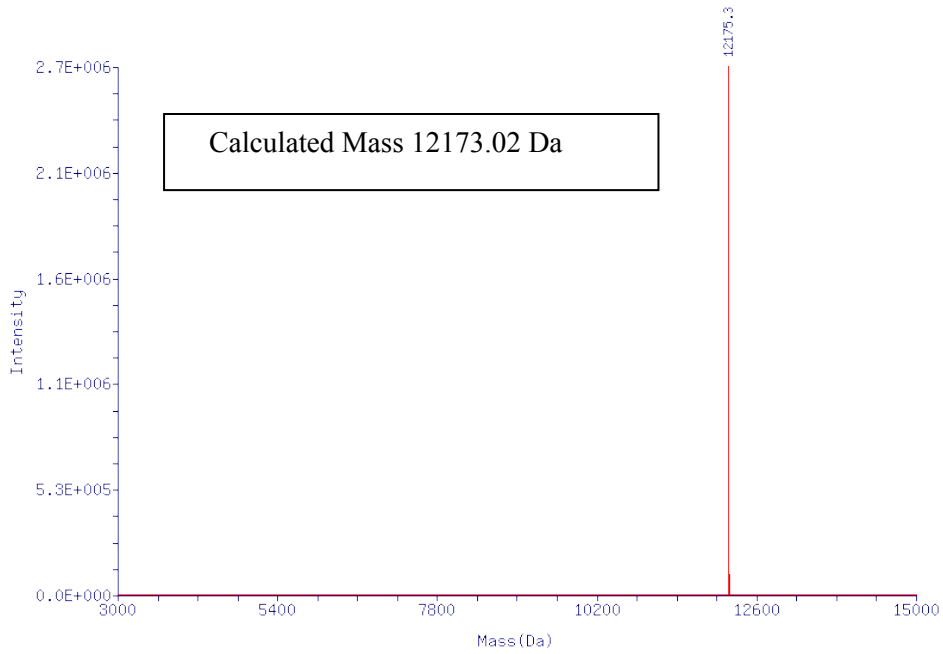
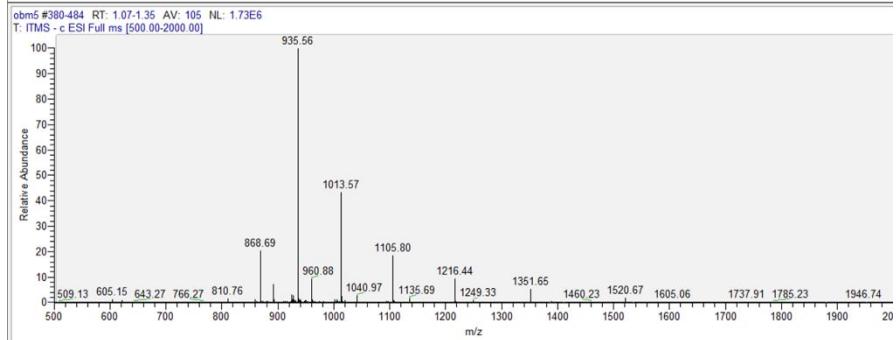
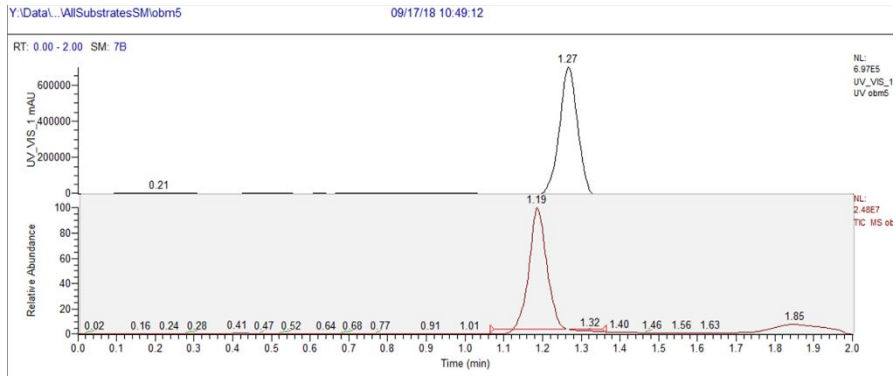
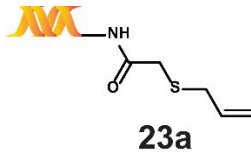
obm15 #392-521 RT: 1.11-1.46 AV: 130 NL: 1.93E5
T: ITMS - c ESI Full ms [500.00-2000.00]

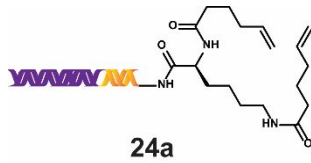




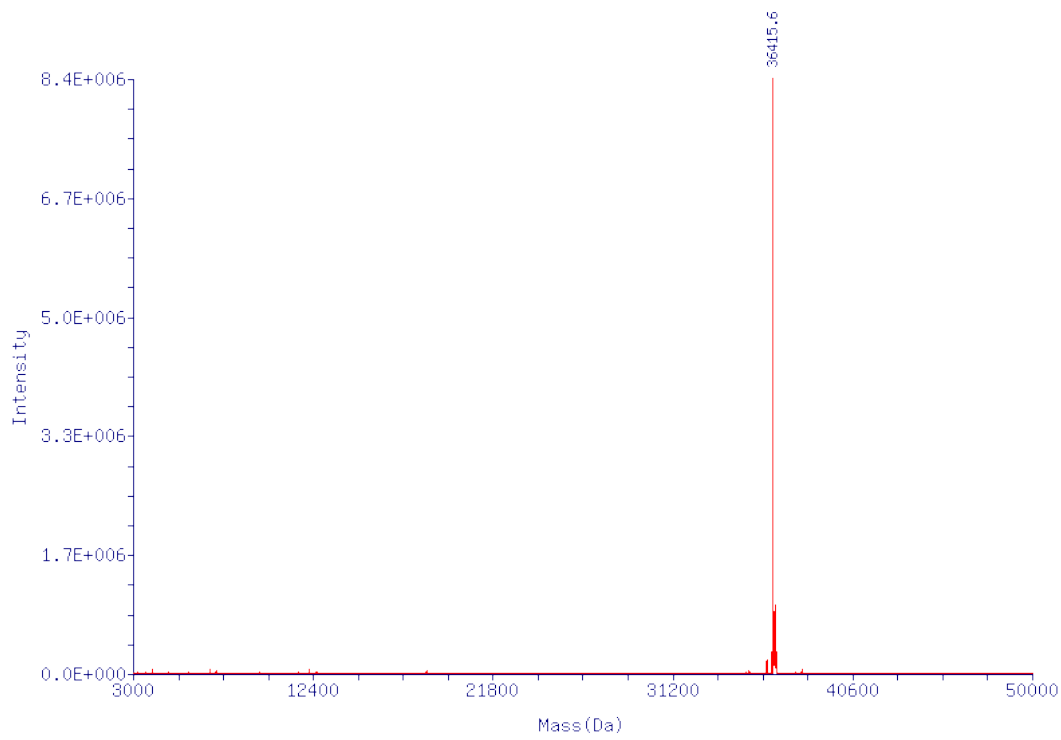
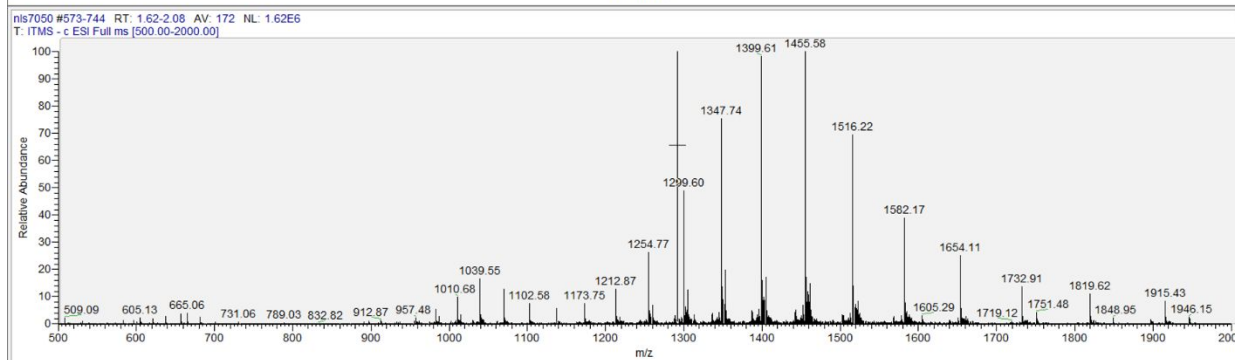
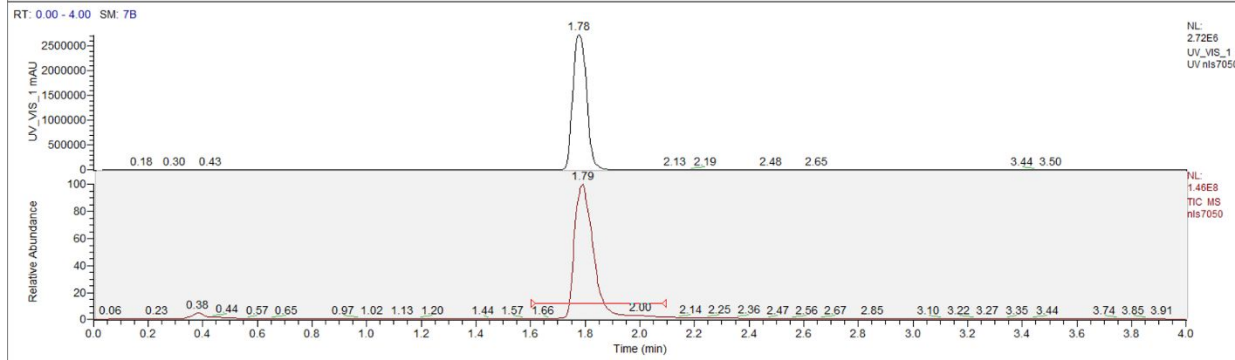


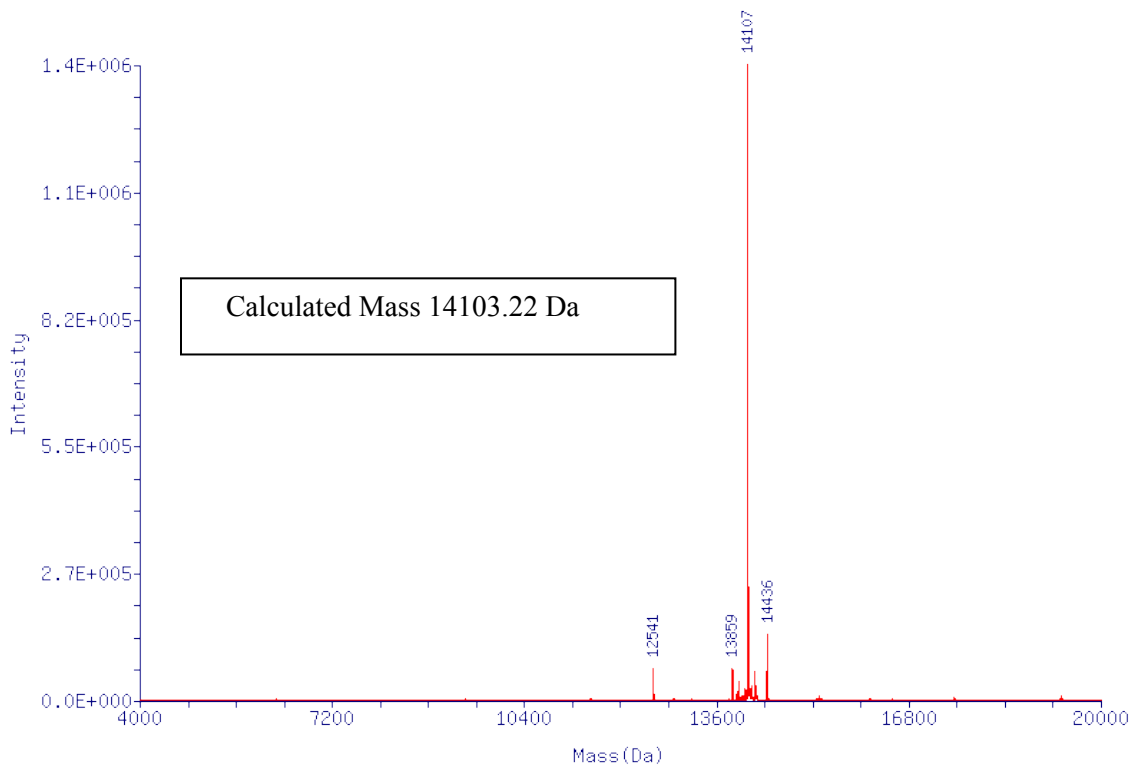
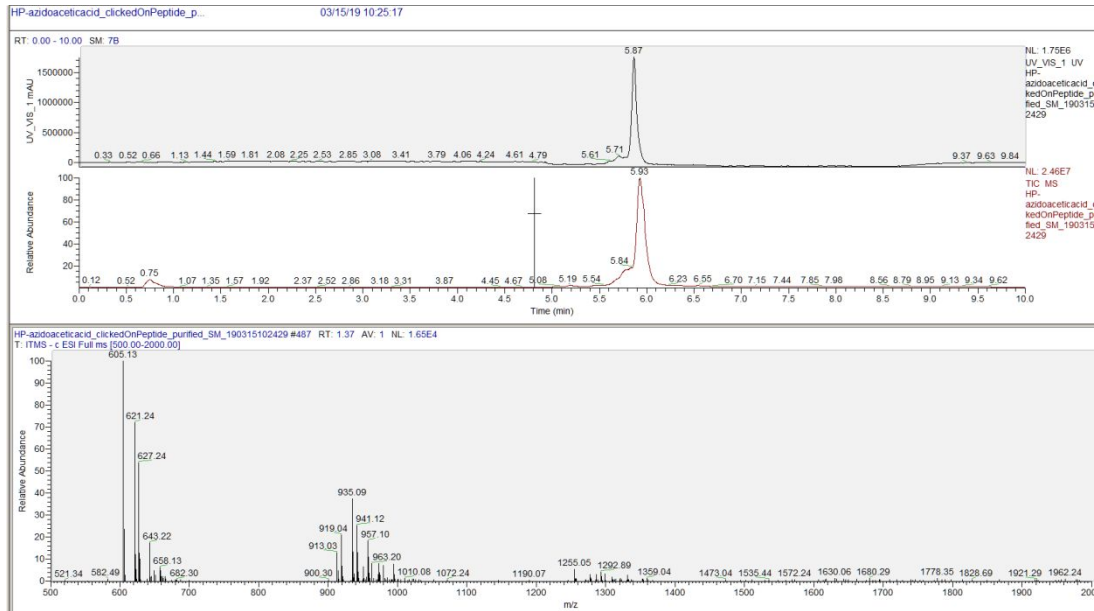
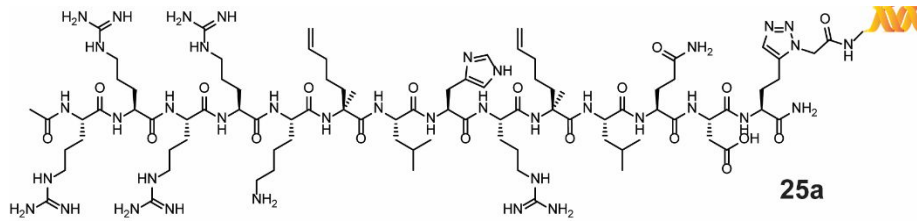


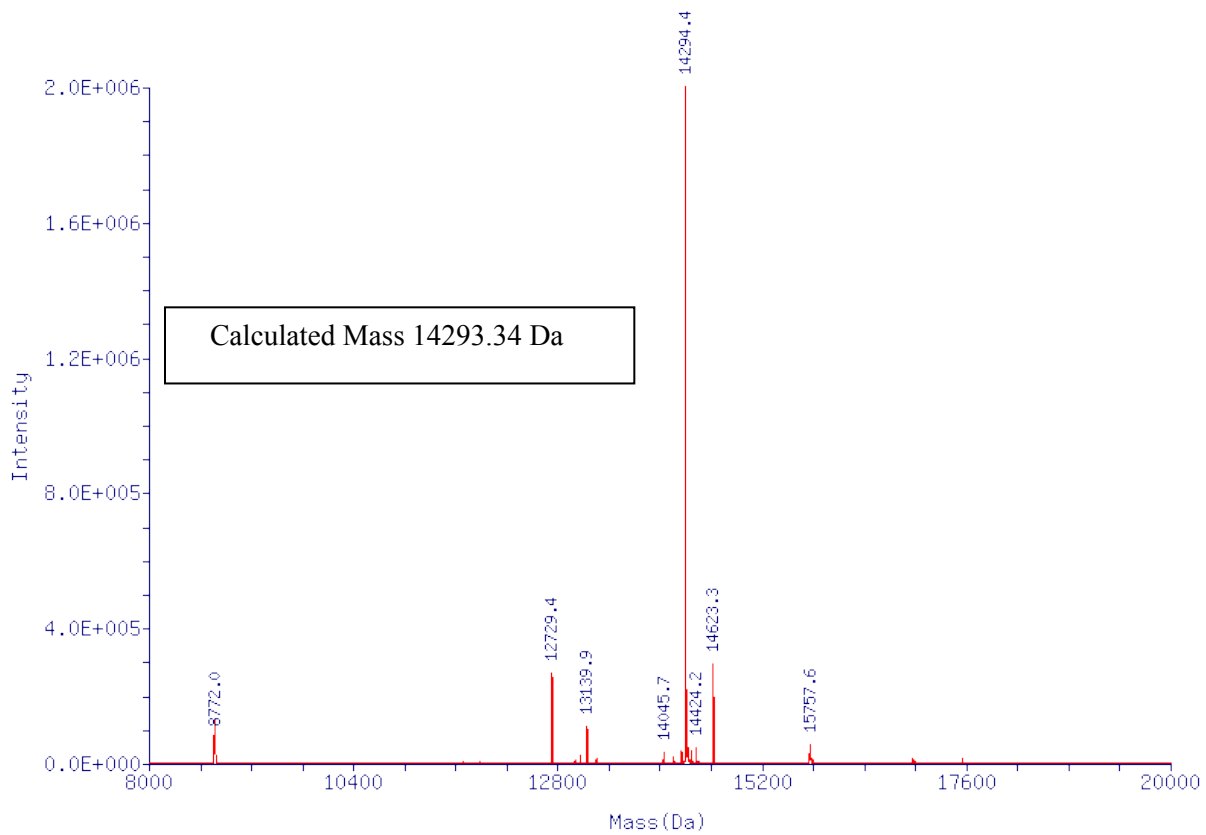
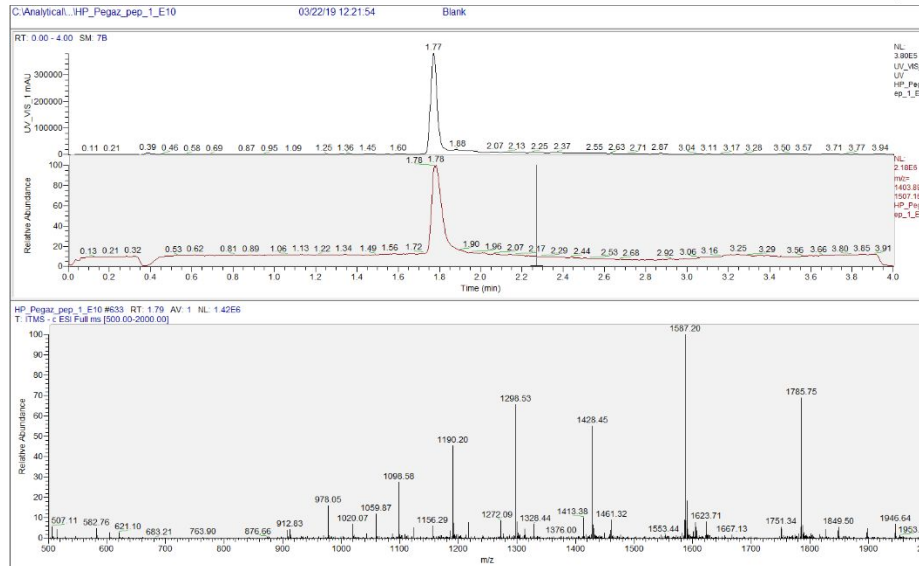
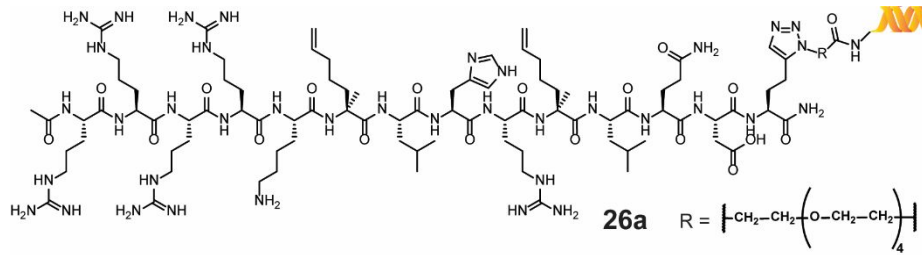


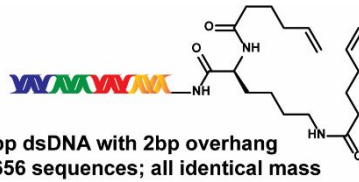


C:\Analytical...ISMTraces\ms7050 01/22/19 11:01:41 DW672_PA4_A1

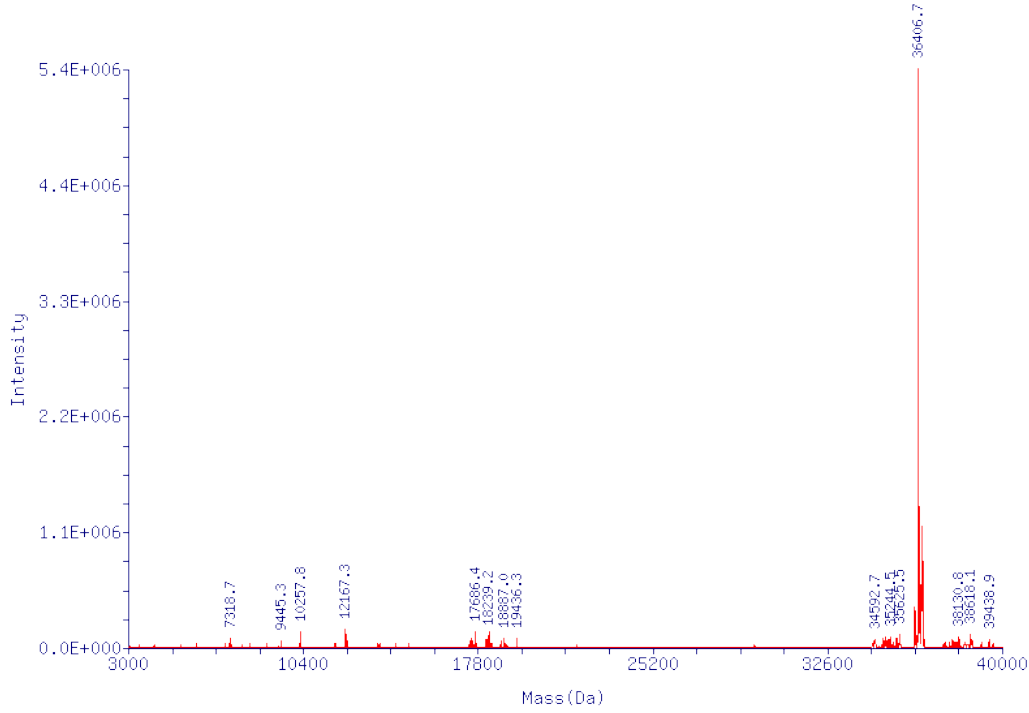
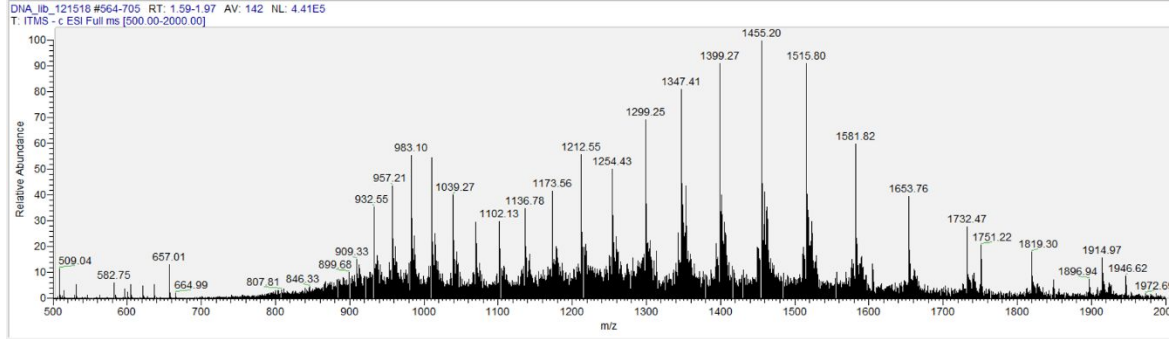
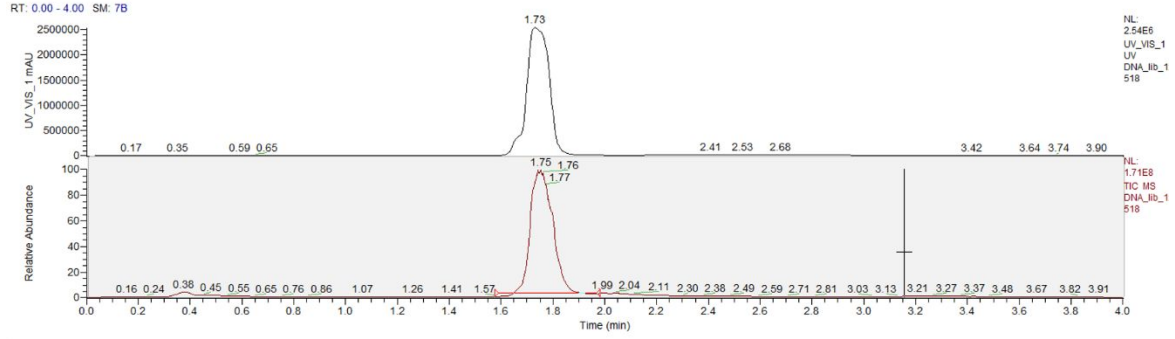








C:\Analytical\DNA_lib_121518 12/15/18 12:55:33 DW672_PA4_A1

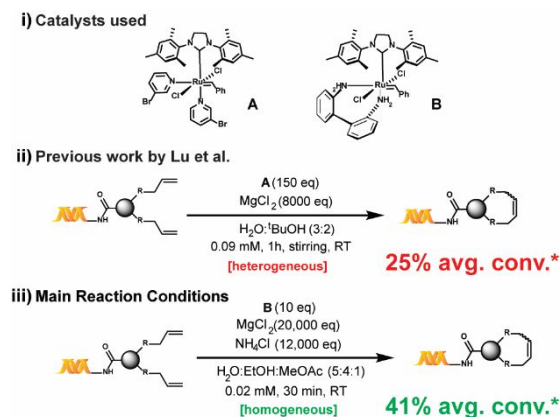


Post-Metathesis (RCM and CM) MS Deconvolution Spectra

Except for the one-substrate (DNA ‘only’) library, the spectra are arranged in quadrants, each of which is labeled with the reaction conditions used. The labels refer to schemes and figures used in the main paper, although those have been included here as well.

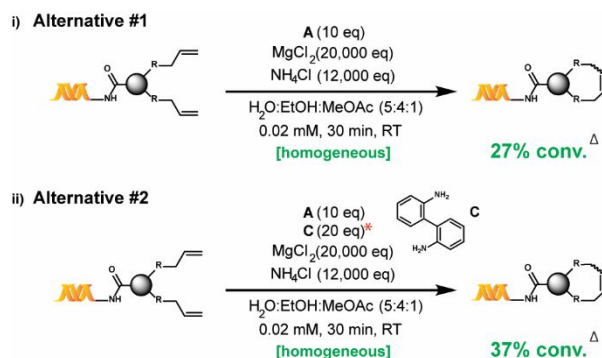
Note that two LCMS instruments were used for analysis, one which gives M+2 masses while the other M masses. Most runs were carried out on the latter.

Scheme 1. Overall Performance of Reaction Conditions Developed Herein Relative to that of Previously Reported Work.

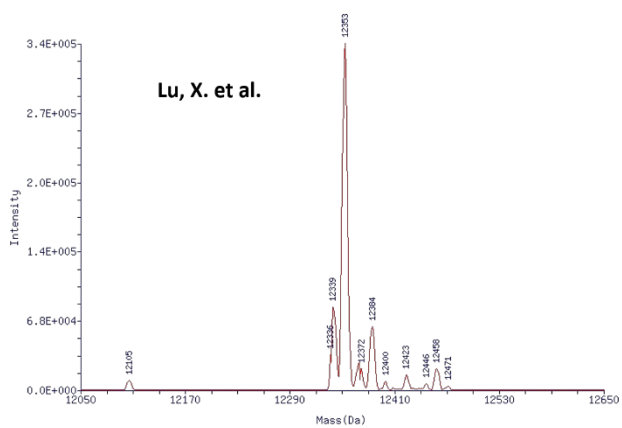
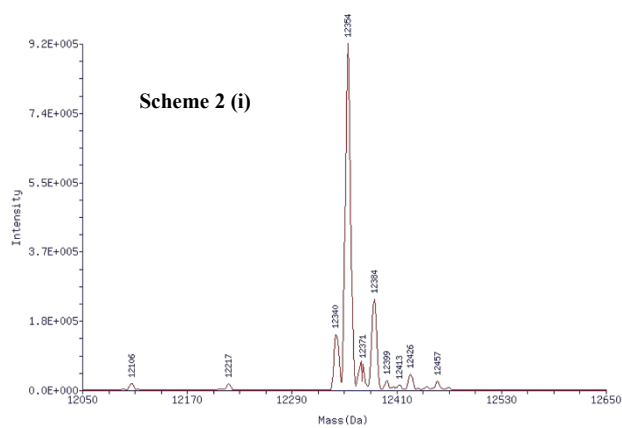
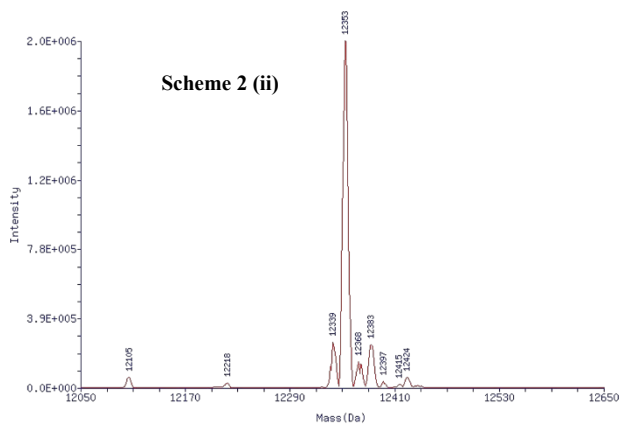
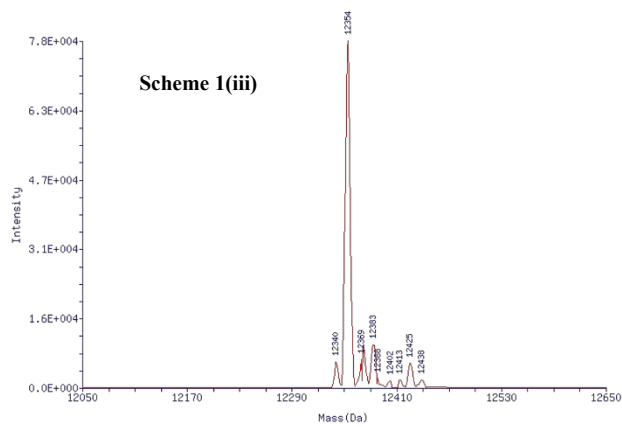
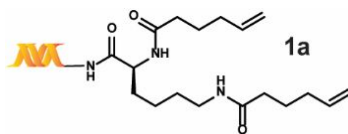


i) Grubbs 3rd generation catalyst **A** and its 2,2'-biphenyldiamine derivative **B**; ii) Previously reported conditions for the on-DNA RCM and CM reactions; iii) Our main conditions for the on-DNA RCM and CM reactions; *Average percent conversion for the investigated substrate scope (22 substrates).

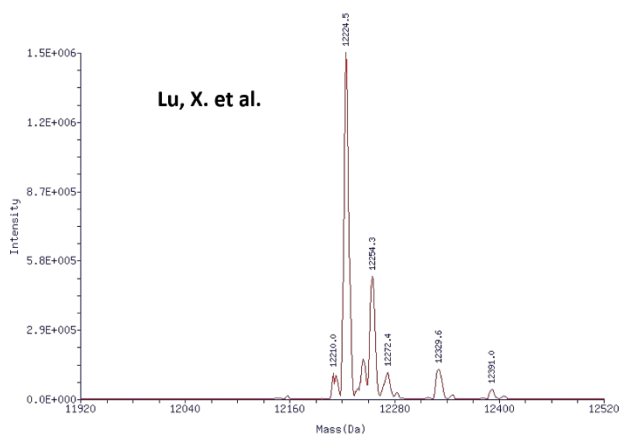
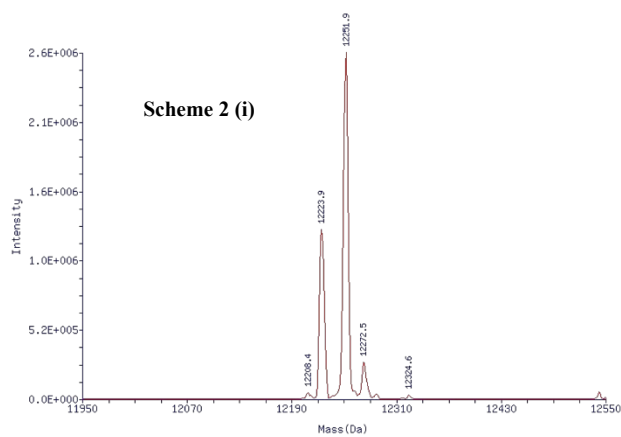
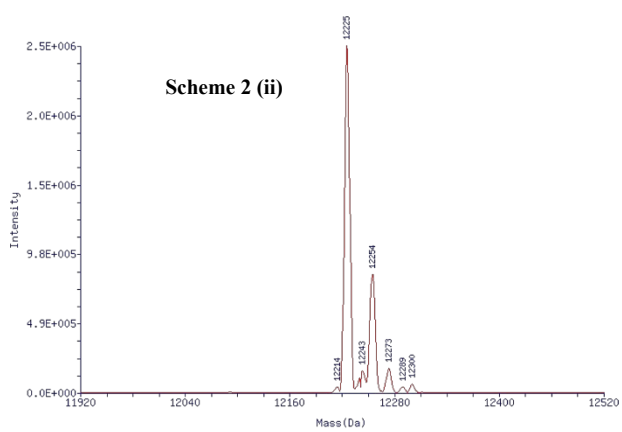
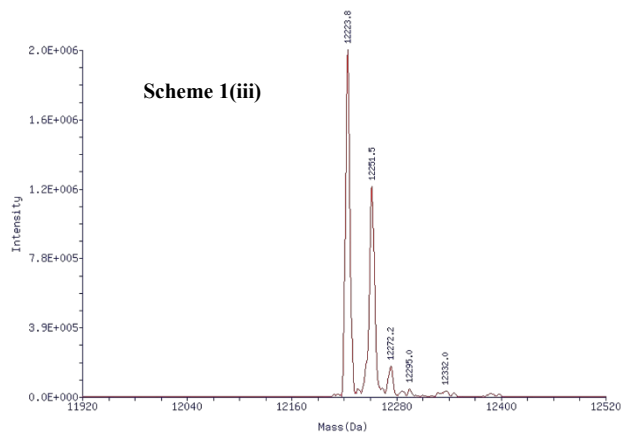
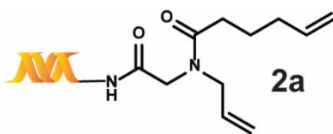
Scheme 2. Alternative Reaction Conditions with Catalyst A



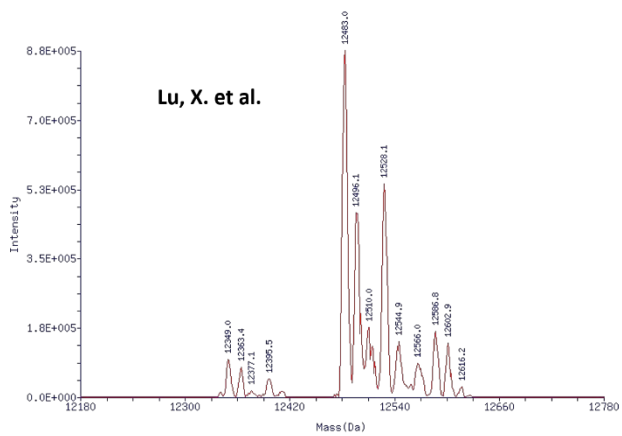
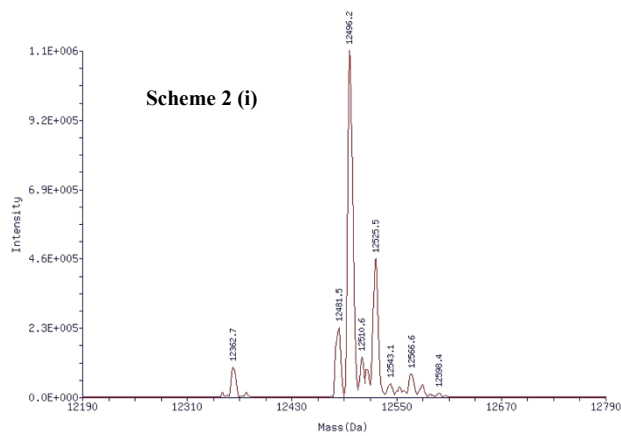
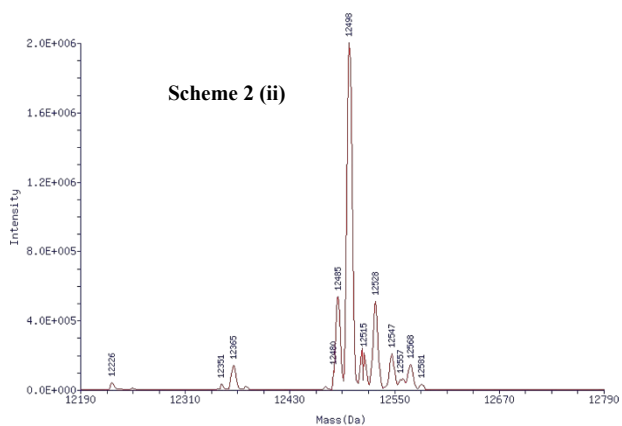
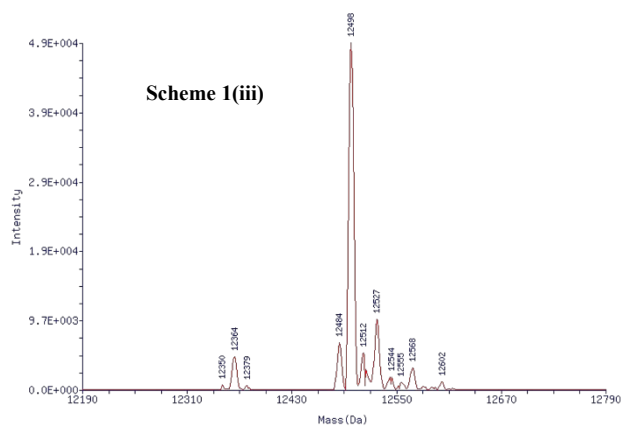
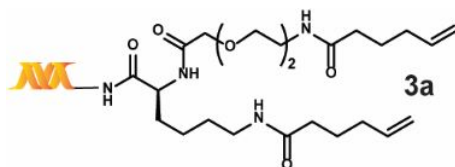
Calculated Product Mass
12351.19 Da



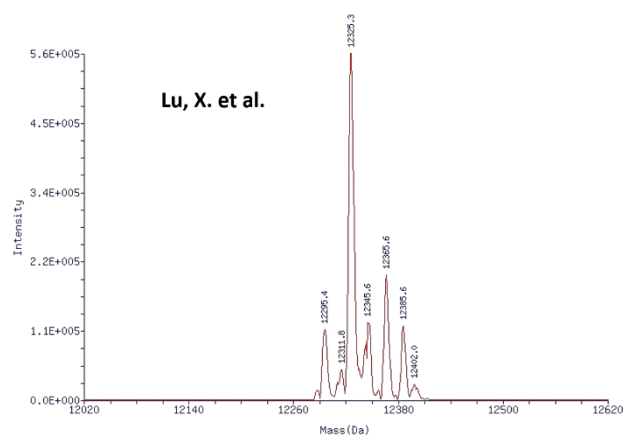
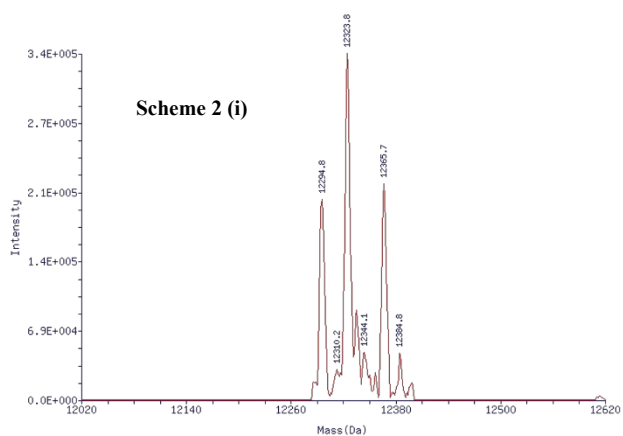
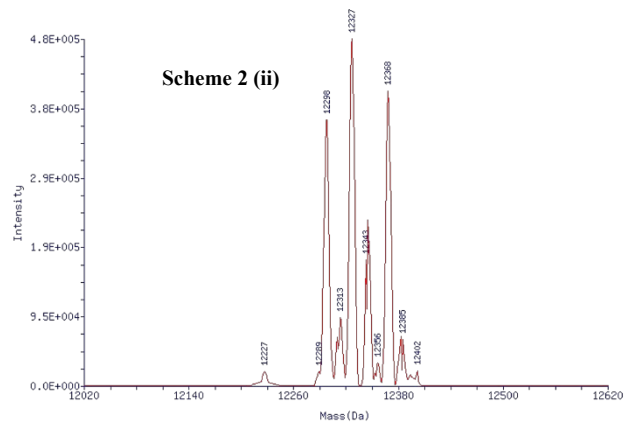
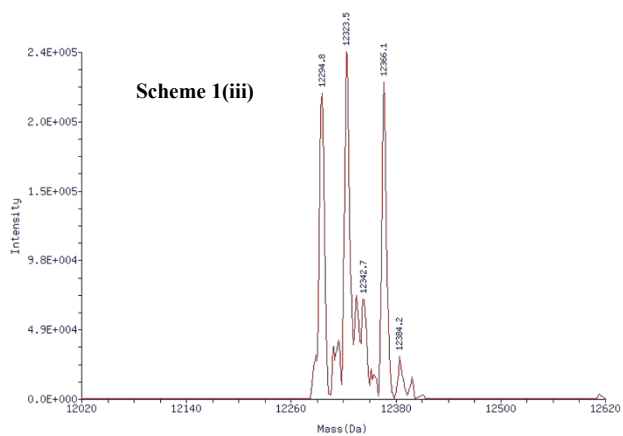
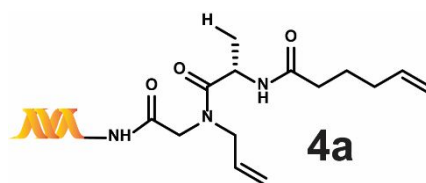
Calculated Product Mass
12224.09 Da



Calculated Product Mass
12496.26 Da

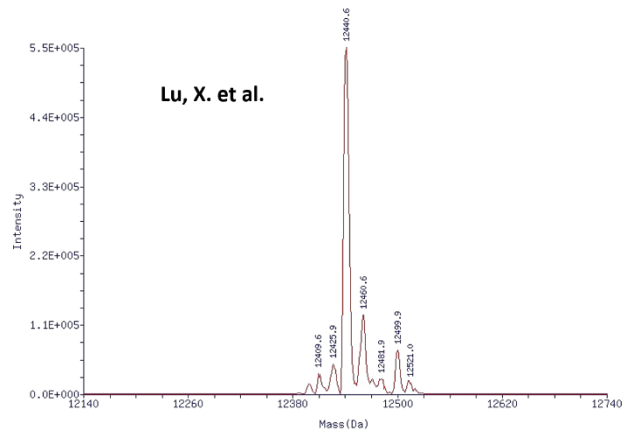
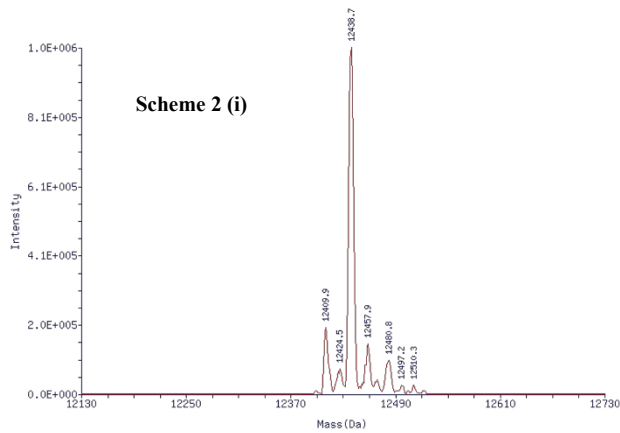
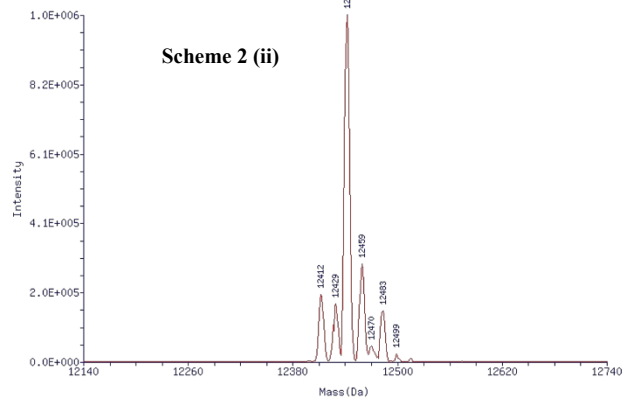
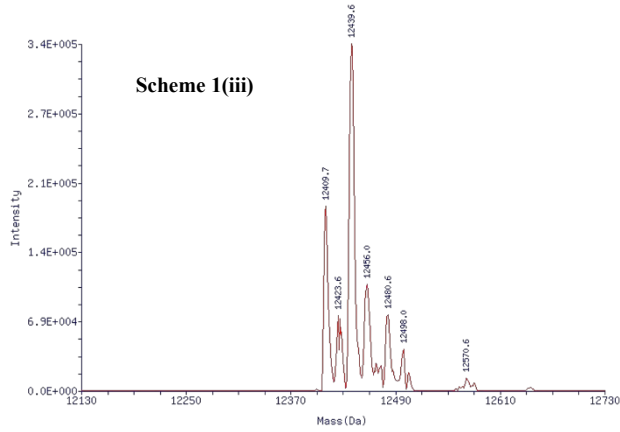
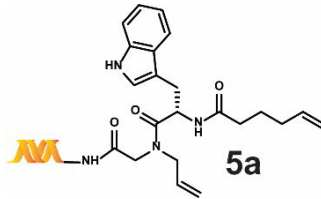


Calculated Product Mass
12295.13 Da



Calculated Product Mass

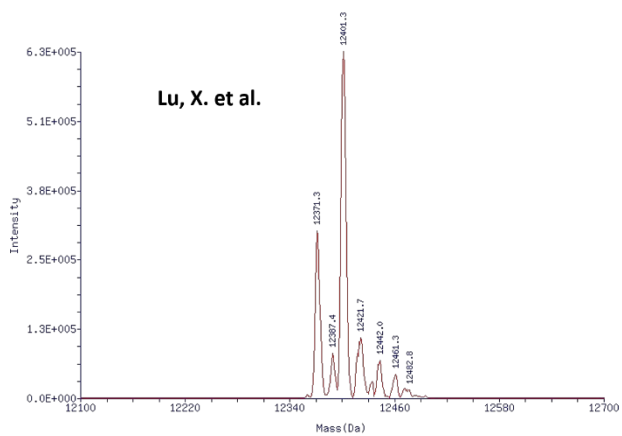
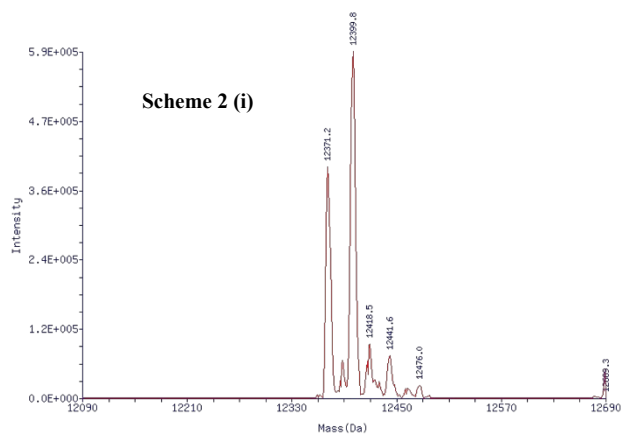
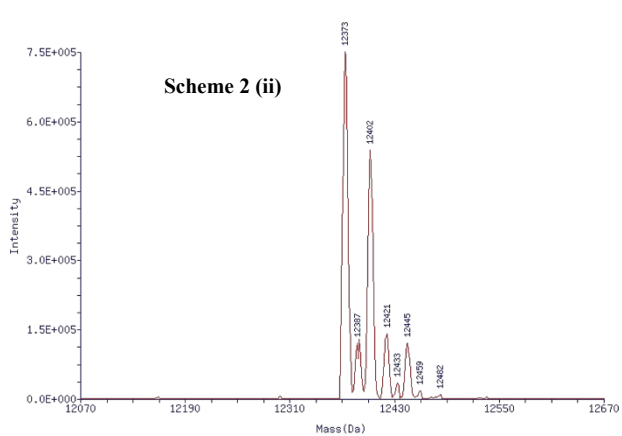
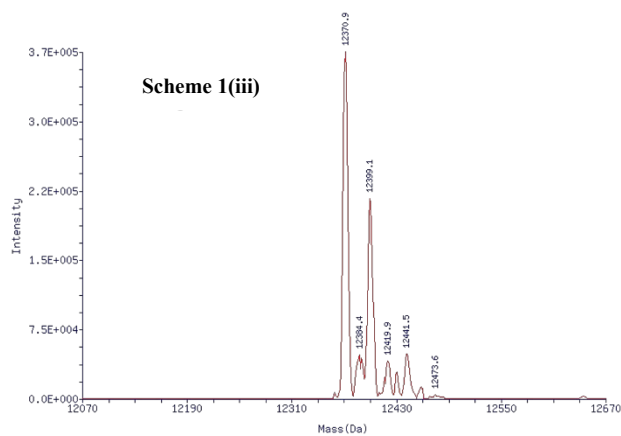
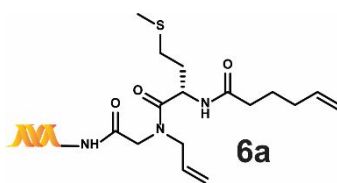
12410.17 Da



Calculated Product Mass

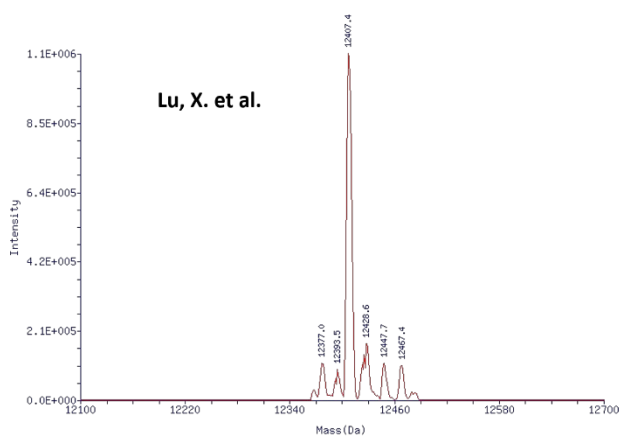
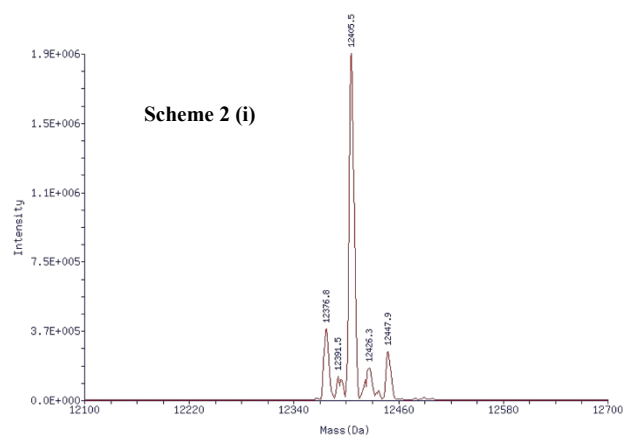
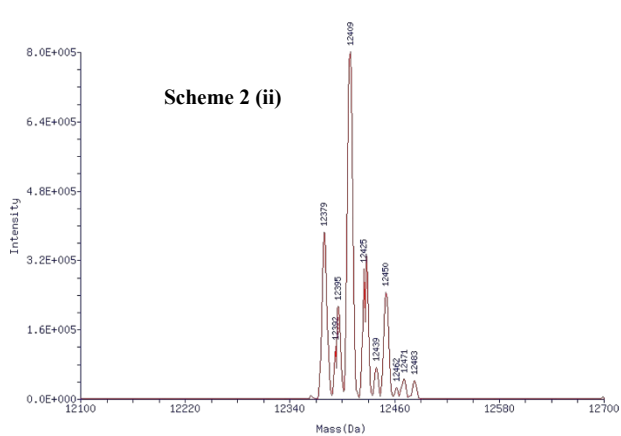
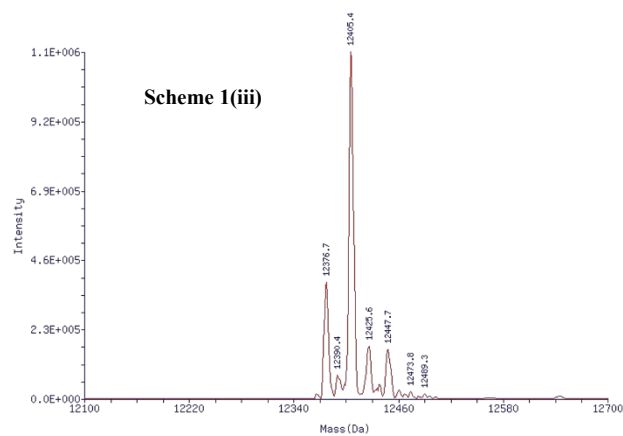
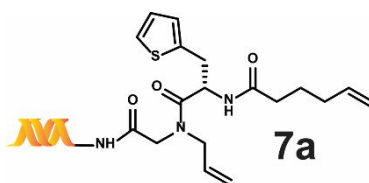
12373.14 Da

(H₂O adduct)

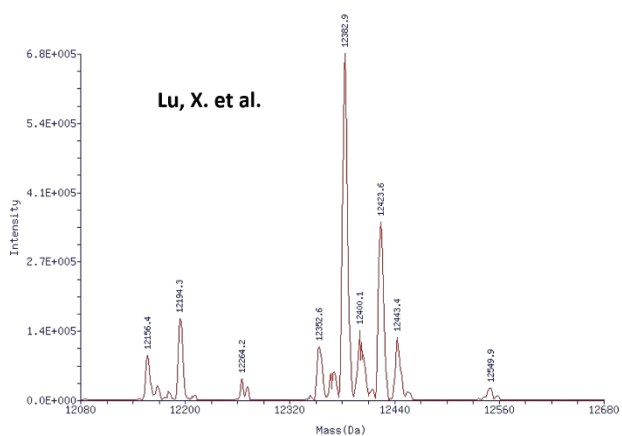
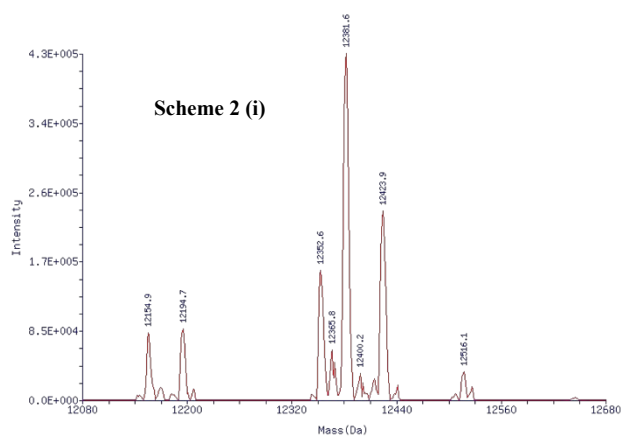
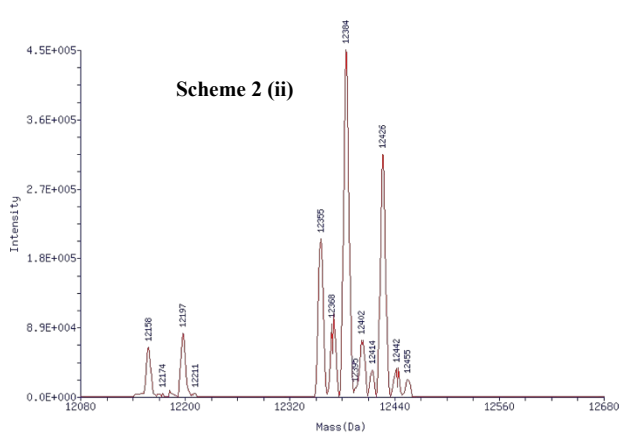
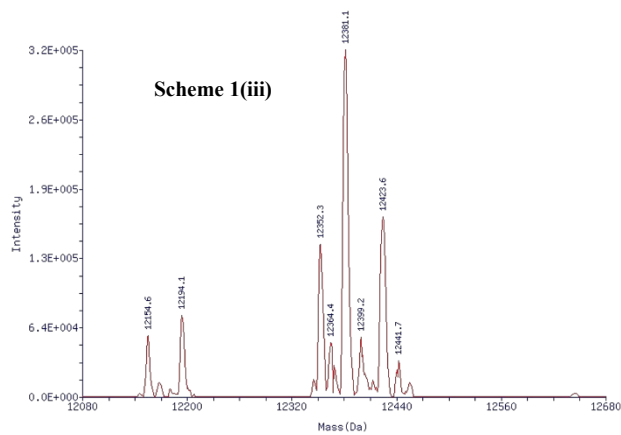
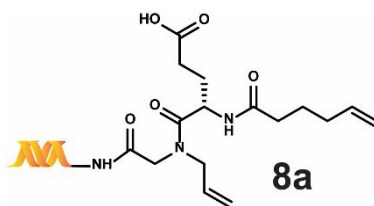


Calculated Product Mass

12377.11 Da

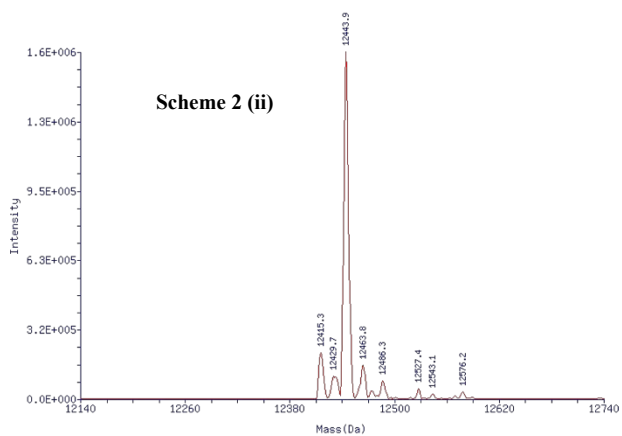
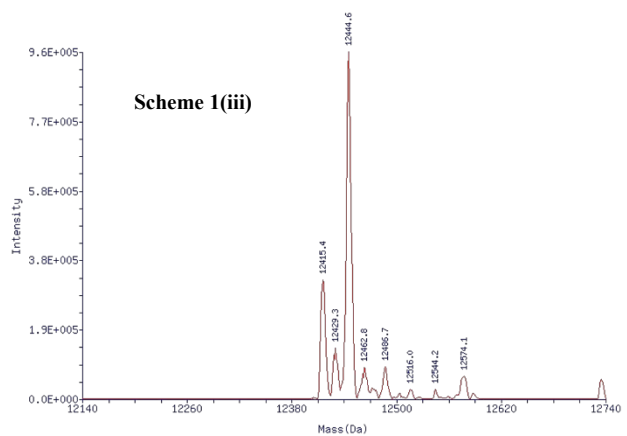
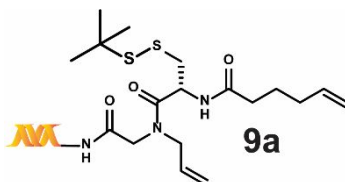


Calculated Product Mass
12353.13 Da

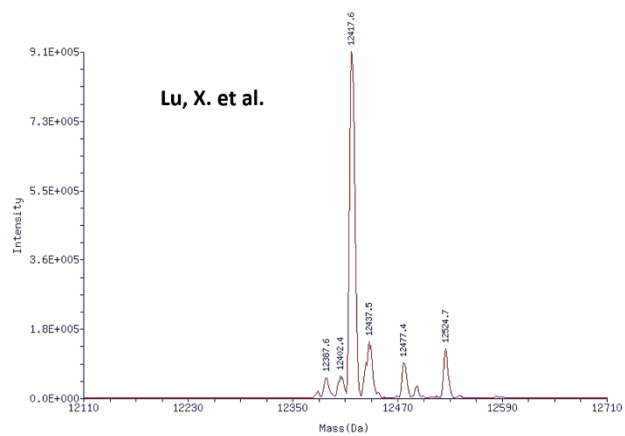
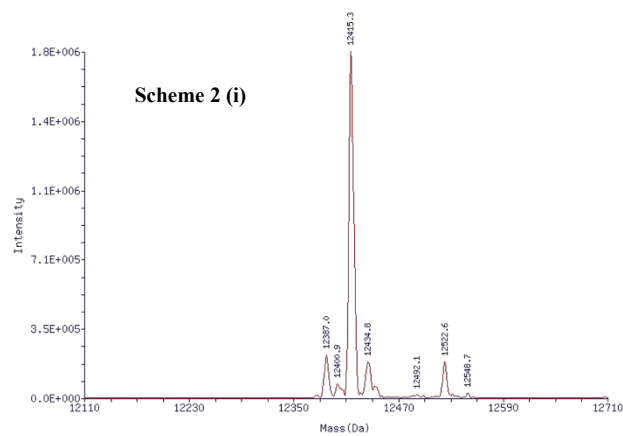
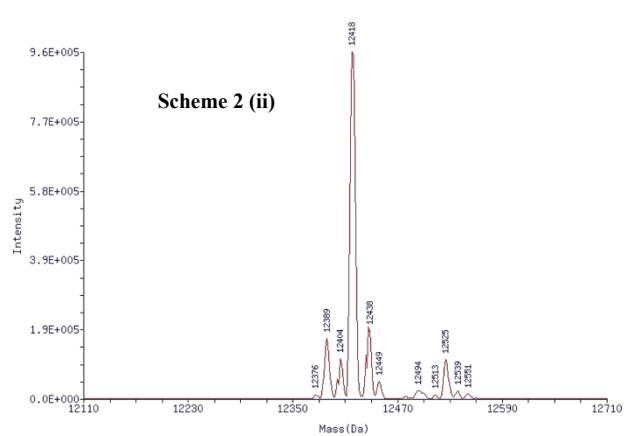
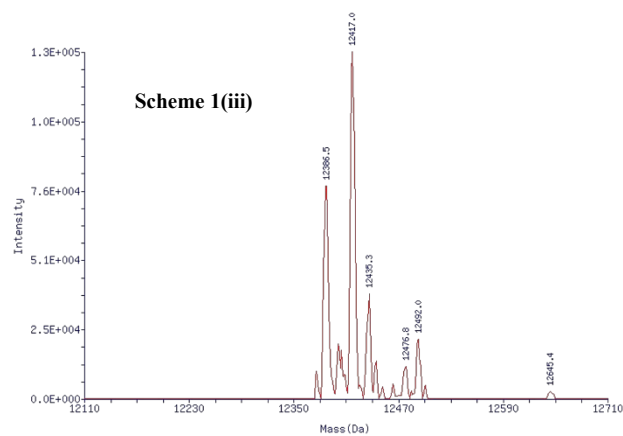
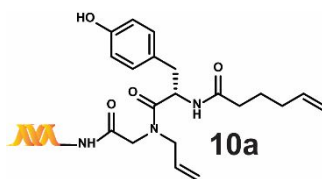


Calculated Product Mass

12415.13 Da

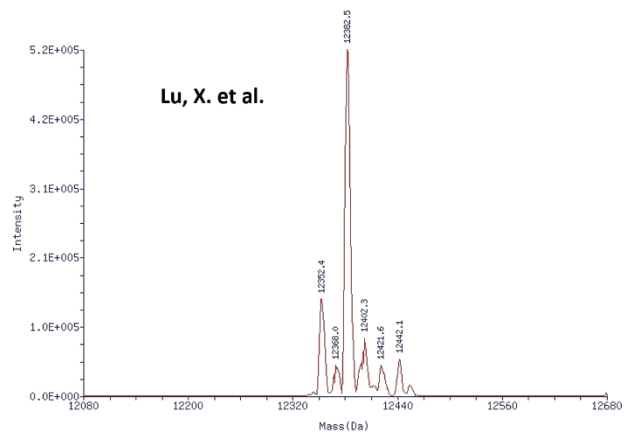
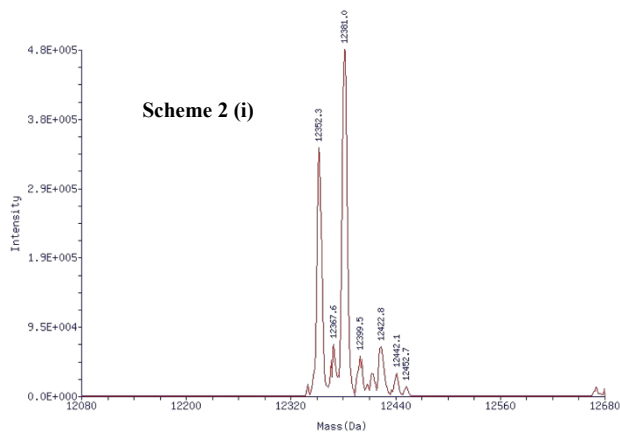
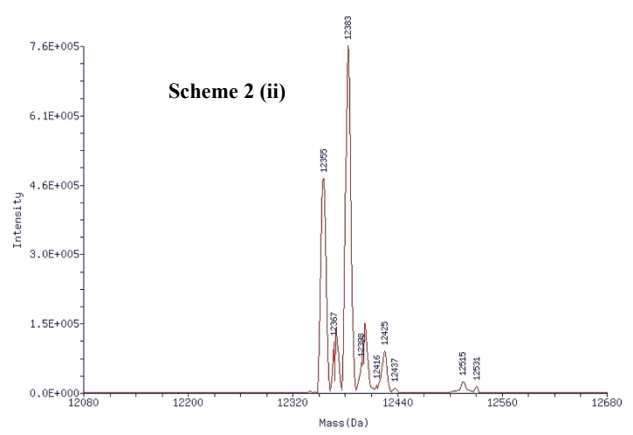
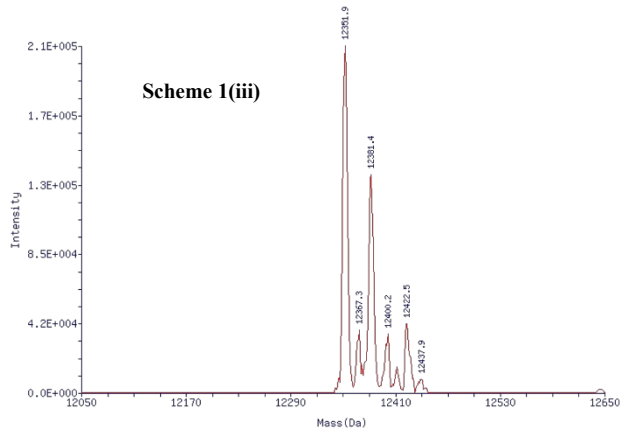
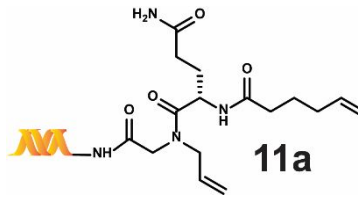


Calculated Product Mass
12387.15 Da



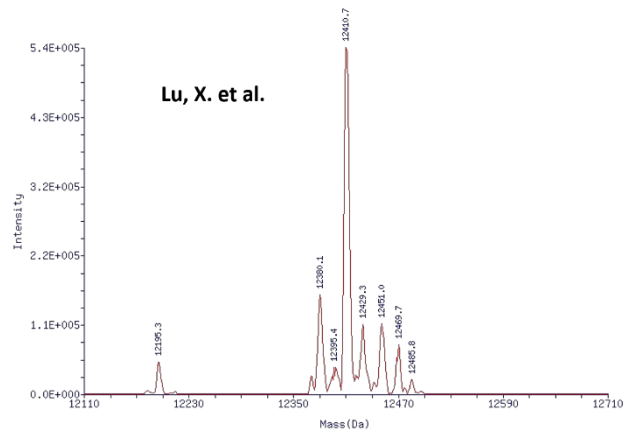
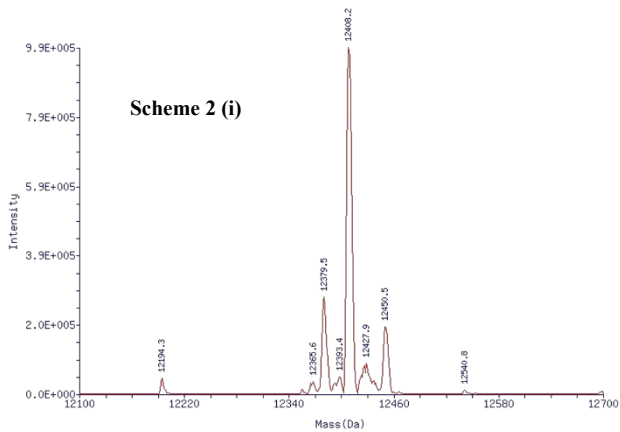
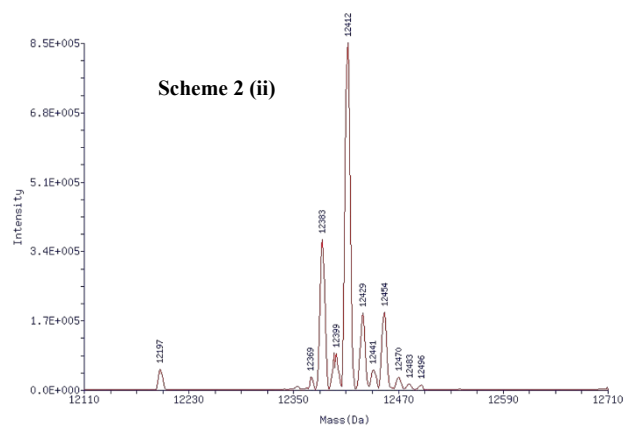
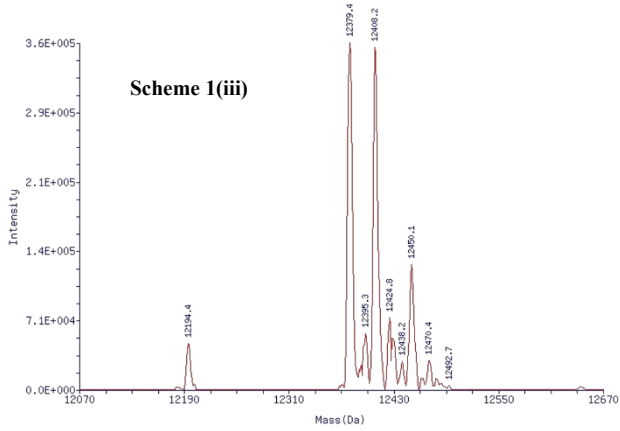
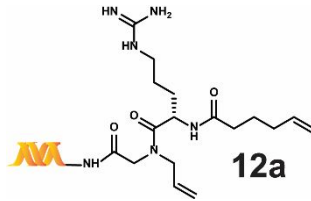
Calculated Product Mass

12352.15 Da



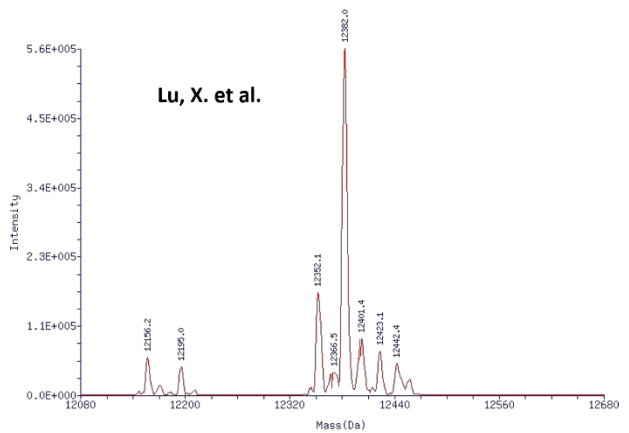
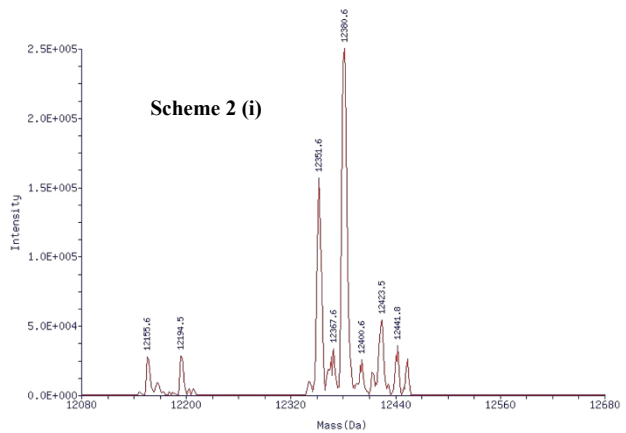
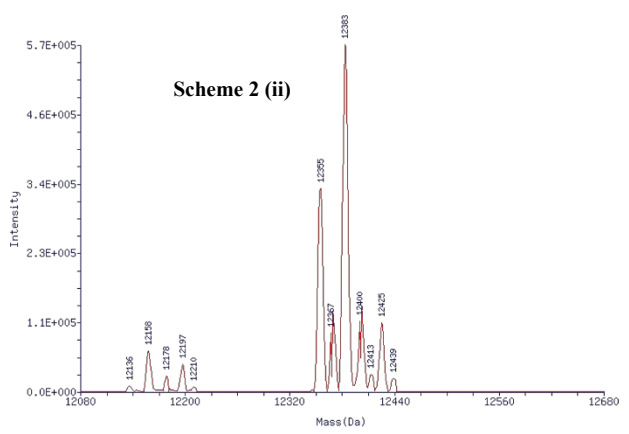
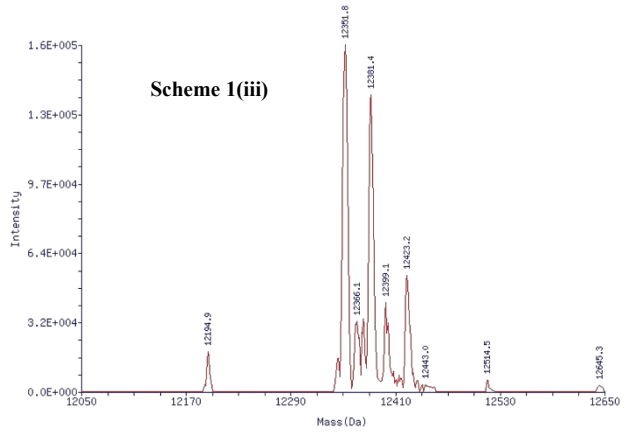
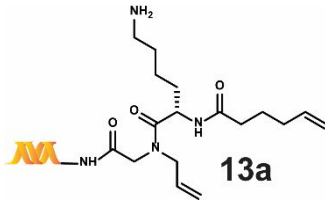
Calculated Product Mass

12380.19 Da



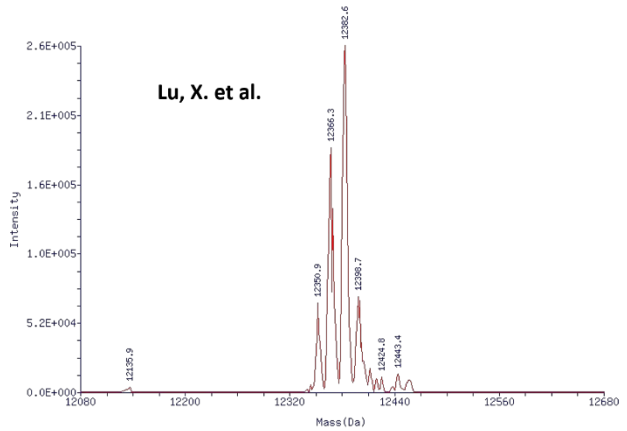
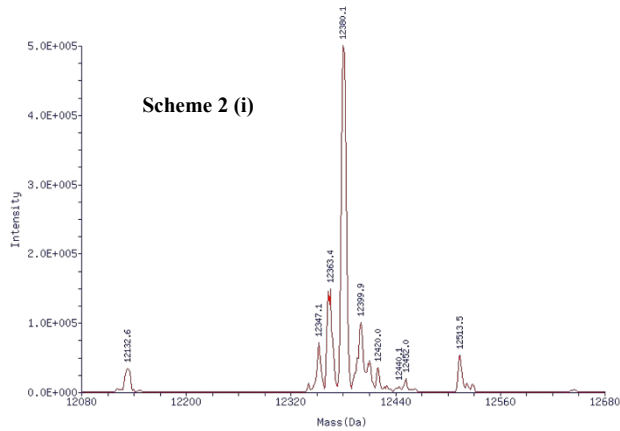
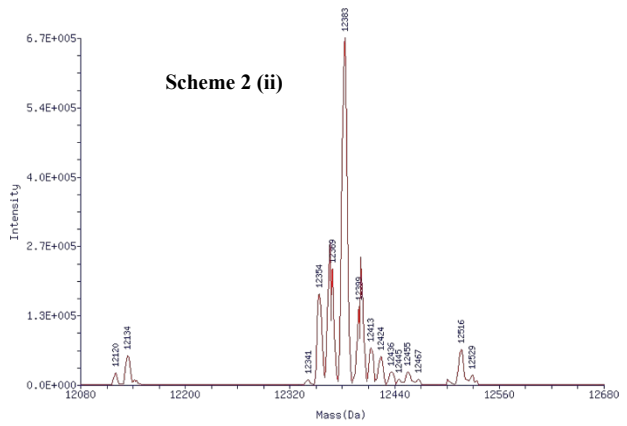
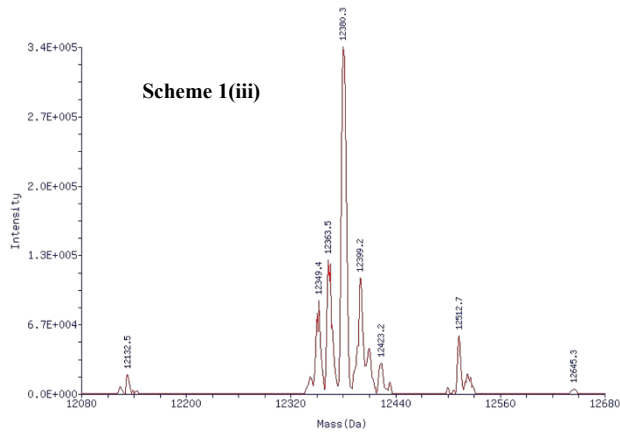
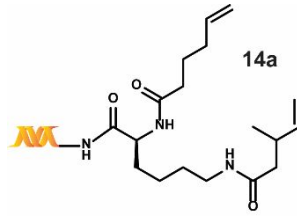
Calculated Product Mass

12352.18 Da



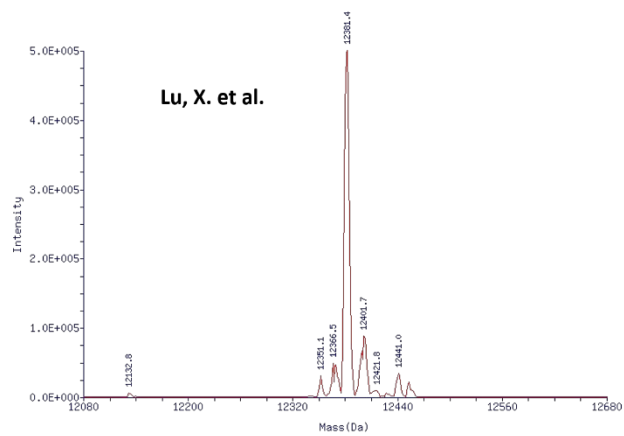
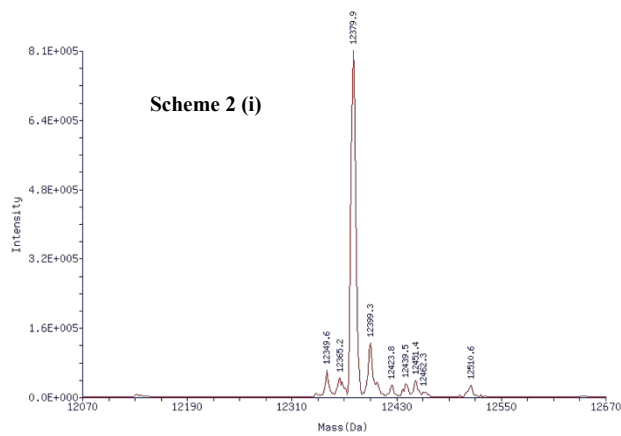
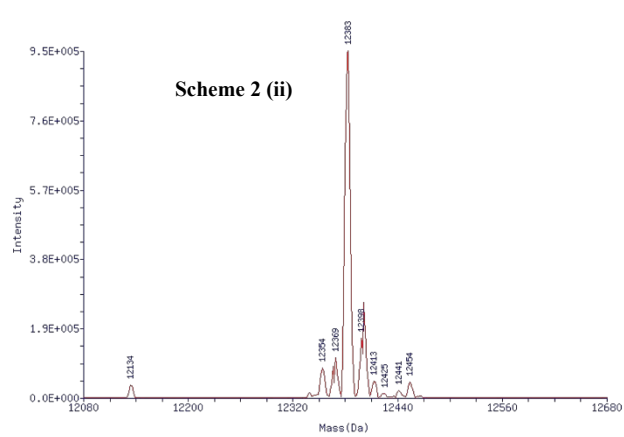
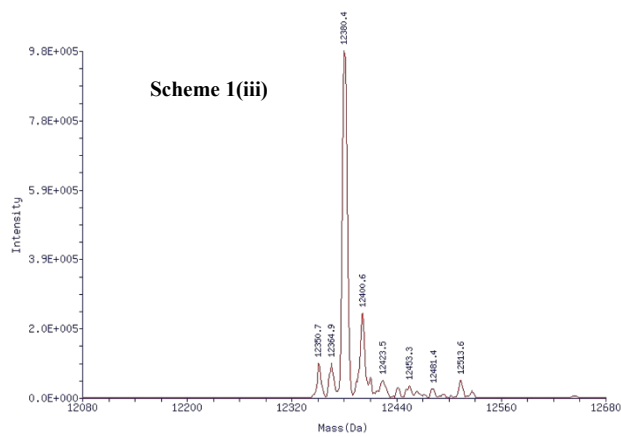
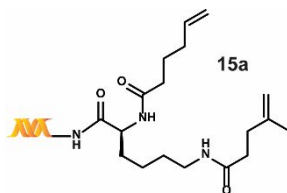
Calculated Product Mass

12351.19 Da



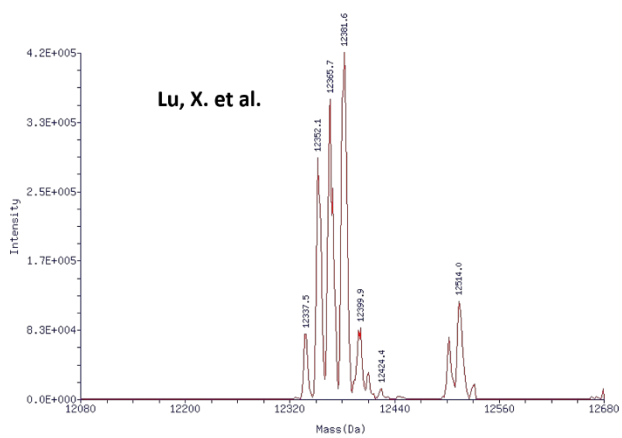
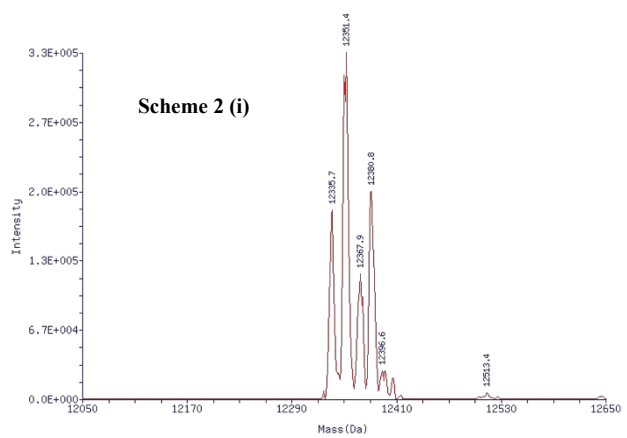
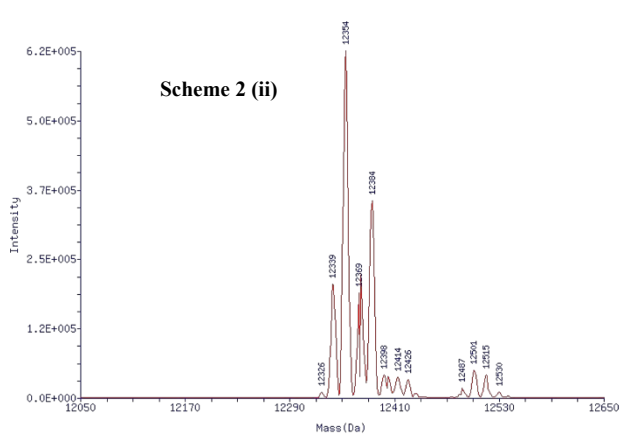
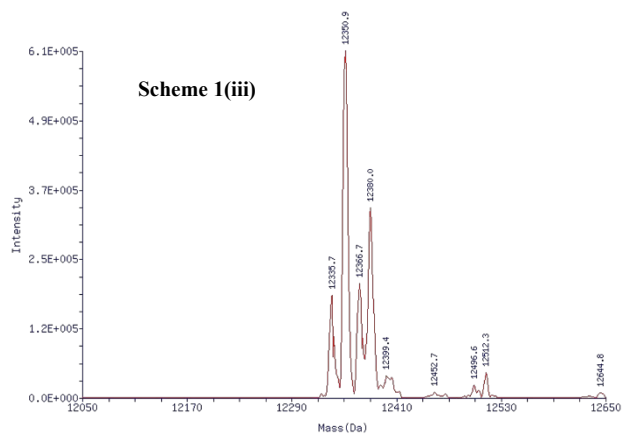
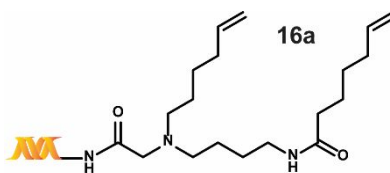
Calculated Product Mass

12351.19 Da



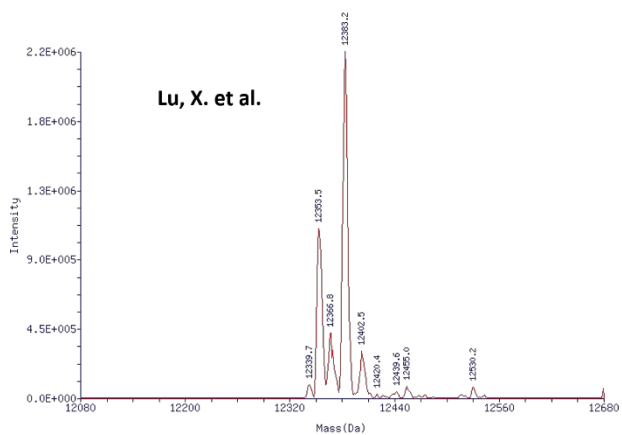
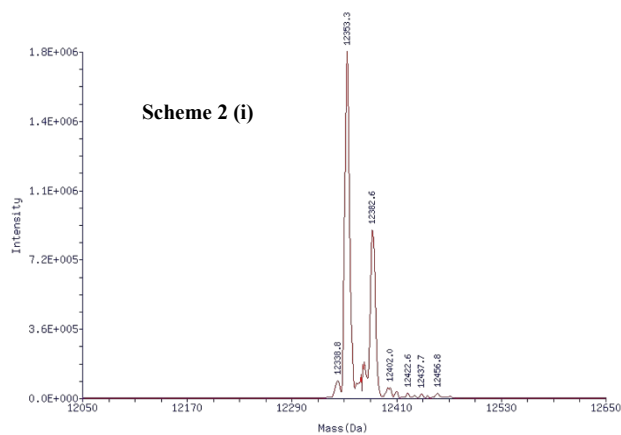
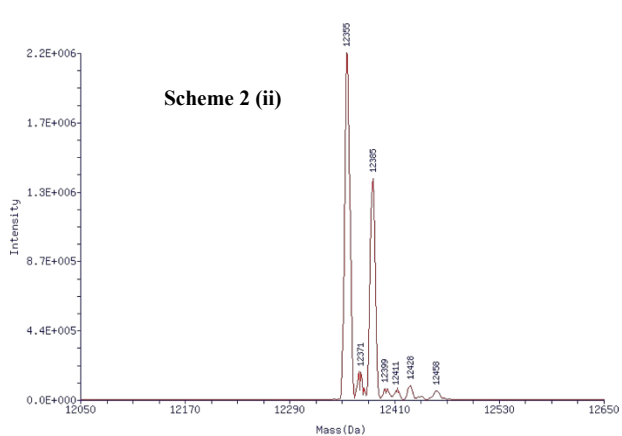
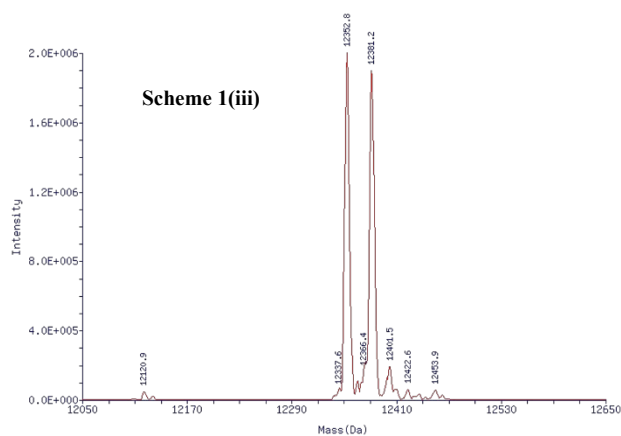
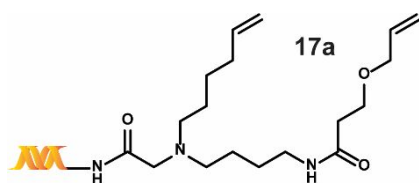
Calculated Product Mass

12351.22 Da

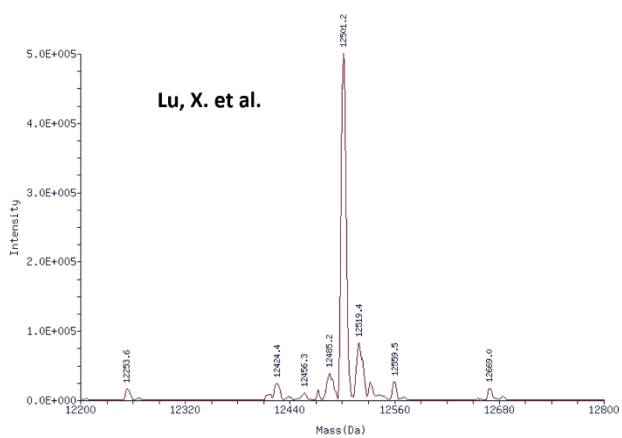
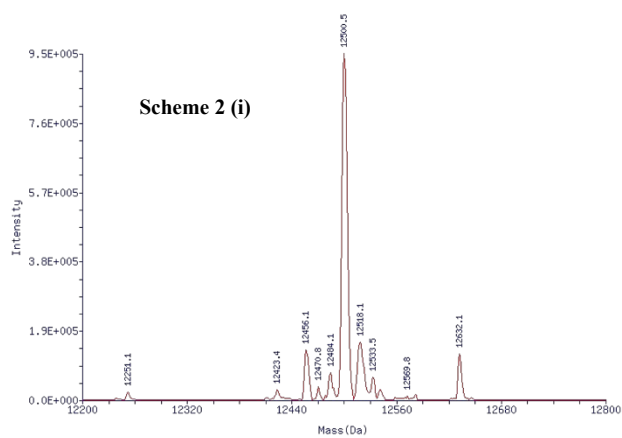
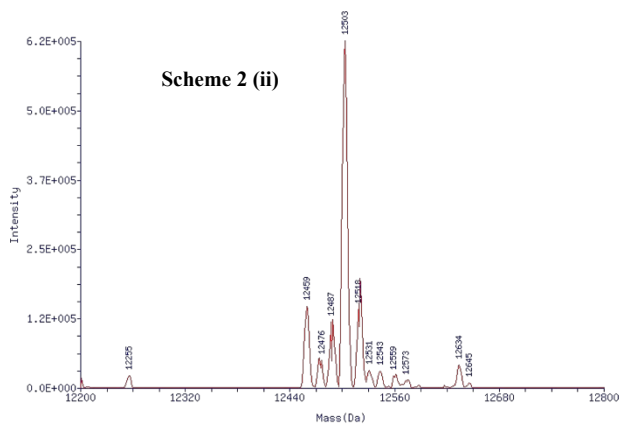
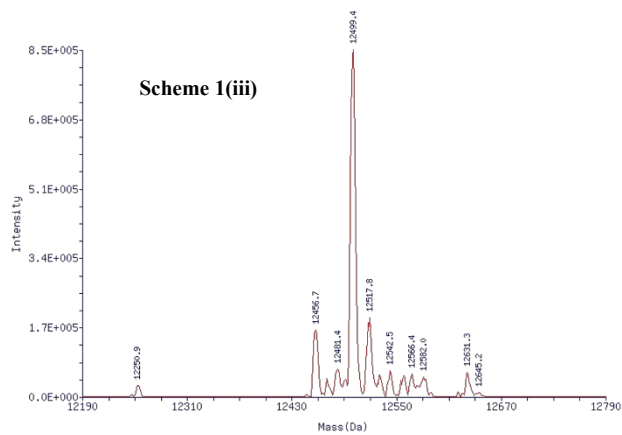
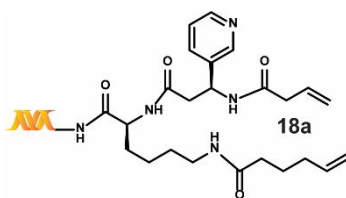


Calculated Product Mass

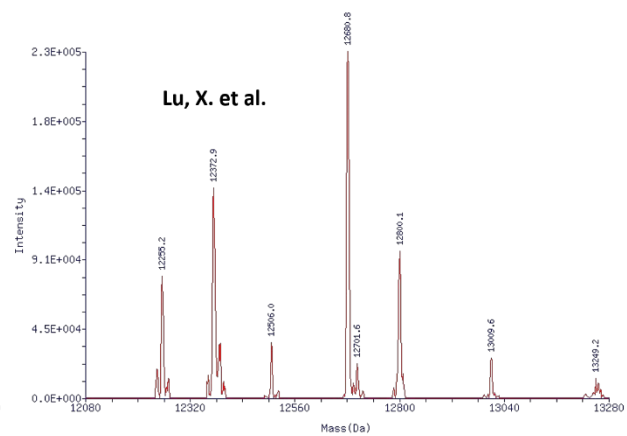
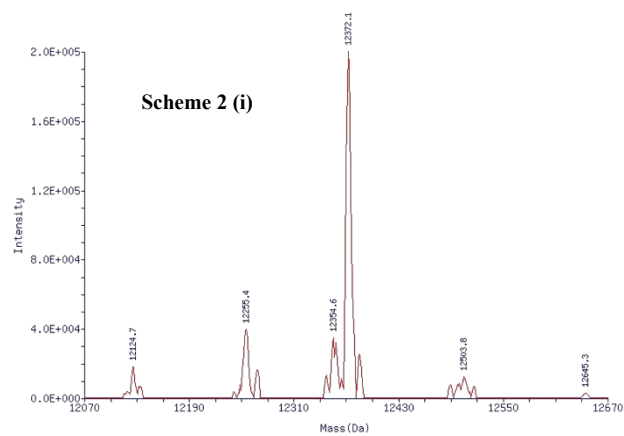
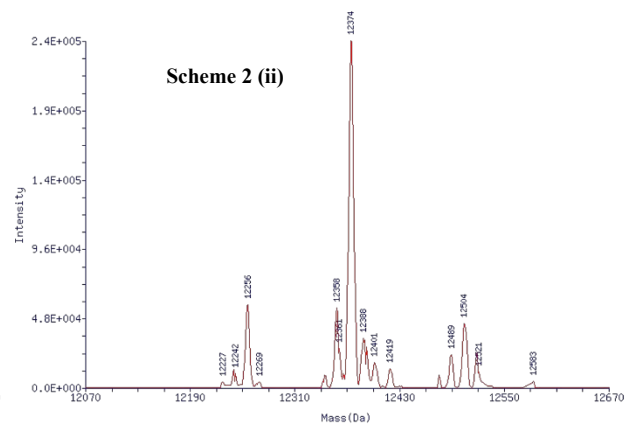
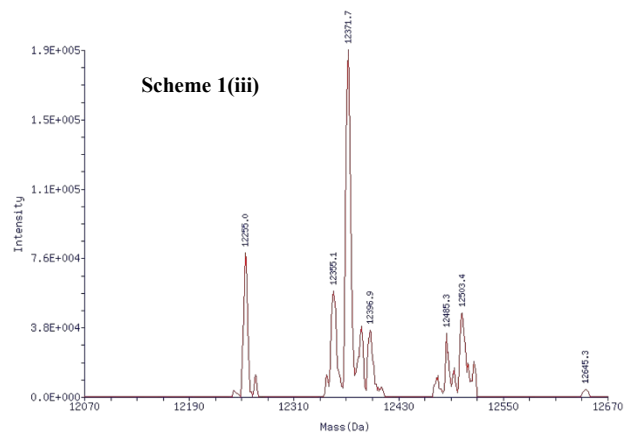
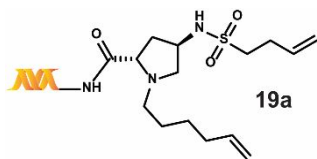
12353.20 Da



Calculated Product Mass
12471.22 Da

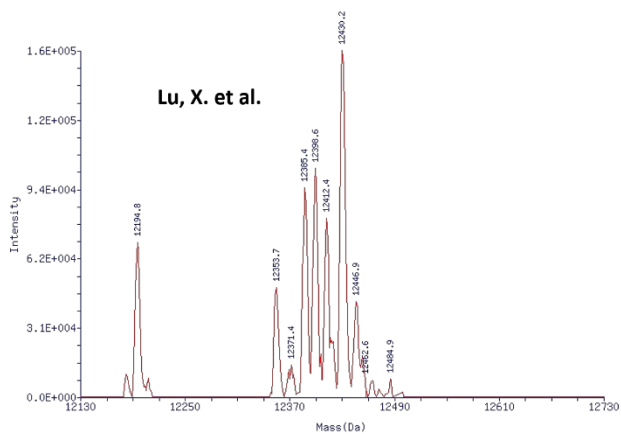
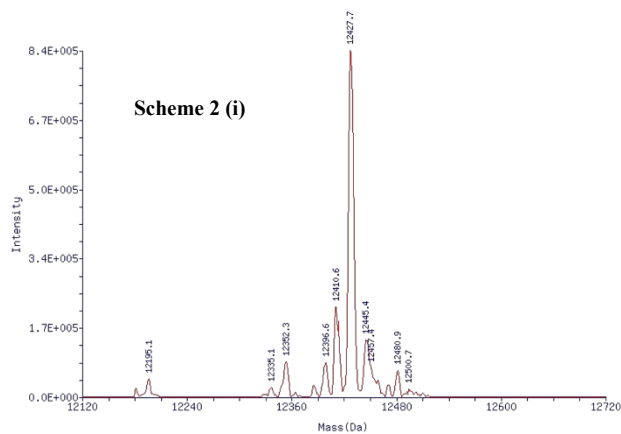
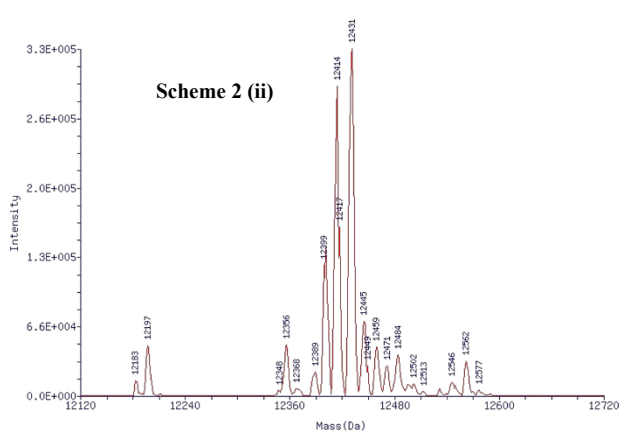
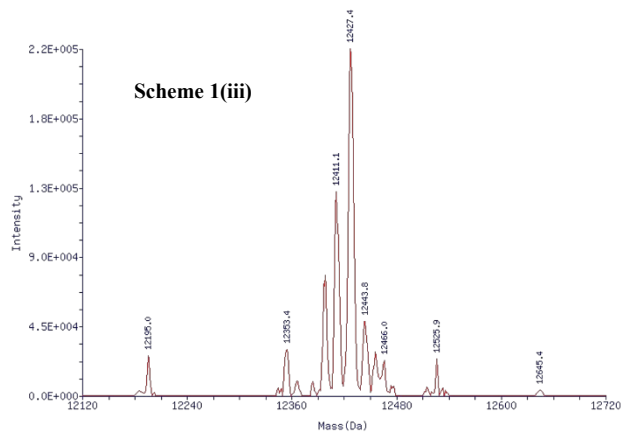
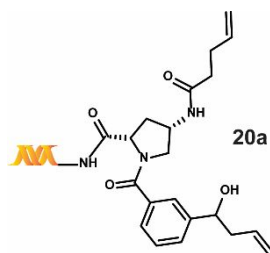


Calculated Product Mass
12343.13 Da

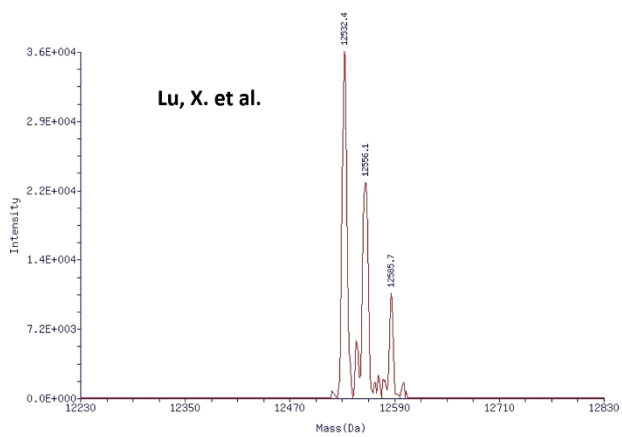
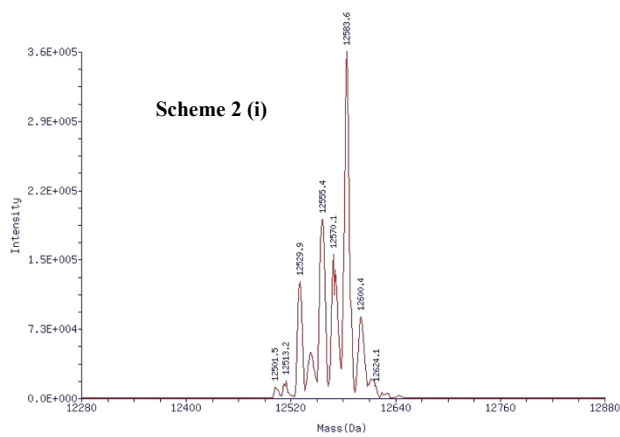
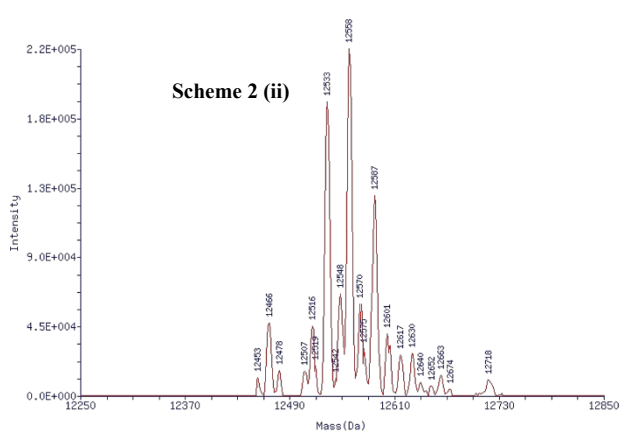
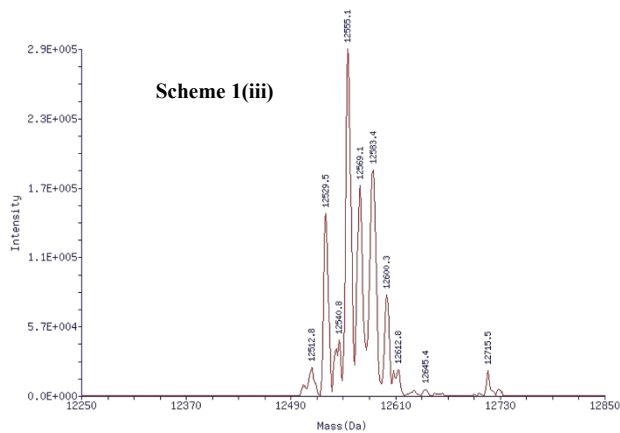
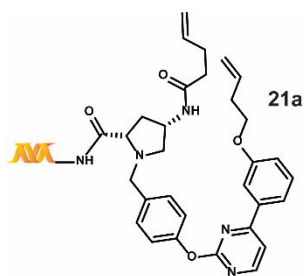


Calculated Product Mass

12399.15 Da

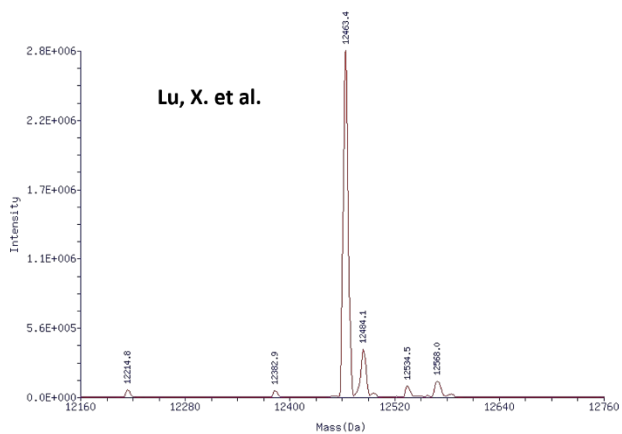
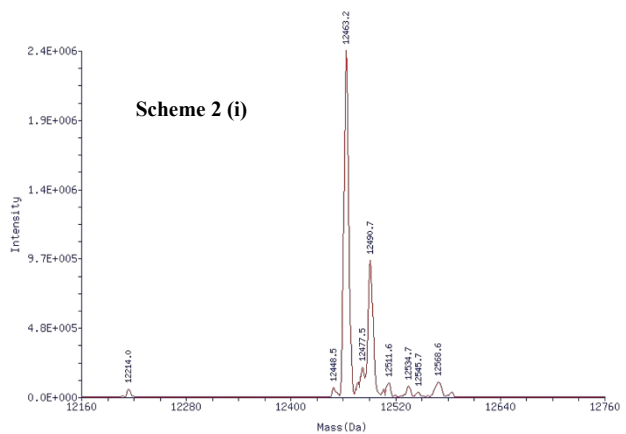
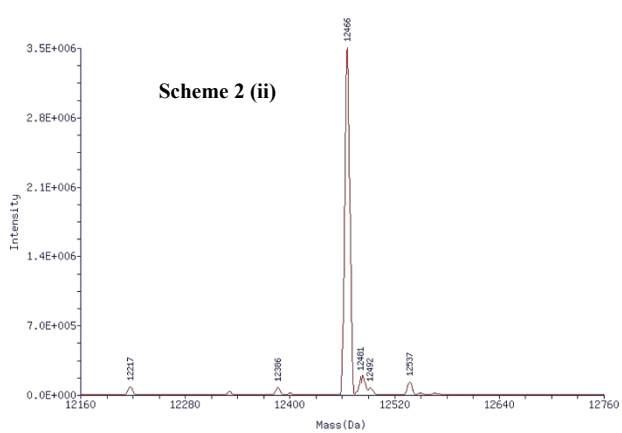
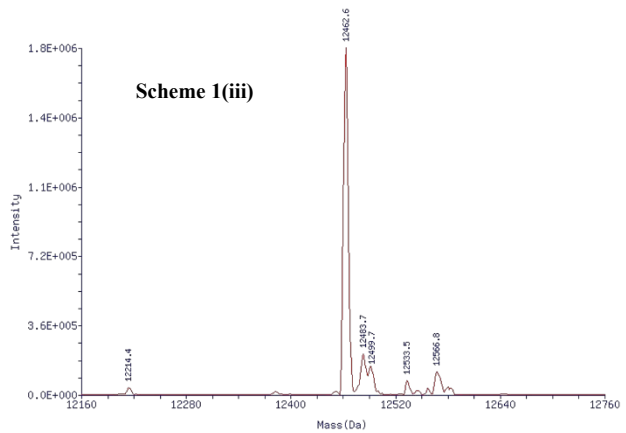
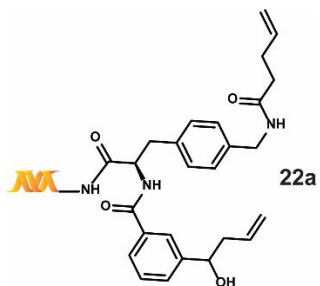


Calculated Product Mass
12555.22 Da

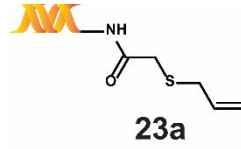


Calculated Product Mass

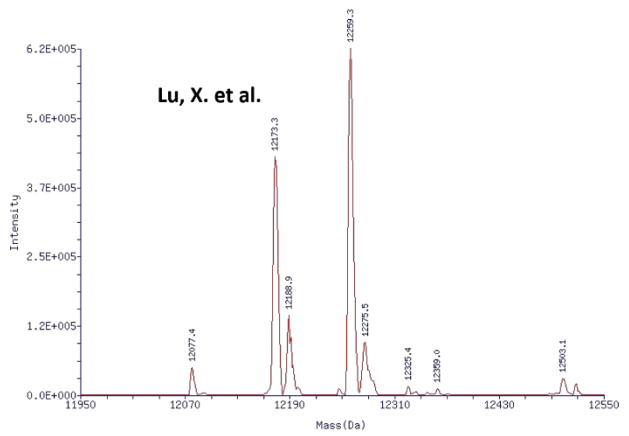
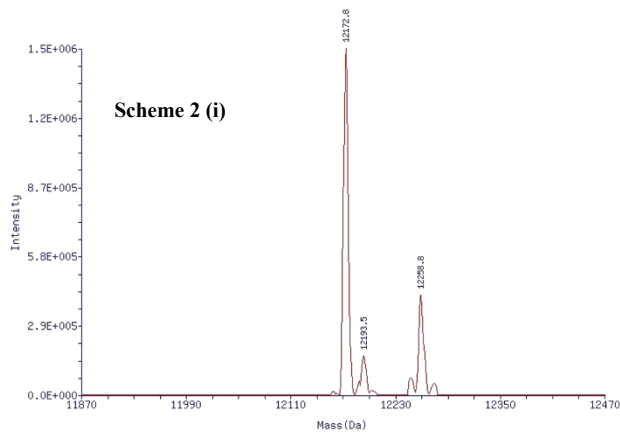
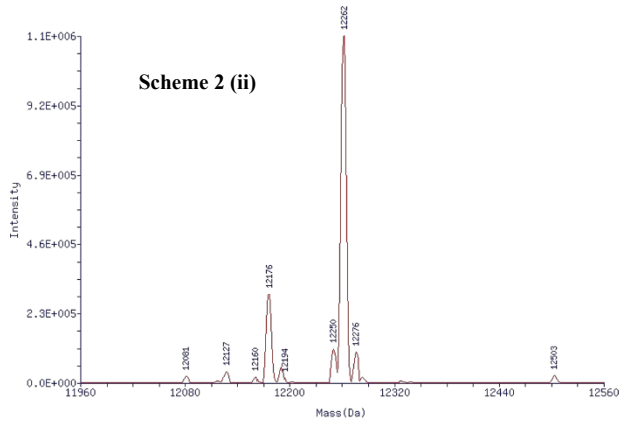
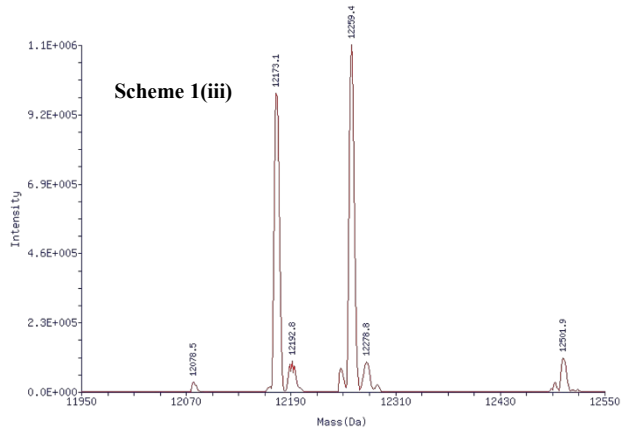
12463.18 Da

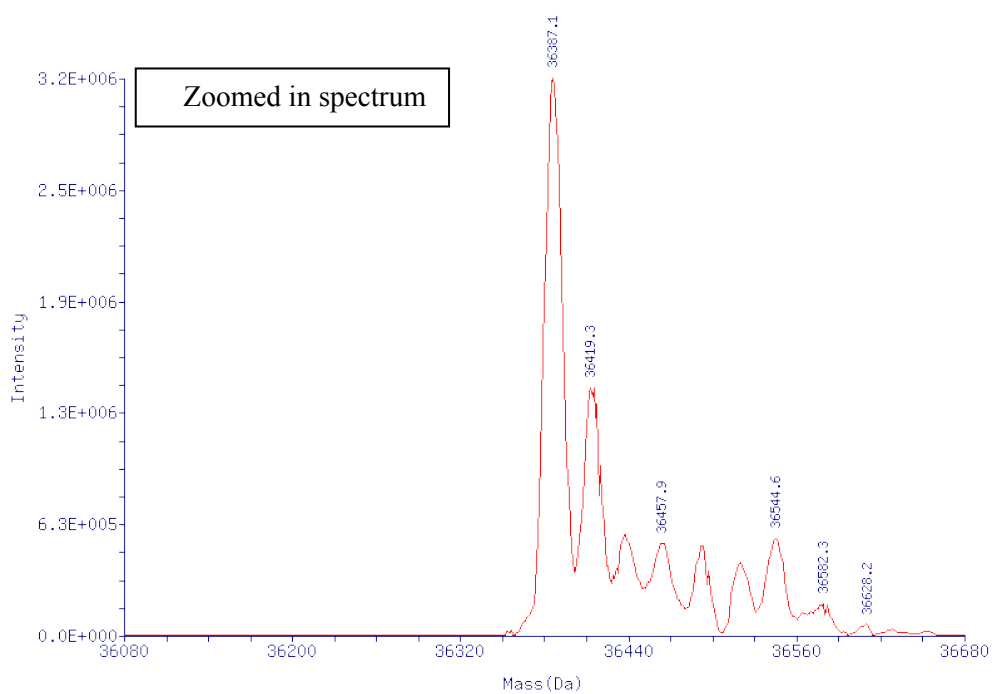
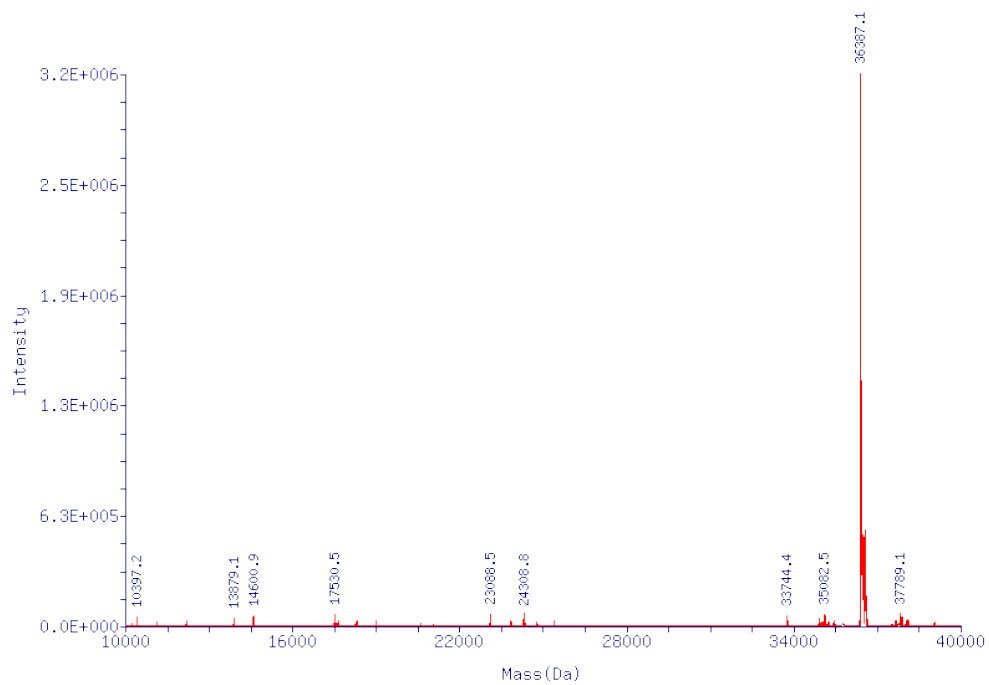
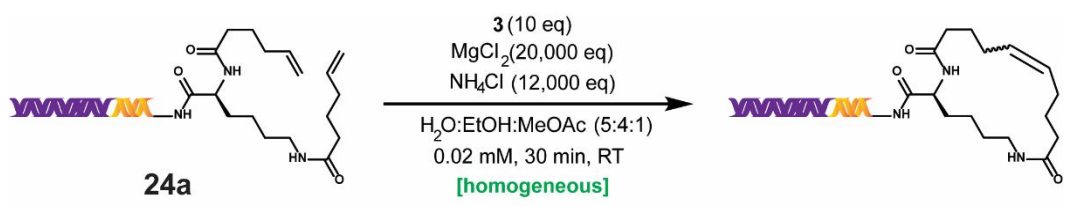


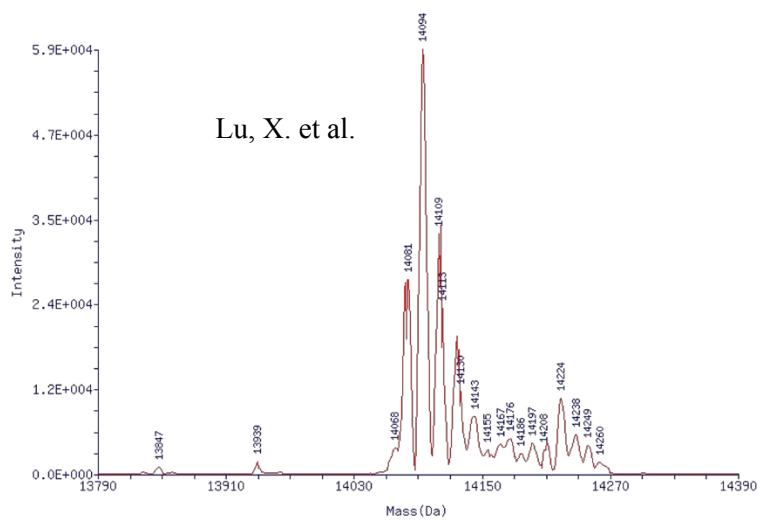
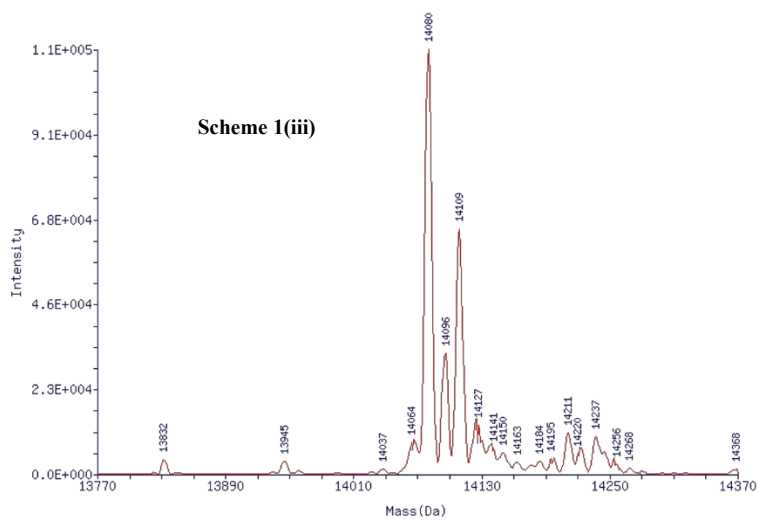
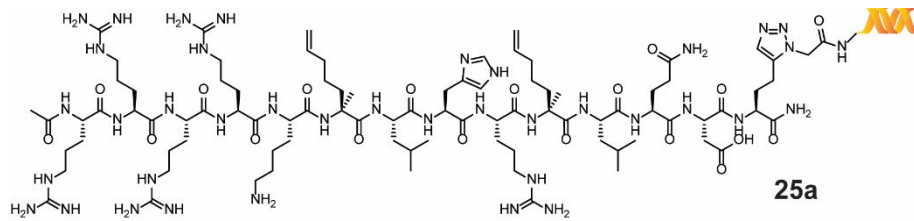
Calculated Product Mass
12259.05 Da



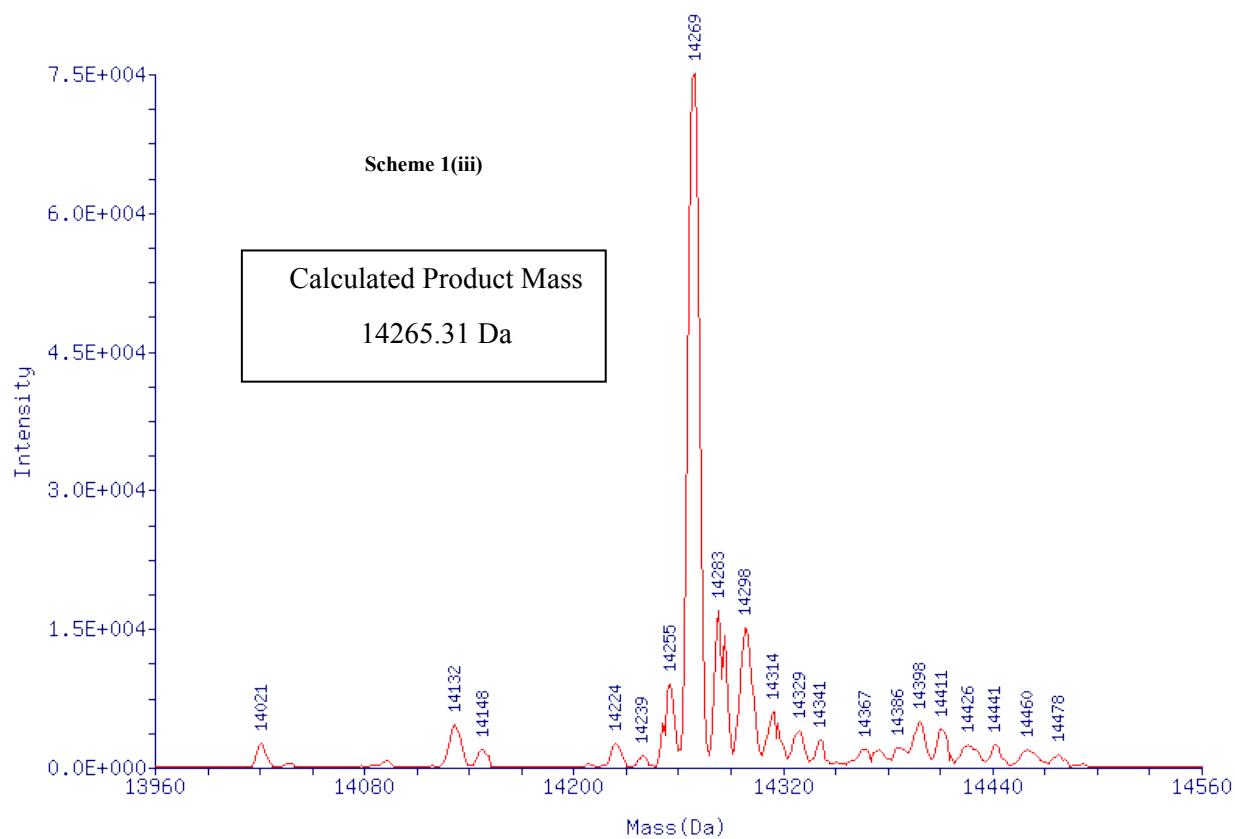
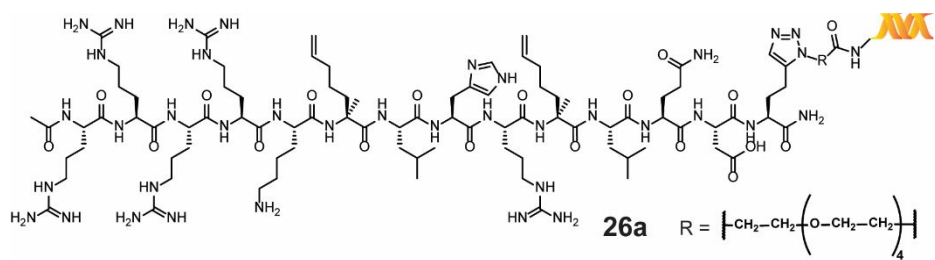
CM with
5-hexenoic acid (+86.03 Da)

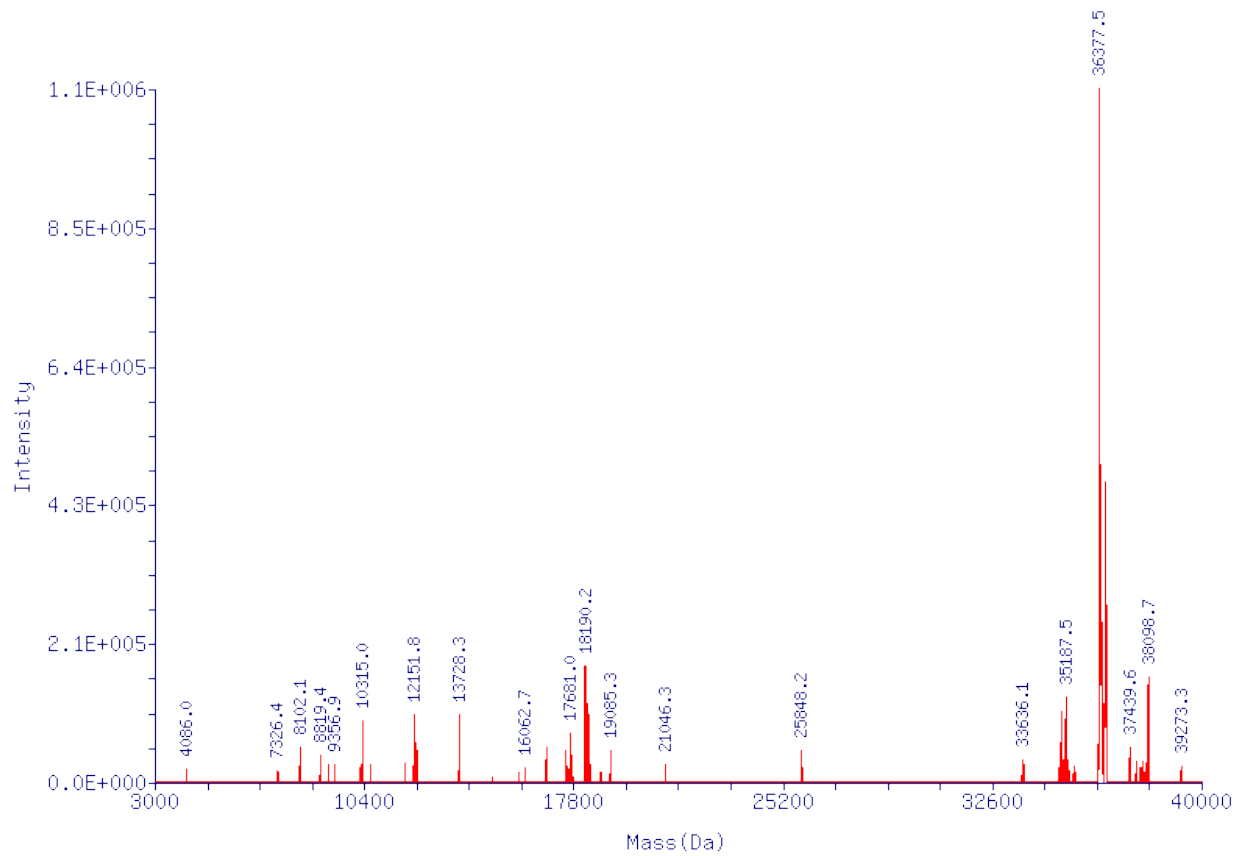
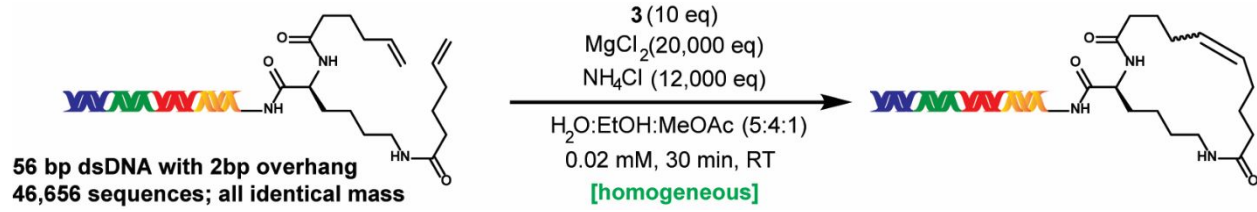






Calculated Product Mass
14075.19 Da

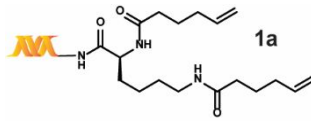




Post-Metathesis (RCM and CM) LCMS Traces for Reactions Ran under Scheme 1(iii) Conditions

When inspecting the LC/MS traces, please keep in mind the following:

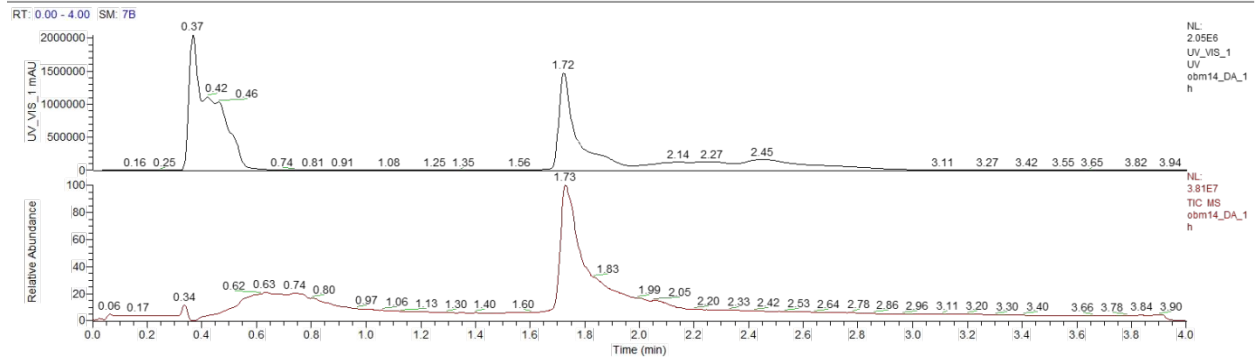
1. All but one of the traces was taken directly after the quenching procedure (see Quenching the RCM Reaction and Preparation for LC/MS Analysis). This affected the quality/cleanliness of signal without sacrificing the accuracy of the analysis.
2. As an example of the improvement in the signal quality after precipitation, the single-substrate DECL was precipitated using the General DNA Precipitation Procedure and then analyzed once more via LCMS. The improvement in the signal quality is illustrated below.
3. Reactions were ran on 1-2 nmol scale. Due to the lack of precipitation before LC/MS analysis, the signal was often weakened closer to the noise level.



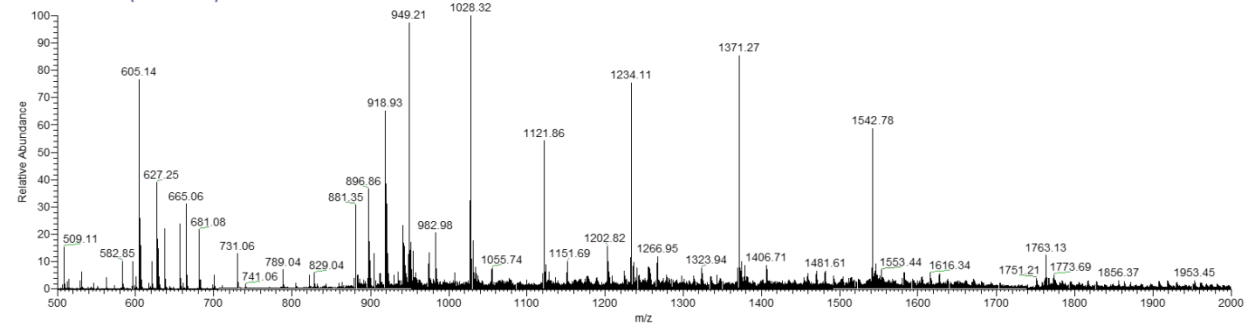
C:\Analytical\lobm14_DA_1h

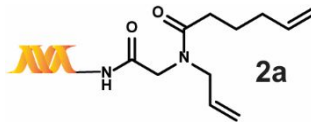
01/21/19 16:29:37

DW672_PA4_A1



obm14_DA_1h #573-777 RT: 1.60-2.16 AV: 205 NL: 2.36E5
T: ITMS - c ESI Full ms [500.00-2000.00]



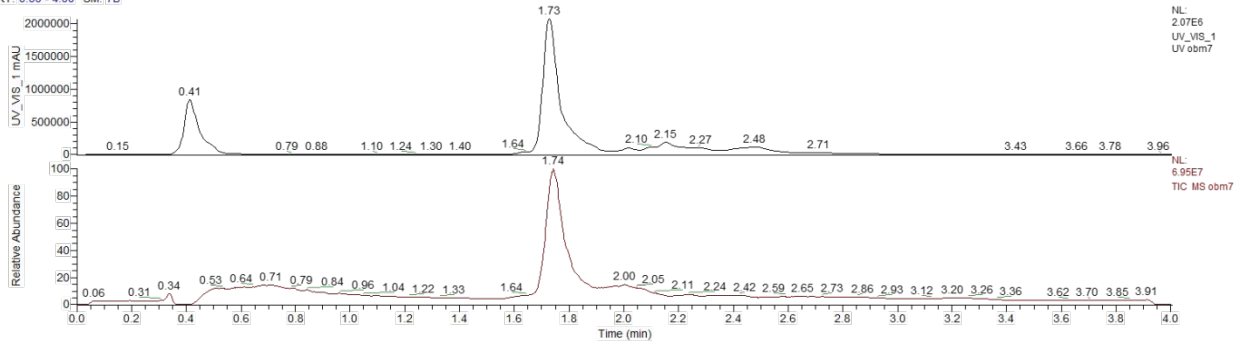


C:\Analytical\...obm7

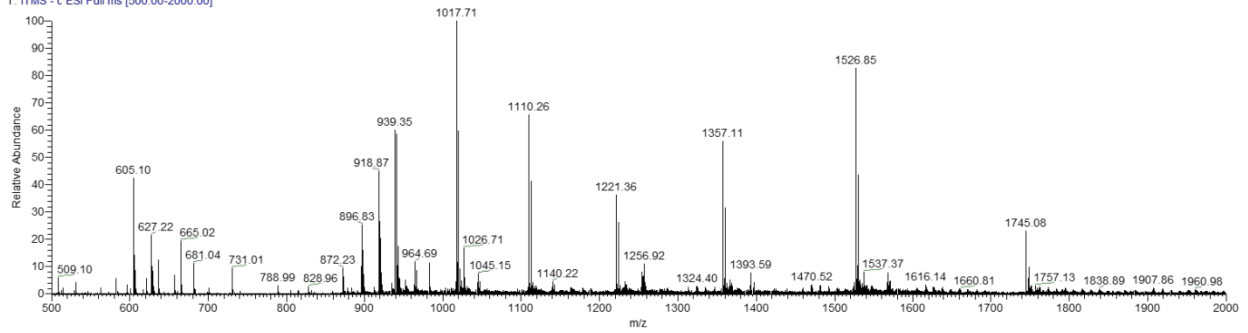
12/11/18 11:09:44

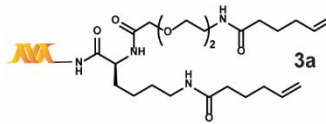
DW672_PA4_A1

RT: 0.00 - 4.00 SM: 7B



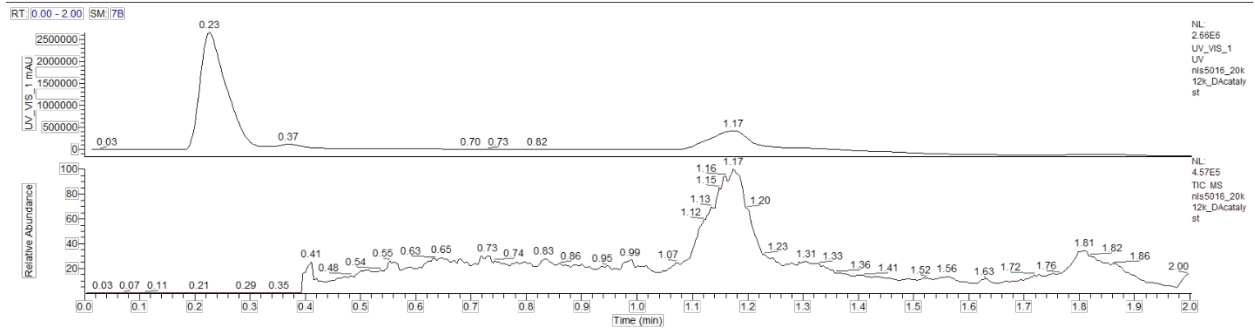
obm7 #571-777 RT: 1.58-2.14 AV: 207 NL: 5.64E5
T: ITMS - c ESI Full ms [500.00-2000.00]



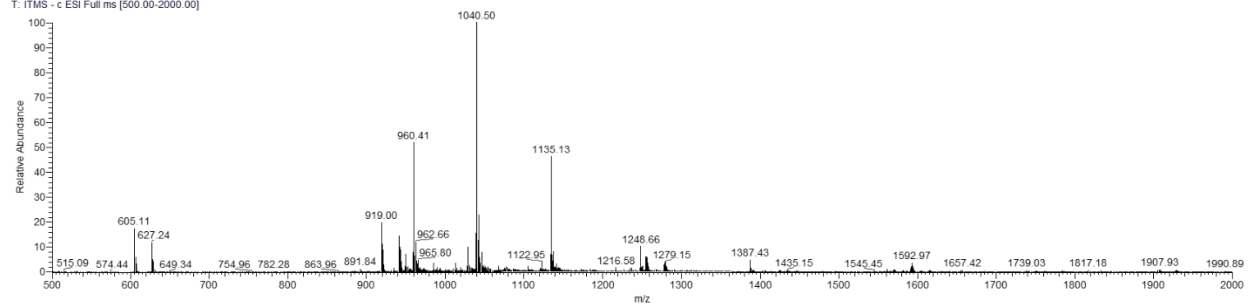


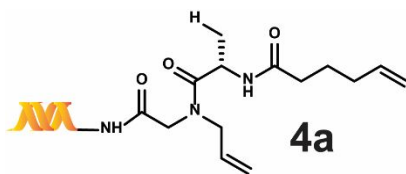
nl5016_20k12k_DAcatalyst

11/13/18 15:41:34



nl5016_20k12k_DAcatalyst #365-455 RT: 1.03-1.28 AV: 91 NL: 2.29E4
T: ITMS - c ESI Full ms [500.00-2000.00]

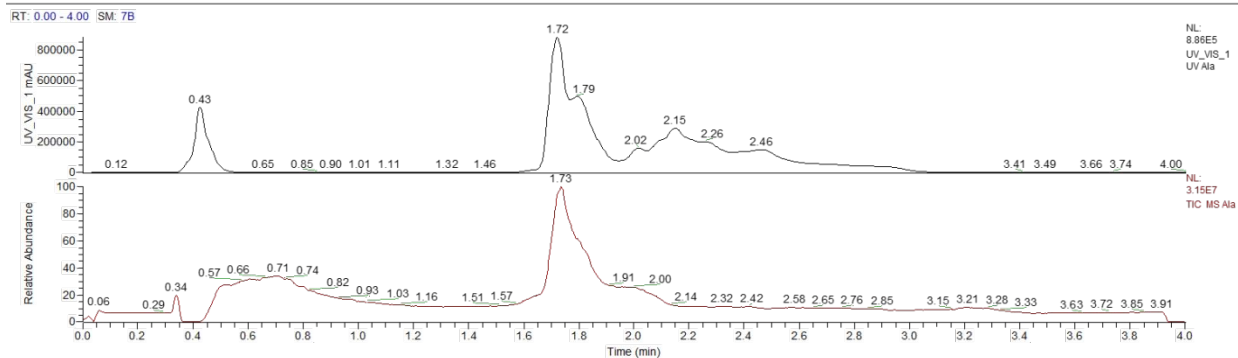




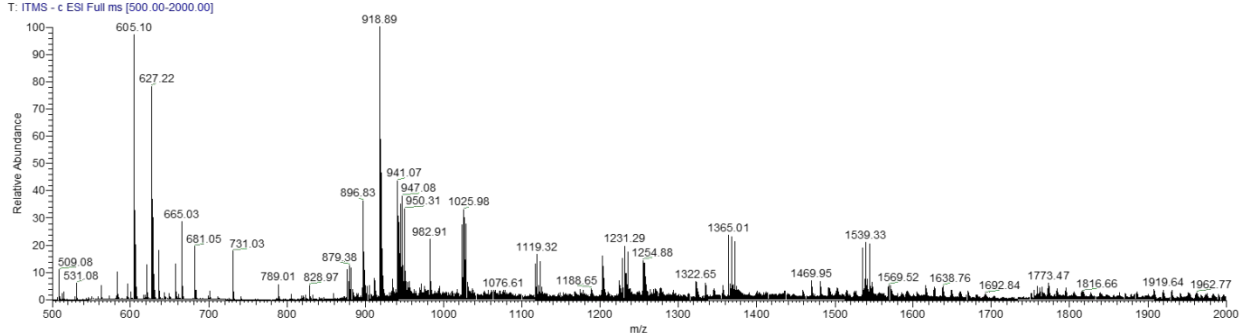
C:\Analytical\...IDACatAlisSubstrates\Ala

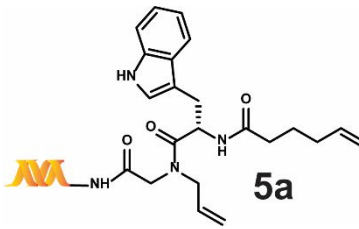
12/11/18 12:06:45

DW672_PA4_A1



Ala #542-797 RT: 1.50-2.19 AV: 256 NL: 2.41E5
T: ITMS - c ESI Full ms [500.00-2000.00]

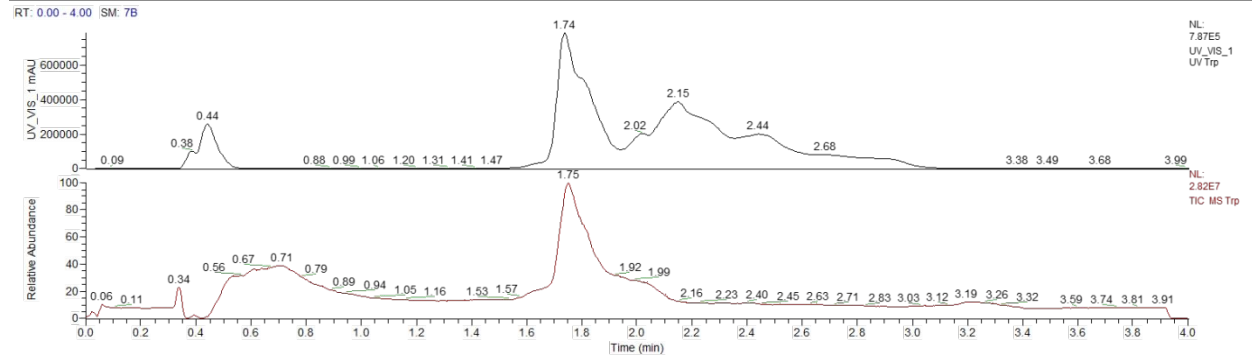




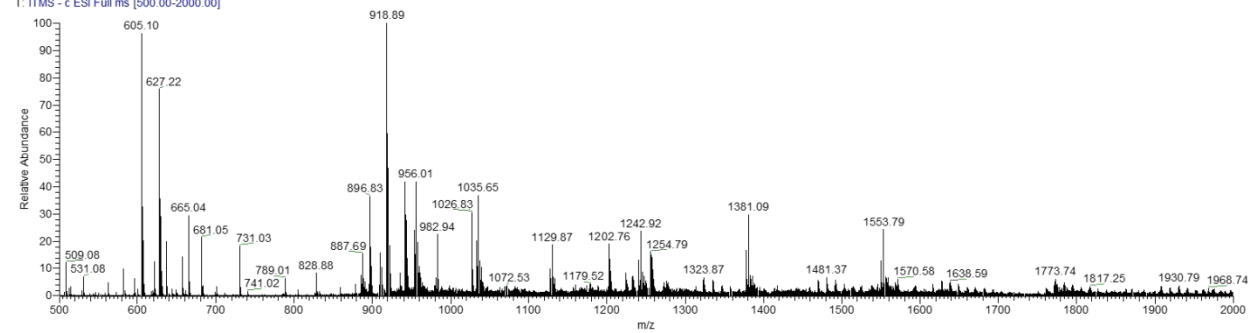
C:\Analytical...IDACat\AllSubstrates\Trp

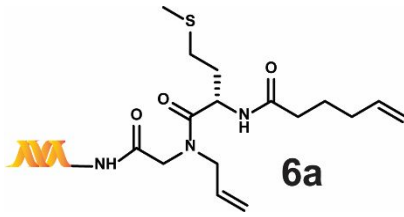
12/11/18 12:30:30

DW672_PA4_A1



Trp #553-788 RT: 1.53-2.17 AV: 236 NL: 2.44E5
T: ITMS - c ESI Full ms [500.00-2000.00]

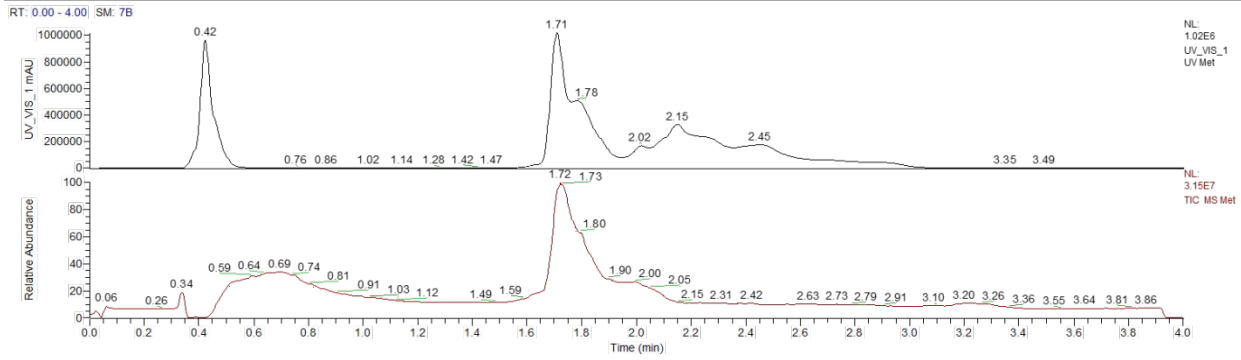




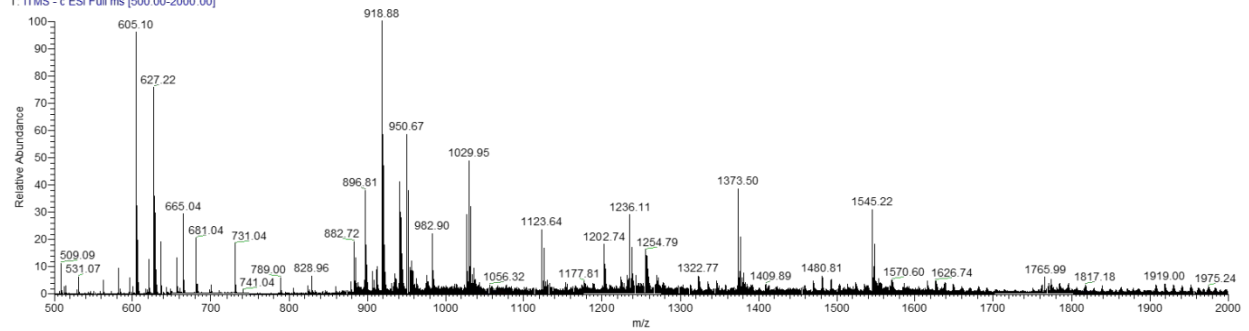
C:\Analytical\...DACat\Substrates\Met

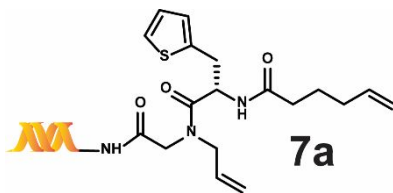
12/11/18 12:11:29

DW672_PA4_A1



Met #548-795 RT: 1.52-2.19 AV: 248 NL: 2.45E5
T: ITMS - c ESI Full ms [500.00-2000.00]



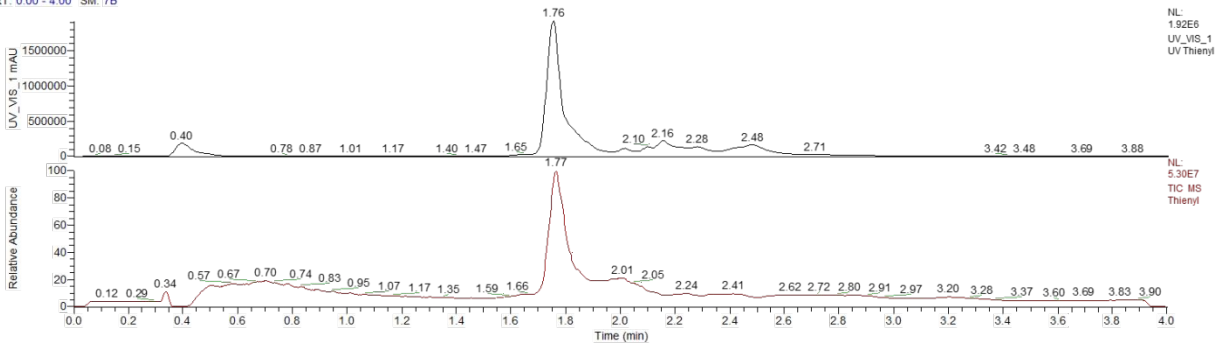


C:\Analytical\...Thienyl

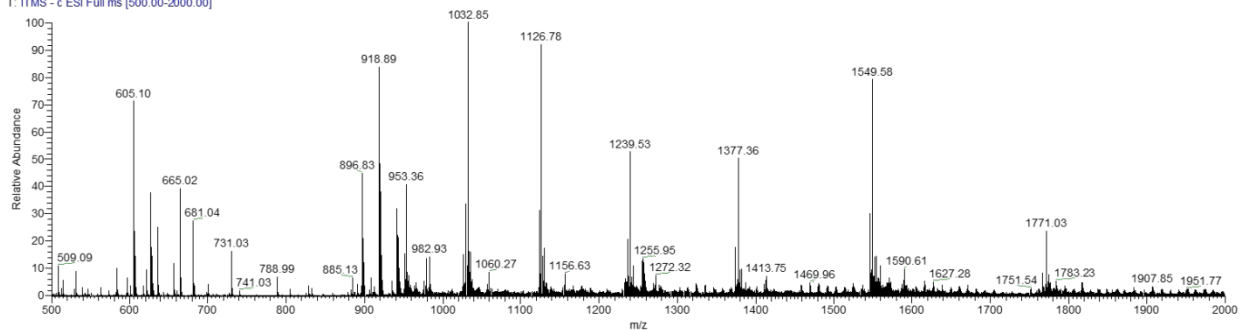
12/11/18 10:44:31

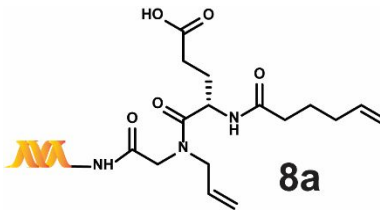
DW672_PA4_A1

RT: 0.00 - 4.00 SM: 7B



Thienyl #581-778 RT: 1.61-2.14 AV: 198 NL: 3.53E5
T: ITMS - c ESI Full ms [500.00-2000.00]

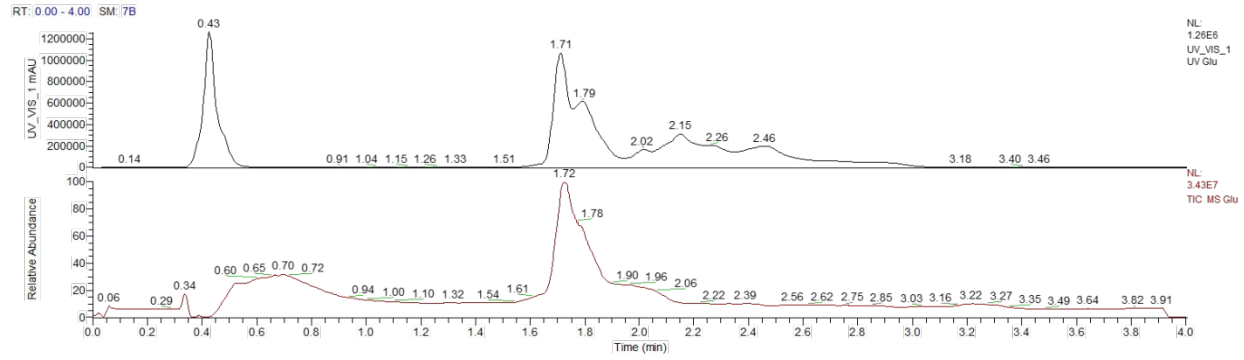




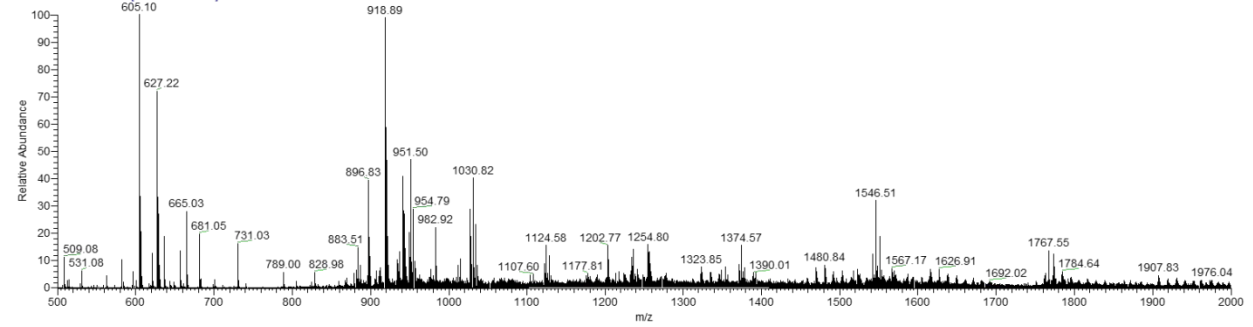
C:\Analytical...IDACat\Substrates\Glu

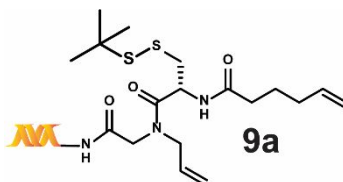
12/11/18 12:21:00

DW672_PA4_A1



Glu #537-786 RT: 1.49-2.16 AV: 250 NL: 2.37E5
T: ITMS - c ESI Full ms [500.00-2000.00]

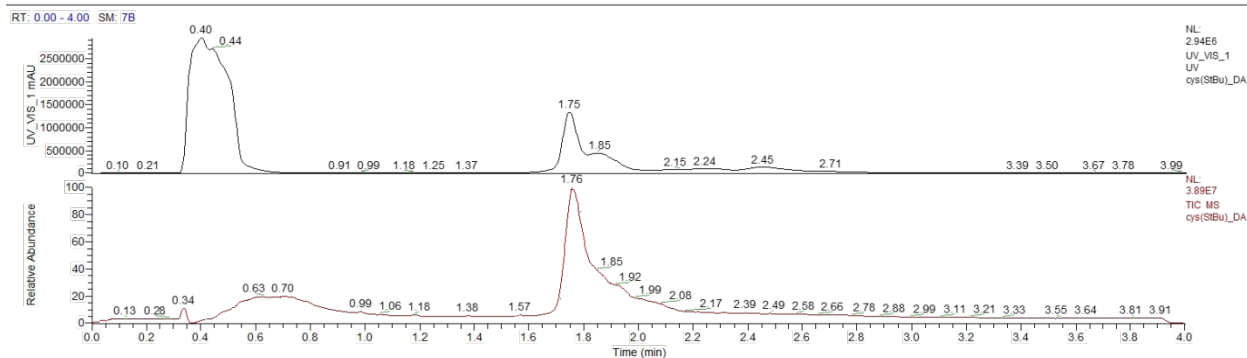




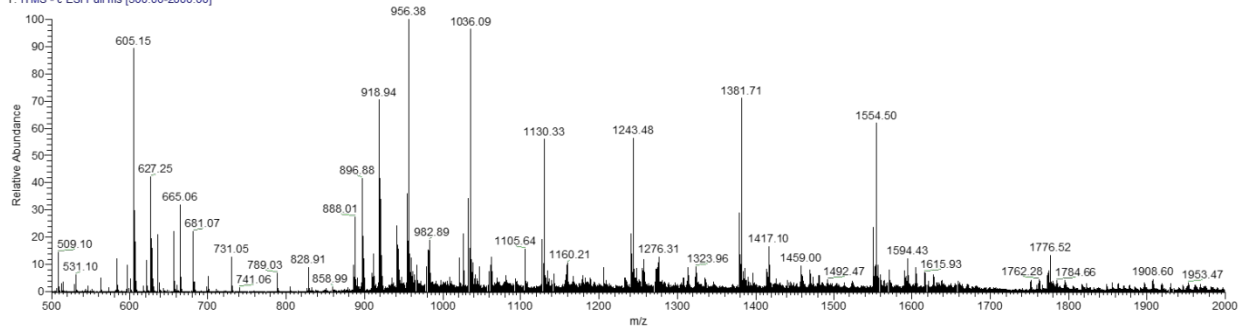
C:\Analytical\...cys(SiBu)_DA

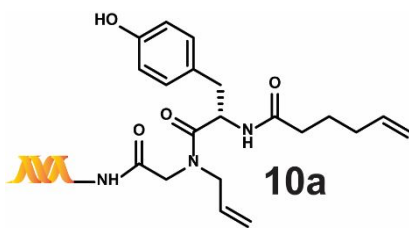
01/21/19 16:15:30

DW672_PA4_A1



cys(SiBu)_DA #557-820 RT: 1.56-2.28 AV: 264 NL: 1.96E5
T: ITMS - c ESI Full ms [500.00-2000.00]



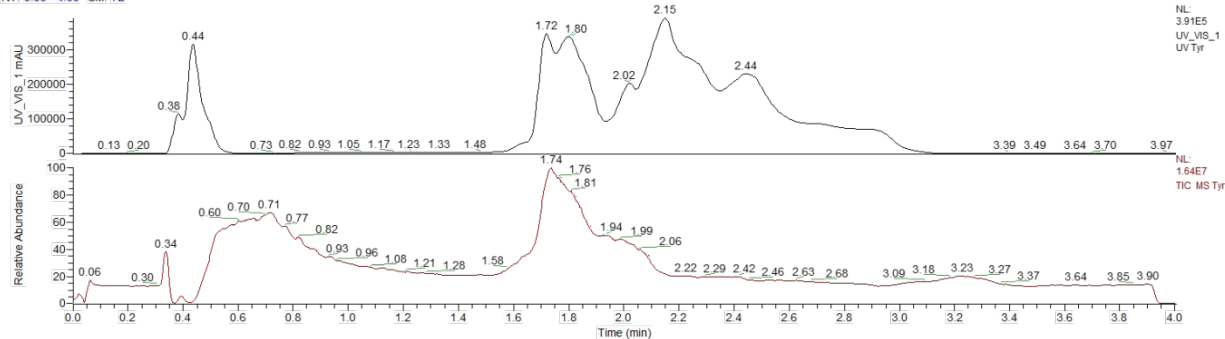


C:\Analytical\...DACatAllSubstrates\Tyr

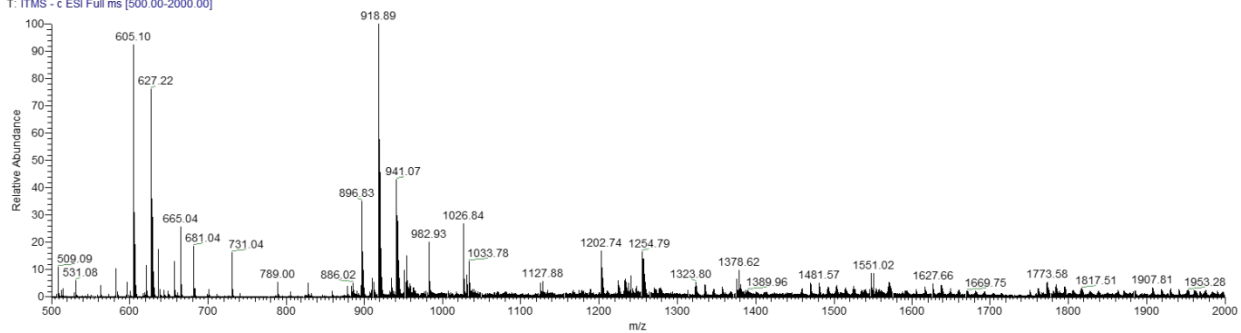
12/11/18 12:35:16

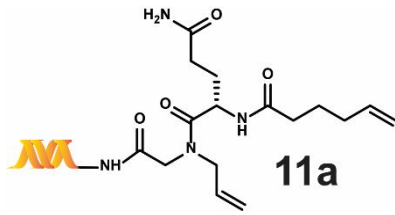
DW672_PA4_A1

RT: 0.00 - 4.00 SM: 7B



Tyr #550-784 RT: 1.52-2.16 AV: 235 NL: 2.70E5
T: ITMS - c ESI Full ms [500.00-2000.00]

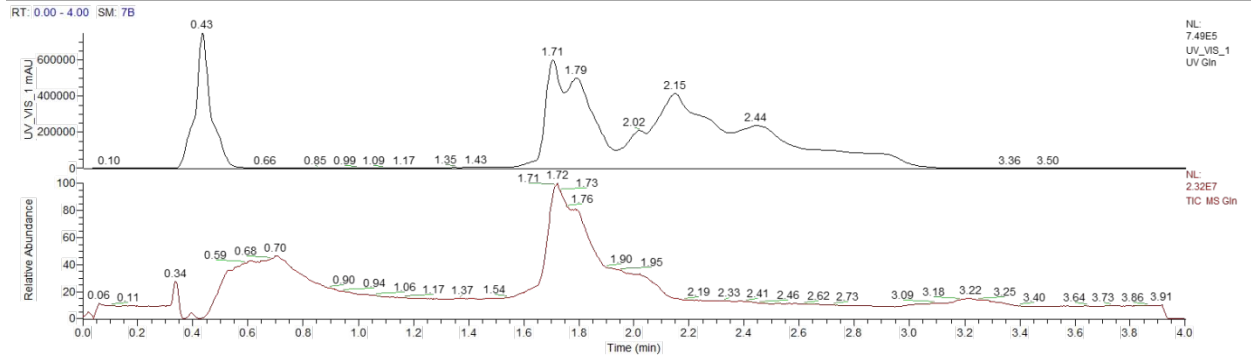




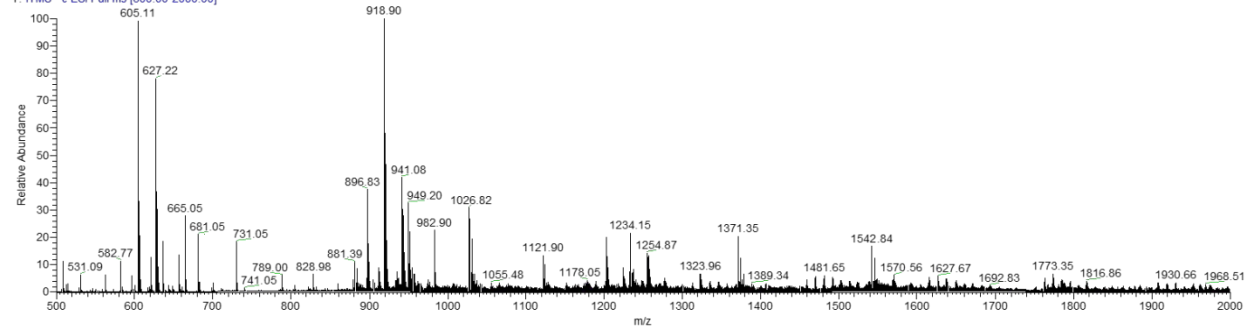
C:\Analytical\...DACatAllSubstrates\Gln

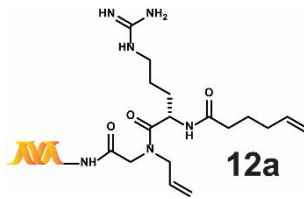
12/11/18 12:39:59

DW672_PA4_A1



Gln #538-809 RT: 1.49-2.23 AV: 272 NL: 2.29E5
T: ITMS - c ESI Full ms [500.00-2000.00]

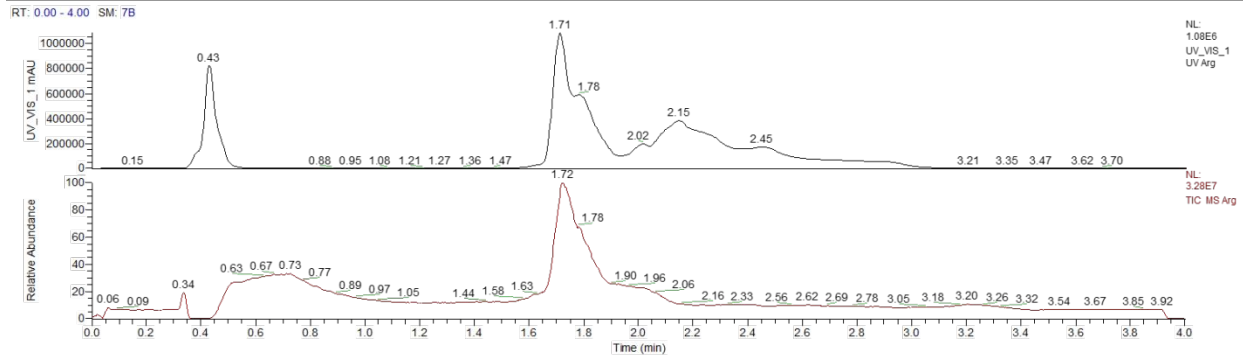




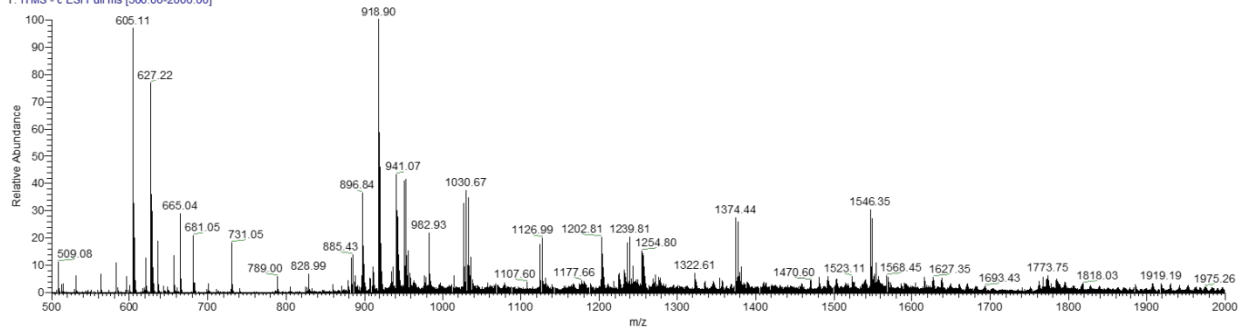
C:\Analytical\IDACat\Substrates\Arg

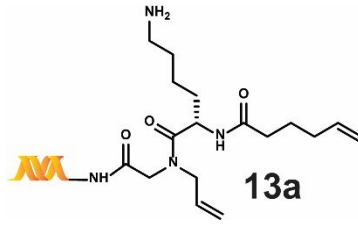
12/11/18 12:25:46

DW672_PA4_A1



Arg #543-814 RT: 1.50-2.24 AV: 272 NL: 2.25E5
T: ITMS - c ESI Full ms [500.00-2000.00]

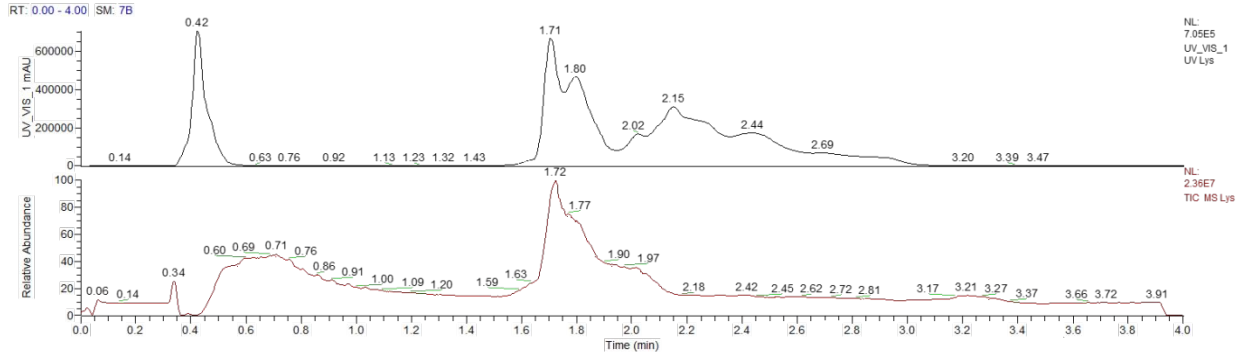




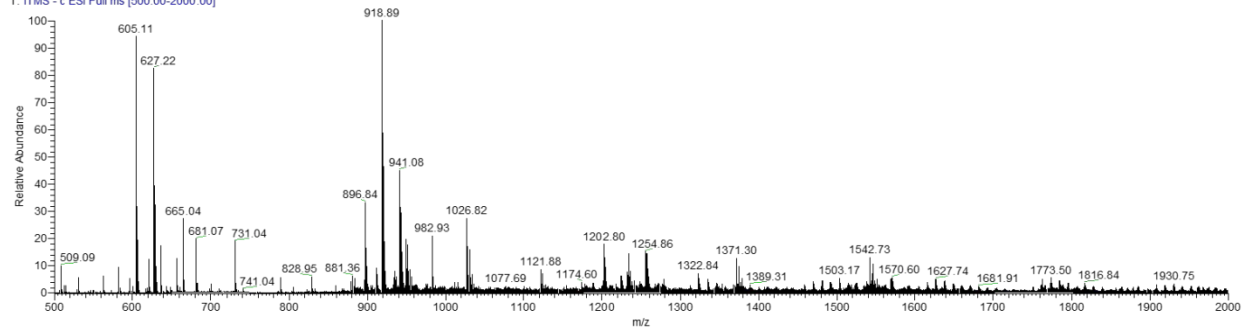
C:\Analytical...IDAcat\Substrates\Lys

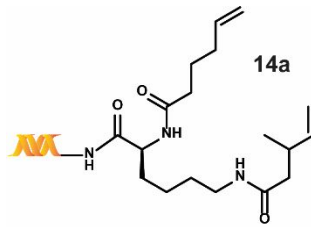
12/11/18 12:16:15

DW672_PA4_A1



Lys #647-831 RT: 1.51-2.29 AV: 265 NL: 2.41E5
T: ITMS - c ESI Full ms [500.00-2000.00]

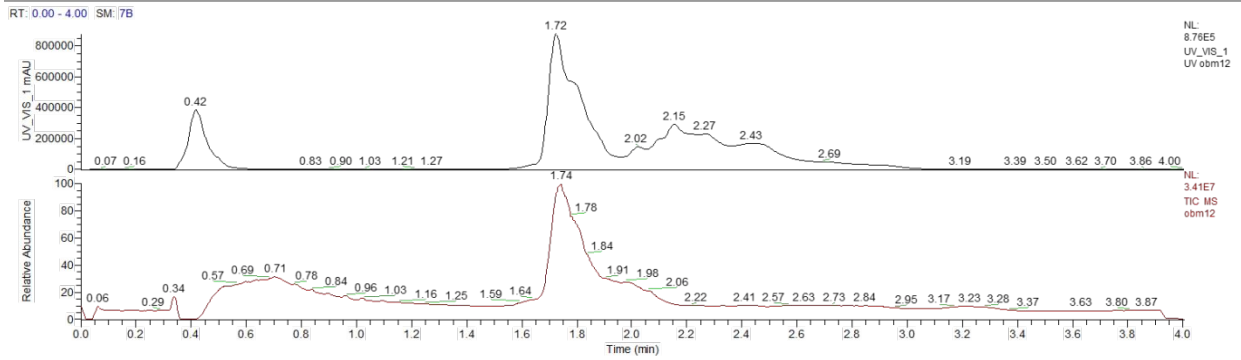




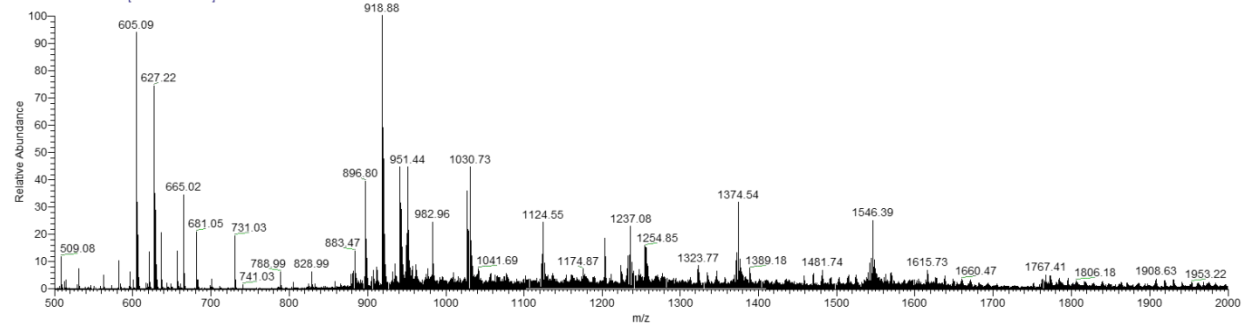
C:\Analytical\lobm12

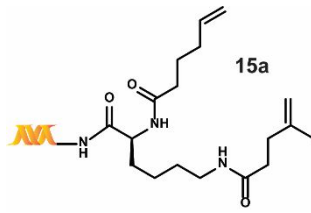
12/11/18 11:28:45

DW672_PA4_A1



obm12 #567-794 RT: 1.57-2.18 AV: 228 NL: 2.45E5
T: ITMS - c ESI Full ms [500.00-2000.00]

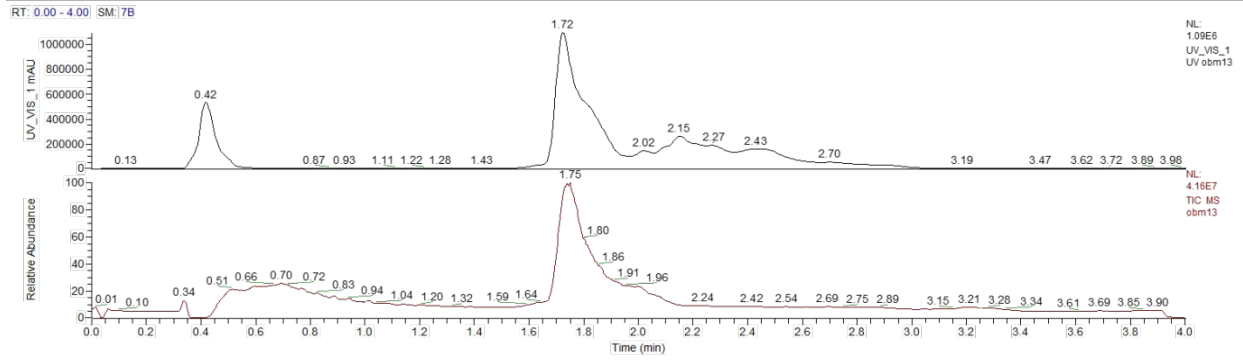




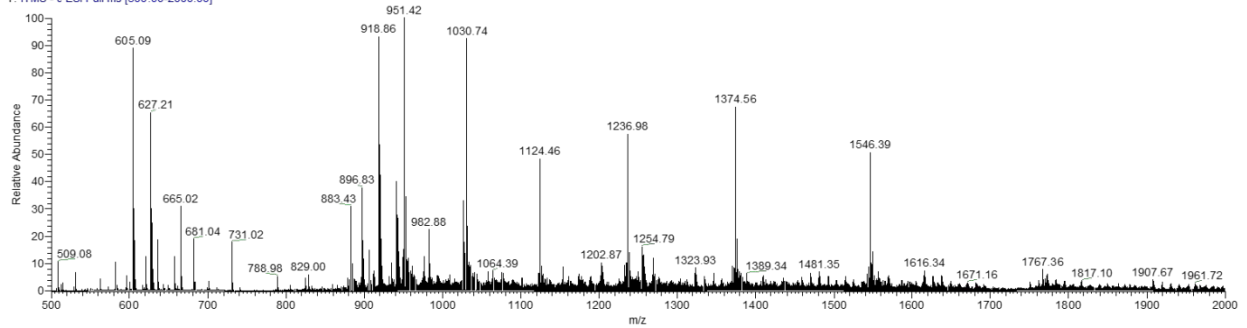
C:\Analytical_lobm13

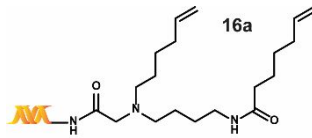
12/11/18 11:33:31

DW672_PA4_A1



obm13 #549-786 RT: 1.52-2.16 AV: 238 NL: 2.56E5
T: ITMS - c ESI Full ms [500.00-2000.00]



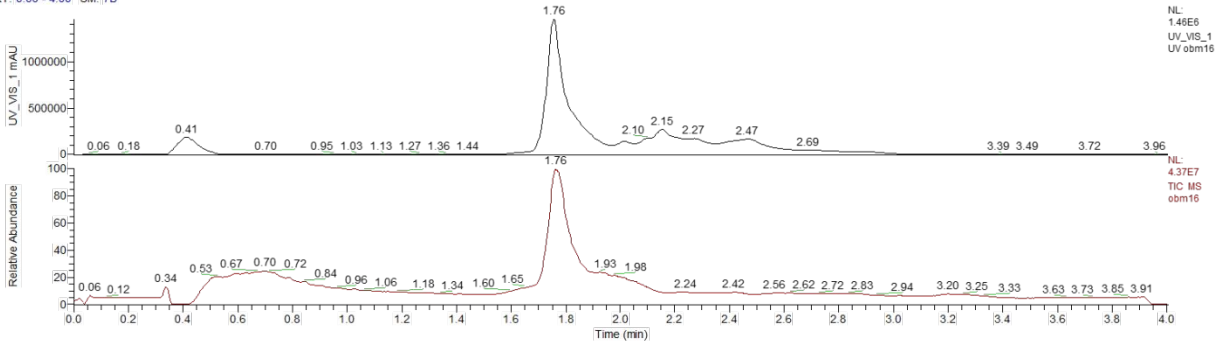


C:\Analytical\lobm16

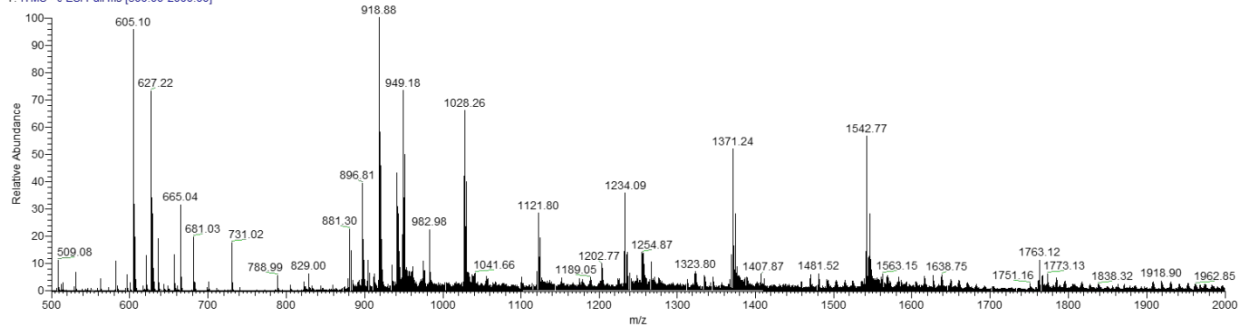
12/11/18 11:43:00

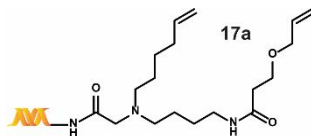
DW672_PA4_A1

RT: 0.00 - 4.00 SM: 7B



obm16 #557-787 RT: 1.54-2.16 AV: 231 NL: 2.51E5
T: ITMS - c ESI Full ms [500.00-2000.00]

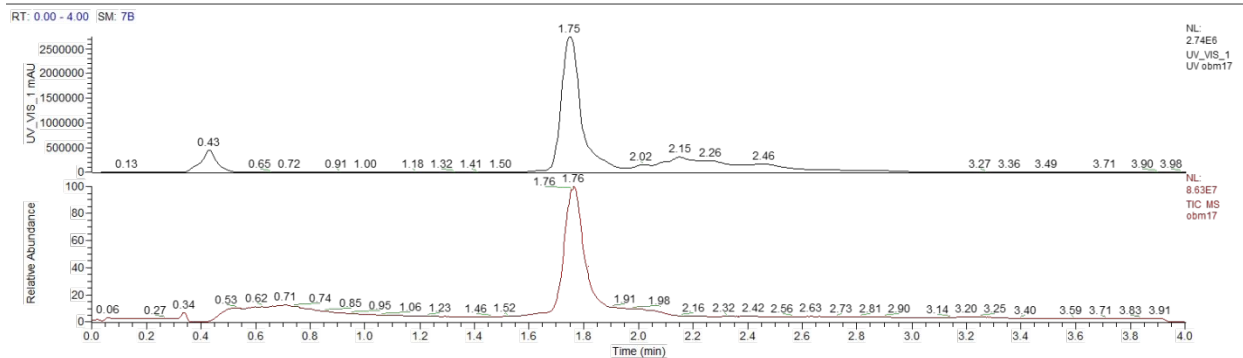




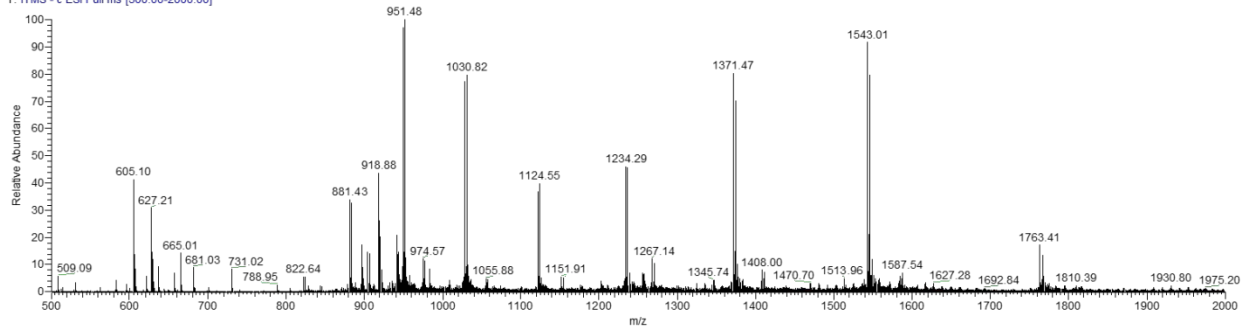
C:\Analytical_lobm17

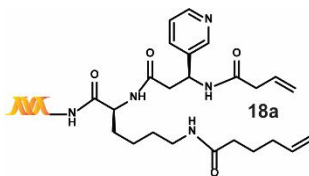
12/11/18 11:47:44

DW672_PA4_A1



obm17 #551-755 RT: 1.52-2.07 AV: 205 NL: 5.53E5
T: ITMS - c ESI Full ms [500.00-2000.00]

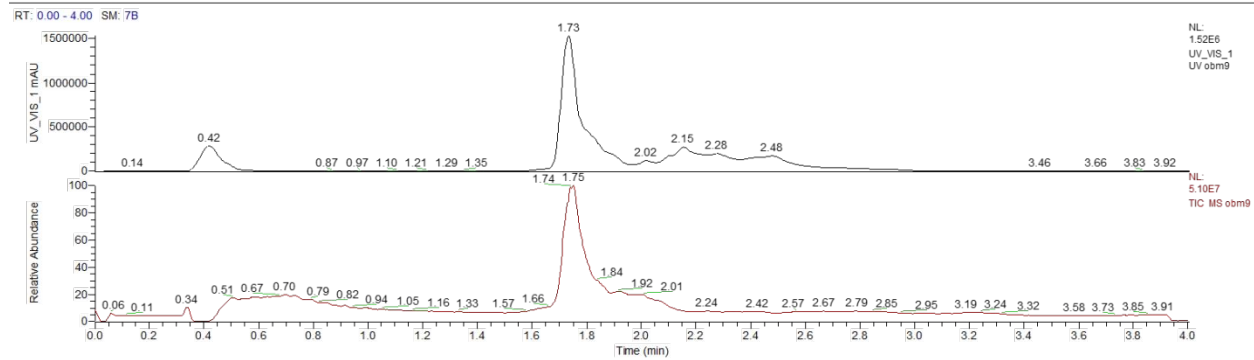




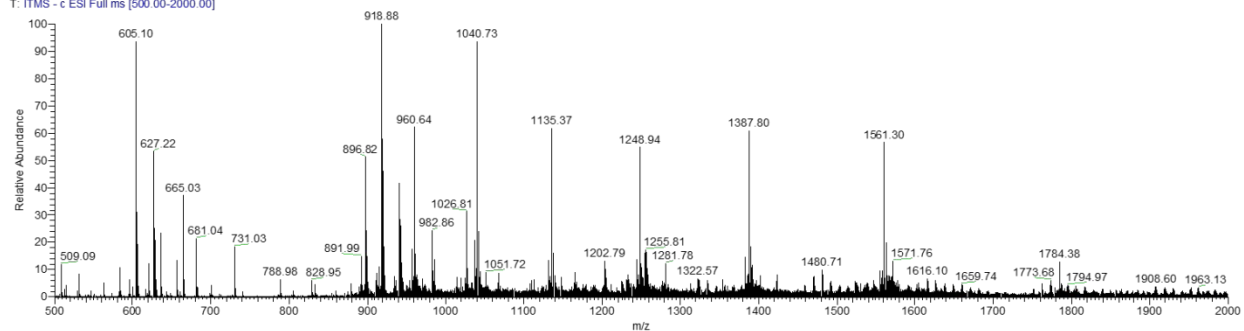
C:\Analytical...lobm9

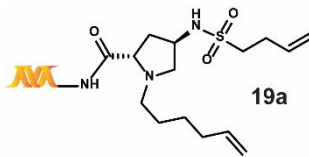
12/11/18 11:14:30

DW672_PA4_A1



obm9 #545-793 RT: 1.51-2.18 AV: 249 NL: 2.47E5
T: ITMS - c ESI Full ms [500.00-2000.00]

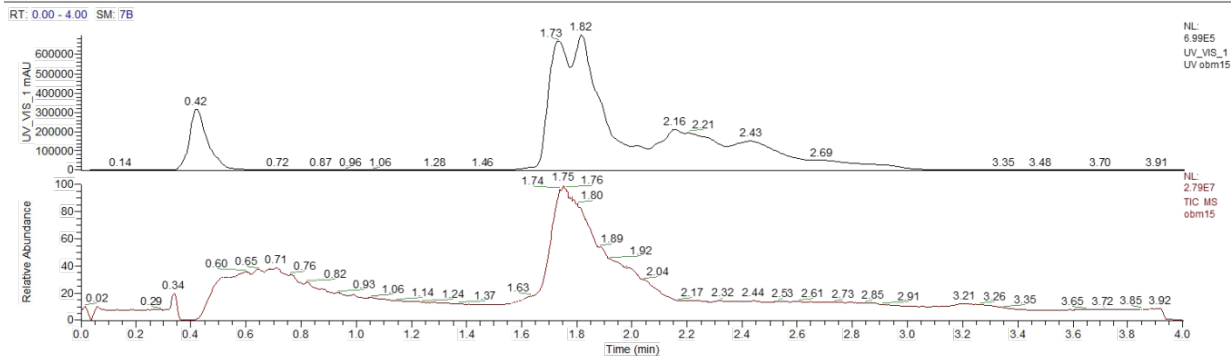




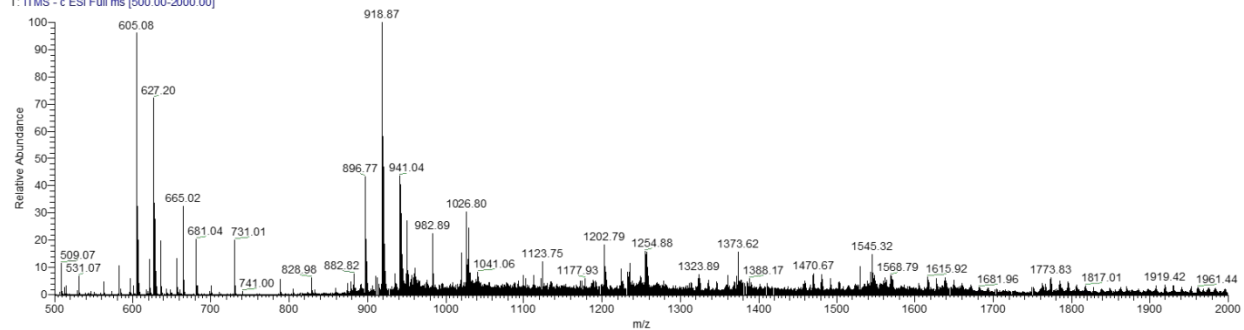
C:\Analytical\lobm15

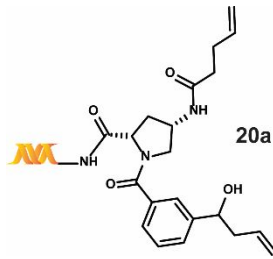
12/11/18 11:38:14

DW672_PA4_A1



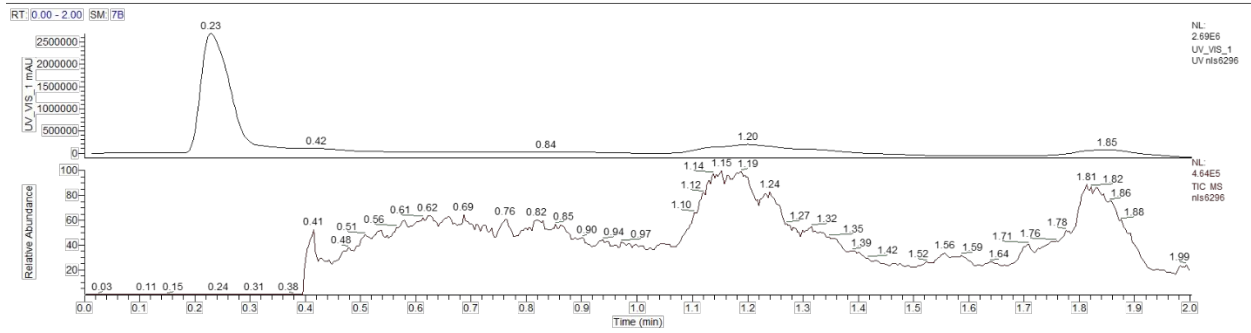
obm15 #553-795 RT: 1.53-2.19 AV: 243 NL: 2.42E5
T: ITMS - c ESI Full ms [500.00-2000.00]



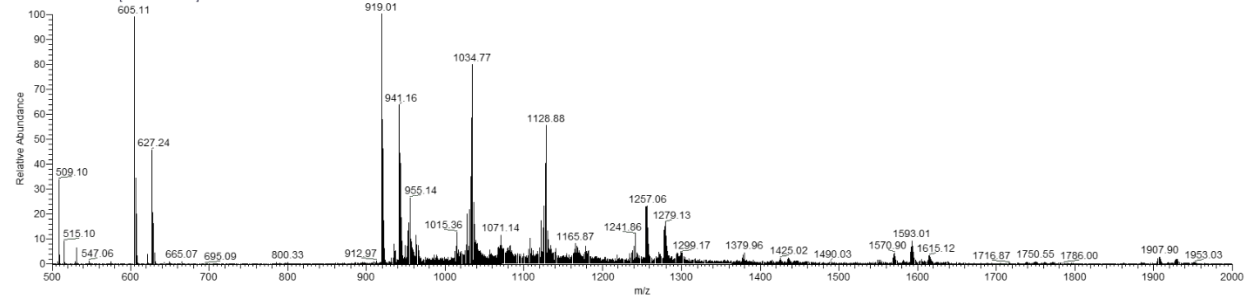


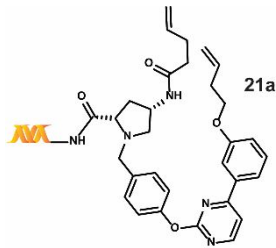
C:\Analytical\...nls6296

12/10/18 17:18:53



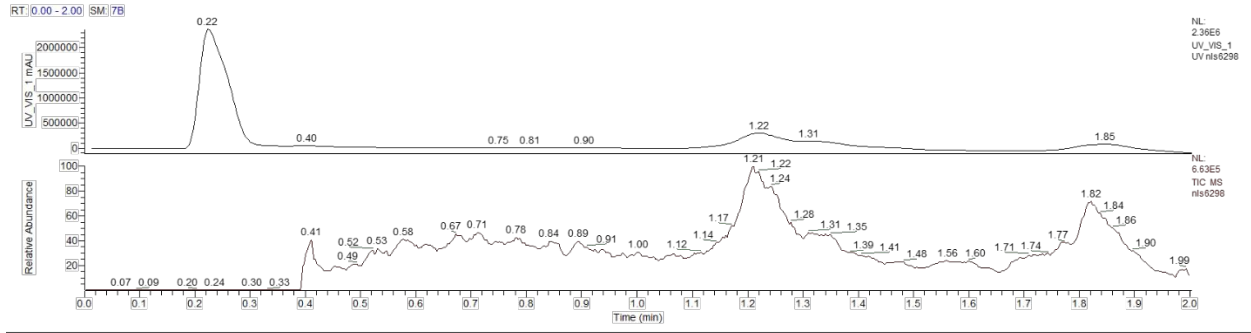
nls6296 #370-496 RT: 1.04-1.37 AV: 117 NL: 9.77E3
T: ITMS - c ESI Full ms [500.00-2000.00]



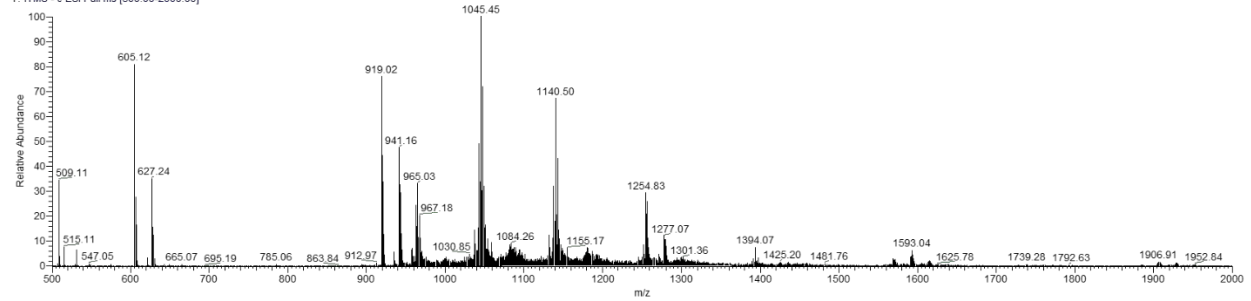


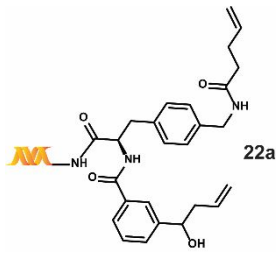
C:\Analytical\...nls6298

12/10/18 17:24:29



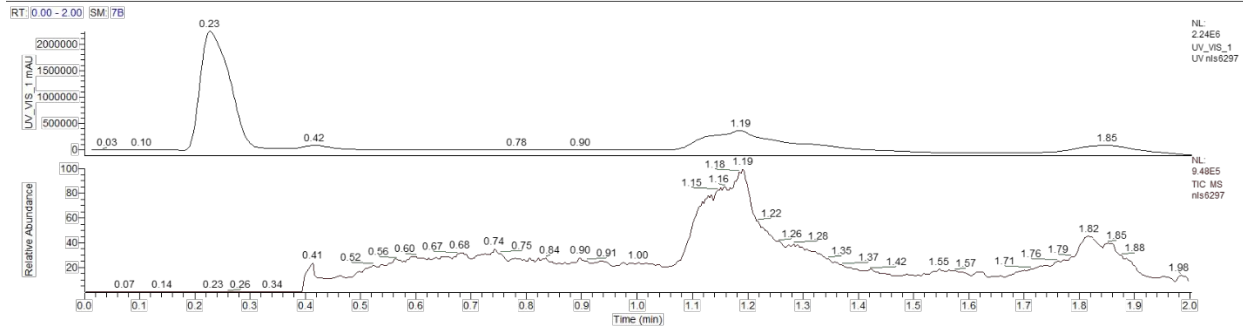
nls6298 #387-497 RT: 1.09-1.40 AV: 111 NL: 1.17E4
T: ITMS - c ESI Full ms [500.00-2000.00]



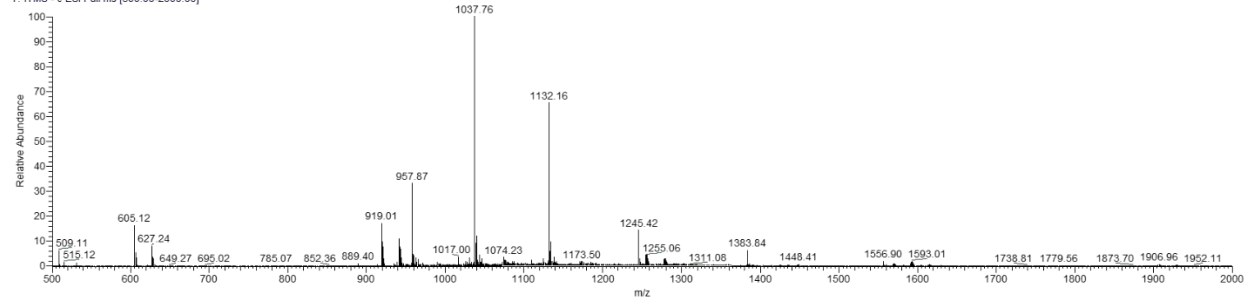


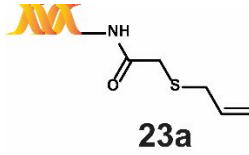
C:\Analytical\...nls6297

12/10/18 17:21:44



nls6297 #371-490 RT: 1.05-1.38 AV: 120 NL: 5.79E4
T: ITMS - c ESI Full ms [500.00-2000.00]

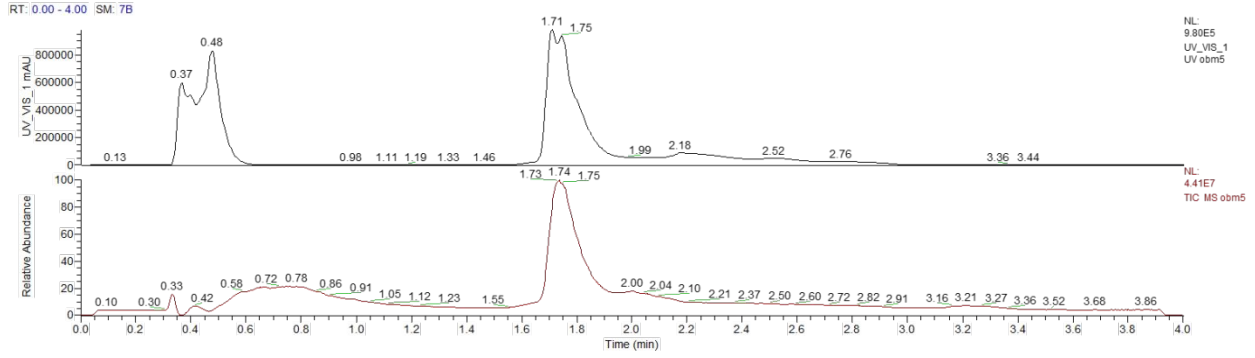




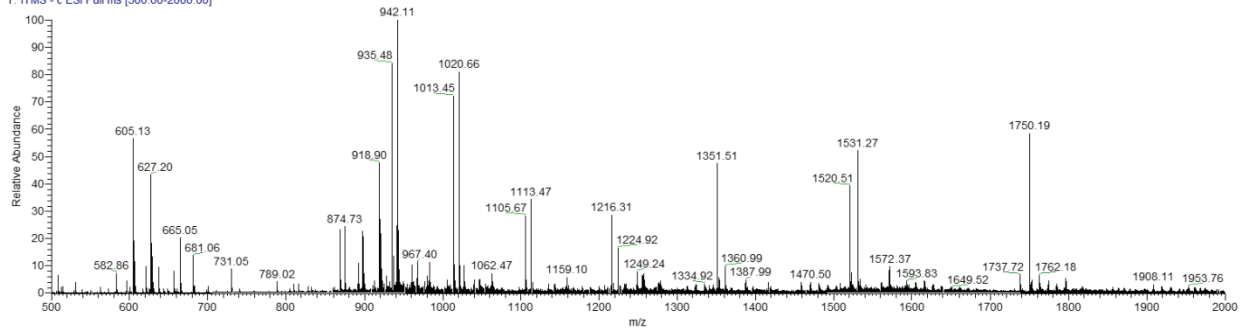
C:\Analytical...lobm5

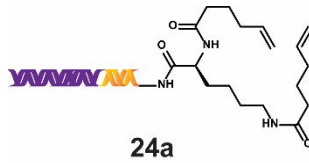
12/13/18 17:04:35

DW672_PA4_A1



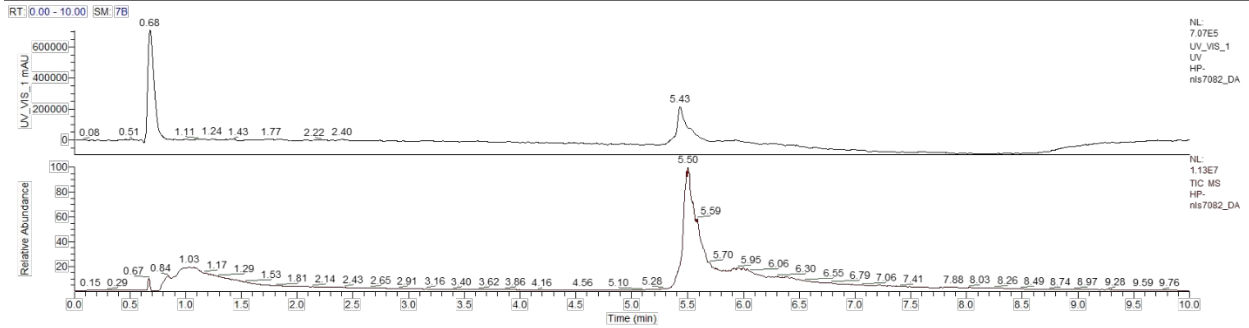
obm5 #549-798 RT: 1.53-2.20 AV: 250 NL: 3.72E5
T: ITMS - c ESI Full ms [500.00-2000.00]



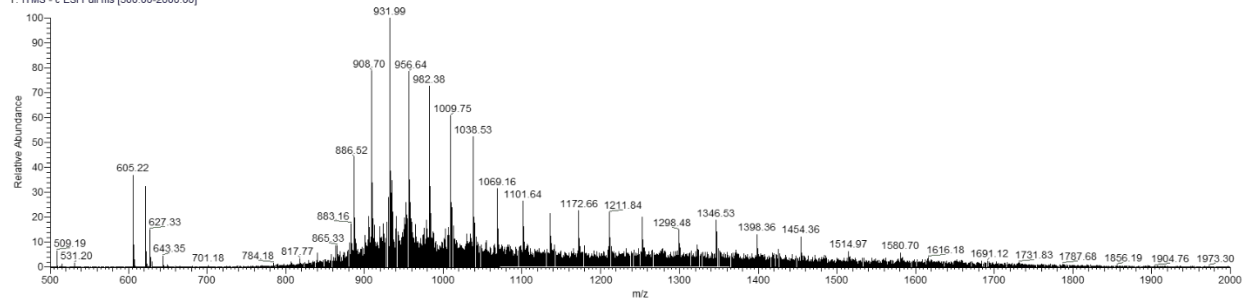


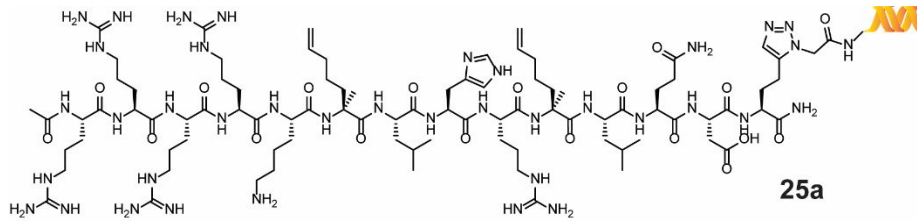
C:\Analytical\..HP-nis7082_DA

03/12/19 17:57:01



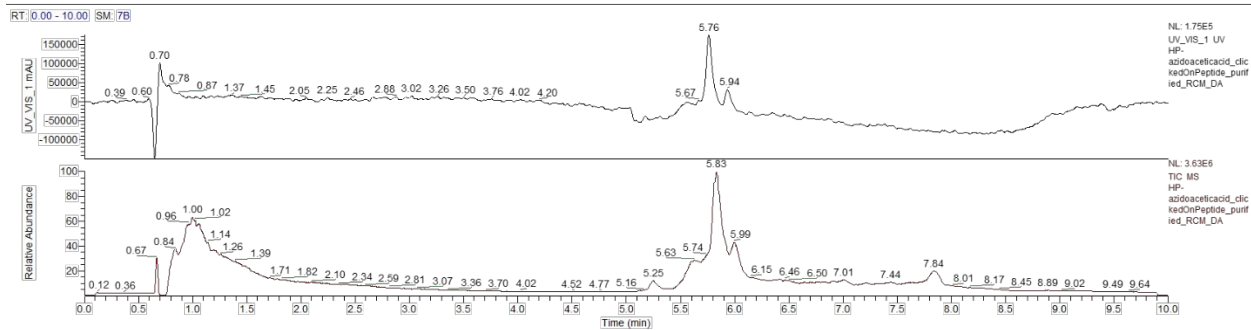
HP-nis7082_DA #1649-2340 RT: 5.20-6.54 AV: 492 NL: 3.63E4
T: ITMS - c ESI Full.ms [500.00-2000.00]





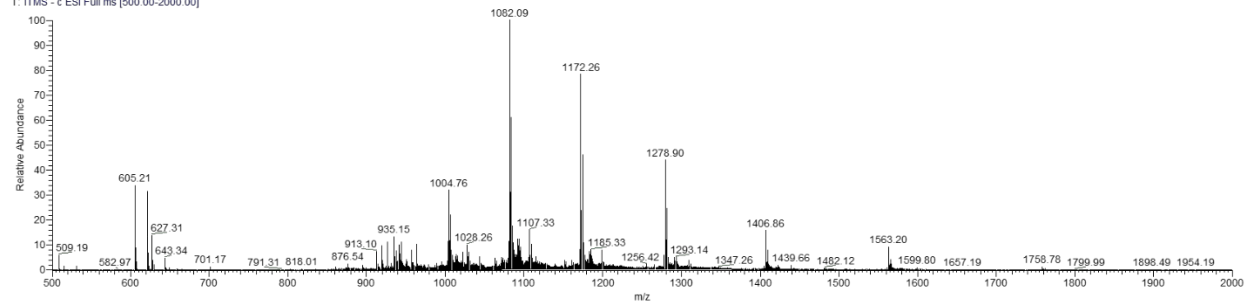
HP-azidoaceticacid_clickedOnPeptide_p...

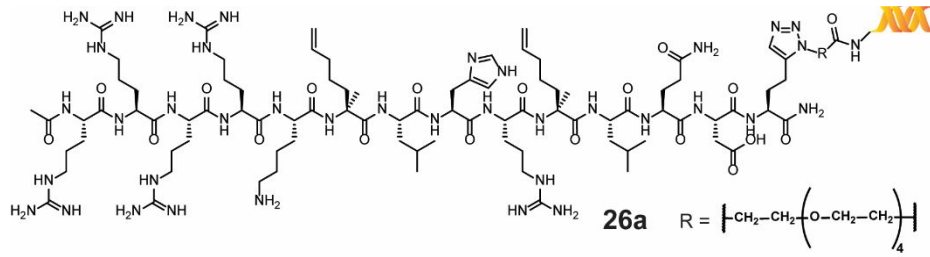
03/12/19 17:35:17



HP-azidoaceticacid_clickedOnPeptide_purified_RCM_DA #1617-2321 RT: 5.11-6.52 AV: 505 NL: 4.05E4

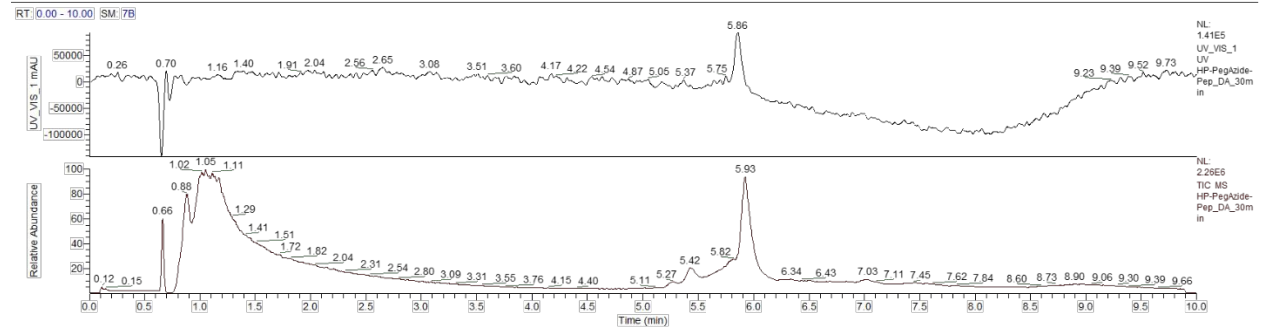
T: ITMS - c ESI Full ms [500.00-2000.00]



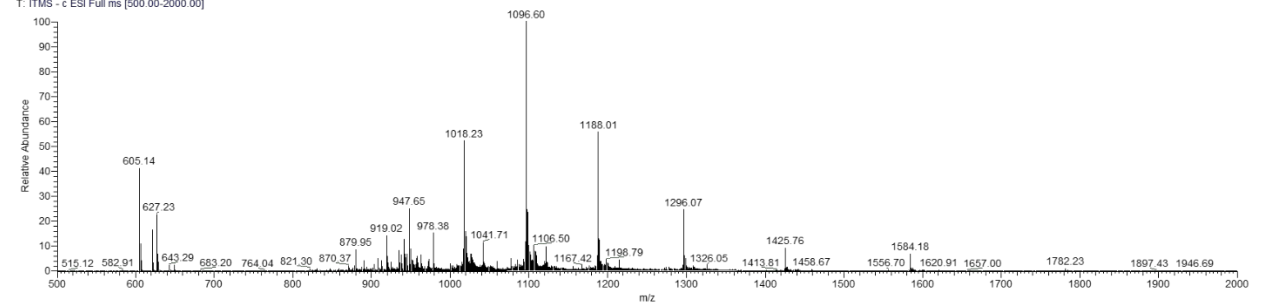


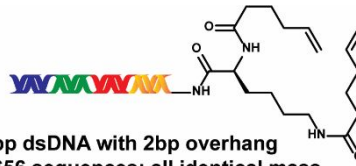
HP-PegAzide-Pep_DA_30min

03/27/19 12:51:03

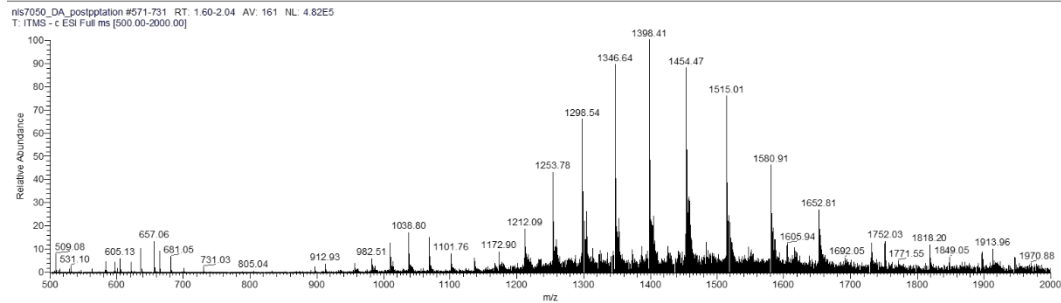
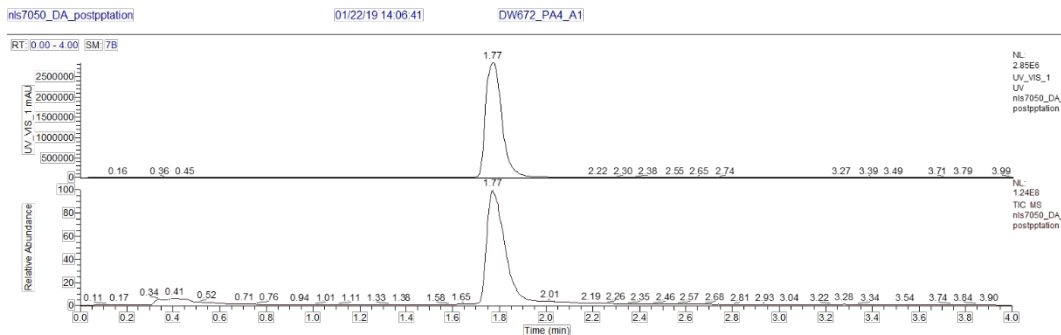
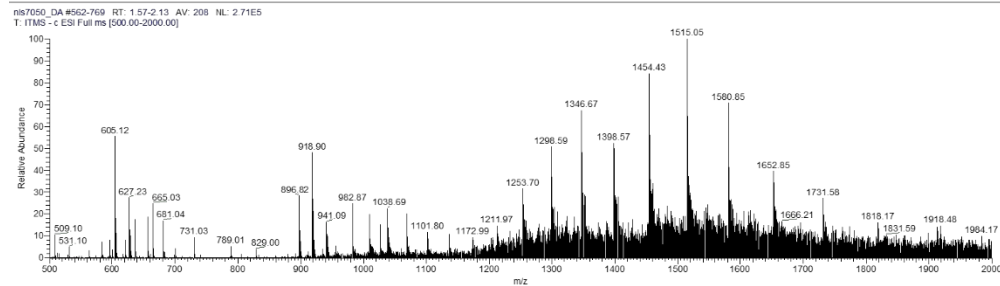
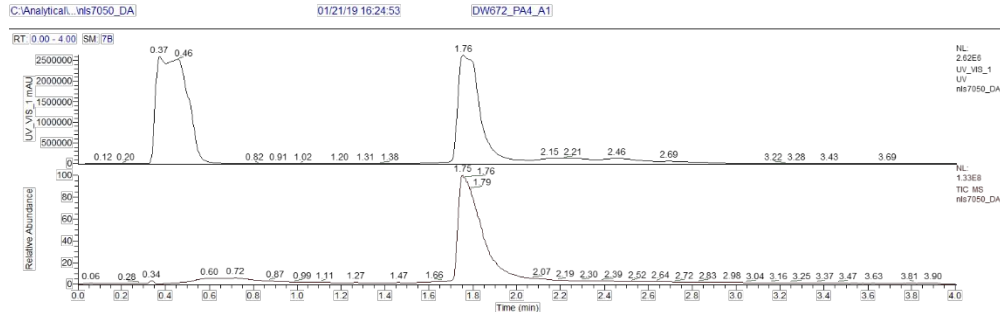


HP-PegAzide-Pep_DA_30min #1809-2244 RT: 5.08-6.31 AV: 436 NL: 3.10E4
T: ITMS - c ESI Full ms (500.00-2000.00)





Post RCM reaction



Post ethanol
precipitation of RCM
reaction

References

- (1) Du, H. C.; Simmons, N.; Faver, J. C.; Yu, Z.; Palaniappan, M.; Riehle, K.; Matzuk, M. M. A Mild, DNA-compatible nitro reduction using $B_2(OH)_4$. *Org. Lett.* **2019**, *21* (7), 2194–2199.
- (2) Li, J. Y.; Huang, H. Development of DNA-Compatible Suzuki-Miyaura Reaction in Aqueous Media. *Bioconjug. Chem.* **2018**, *29* (11), 3841–3846.
- (3) Du, H. C.; Huang, H. DNA-Compatible Nitro Reduction and Synthesis of Benzimidazoles. *Bioconjug. Chem.* **2017**, *28* (10), 2575–2580.
- (4) Du, H. C.; Bangs, M. C.; Simmons, N.; Matzuk, M. M. Multistep synthesis of 1,2,4-Oxadiazoles via DNA-Conjugated aryl nitrile substrates. *Bioconjug. Chem.* **2019**, *30* (5), 1304–1308.
- (5) Higman, C. S.; Nascimento, D. L.; Ireland, B. J.; Audörsch, S.; Bailey, G. A.; McDonald, R.; Fogg, D. E. Chelate-Assisted Ring-Closing Metathesis: A Strategy for Accelerating Macrocyclization at Ambient Temperatures. *J. Am. Chem. Soc.* **2018**, *140* (5), 1604–1607.
- (6) William, A. D.; Lee, A. C. H.; Poulsen, A.; Goh, K. C.; Madan, B.; Hart, S.; Tan, E.; Wang, H.; Nagaraj, H.; Chen, D.; et al. Discovery of the macrocycle (9 E)-15-(2-(Pyrrolidin-1-yl)ethoxy)-7,12,25-trioxa-19,21,24-triaza-tetracyclo[18.3.1.1(2,5).1(14,18)]hexacos-1(24),2,4,9,14(26),15,17,20,22-nonaene (SB1578), a potent inhibitor of Janus kinase 2/Fms-liketyrosine kinase-3 (JAK2/FLT3) for the treatment of rheumatoid arthritis. *J. Med. Chem.* **2012**, *55* (6), 2623–2640.
- (7) Faver, J. C.; Riehle, K.; Lancia, D. R.; Milbank, J. B. J.; Kollmann, C. S.; Simmons, N.; Yu, Z.; Matzuk, M. M. Quantitative Comparison of Enrichment from DNA-Encoded Chemical Library Selections. *ACS Comb. Sci.* **2019**, *21* (2), 75–82.
- (8) Dinger, M. B.; Mol, J. C. Degradation of the second-generation Grubbs metathesis catalyst with primary alcohols and oxygen - Isomerization and hydrogenation activities of monocarbonyl complexes. *Eur. J. Inorg. Chem.* **2003**, *2003* (15), 2827–2833.
- (9) Lu, X.; Fan, L.; Phelps, C. B.; Davie, C. P.; Donahue, C. P. Ruthenium Promoted On-DNA Ring-Closing Metathesis and Cross-Metathesis. *Bioconjug. Chem.* **2017**, *28* (6), 1625–1629.