Supporting Information for:

A Chemoselective Rapid Azo-Coupling Reaction (CRACR) for Unclickable Bioconjugation

Partha Sarathi Addy, Sarah B. Erickson, James S. Italia, and Abhishek Chatterjee*

*Department of Chemistry, Boston College, 2609 Beacon Street, 246B Merkert Chemistry Center, Chestnut Hill, Massachusetts 02467, United States. Tel: +1-617-552-1778

email: abhishek.chatterjee@bc.edu

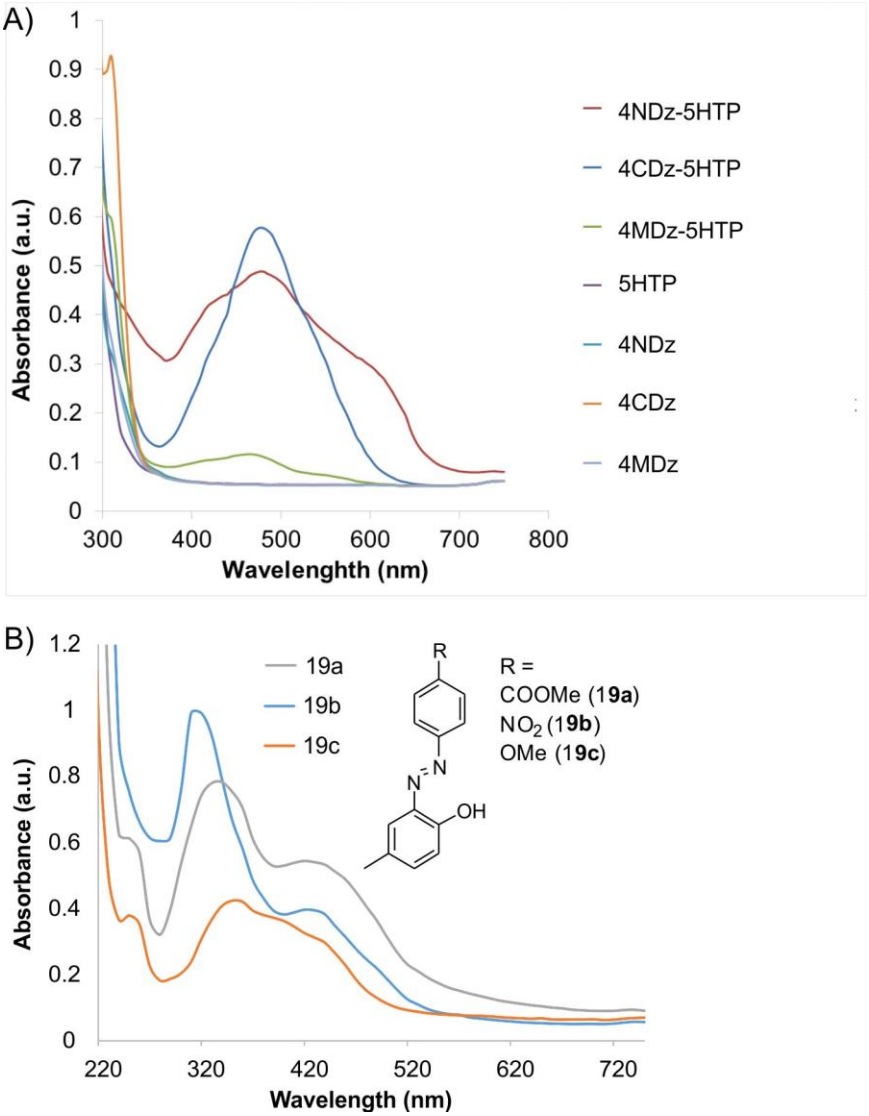


Figure S1: Azo conjugates, but not their precursors, uniquely absorb at 450 nm, allowing facile spectrophotometric monitoring of their formation.

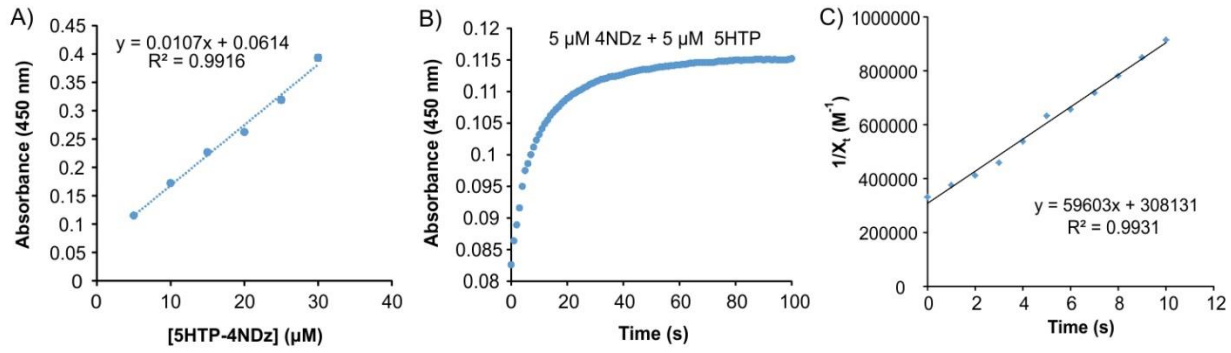


Figure S2: Evaluating the kinetic parameters of the reaction between 4NDz and 5HTP. Due to the very fast nature of this reaction, its rate could only be measured under low concentrations (<10 μM) of both reagents. To measure the kinetic parameters of this bimolecular reaction, we used equal concentrations of both reactants (5 μM each in 100 mM phosphate buffer at pH 7), where the second order rate equation reduces to $1/[X]_t = 1/[X]_0 + k_2t$ ($[X]_t$: concentration of either reactant at time t; $[X]_0$: initial concentration of either reactant; k_2: rate constant). To obtain the concentration of the product (4NDz-5HTP) at a given time from the observed absorbance at 450 nm, a standard curve was first generated (A). This was used to obtain X_t from the observed increase in absorbance at 450 nm (B). Plotting 1/Xt against time yielded the rate constant (C). The reported value (63287 ± 6512 M^{-1} s^{-1}) represents an average of three independent experiments; error represents standard deviation.

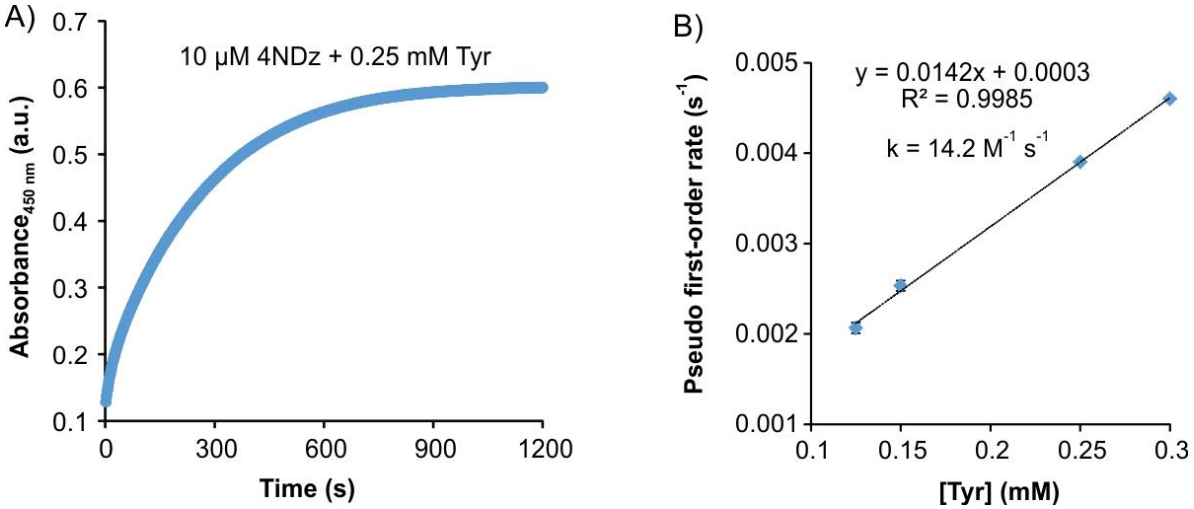


Figure S3: Evaluating the kinetic parameters of the reaction between 4NDz and tyrosine. Rate of product formation was measured under pseudo-first order conditions, where 10 μM 4NDz was mixed with an excess of tyrosine (0.1-0.3 mM) in 100 mM phosphate buffer at pH 7. A representative example is shown (A). The second-order rate constant was obtained by plotting the pseudo first-order rates against the concentration of tyrosine used (B). The reported value (14.2 ± 0.1 M^{-1} s^{-1}) represents an average of three independent experiments; error represents standard deviation.

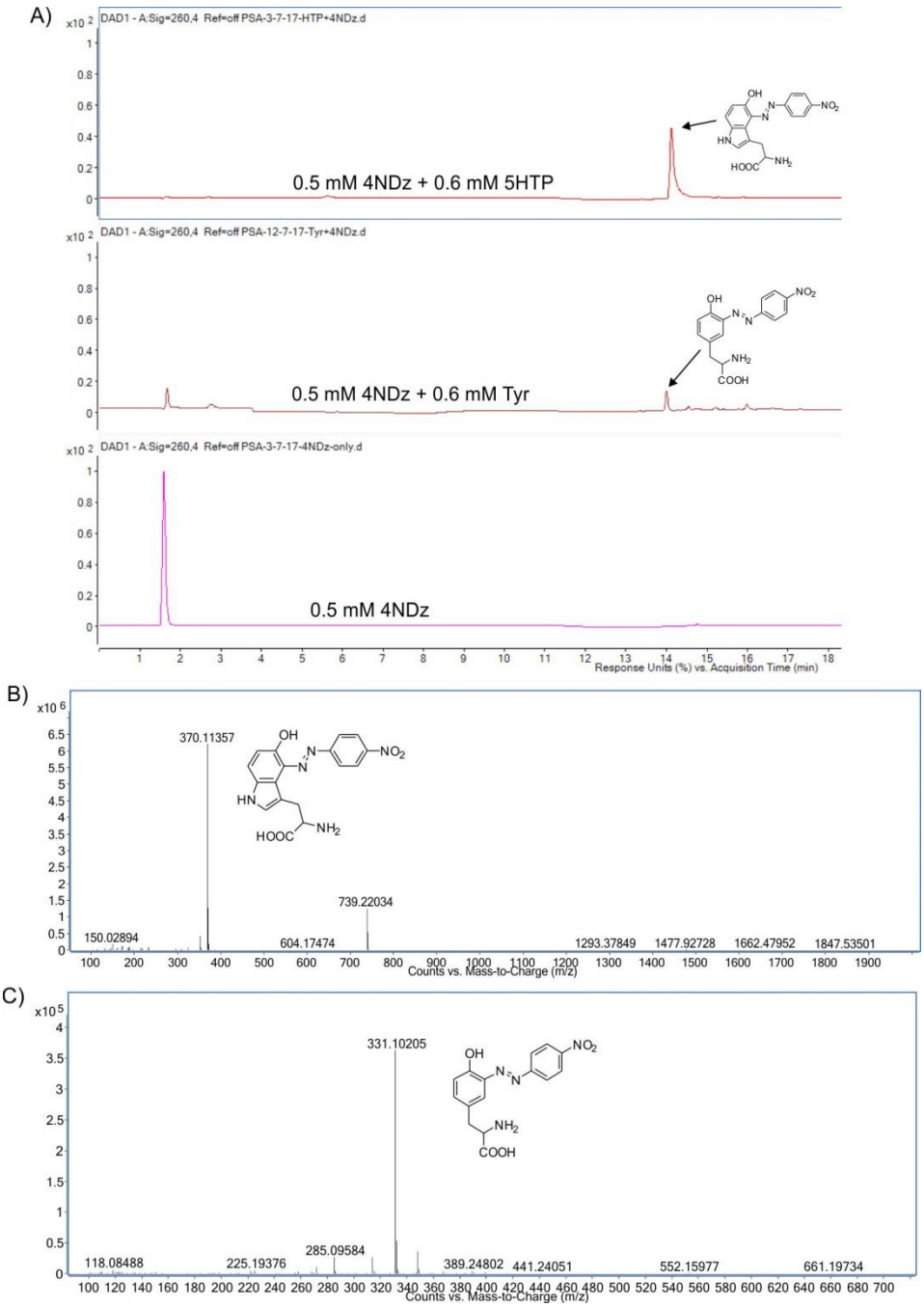


Figure S4: HPLC-MS (ESI-TOF) analysis of the azo-coupling reactions of 4NDz with 5HTP and Tyr. Indicated reactions were allowed to continue at room temperature for 15 minutes, and then analyzed by HPLC-MS. Panel A shows absorption traces at 260 nm; Panel B and C show MS scans of the azo-coupling product peaks.

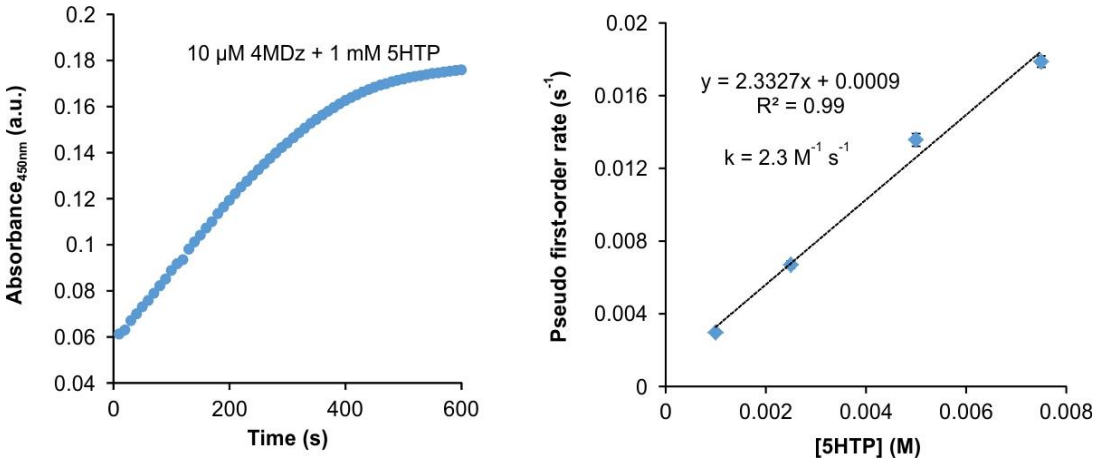


Figure S5: Evaluating the kinetic parameters of the reaction between 4MDz and 5HTP. Rate of product formation was measured under pseudo-first order conditions, where 10 µM 4MDz was mixed with an excess of 5HTP (1-8 mM) in 100 mM phosphate buffer at pH 7. A representative example is shown (A). The second-order rate constant was obtained by plotting the pseudo first-order rates against the concentration of 5HTP used (B). The reported value (2.3 ± 0.1 M^{-1} s^{-1}) represents an average of three independent experiments; error represents standard deviation.

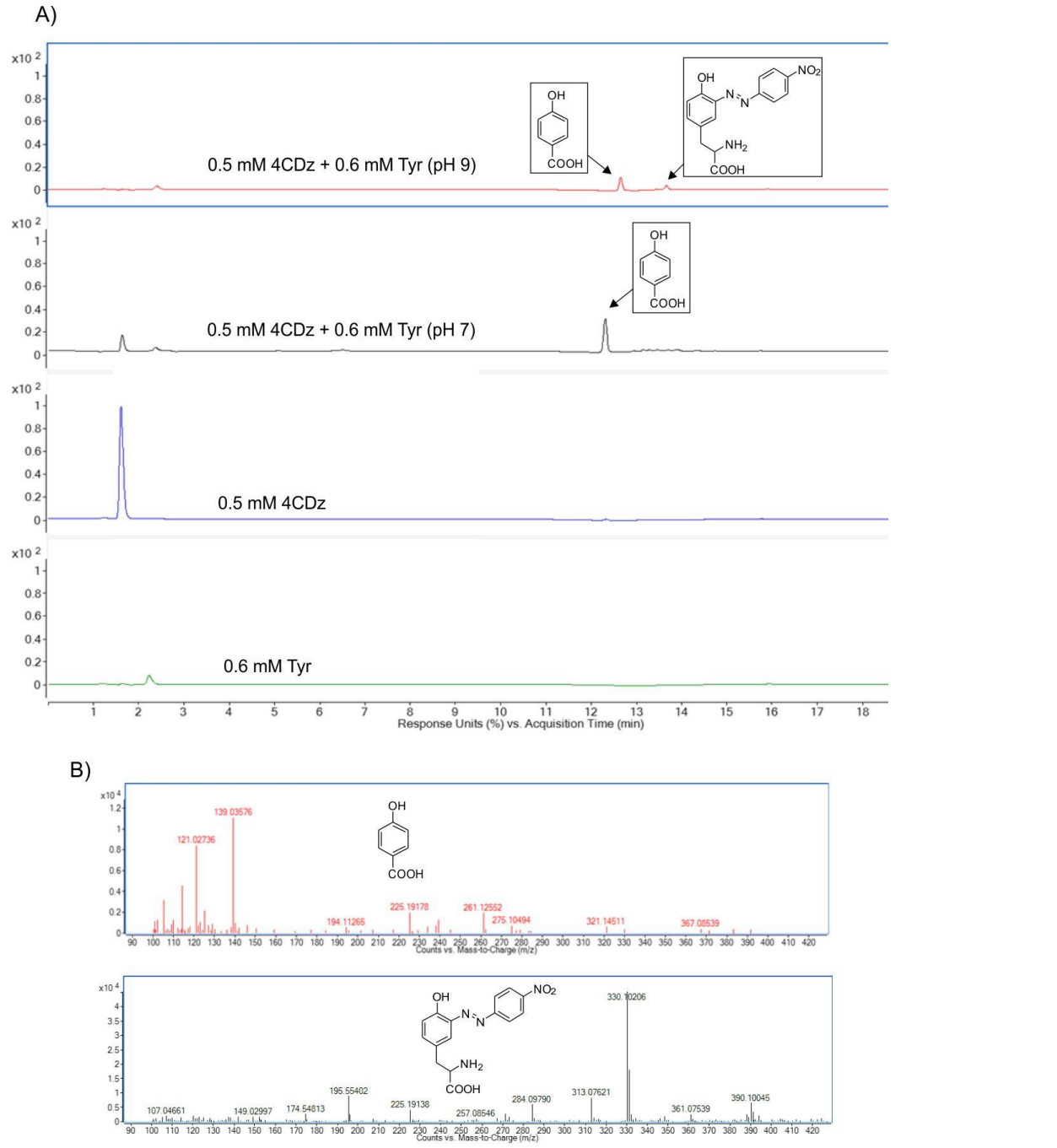


Figure S6: A) HPLC traces collected at 260 nm using a DAD detector for reaction between 4CDz (0.5 mM) with Tyr (0.6 mM) in 100 mM phosphate buffer (pH 7) show no formation of the azo-coupling product even after 30 min incubation, but some hydrolysis of 4CDz. However, at pH 9, some azo-coupling product is observed along with the hydrolysis of 4CDz. B) MS-scans of the hydrolysis product of 4CDz and the azo-coupling product observed at pH 9, collected using an in-line ESI-TOF.

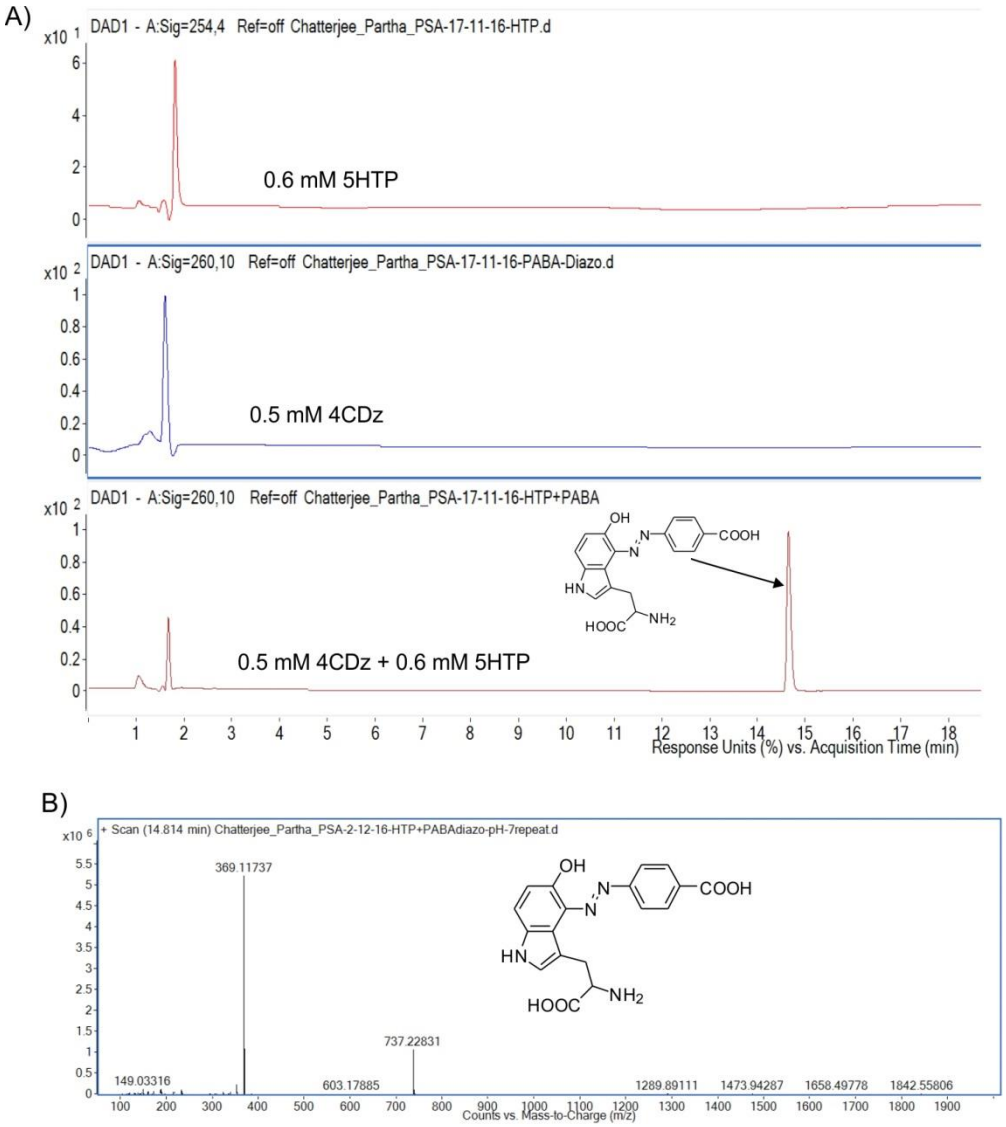


Figure S7: A) HPLC traces collected at 260 nm using a DAD detector for reaction between 4CDz and 5HTP. Indicated reactions in 100 mM phosphate buffer (pH 7) were allowed to continue at room temperature for 10 minutes, and then analyzed by HPLC-MS. B) MS scans of the azo-coupling product peak, collected using an in-line ESI-TOF, is shown.

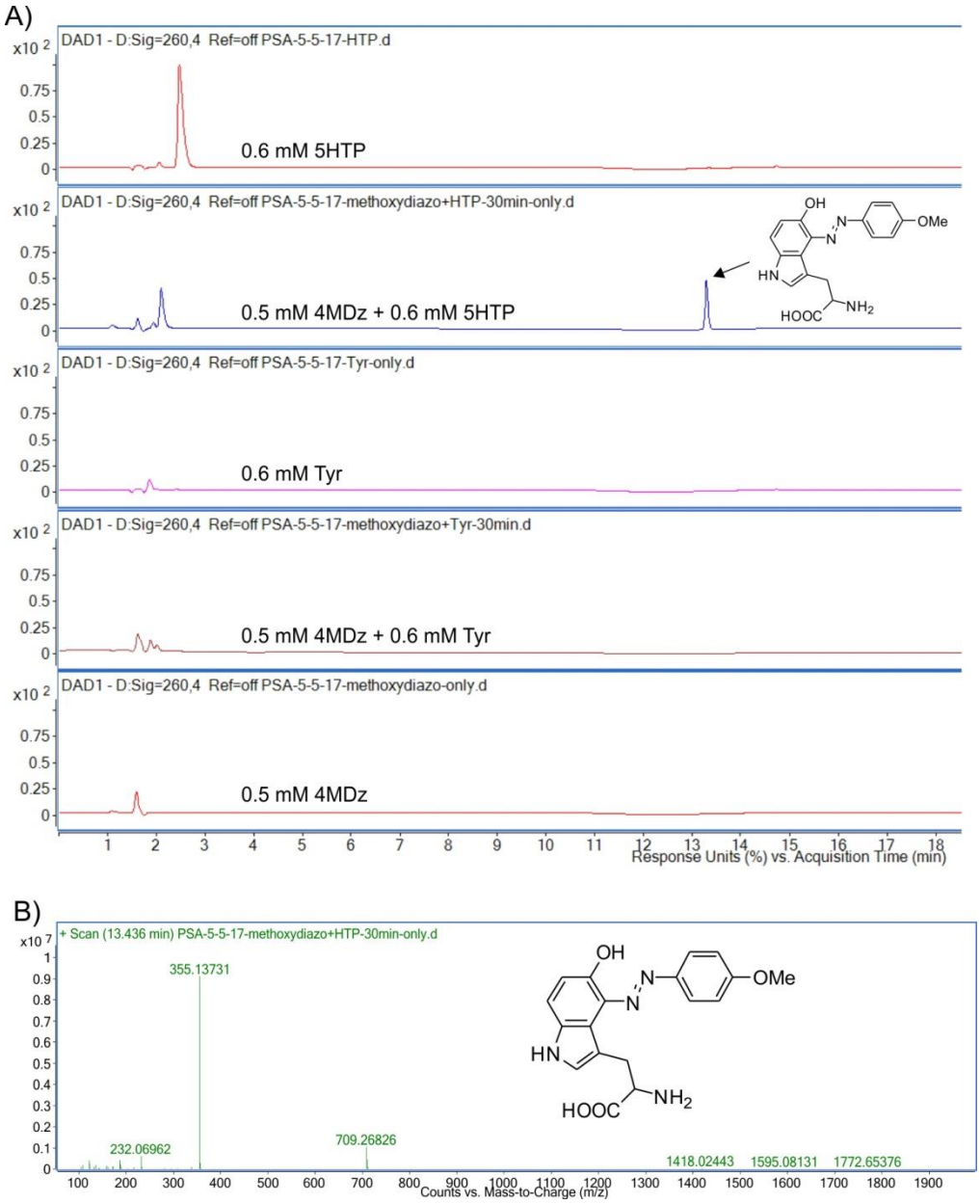


Figure S8: A) HPLC traces collected at 260 nm using a DAD detector for reaction of 4MDz with 5HTP and Tyr. Indicated reactions in 100 mM phosphate buffer (pH 7) were allowed to continue at room temperature for 30 minutes, and then analyzed by HPLC-MS. While formation of the azo-coupling product of 5HTP-4MDz was observed, no reaction was observed with Tyr. B) MS scans of the azo-coupling product peak, collected using an in-line ESI-TOF, is shown.

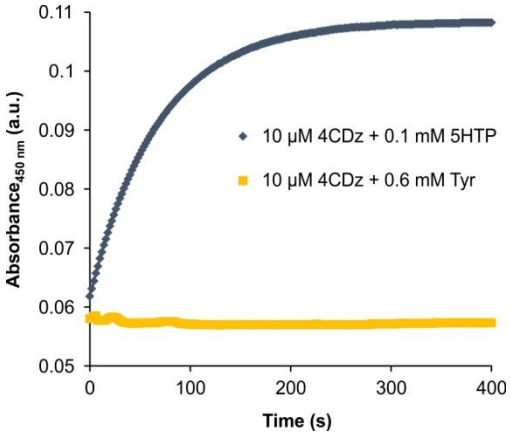


Figure S9: 4CDz reacts selectively with 5HTP. When 10 µM 4CDz was reacted incubated with 0.1 mM of 5HTP in 100 mM phosphate buffer (pH 7), the reaction reached near completion in minutes, while the corresponding reaction with 0.6 mM Tyr showed no progress.

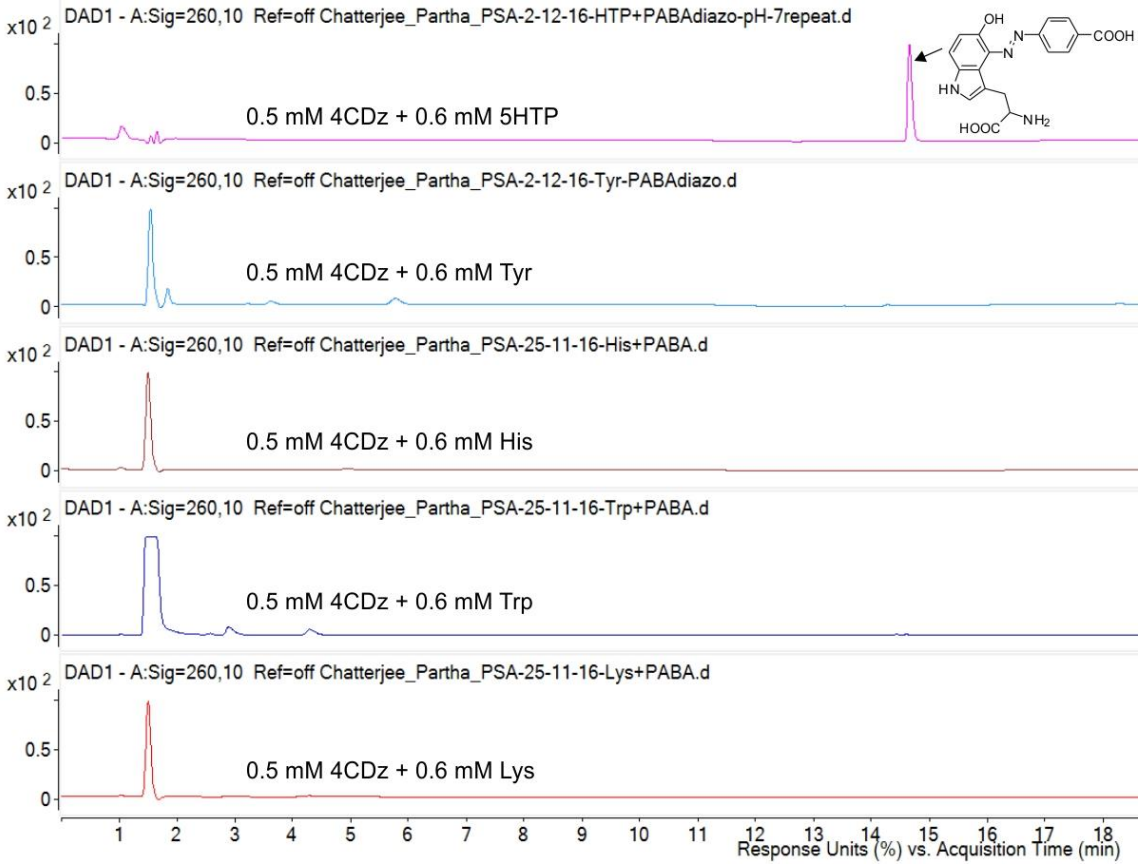


Figure S10: HPLC-MS (ESI-TOF) analysis of the azo-coupling reactions of 4CDz with 5HTP and various canonical amino acids. Indicated reactions in 100 mM phosphate buffer (pH 7) were allowed to continue at room temperature for 30 minutes, and then analyzed by HPLC-MS. Only 5HTP was found to react with 4CDz under these conditions.

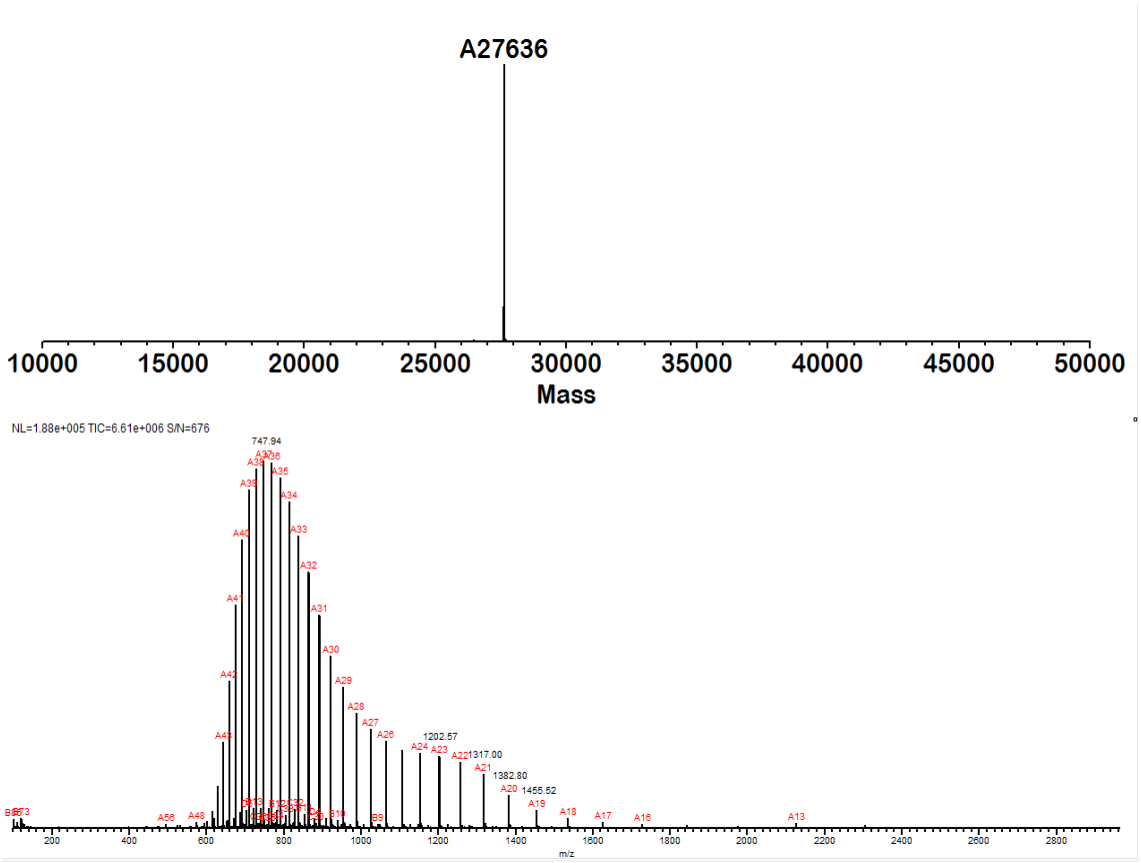


Figure S11: ESI-MS analysis of sfGFP-151-5HTP expressed in ATM(Trp) *E. coli*.[1] Bottom and top panel show raw and deconvoluted spectra, respectively. The observed mass matches expected mass (27636 Da).

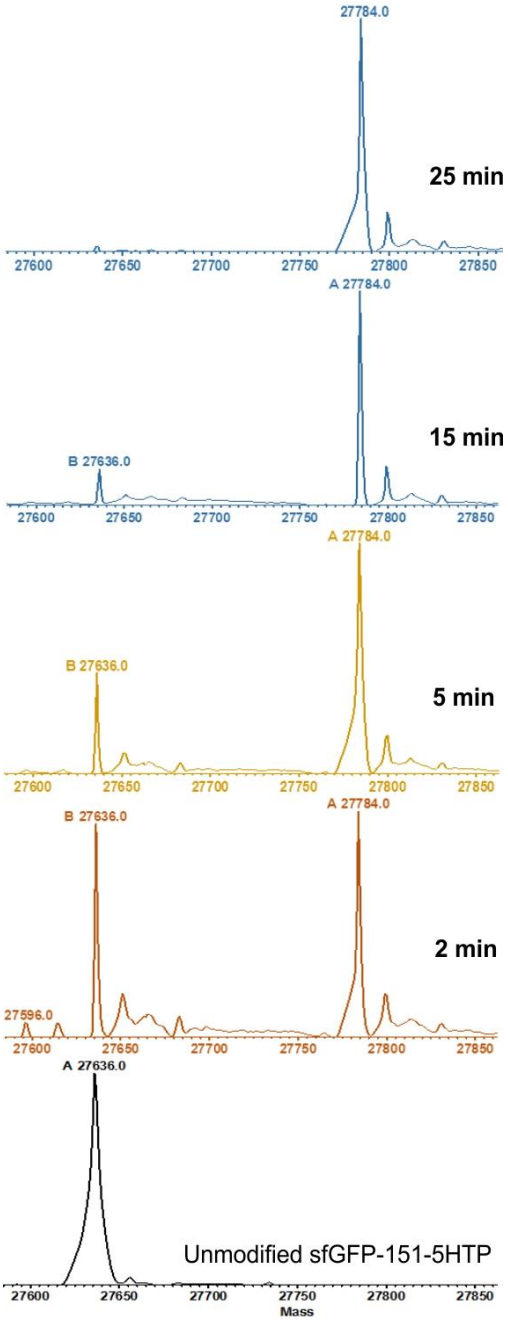


Figure S12: Rapid labeling of sfGFP-151-5HTP with 4CDz. A reaction between 10 μM sfGFP-151-5HTP and 40 μM 4CDz (at room temperature; in 100 mM phosphate buffer, pH 7) was quenched at different time points by adding excess 5HTP and the degree of protein labeling (expected mass 27784 Da) was monitored by ESI-MS over time.

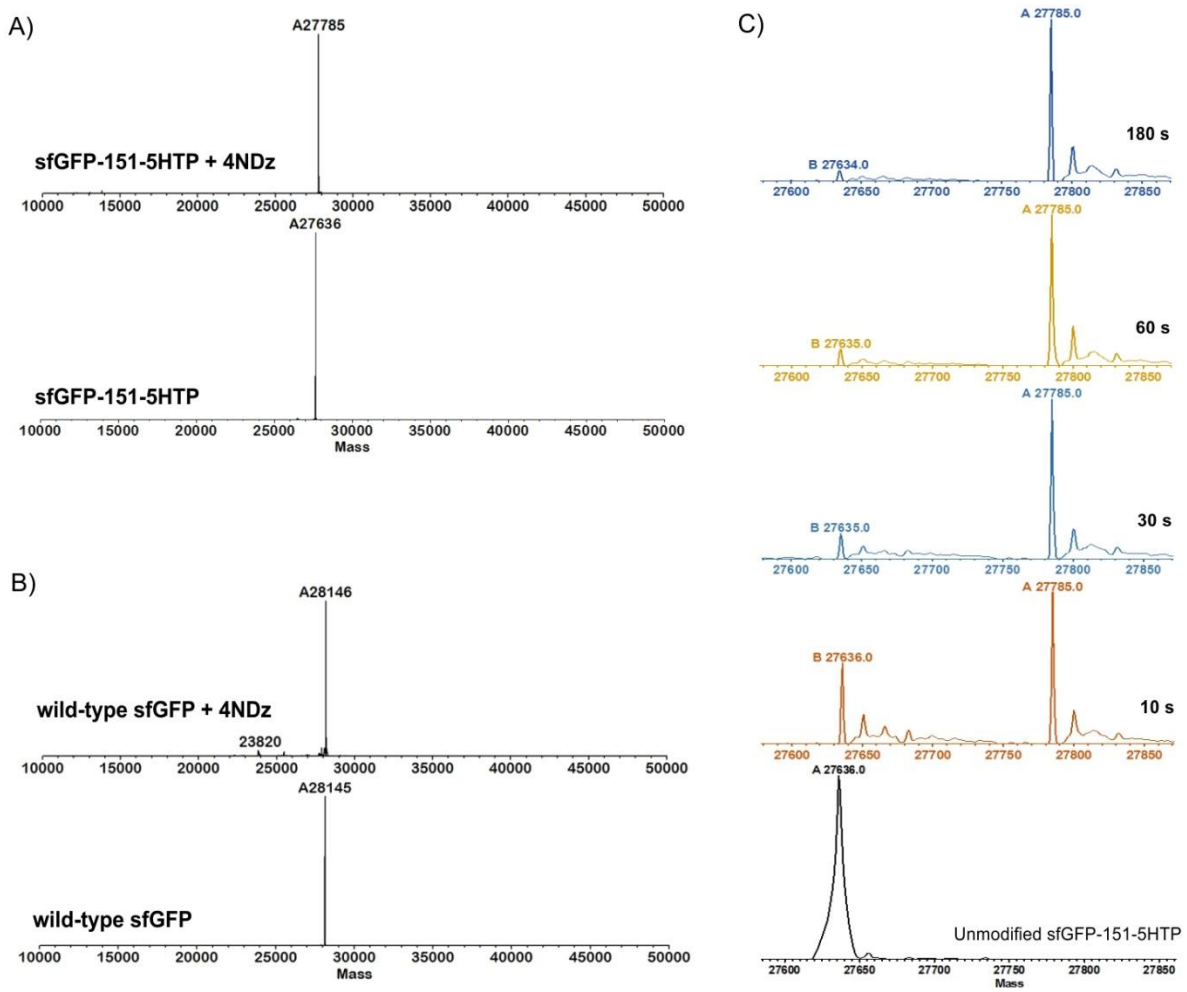


Figure S13: A) When sfGFP-151-5HTP (10 μM) was incubated with 40 μM 4NDz in 100 mM phosphate buffer (pH 7) at room temperature, complete labeling (expected mass 27785 Da) was observed in minutes. B) Identical treatment of wild-type sfGFP resulted in no protein modification. C) The reaction between sfGFP-151-5HTP and 5HTP was quenched with excess 5HTP at different points of time and degree of protein modification was monitored by ESI-MS.

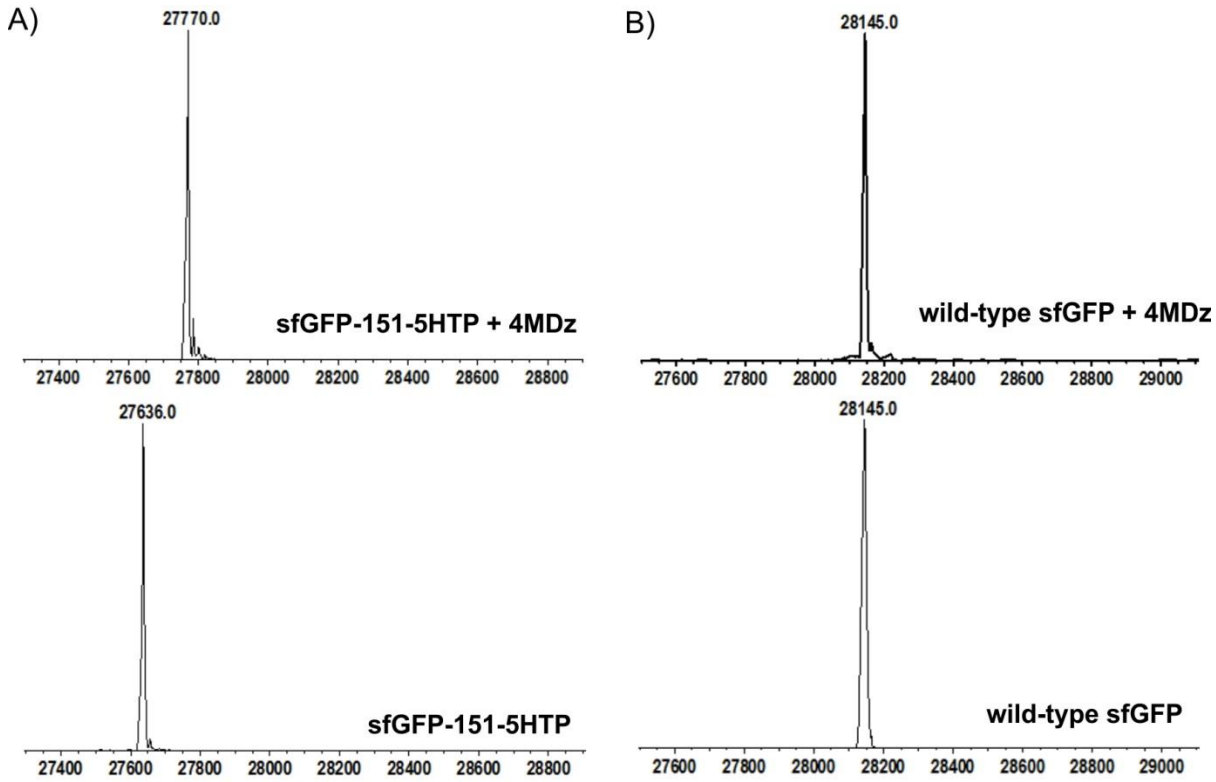


Figure S14: A) When sfGFP-151-5HTP (10 μM) was incubated with 250 μM 4MDz in 100 mM phosphate buffer (pH 7) at room temperature, complete labeling (expected mass 27770 Da) was observed within 45 minutes. B) Wild-type sfGFP was not labeled under these conditions.

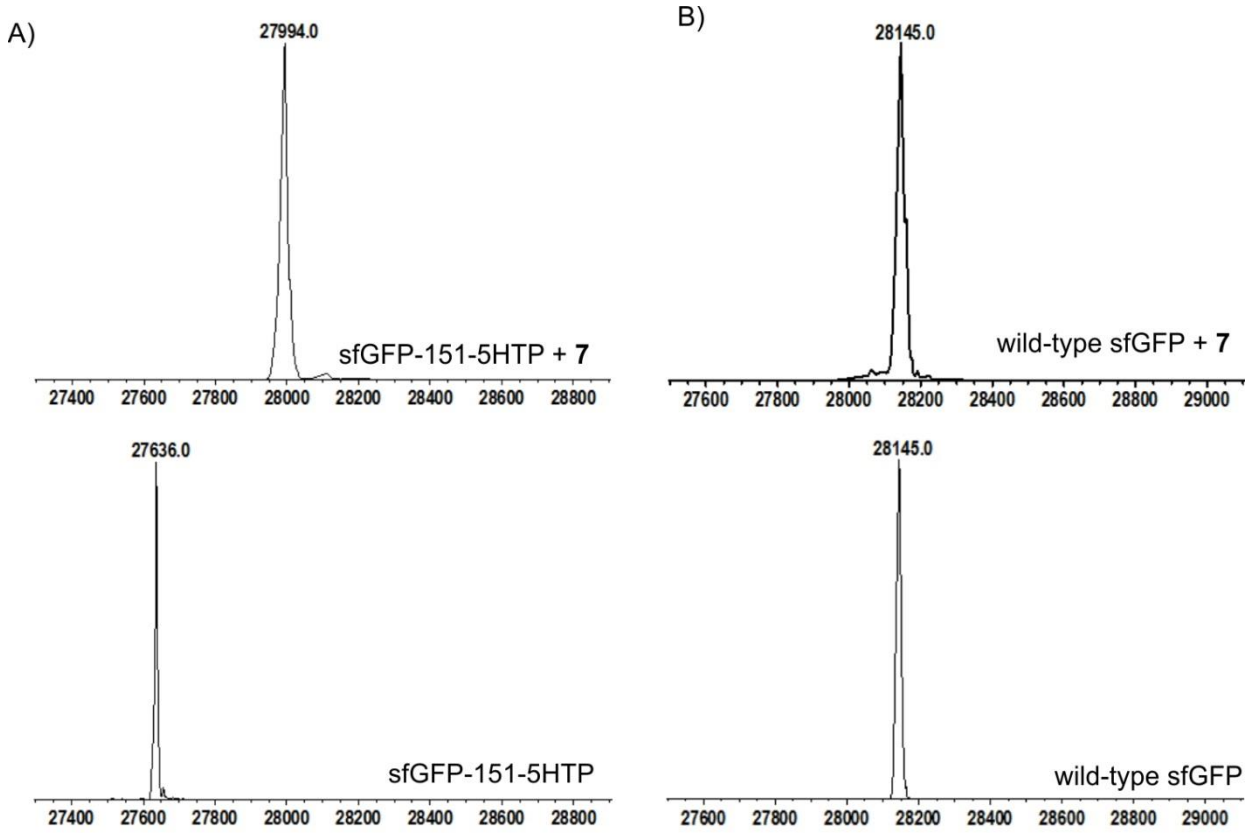


Figure S15: A) Treatment of sfGFP-151-5HTP (10 μM) with 150 μM **7** (Figure 3C) in 100 mM phosphate buffer (pH 7) on ice, resulted in complete labeling (expected mass 27994 Da). B) Identical treatment of wild-type sfGFP resulted in no protein modification.

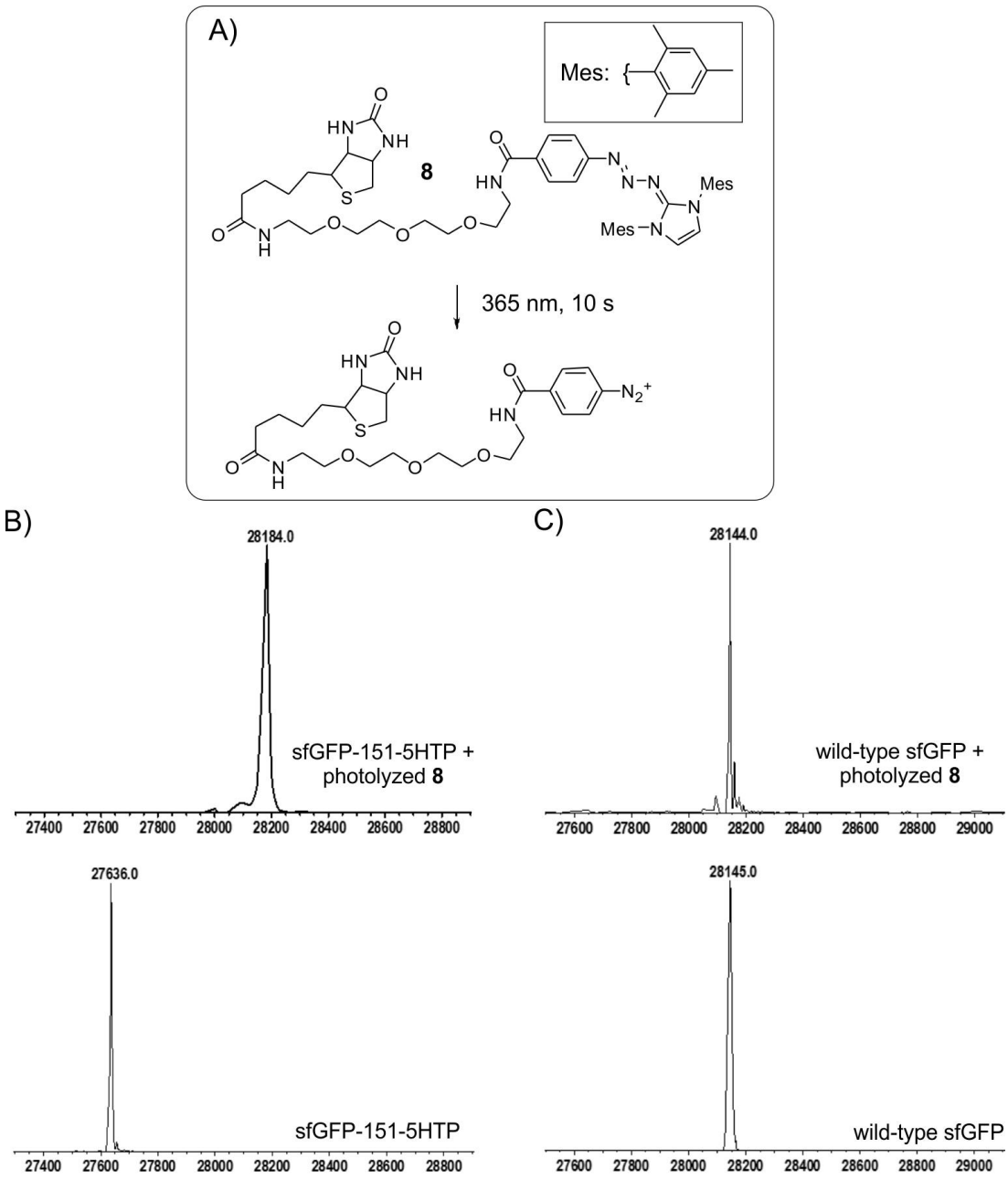


Figure S16: A) The biotin-triazabutadiene conjugate can be rapidly photolyzed to release the biotin-4CDz conjugate, as shown previously. B) Treatment of sfGFP-151-5HTP (10 μM) with 150 μM photolyzed **8** in 100 mM phosphate buffer (pH 7) on ice, results in complete labeling (expected mass 28184 Da). C) Identical treatment of wild-type sfGFP resulted in no protein modification.

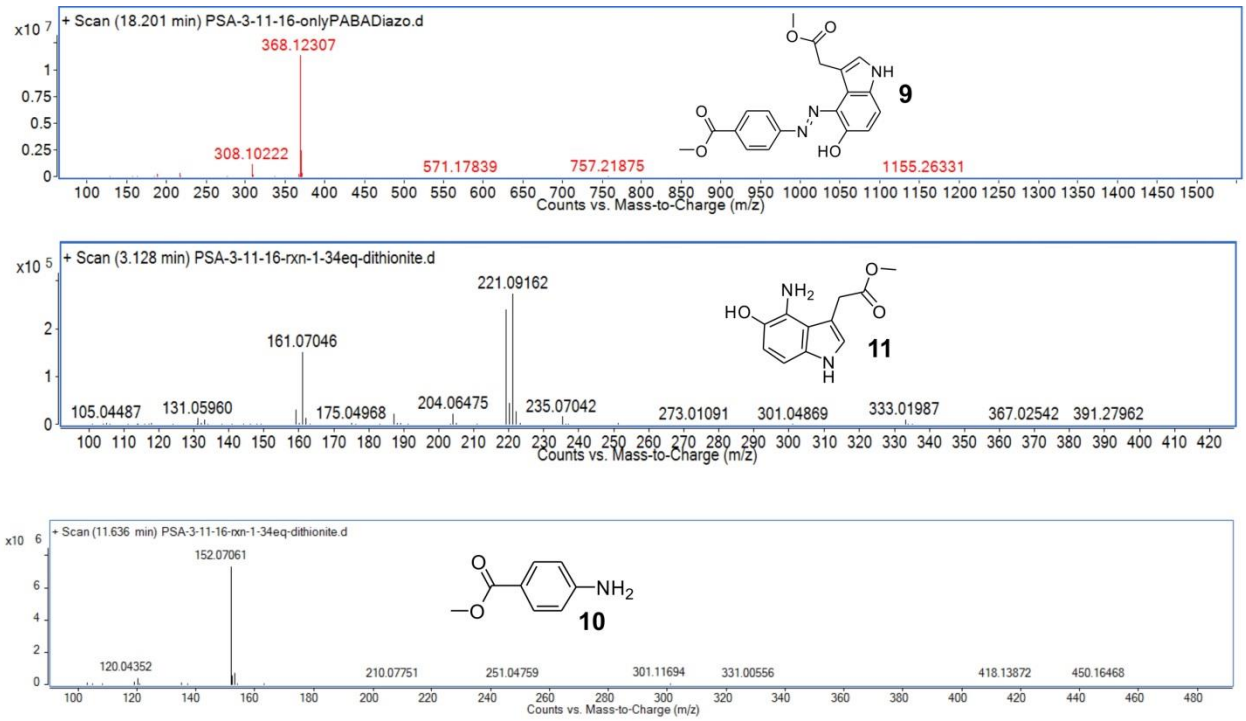


Figure S17: MS scans of the peaks found in the HPLC analysis described in Figure 5A confirms the dithionite mediated cleavage of **9** to **10** and **11**.

Materials and Methods:

All chemicals and solvents were purchased from Sigma-Aldrich or Fischer Scientific and used without further purification. In particular, 4-nitrobenzenediazonium (4NDz) and 4-methoxybenzenediazonium (4MDz) salts were purchased from Fisher chemicals. 4-carboxydiazonium tetrafluoroborate (4CDz) was synthesized starting from 4-aminobenzoic acid under standard diazotization conditions (described later). Oligonucleotides for PCR were purchased from Integrated DNA technologies. PCR reactions were performed using Phusion DNA polymerase (Thermo Fisher) according to manufacturer's protocol. The DH10b strain was used for plasmid propagation and cloning purposes. Standard LB medium was used to grow bacteria either in liquid culture or Agar plates, supplemented with necessary antibiotics (95 μg/mL spectinomycin, 100 μg/mL ampicillin, 30 μg/mL kanamycin, 15 μg/mL zeocin, 10 μg/mL gentamycin). Sanger sequencing services were provided by Eton Bioscience. The ATM(Trp) *E. coli* strain was used to express protein incorporating 5HTP via TGA-suppression as described previously.[1] All HPLC-coupled mass-spectrometry analyses were performed using an Agilent Technologies 1260 Infinity ESI-TOF instrument.

Analysis azo-coupling reaction of free amino acids using HPLC-MS:

Using a freshly prepared stock solution of the aryl-diazonium compound, reactions were set up in 100 mM phosphate buffer (pH 7) containing 0.5 mM of the aryl-diazonium and 0.6 mM of the amino acid. Upon the desired length of incubation at room temperature, the reactions were analyzed by HPLC-MS. A C18-reversed phase column (Phenomenex, WIDEPORE 3.6 u XB-C18, 150 x 2.1 mm) was used and the following gradient was used for separation (flow rate 0.2 mL/min). Mobile phase A is 0.1 % formic acid in HPLC grade water and B is 0.1 % formic acid in HPLC grade acetonitrile. [5% B until 0-5 min; 5%-95 % linear gradient of B from 5-23 min; 95%-5 % linear gradient of B from 23-28 min].

To facilitate the purification and characterization of the resulting azo-compounds, we used the non-polar substrates described in Figure 2. These reactions were performed in a 1:1 mixture of methanol and 100 mM aqueous phosphate buffer, followed by an ethyl acetate/water extraction to isolate the product in the organic layer (Described in detail under the synthesis section).

Determining the kinetic parameters of various azo-coupling reactions:

As shown in Figure S1, azo conjugates uniquely absorb at 450 nm, which was used to spectrophotometrically monitor their formation. Except for the reaction between 5HTP and 4NDz, the rates for all other reactions were measured under pseudo-first order reaction conditions, where the aryl-diazonium was used as the limiting reagent (10 μM) and varying concentrations of the amino acid was used in significant excess (at least 10-fold). Reactions were performed in 100 mM phosphate buffer (pH 7) at room temperature. Pseudo-first order rate (k_B)constant for each different amino acid concentration used according to the following equation:

$\ln[A] = \ln[A_0] - k_B t$

Plotting each pseudo-first order rate constant against the corresponding amino acid concentration yielded the second-order rate constant (k_2) according to the following equation:

$k_B = k_2 * [B]$

Due to the extremely rapid reaction between 4NDz and 5HTP, it was not possible to measure its rate similarly under pseudo-first order conditions (we do not have access to a stopped-flow equipment). The rate of this reaction could only be measured under very low concentrations (<10 μM) of both reagents. To measure the kinetic parameters of this bimolecular reaction, we used equal concentrations of both reactants (5 μM each in 100 mM phosphate buffer at pH 7), where the second order rate equation reduces to:

$1/[X]_t = 1/[X]_0 + k_2 t$ ($[X]_t$: concentration of either reactant at time t; $[X]_0$: initial concentration of either reactant; k_2: second-order rate constant).

To obtain the concentration of the product (4NDz-5HTP) at a given time from the observed absorbance at 450 nm, a standard curve was first generated (Figure S2A). This was used to obtain $[X]_t$ from the observed increase in absorbance at 450 nm (Figure S2B). Plotting 1/[X]t against time yielded the rate constant.

For each reaction three independent experiments were performed, and the resulting rates were averaged (associated standard deviation was reported as error).

Expression and purification of sfGFP-151-5HTP:

sfGFP-151-5HTP was expressed and purified as described previously.[1] Briefly, the ATM(Trp) E. coli strain harboring pET22b-sfGFP-151-TGA and pEvol-tac-EcTrpRS-h14 plasmids was grown in shake-flasks at 37 °C in L.B. medium to an $O.D._{600}$ of 0.6, at which point sfGFP and EcTrpRS-h14 expression was induced by adding 1 mM each of IPTG and 5HTP. After 12 hours of protein expression at 30 °C, cells were harvested and subjected to protein purification by Ni-NTA affinity chromatography. All proteins were analyzed by SDS-PAGE followed by coomassie staining, as well as by ESI mass spectrometry.

Labeling 5HTP-residues on proteins with 4NDz, 4CDz, and 4MDz:

For 4NDz, and 4CDz, 10 μM protein in 100 mM phosphate buffer was incubated with 40 μM of aryl-diazonium compound and incubated at room temperature for 2-30 min. For 4NDz and 4CDz, complete labeling was observed in 2 min and 30 min, respectively. 4MDz was used at 250 μM for 45 min to achieve complete protein labeling. Next, free 5HTP (final concentration 200 μM) was added to quench the unreacted diazonium. The protein was desalted by repeated dilution with 100 mM phosphate buffer, followed by concentration using a concentrator (10 kDa

molecular weight cutoff, Amicon Ultra-0.5 mL, centrifugal filters). Concentration of the desalted protein was adjusted back to 10 µM, followed by ESI-MS and SDS-PAGE analysis.

Fluorescein-labeling of proteins using CRACR:

Preparation of fluorescein-diazonium (7): 6-aminofluorescein was dissolved in 10 mM HCl at a concentration of 10 mM. Separately a 60 mM sodium nitrite solution was prepared and both solutions were cooled on ice. 20 µL $NaNO_2$ (60 mM) was added to 100 µL of fluorescein (10 mM) and vortexed for 10 sec then kept on ice to generate a 8.3 mM fluorescein diazonium solution. This solution was further diluted to 1 mM with water and used for the labeling studies.

Fluorescein-labeling protocol: 10 µL of 10 µM protein in 100 mM phosphate buffer was incubated on ice with 150 µM of **7** for 10 min. Next, 2 µl of 6 mM 5HTP was added to quench the unreacted diazonium. The protein was desalted by repeated dilution with 100 mM phosphate buffer, followed by concentration using a concentrator (10 kDa molecular weight cutoff, Amicon Ultra-0.5 mL, centrifugal filters). Concentration of the desalted protein was adjusted back to 10 µM, followed by ESI-MS and SDS-PAGE analysis.

Biotinylation of proteins with photolyzed 8:

Compound **8** was synthesized as described in the synthesis section.

8 was photolyzed for 10 s using a 120 W LED-array that emits at 365 nm (Larson Electronics), in which time the conversion of the triazabutadiene to diazonium was confirmed to be complete. 10 µM protein in 100 mM phosphate buffer (pH 7) was incubated with 150 µM of freshly photolyzed **8** for 15 min on ice, followed by the addition of 5HTP (1mM) to quench the unreacted diazonium. The protein was desalted by repeated dilution with 100 mM phosphate buffer, followed by concentration using a concentrator (10 kDa molecular weight cutoff, Amicon Ultra-0.5 mL, centrifugal filters). Concentration of the desalted protein was adjusted back to 10 µM, followed by ESI-MS and SDS-PAGE analysis.

Dithionite-cleavage of the azo-conjugate between sfGFP-151-5HTP and photolyzed 8:

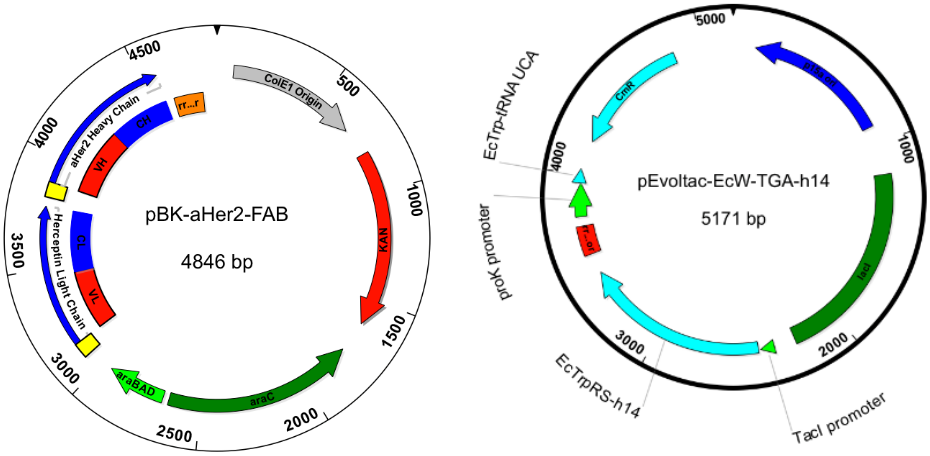
The biotinylated protein, prepared as described above was desalted into 0.5 M phosphate buffer, and its concentration was adjusted to 10 µM. 2 µL of freshly prepared dithionite (1.2 M stock solution; the final concentration 200 mM) was added to 10 µL of 10 µM biotinylated protein and incubated for 30 min at room temperature (22 ˚C), followed by SDS-PAGE and western blot analysis.

Anti-Biotin western Blot:

12% SDS-PAGE gels were used to resolve proteins, which were transferred from the gel to a PVDF membrane (Life Technologies) using a Trans-Blot Turbo Transfer System 15 (Bio-Rad). After transferring, the membrane was blocked with blocking solution [BSA in PBS (Thermo

Scientific)] overnight at 4 ˚C with gentle agitation. On the next day, blocking solution was removed, and membrane was incubated with Streptavidin-HRP (Pierce; at 1:2500 dilution) for 2 hours in fresh blocking solution with gentle shaking at room temperature. The membrane was washed 7 times (10-minute incubation with agitation) with wash solution (0.1% Tween 20 in TBS). The membrane was developed using SuperSignal West Dura Kit (Fisher Scientific) and incubated for 2 minutes before signal detection by the ChemiDoc MP imaging system (BioRad).

<u>Expression, purification and CRACR labeling of anti-Her2 Fab</u>:



Map of pBK-aHer2-FAB, the plasmid that was used for expression of Herceptin FAB. The previously described pEvoltac-EcW-TGA-h14 contains the EcTrpRS-h14 and $tRNA_{UCA}^{EcTrp}$. These plasmids were co-transformed into ATMW1 strain.

Antibody Purification

Herceptin-FAB sequence was obtained from literature[2] and cloned into a pBK vector behind the STII secretion signal. A TGA stop codon was introduced into the Fab at site K169 in the light chain using site-directed mutagenesis. pBK-aHer2-K169TGA was co-transformed with pEvoltac-EcW-5HTP, which encode EcTrpRS-h14/$tRNA_{UCA}^{EcTrp}$ pair that can specifically charge 5HTP in response to TGA in the recently developedATMW1 *E. coli* strain[2]. The cells were grown in LB supplemented with 35 µg/mL chloramphenicol, 10 µg/mL kanamycin, 100 µg/mL spectinomycin. Upon reaching an OD_{600} of 0.6, protein expression was induced with 0.02% arabinose and 1 mM IPTG and 1 mM 5HTP (Chem Impex Int'l. Inc) was added to the media. Cells were allowed to grow for 16 hours at 30 ˚C in shake flasks, harvested by centrifugation, and resuspended in a periplasmic lysis buffer (20% sucrose, 30 mM Tris, pH 8, 1 mM EDTA, 0.2 mg/mL lysozyme, and Halt protease inhibitor) for 30 minutes at 37 ˚C. Lysate was then diluted 1:1 with binding buffer (50 mM NaOAc, pH 5.2) and clarified by centrifugation at 1700 rpm for 30 minutes. Herceptin-FAB was then purified using Pierce Protein G agarose according to the manufacturer's instructions. The yield for wild-type Fab was 0.35 mg/L, while the same for the 5HTP mutant was 0.17 mg/L.

Antibody labelling:

4 µM Fab in 0.1 M glycine buffer (pH 2.8) was neutralized using 1 M Tris-HCl (pH 8) buffer, and incubated with 80 µM of freshly prepared **7** on ice for 15 min; then, the excess **7** was quenched by adding 5HTP to a final concentration of 0.5 mM. The labeled antibody was dialyzed overnight against PBS buffer at 4 ˚C, and used for SDS-PAGE, ESI-MS and FACS analysis.

Cell Culture:

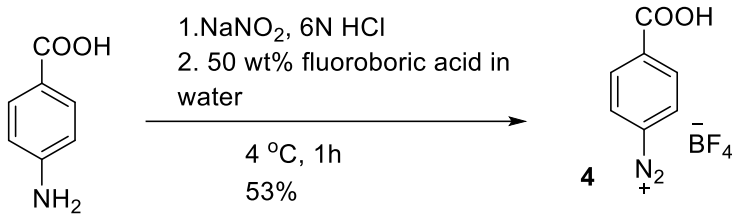
Adherent SK-BR-3 cell lines were purchased from ATCC and grown in DMEM-high glucose (HyClone) supplemented with 100 U/mL penicillin/streptomycin (HyClone) and 10% fetal bovine serum (FBS; Corning). They were maintained at 37 ˚C (humidified) and 5% CO_2.

Association of fluorescein-labeled anti-Her2 Fab to SK-BR-3 cells:

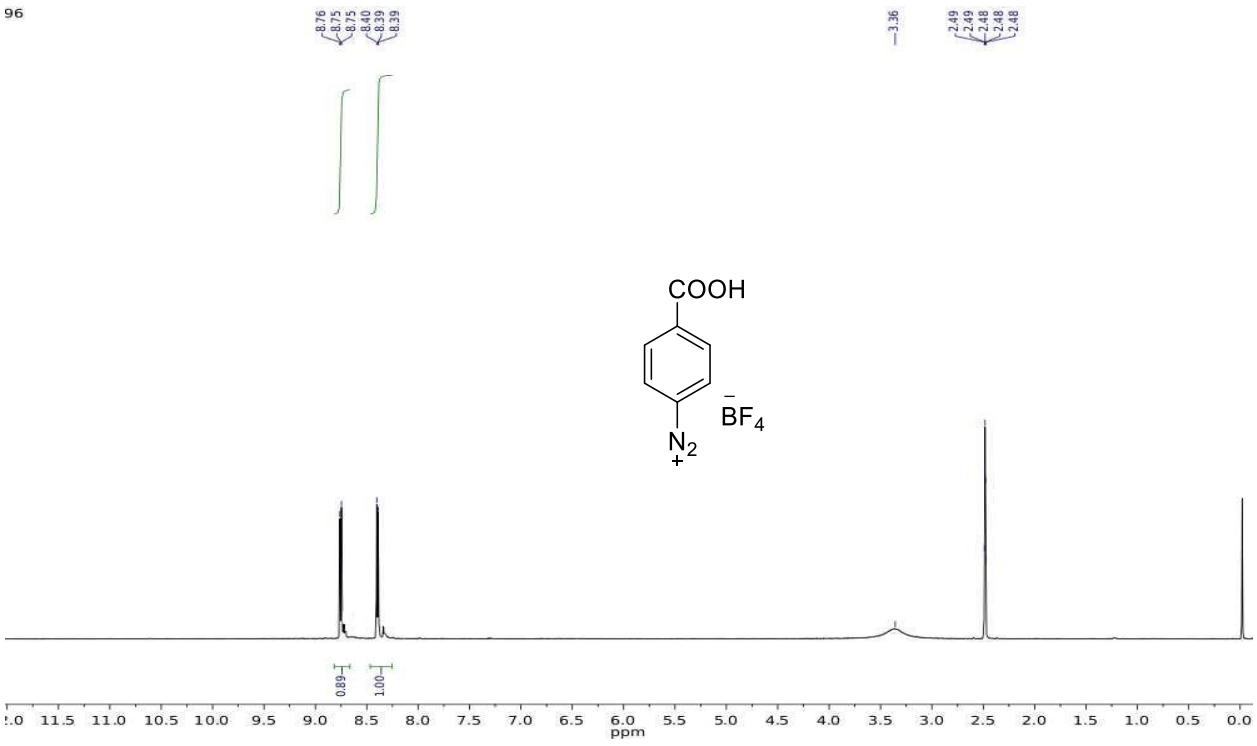
Upon reaching confluency, SK-BR-3 cells were detached from dish using 0.25% trypsin (HyClone) for 2 minutes at 37 ˚C. Trypsin was quenched with ice-cold DMEM supplemented with 10 % FBS, and the cells were washed once with ice-cold DMEM+10% FBS, centrifuged at 2,000 xg for 5 minutes, and re-suspended in ice-cold PBS. Fluorescein-labeled Herceptin-Fab (as prepared above) was added to SK-BR-3 cells at 4 ˚C with gentle agitation for one hour. Cells were washed once using PBS, re-suspended in cold PBS, and then analyzed by flow cytometry using a Bio-Rad S3e cell sorter. All flow cytometry plots were generated in ProSort (Bio-Rad).

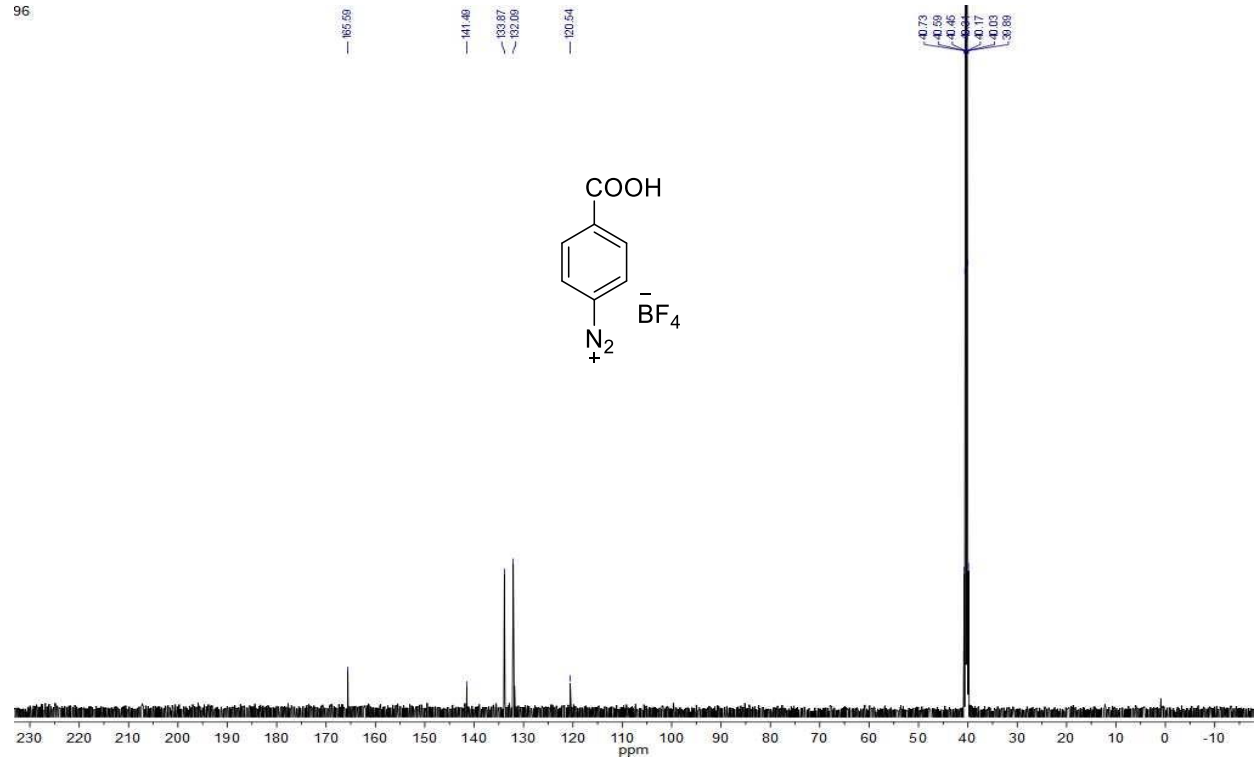
Synthesis of various compounds:

Synthesis of 4CDZ, 4:

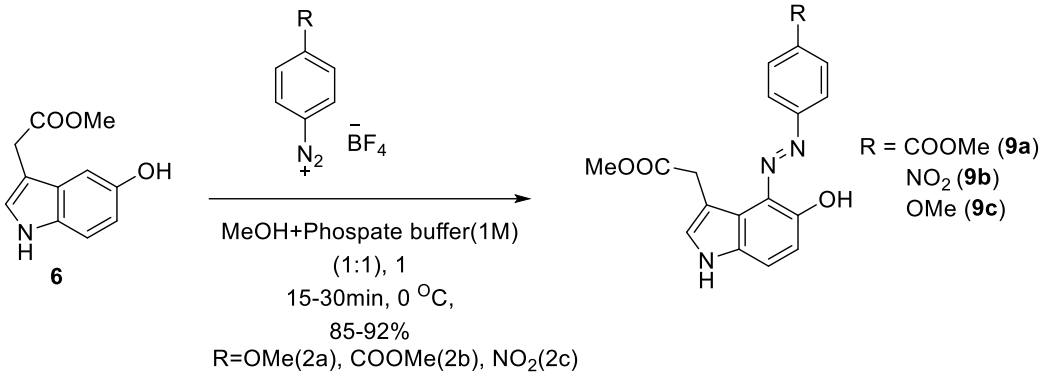


A solution of 4-aminobenzoic acid (1 gm, 7.3 mmol) in 10 mL 6 N HCl was kept on ice for 15 min. To this, a solution of sodium nitrite (590 mg, 8.75 mmol) was added and stirred for 30 min on ice. After 30 min, 10 mL 50 wt% fluoroboric acid in water was added to it and stirred for another 30 min on ice. A yellowish precipitate appeared which was filtered and dried on vacuum. 0.92 g pure product was obtained. Yield 53%; ^{1}H NMR (600 MHz, DMSO-d_6) δ 8.82 – 8.66 (m, 2H), 8.40 (d, J = 8.9 Hz, 2H);[3] ^{13}C NMR (151 MHz, DMSO-d_6) δ 165.59, 141.49, 133.87, 132.09, 120.54; ^{19}F NMR (565 MHz, DMSO-d_6) -148.33; ^{11}B NMR (192 MHz, DMSO-d_6): -1.33; HRMS (ESI-TOF-MS) (m/z): $[M]^+$ Calculated for $C_7H_5N_2O_2^+$ 149.0346 found 149.0340.


96
8.76
8.75
8.75
8.40
8.39
8.39
3.36
2.49
2.49
2.48
2.48
2.48
COOH
BF4
N2
0.89
1.00
ppm


96
165.59
141.49
133.87
132.09
120.54
40.73
40.59
40.45
40.31
40.17
40.03
39.89
COOH
BF4
N2
ppm

Synthesis of non-polar azo-compounds from 6:

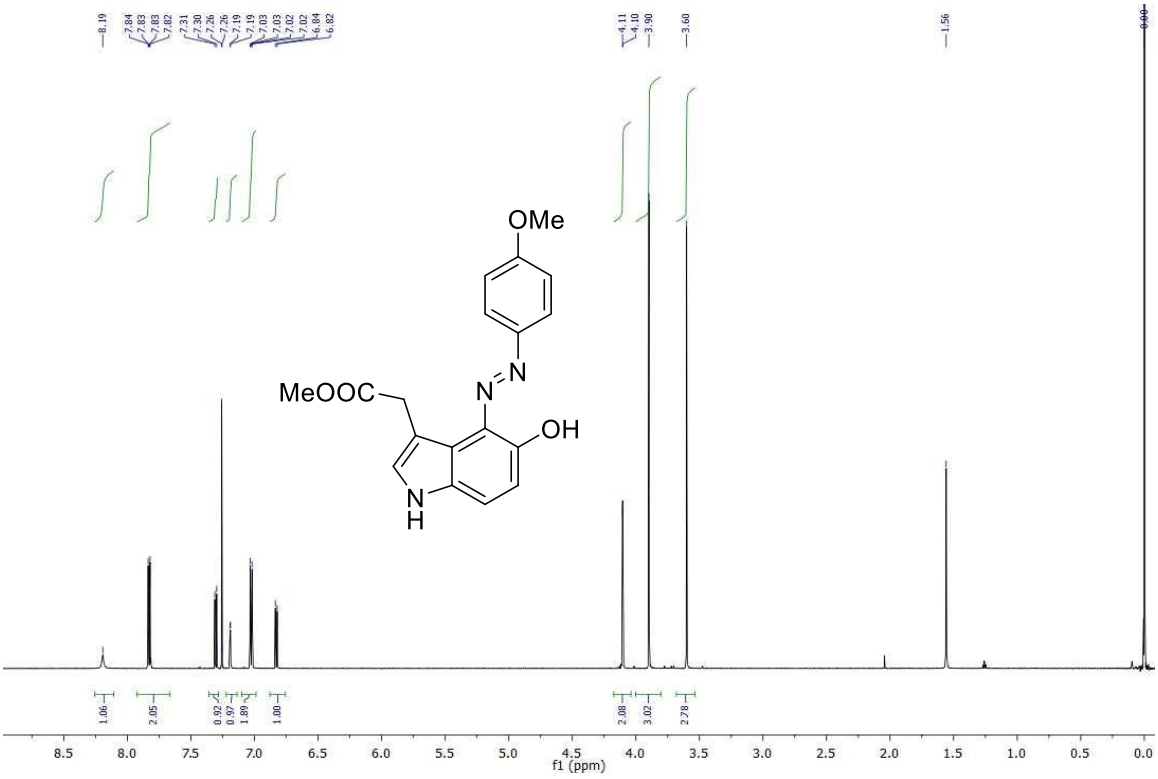


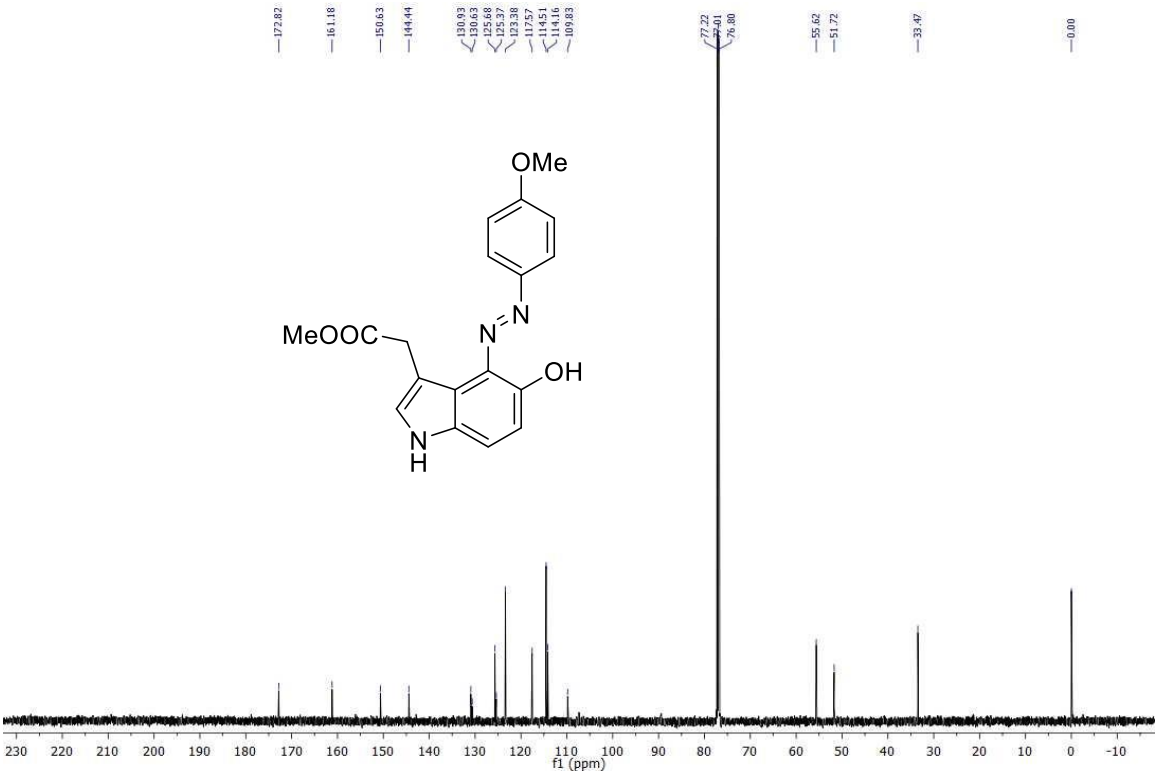
Methyl 2-(5-hydroxy-1H-indol-3-yl)acetate, **6** (0.1 g, 0.5 mmol) was dissolved in 30 mL methanol : 1 M phosphate buffer (1:1). The solution was kept on ice for 20 min and after that a solution of the aryl-diazonium (0.75 mmol) in water was added to it. Formation of the azo product could be monitored by the appearance of bright orange color. The reaction was continued for 30 min, and 30 mL ethyl acetate was added to it. The product was extracted in ethyl acetate, and the organic layer was washed alternatingly with brine and water three times. Finally, the solvent was evaporated and the compounds were characterized by NMR and mass spectrometry.

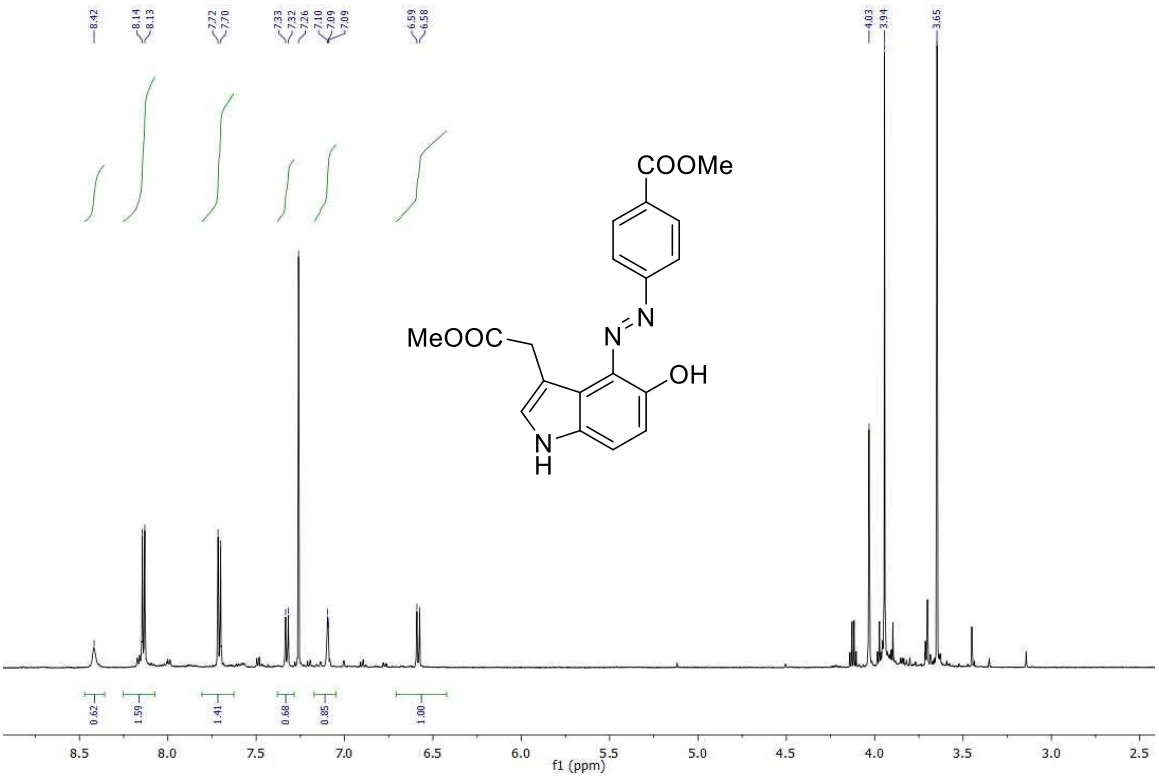
Methyl (E)-2-(5-hydroxy-4-((4-methoxyphenyl)diazenyl)-1H-indol-3-yl)acetate (9c): Yield 85%; R_f : 0.5 in 2:1 Hexane-Ethylacetate; ^{1}H NMR (600 MHz, Chloroform-*d*): δ 8.19 (s, 1H), 7.83 (d, *J* = 9.0 Hz, 2H), 7.31 (d, *J* = 8.8 Hz, 1H), 7.19 (d, *J* = 2.5 Hz, 1H), 7.03 (d, *J* = 9.0 Hz, 2H), 6.83 (d, *J* = 8.6 Hz, 1H), 4.10 (s, 2H), 3.90 (s, 3H), 3.60 (s, 3H); ^{13}C NMR (150 MHz, Chloroform-*d*) δ 172.82, 161.18, 150.63, 144.44, 130.93, 130.63, 125.68, 125.37, 123.38, 117.57, 114.51, 114.16, 109.83, 55.62, 51.72, 33.47. HRMS (ESI-TOF-MS) (m/z): $[M+H]^+$ Calculated for $C_{18}H_{18}N_3O_4$ 340.1292, found 340.1289.

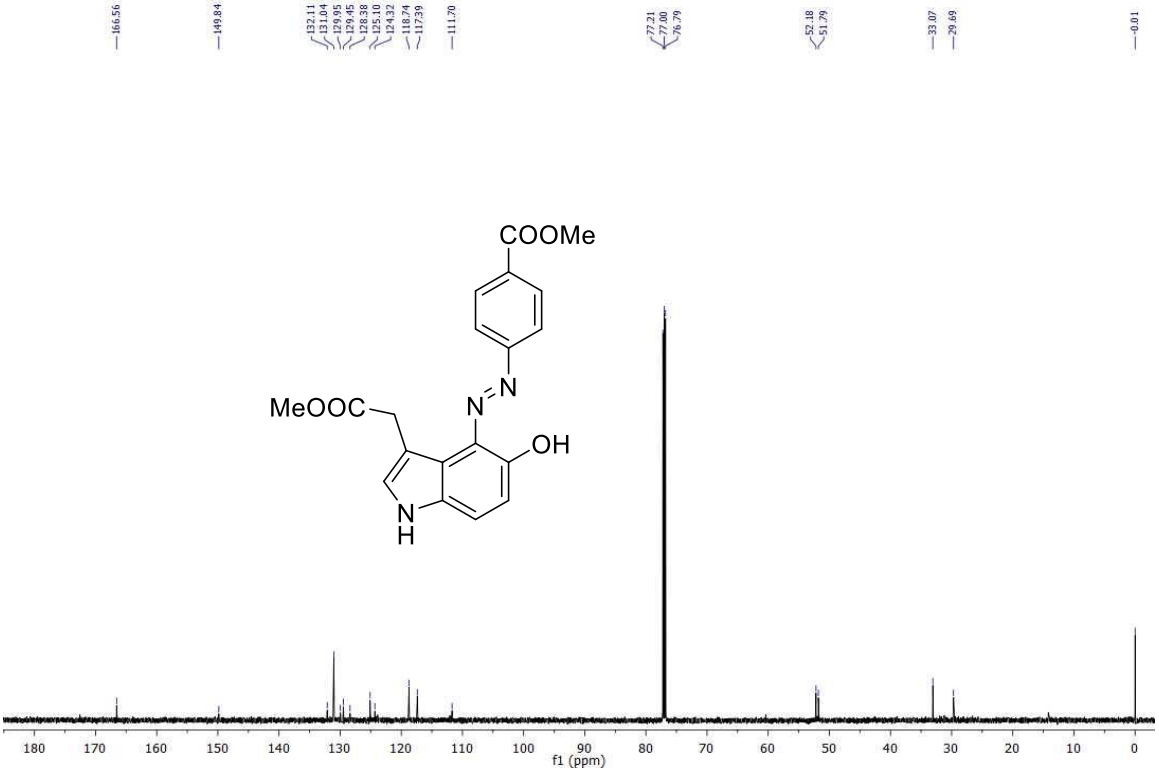
Methyl (E)-4-((5-hydroxy-3-(2-methoxy-2-oxoethyl)-1H-indol-4-yl)diazenyl)benzoate (9a): Yield 90%; R_f : 0.4 in 1:1 Hexane-Ethylacetate; ^{1}H NMR (600 MHz, Chloroform-*d*) δ 8.42 (s, 1H), 8.14 (d, *J* = 8.7 Hz, 2H), 7.71 (d, *J* = 8.6 Hz, 2H), 7.33 (d, *J* = 8.9 Hz, 1H), 7.09 (d, *J* = 2.9 Hz, 1H), 6.58 (d, *J* = 9.0 Hz, 1H), 4.03 (s, 2H), 3.94 (s, 3H), 3.65 (s, 3H); ^{13}C NMR (150 MHz, Chloroform-*d*) δ 166.56, 149.84, 132.11, 131.04, 129.95, 129.45, 128.38, 125.10, 124.32, 118.74, 117.39, 111.70, 52.18, 51.79, 33.07, 29.69; HRMS (ESI-TOF-MS) (m/z): $[M+H]^+$ Calculated for $C_{19}H_{18}N_3O_5$ 368.1241 found 368.1230.

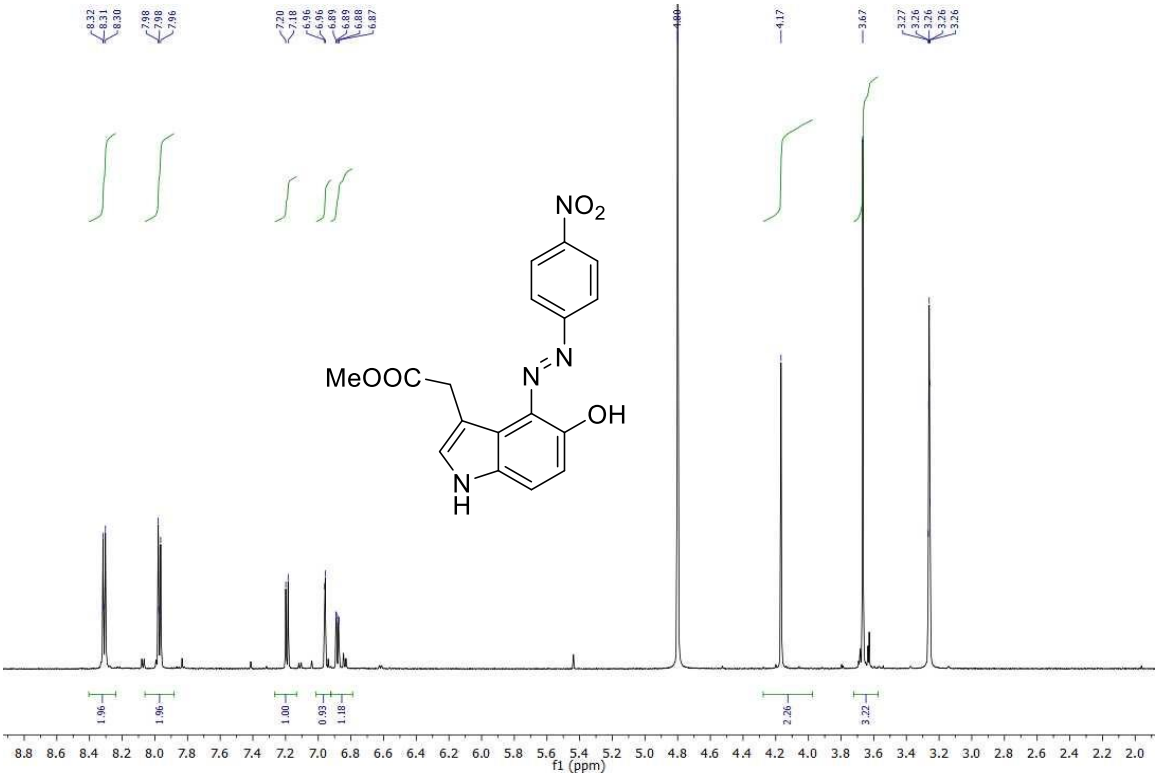
Methyl (E)-2-(5-hydroxy-4-((4-nitrophenyl)diazenyl)-1H-indol-3-yl)acetate (9b): Yield 95%; R_f : 0.3 in 1:1 Hexane-Ethylacetate; ^{1}H NMR (600 MHz, Methanol-d_4) δ 8.31 (d, *J* = 8.9 Hz, 2H), 7.97 (d, *J* = 8.9 Hz, 2H), 7.19 (d, *J* = 8.8 Hz, 1H), 6.96 (d, *J* = 2.3 Hz, 1H), 6.88 (dd, *J* = 8.8, 2.3 Hz, 1H), 4.17 (s, 2H), 3.67 (s, 3H); ^{13}C NMR (150 MHz, Methanol-d_4) δ 174.81, 159.34, 153.89, 150.47, 135.08, 131.03, 127.00, 125.13, 121.54, 115.54, 106.01, 53.91, 31.87; HRMS (ESI-TOF-MS) (m/z): $[M+H]^+$ Calculated for $C_{17}H_{15}N_4O_5$ 355.1037 found 355.1029.

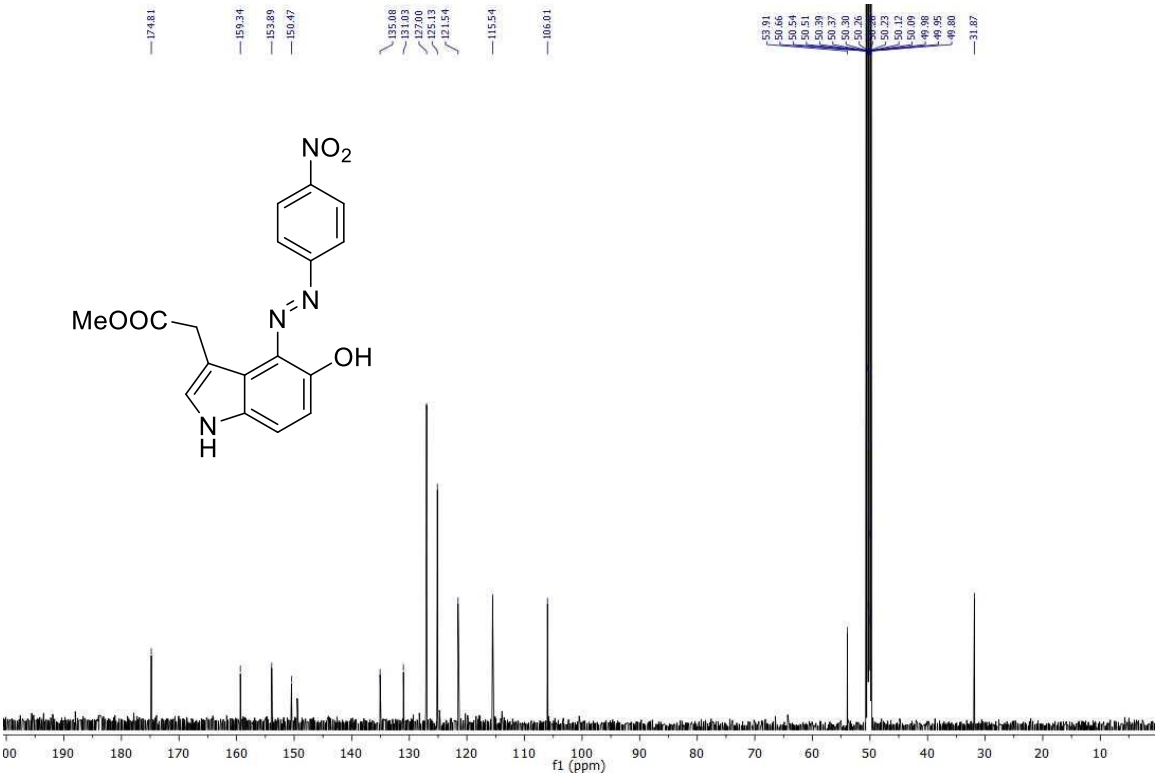

OMe
MeOOC
N
N
OH
N
H
f1 (ppm)


OMe
MeOOC
N
N
OH
N
H
f1 (ppm)

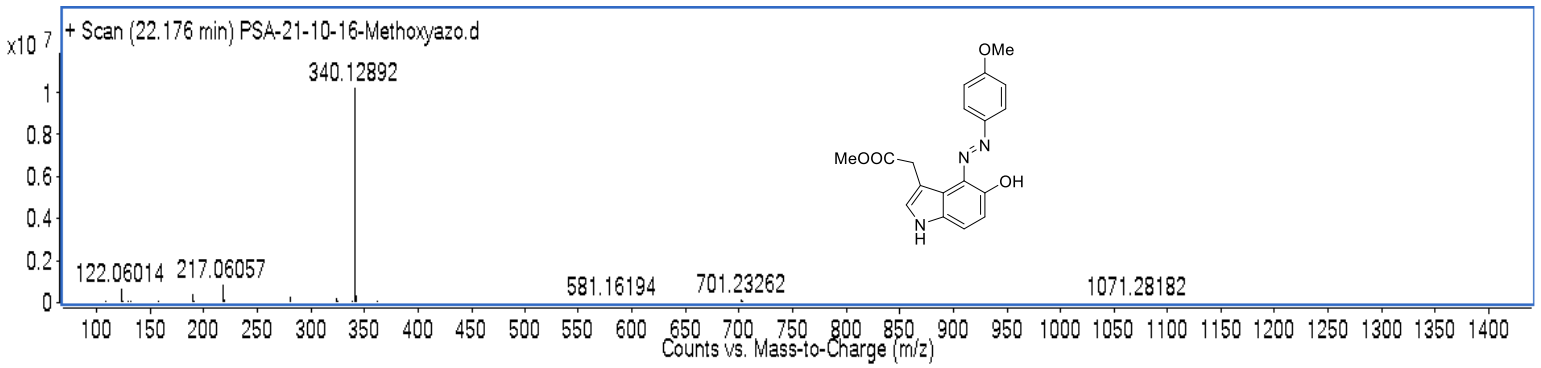

8.42
8.14
8.13
7.72
7.70
7.33
7.32
7.26
7.10
7.09
7.09
6.59
6.58
4.03
3.94
3.65
COOMe
MeOOC
N
N
OH
N
H
0.62
1.59
1.41
0.68
0.85
1.00
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
f1 (ppm)

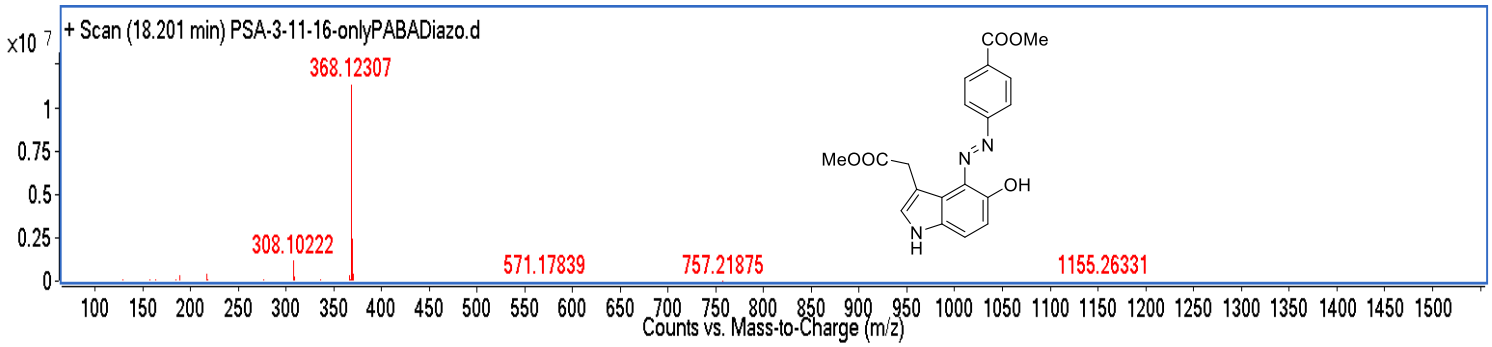

166.56
149.84
132.11
131.04
129.95
129.45
128.38
125.10
124.32
118.74
117.39
111.70
77.21
77.00
76.79
52.18
51.79
33.07
29.69
0.01
COOMe
MeOOC
N
N
OH
N
H
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

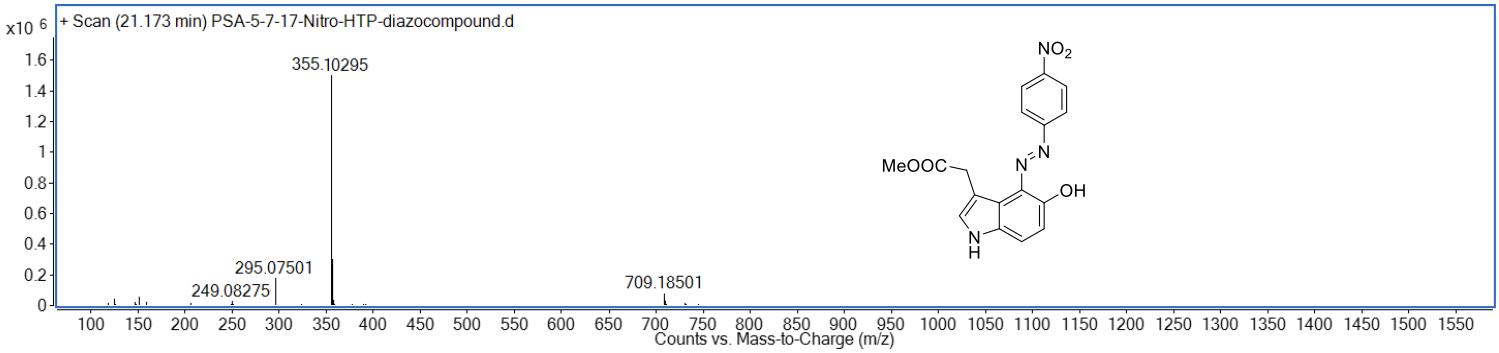

8.32
8.31
8.30
7.98
7.98
7.96
7.20
7.18
6.96
6.96
6.89
6.89
6.88
6.87
4.17
3.67
3.27
3.26
3.26
3.26
NO2
N
N
MeOOC
OH
N
H
1.96
1.96
1.00
0.93
1.18
2.26
3.22
f1 (ppm)


174.81
159.34
153.89
150.47
135.08
131.03
127.00
125.13
121.54
115.54
106.01
53.91
50.66
50.54
50.51
50.39
50.37
50.30
50.26
50.23
50.12
50.09
49.98
49.95
49.80
31.87
NO2
N
N
MeOOC
OH
N
H
f1 (ppm)

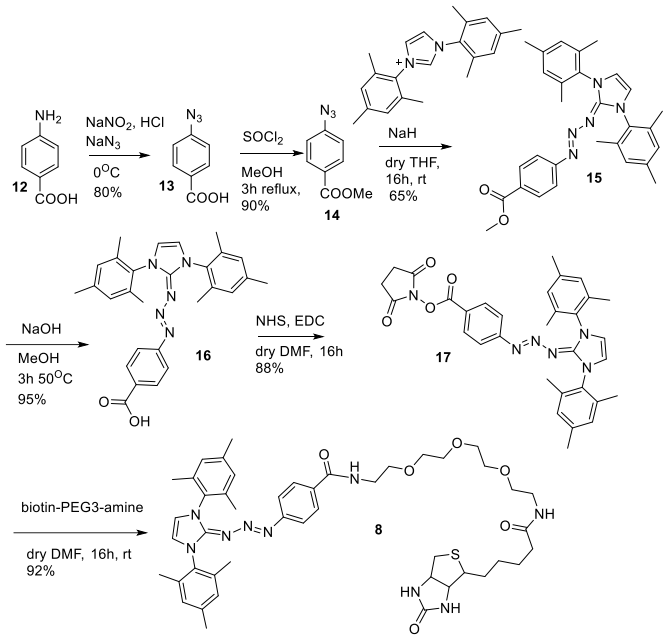
Mass Spectra:







Synthesis of photo-activable diazo biotin[4]:



methyl 4-azidobenzoate (14): 4-aminobenzoic acid (5g, 36 mmol) was dissolved in a mixture of 25 mL water and 8 mL concentrated H_2SO_4, and the resulting solution was cooled to 4 ˚C on ice for 15 min. To this was added an ice-cold solution of sodium nitrite (2.6 g, 38 mmol; in 5 mL water) in a drop-wise manner. Finally, a solution of sodium azide (2.9 g, 45 mmol; in 7 mL water) was added to this reaction mixture. Immediately after the addition of sodium azide, a white precipitate appeared. After stirring the reaction for 15 min, the product was filtered, washed with water, and air-dried to obtain **13** as a white powder (yield 4.8 g; 80%). This compound was directly used for the next step. Compound **13** (2 g, 12 mmol) was dissolved in 50 mL dry methanol and cooled on ice. Next, thionyl chloride (3.5 mL, 48 mmol) was added in a drop-wise manner and the reaction was allowed to continue for 20 h gradually warming up to room temperature. Upon completion (TLC), the solvent was evaporated using rotatory evaporator and 50 mL ethyl acetate was added to the residue. The organic layer was washed with 10 mL saturated sodium bicarbonate solution, dried on anhydrous sodium sulfate and finally evaporated to obtain a white solid 1.8 g. Yield: 90%; R_f: 0.5 in hexane-ethylacetate (3:1); ^{1}H NMR (600 MHz, Chloroform-*d*) δ 8.02 (d, *J*=8.6 Hz,2H), 7.05 (d, *J*=8.6 Hz,2H), 3.90 (s, 1H); ^{13}C NMR (150 MHz, Chloroform-*d*) δ 166.2, 144.7, 131.4, 126.7, 118.8, 52.1. HRMS (ESI-TOF-MS) (m/z): $[M]^+$ Calculated for $C_7H_6N_3O_2$ 177.0538 obtained 177.0865.

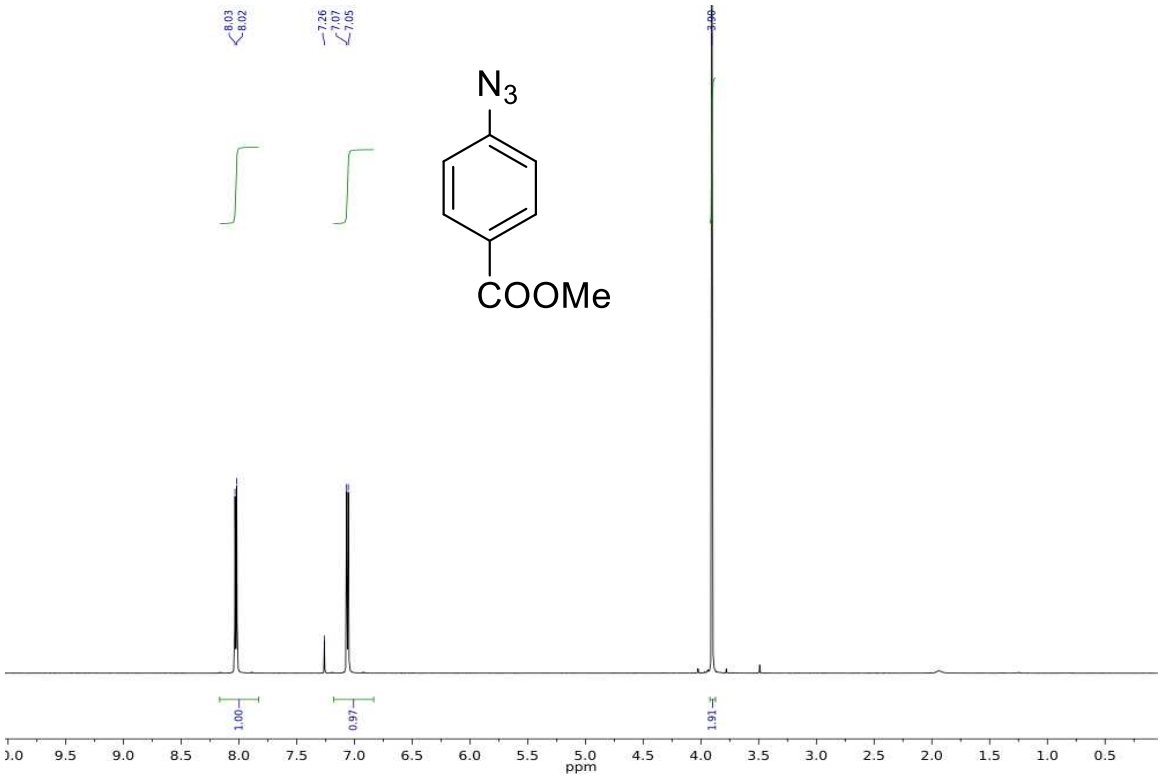
methyl (E)-4-((1,3-dimesityl-1,3-dihydro-2H-imidazol-2-ylidene)triaz-1-en-1-yl)benzoate (15): To a solution of methyl methyl 4-azidobenzoate, **3** (0.3 g, 1.7 mmol) in 20 mL dry THF, 1,3-dimesitylimidazolium chloride (0.58 g, 1.7 mmol) was added and the solution was cooled on ice for 20 min. To this solution, 0.136 g NaH (3.4 mmol, 60% in mineral oil) was added, and the reaction was allowed to continue for 18 h gradually warming from ice to room temperature.

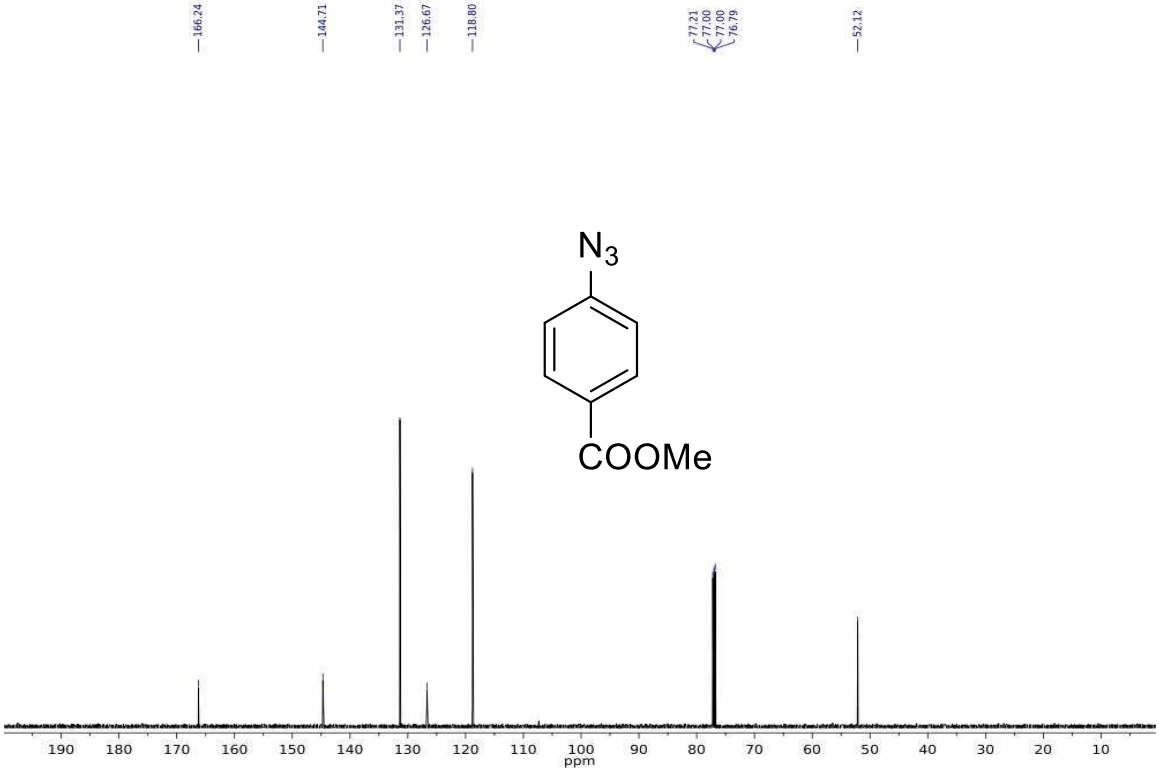
Next, the reaction was quenched with 10 mL water and 30 mL ethyl acetate was added to it. Organic layer washed with 10 mL brine and dried on anhydrous sodium sulfate. Finally the organic layer was evaporated to give 0.57 g bright yellow solid. Yield: 70%. R_f: 0.7 in Hexane-Ethylacetate (1:1); ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.69 (d, *J* = 8.7 Hz, 2H), 7.00 (s, 4H), 6.62 (s, 2H), 6.57 (d, *J* = 8.7 Hz, 2H), 3.85 (s, 3H), 2.37 (s, 6H), 2.15 (s, 12H); ^{13}C NMR (150 MHz, Chloroform-*d*) δ 167.25, 155.30, 151.57, 138.91, 134.90, 133.87, 129.76, 129.37, 126.28, 120.88, 117.23, 51.78, 21.09, 17.91. HRMS (ESI-TOF-MS) (m/z): $[M+H]^+$ Calculated for $C_{29}H_{32}N_5O_2$ 482.2551 obtained 482.2526.

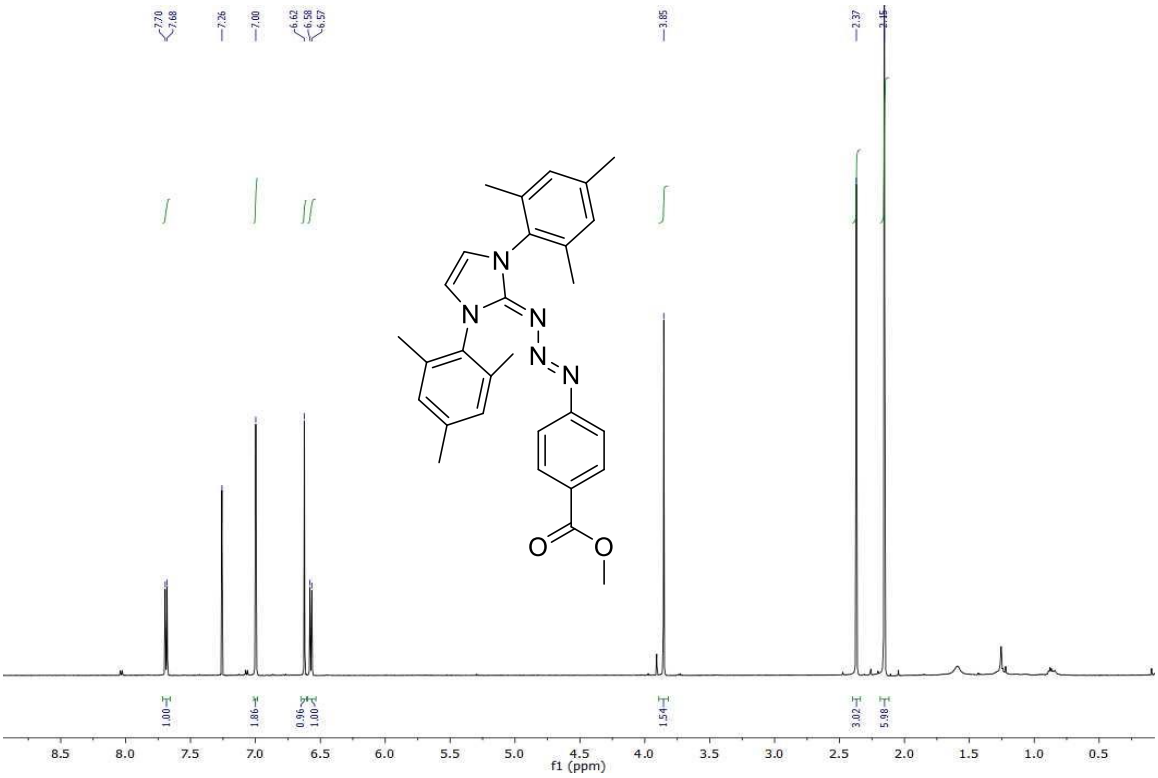
(E)-4-((1,3-dimesityl-1,3-dihydro-2H-imidazol-2-ylidene)triaz-1-en-1-yl)benzoic acid (16): To a methanolic solution of **15** (0.15 g, 0.31 mmol in 7 mL methanol), 1 mL of NaOH solution was added (from a stock solution of 0.1 g NaOH in 5 mL water) in a drop-wise manner and the reaction was stirred overnight. The solution was evaporated under vacuum and lyophilized to get 0.144 g yellow solid. Yield: 95%; R_f: 0.2 in Hexane-Ethylacetate (1:1); ^{1}H NMR (400 MHz, Methanol-d_4) δ 7.48 (d, *J* = 8.7 Hz, 2H), 6.90 (q, *J* = 0.8 Hz, 4H), 6.83 (s, 2H), 6.26 (d, *J* = 8.6 Hz, 2H), 2.22 (s, 6H), 2.00 (s, 12H); ^{13}C NMR (100 MHz, Methanol-d_4) δ 174.0, 152.1, 145.1, 139.0, 134.5, 133.8, 129.1, 128.9, 119.4, 117.7, 109.9, 19.8, 16.6; HRMS (ESI-TOF-MS) (m/z): $[M+H]^+$ Calculated for $C_{28}H_{30}N_5O_2$ 468.2394 obtained 468.2392.

2,5-dioxopyrrolidin-1-yl (E)-4-((1,3-dimesityl-1,3-dihydro-2H-imidazol-2-ylidene)triaz-1-en-1-yl)benzoate (17): 0.144 g (E)-4-((1,3-dimesityl-1,3-dihydro-2H-imidazol-2-ylidene)triaz-1-en-1-yl)benzoic acid, **16** was dissolved in 7 mL dry DCM. To this solution 0.15 g EDC (0.61 mmol) and 0.07 g of NHS (0.61 mmol) and was added and stirred for 10 h. Next, the solvent was evaporated, 30 mL ethyl acetate was added to the residue, and the ethylacetate lair was washed three times with 20 mL water each. The ethyl acetate solution was dried over anhydrous sodium sulfate and evaporated to give 0.145 g yellow product which directly used to next step without further purification. Yield 88%; R_f: 0.4 in Hexane-Ethylacetate (1:1); ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.78-7.68 (m, 2H), 7.01 (s, 4H), 6.86 (s, 2H), 6.61-6.59 (m, 2H), 2.87 (s, 4H), 2.37 (s, 6H), 2.16 (s, 12H); ^{13}C NMR (150 MHz, Chloroform-*d*) δ 189.4, 161.8, 157.0, 139.0, 134.8, 133.7, 132.5, 130.9, 129.4, 121.2, 120.4, 119.3, 117.5, 25.6, 21.0, 17.8; HRMS (ESI-TOF-MS) (m/z): $[M+H]^+$ Calculated for $C_{32}H_{33}N_6O$ 565.2558 obtained 565.2549.

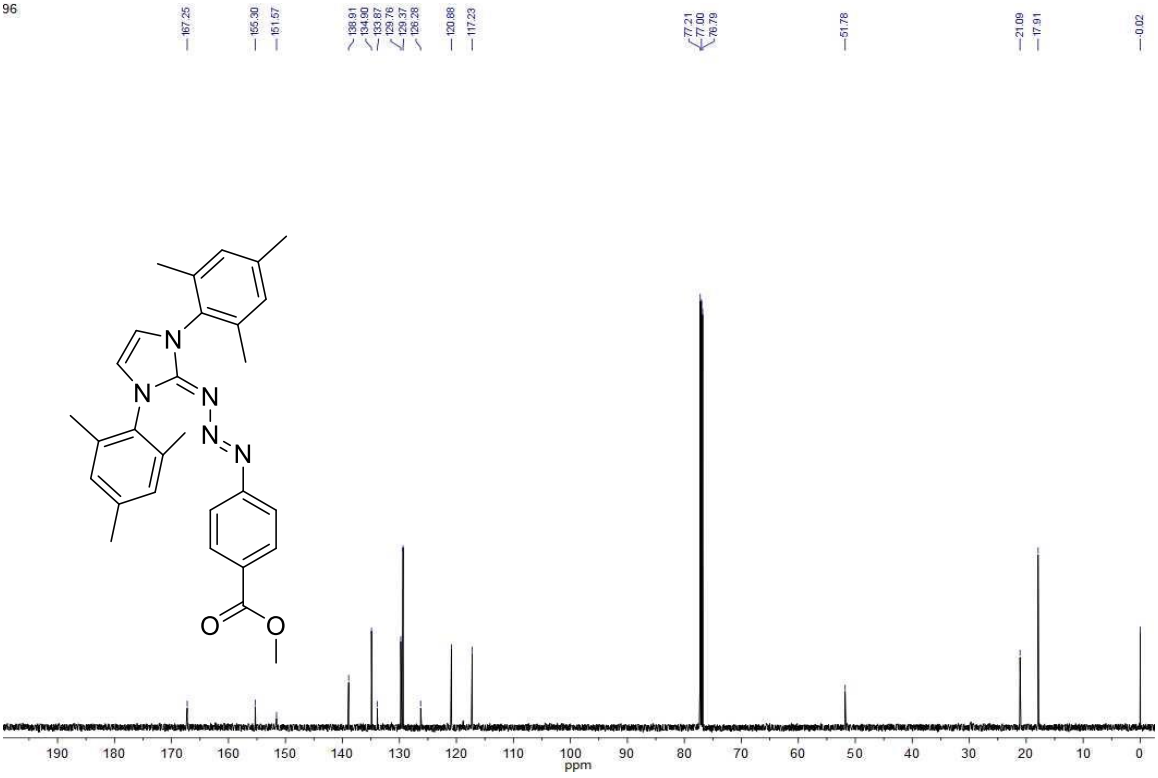
(E)-N-(1-(4-((1,3-dimesityl-1,3-dihydro-2H-imidazol-2-ylidene)triaz-1-en-1-yl)phenyl)-1-oxo-2,6,9,12-tetraoxa-3-azatetradecan-14-yl)-5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide (8): 0.01 g of **17** (0.018 mmol) was dissolved in 5 mL of dry DMF, into which 0.01 g Biotin-PEG_3-amine (0.02 mmol, purchased from Chem Impex Int'l. Inc) was added and stirred for 10 h at room temperature. Next, the product was extracted using 10 mL ethyl acetate, and the organic payer was washed with brine and saturated sodium bicarbonate solutions, dried over sodium sulfate, and evaporated under vacuum in the dark to get 0.014 g yellow solid product (Yield 92%). The purity of the product (**8**) was confirmed by checked by HPLC-MS analysis. HRMS (ESI-TOF-MS) (m/z): $[M+H]^+$ Calculated for $C_{46}H_{62}N_9O_7S$ 868.4538 obtained 868.4531.

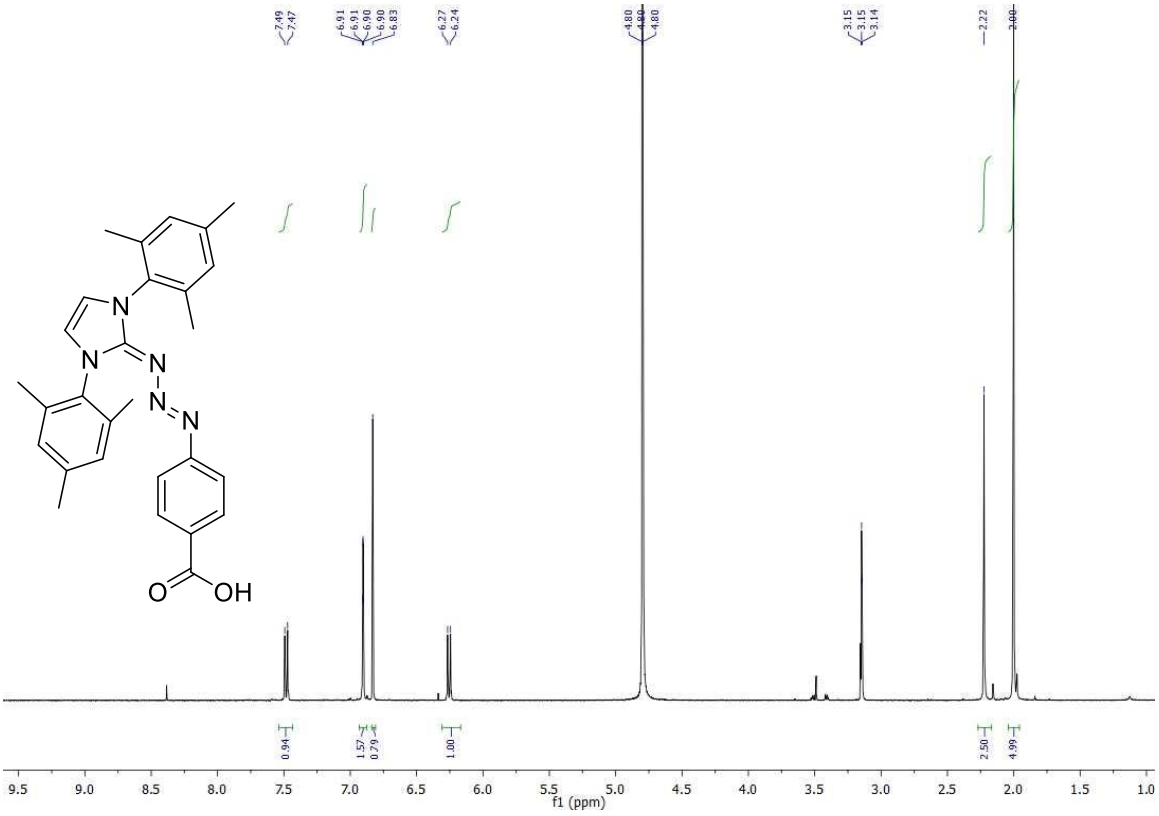

8.03
8.02
7.26
7.07
7.05
3.00
N3
COOMe
1.00
0.97
1.91
ppm

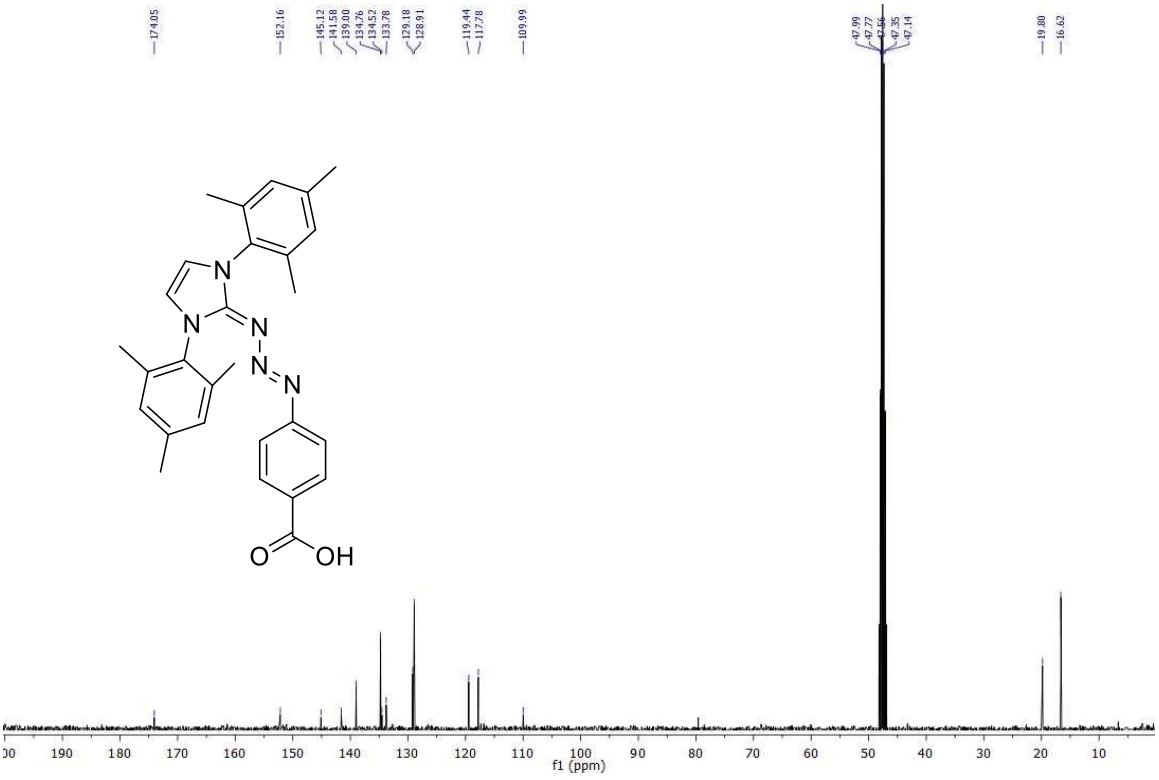

166.24
144.71
131.37
126.67
118.80
77.21
77.00
77.00
76.79
52.12
N3
COOMe
ppm

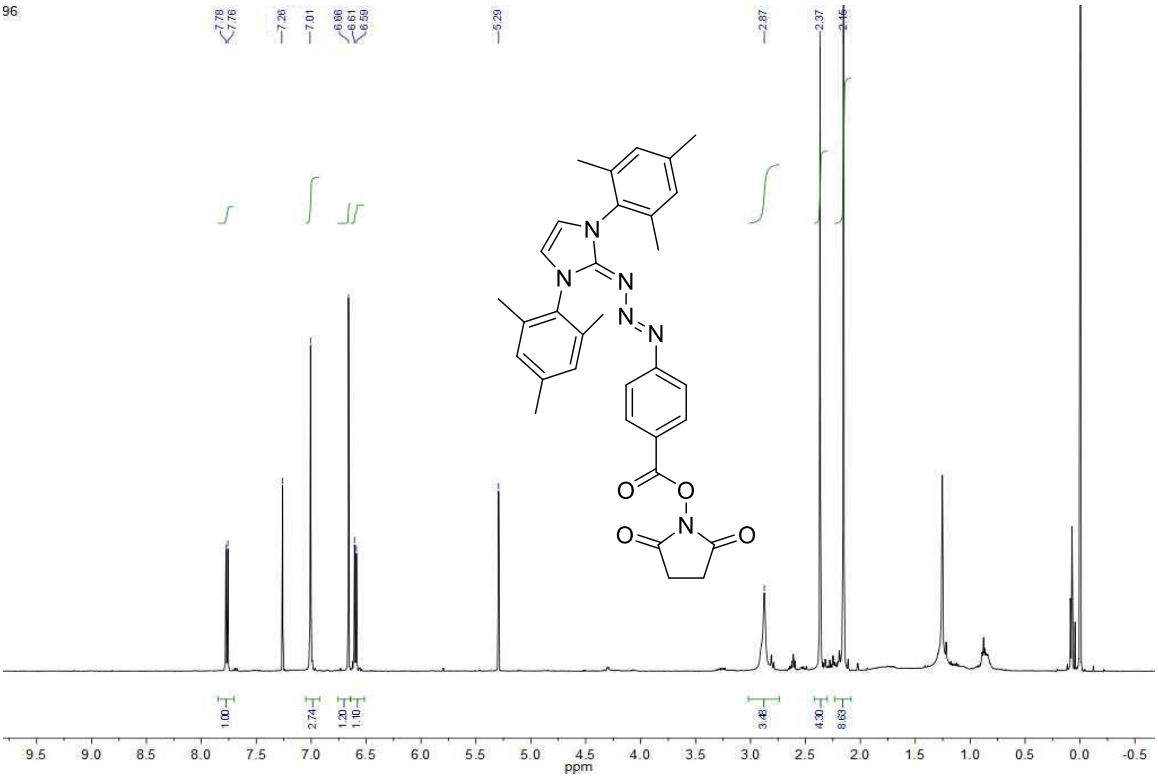

7.70
7.68
7.26
7.00
6.62
6.58
6.57
3.85
2.37
3.45
1.00
1.86
0.96
1.00
1.54
3.02
5.98
f1 (ppm)

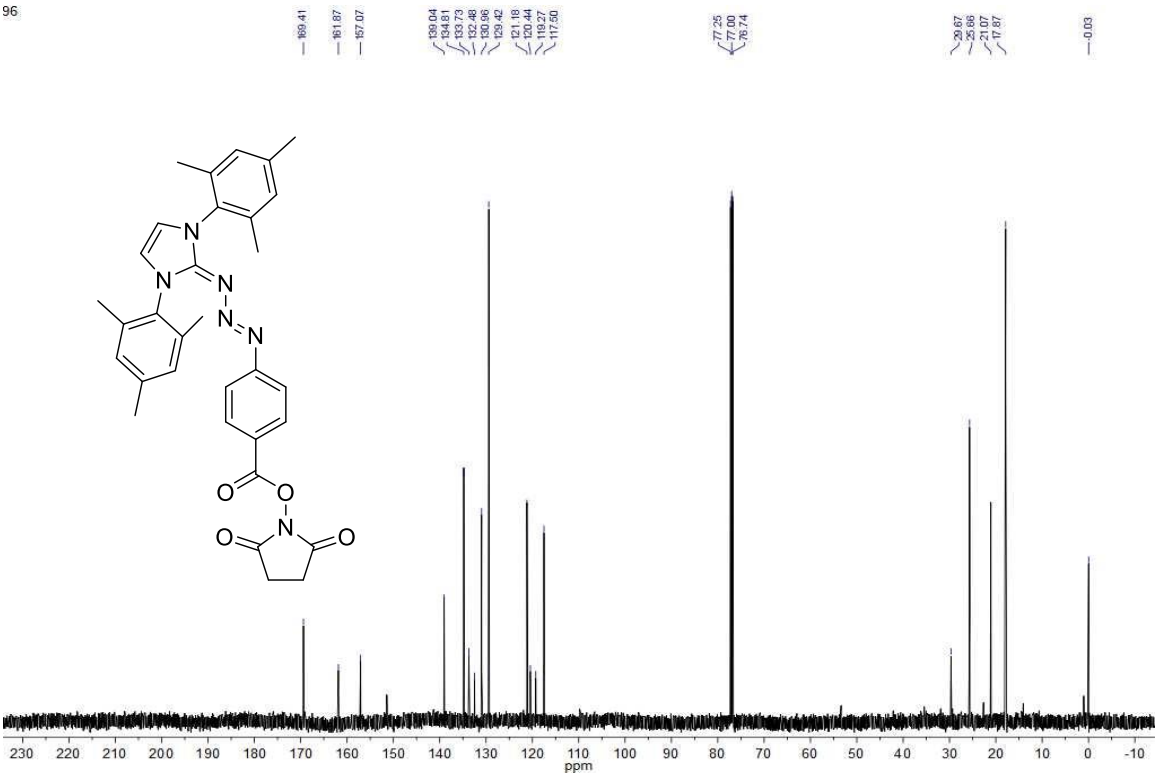
96

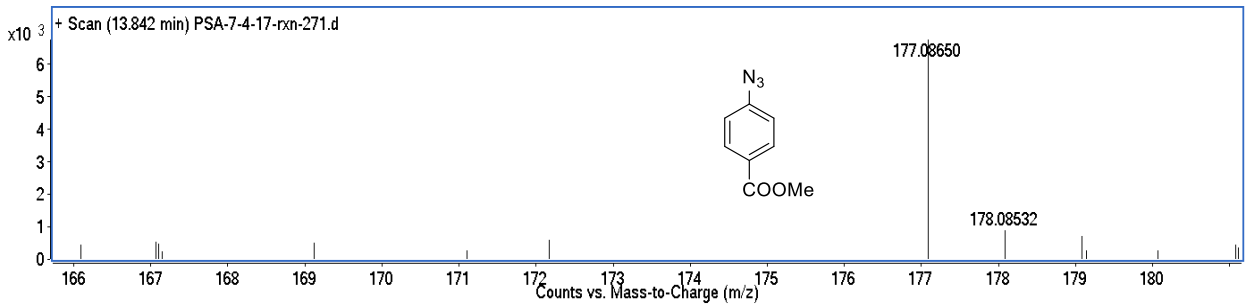

167.25
155.30
151.67
138.91
134.90
133.87
129.76
129.37
126.28
120.88
117.23
77.21
77.00
76.79
61.78
21.09
17.91
0.02
ppm

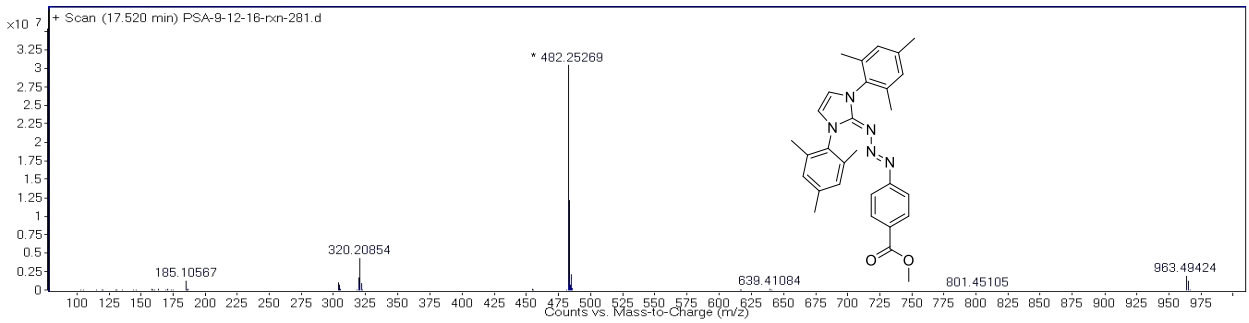

7.49
7.47
6.91
6.91
6.90
6.90
6.83
6.27
6.24
4.80
4.80
4.80
3.15
3.15
3.14
2.22
2.00
0.94
1.57
0.79
1.00
2.50
4.99
f1 (ppm)

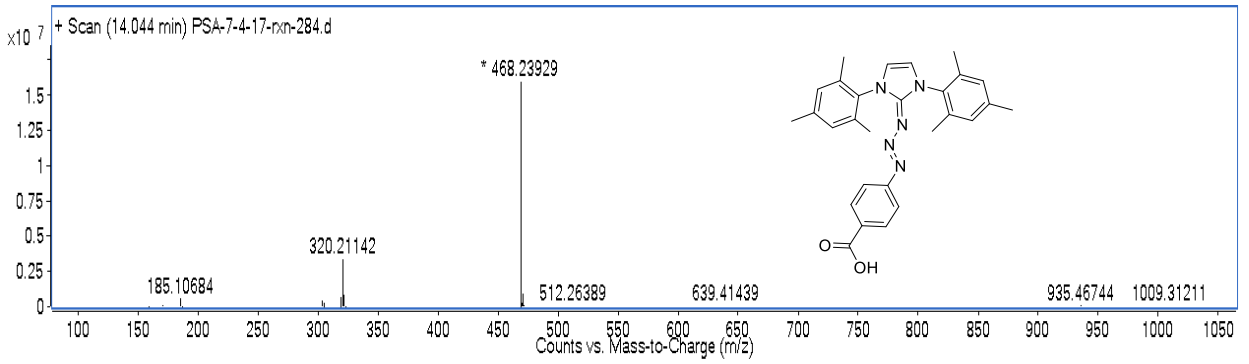

174.05
152.16
145.12
141.58
139.00
134.76
134.52
133.78
129.18
128.91
119.44
117.78
109.99
47.99
47.77
47.56
47.35
47.14
19.80
16.62
f1 (ppm)

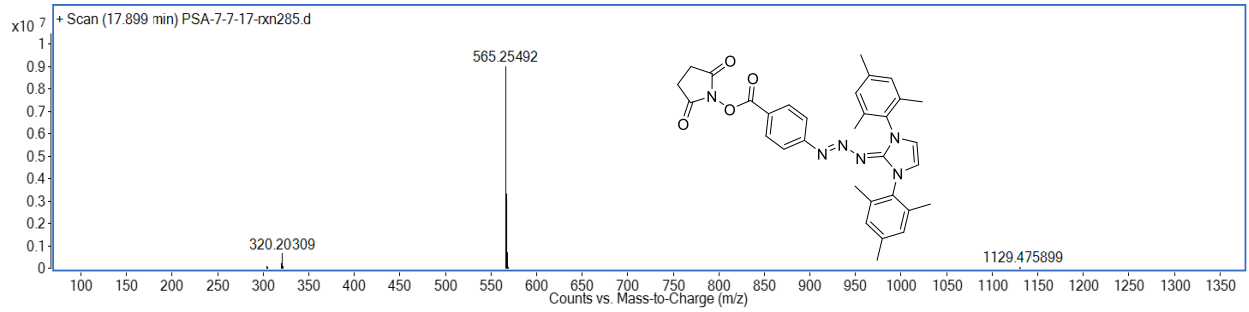

96
7.78
7.76
7.26
7.01
6.66
6.61
6.59
5.29
2.87
2.37
2.46
1.00
2.74
1.20
1.10
3.46
4.00
8.63
ppm

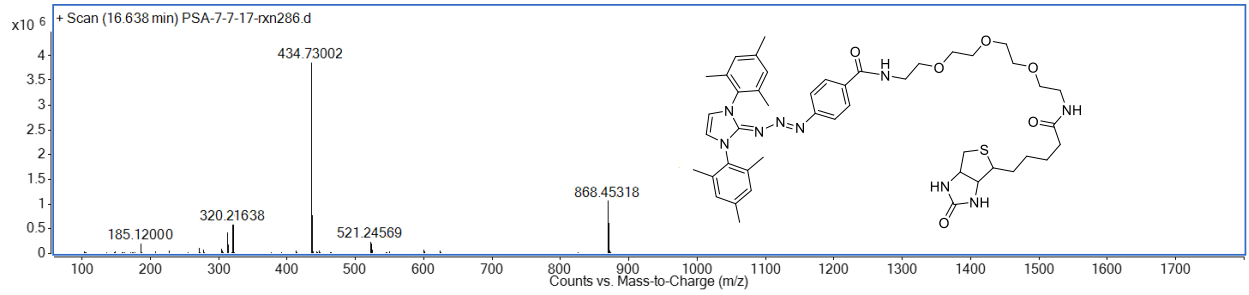

96
169.41
161.87
157.07
138.04
134.81
133.73
132.48
130.96
129.42
121.18
120.44
119.27
117.50
77.25
77.00
76.74
39.67
25.66
21.07
17.87
-0.03
ppm


+ Scan (13.842 min) PSA-7-4-17-rxn-271.d
x10 3
177.08650
178.08532
N3
COOMe
Counts vs. Mass-to-Charge (m/z)

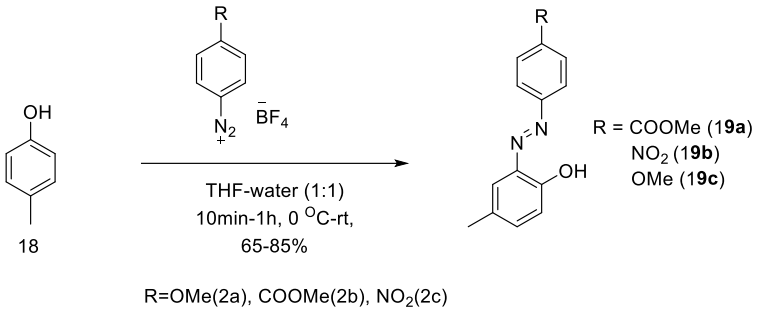

+ Scan (17.520 min) PSA-9-12-16-rxn-281.d
x10 7
* 482.25269
320.20854
185.10567
639.41084
801.45105
963.49424
Counts vs. Mass-to-Charge (m/z)


+ Scan (14.044 min) PSA-7-4-17-rxn-284.d
x10 7
* 468.23929
320.21142
185.10684
512.26389
639.41439
935.46744
1009.31211
OH
Counts vs. Mass-to-Charge (m/z)


+ Scan (17.899 min) PSA-7-7-17-rxn285.d
x10 7
565.25492
320.20309
1129.475899
Counts vs. Mass-to-Charge (m/z)


+ Scan (16.638 min) PSA-7-7-17-rxn286.d
x10 6
434.73002
868.45318
320.21638
185.12000
521.24569
Counts vs. Mass-to-Charge (m/z)

Synthesis of Different p-Cresol derivatives:



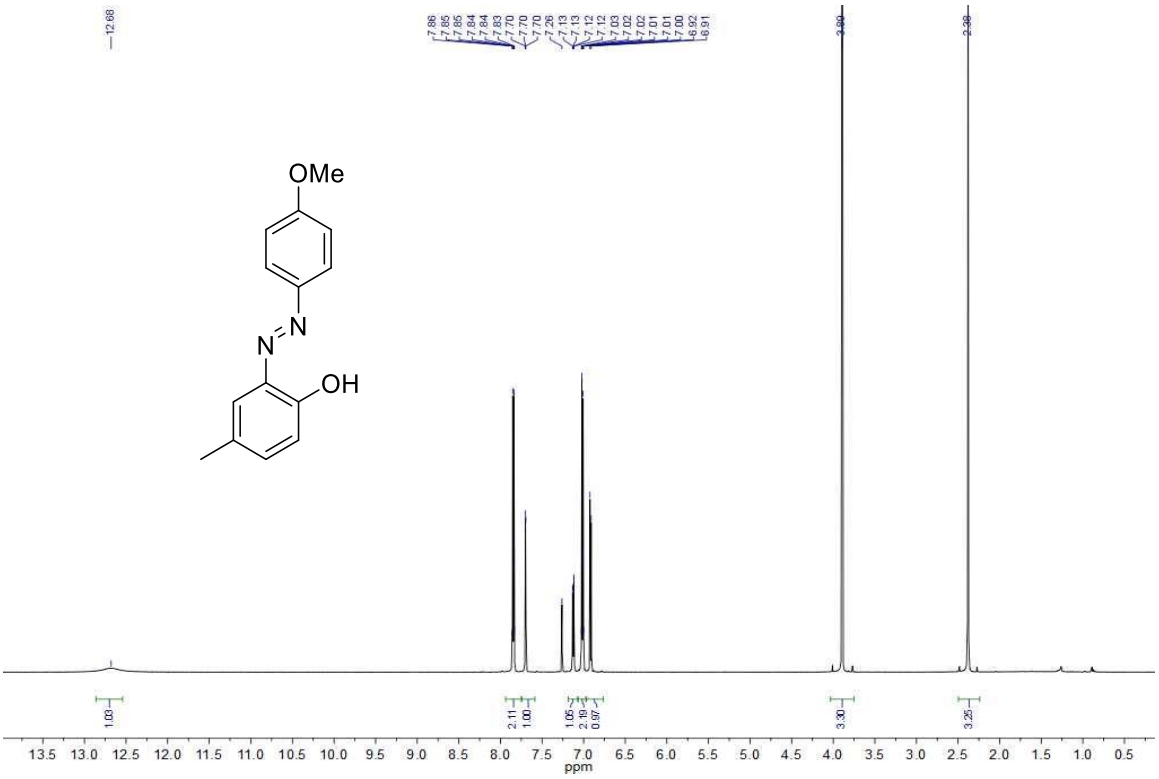
p-Cresol, 18 (0.1 g, 0.92 mmol) was dissolved in 10 mL THF. To this solution 2 mL water and K_2CO_3 (0.127 g, 0.92 mmol) was added and stirred on ice for 15 min. A solution of aryl diazonium (2 mmol) in water was added to the previous solution and the reaction mixture was stirred on room temperature for 1h. The solution then diluted with 30 mL ethyl acetate. The organic layer was washed alternatingly with brine and water three times. Finally, the organic layer evaporated and the compounds were purified by column chromatography. The compounds are characterized by NMR and mass spectrometry.

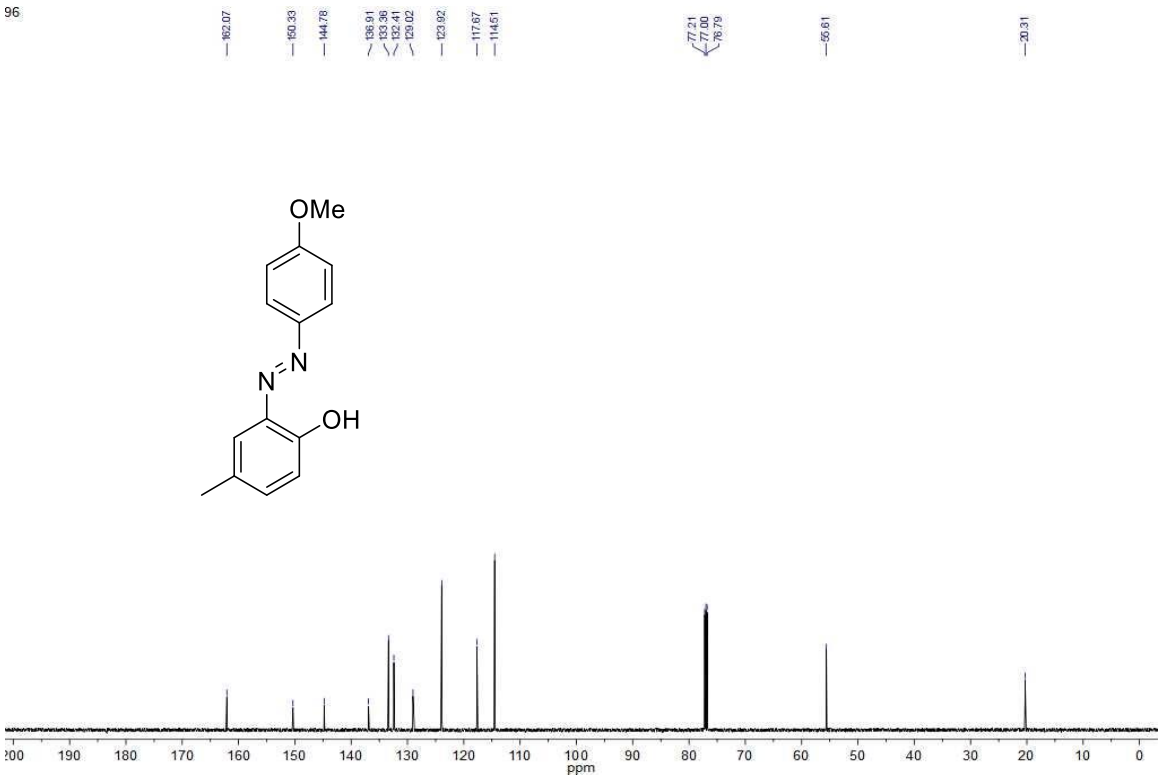
(E)-2-((4-methoxyphenyl)diazenyl)-4-methylphenol (19c): Yield 65%; R_f : 0.7 in 20:1 Hexane-Ethylacetate; ^{1}H NMR (600 MHz, Chloroform-*d*) δ 12.68 (s, 1H), 7.94 – 7.75 (m, 2H), 7.75 – 7.59 (m, 1H), 7.12 (dd, *J* = 8.4, 2.2 Hz, 1H), 7.07 – 6.96 (m, 2H), 6.92 (d, *J* = 8.3 Hz, 1H), 3.89 (s, 3H), 2.38 (s, 3H); ^{13}C NMR (151 MHz, cdcl$_3$) δ 162.07, 150.33, 144.78, 136.91, 133.36, 132.41, 129.02, 123.92, 117.67, 114.51, 77.21, 77.00, 76.79, 55.61, 20.31; HRMS (ESI-TOF-MS) (m/z): $[M+H]^+$ Calculated for $C_{14}H_{15}N_2O_2$ 243.1128, found 243.1121.

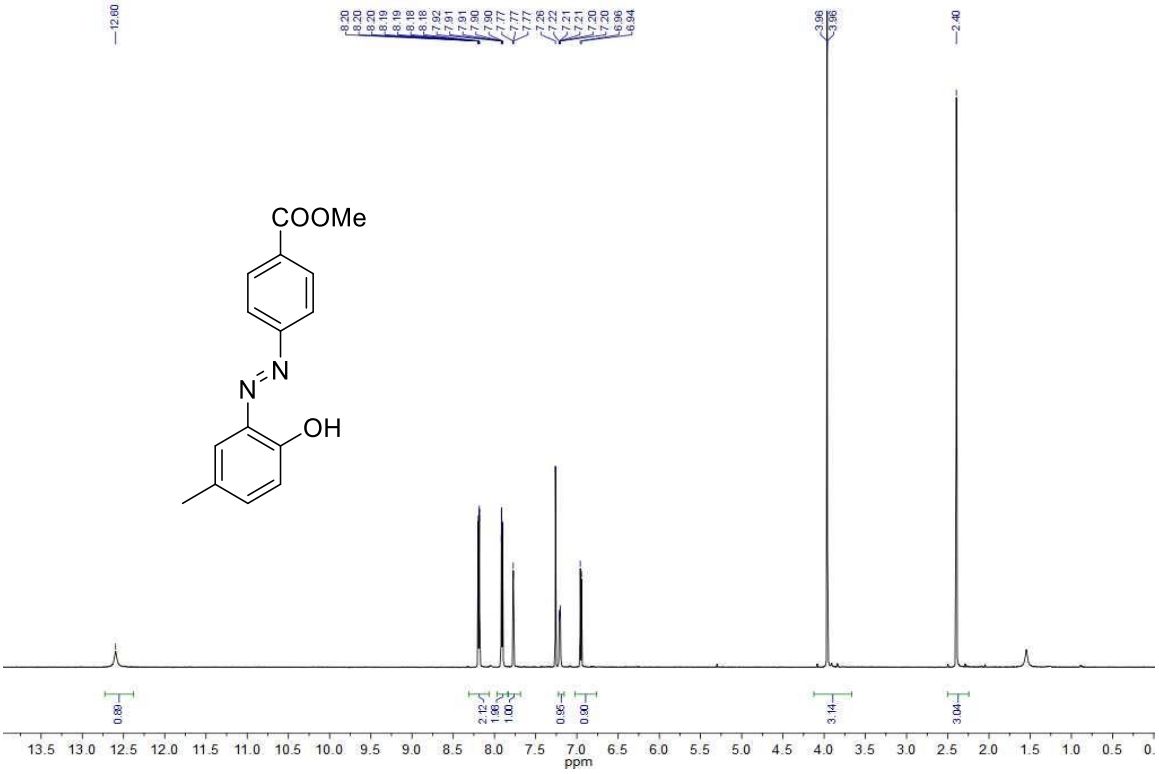
methyl (E)-4-((2-hydroxy-5-methylphenyl)diazenyl)benzoate (19a): Yield 82%; R_f : 0.6 in 20:1 Hexane-Ethylacetate; ^{1}H NMR (600 MHz, Chloroform-*d*) δ 12.60 (s, 1H), 8.31 – 8.06 (m, 2H), 7.91 (dd, *J* = 8.5, 1.2 Hz, 2H), 7.84 – 7.68 (m, 1H), 7.22 – 7.15 (m, 1H), 6.95 (d, *J* = 8.4 Hz, 1H), 3.96 (s, 3H), 2.40 (s, 3H); ^{13}C NMR (151 MHz, cdcl$_3$) δ 166.33, 153.31, 150.70, 137.34, 135.30, 133.36, 131.79, 130.82, 129.50, 121.99, 118.04, 77.21, 77.00, 76.79, 52.38, 20.29; HRMS (ESI-TOF-MS) (m/z): $[M+H]^+$ Calculated for $C_{15}H_{15}N_2O_3$ 271.1077, found 271.1072.

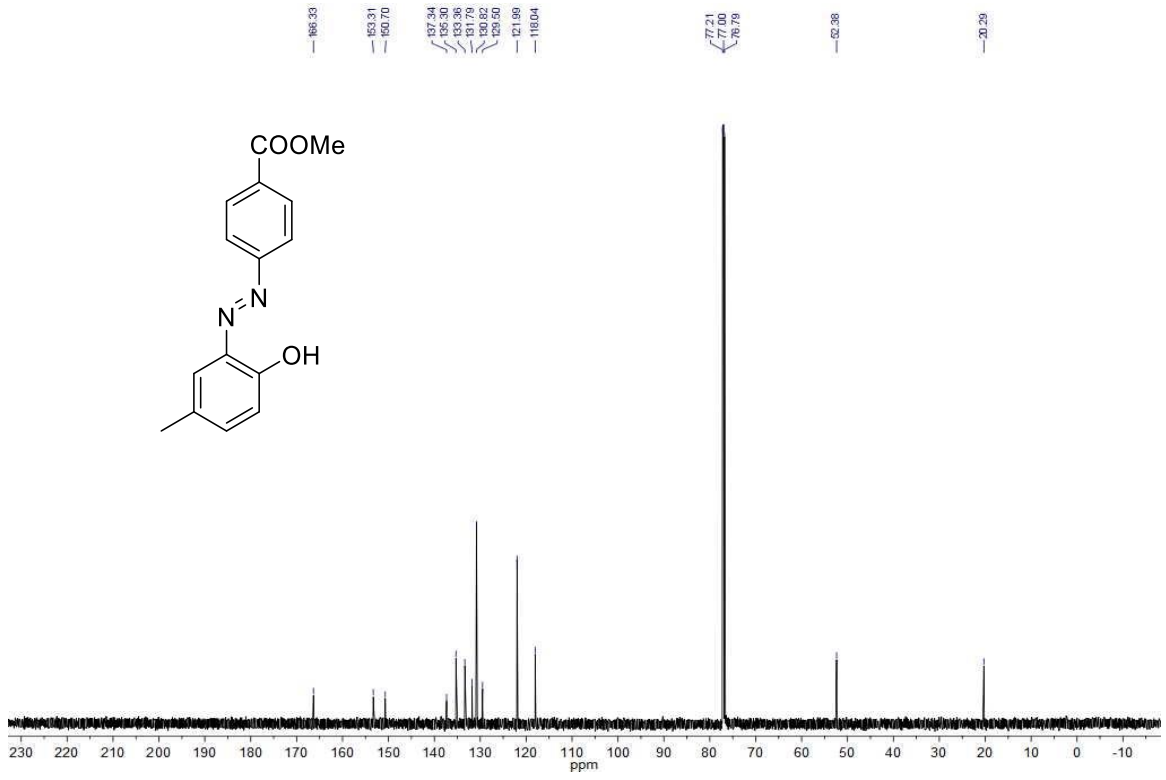
(E)-4-methyl-2-((4-nitrophenyl)diazenyl)phenol (19b): Yield 88%; R_f : 0.5 in 25:1 Hexane-Ethylacetate; ^{1}H NMR (600 MHz, Chloroform-*d*) δ 12.43 (s, 1H), 8.58 – 8.23 (m, 2H), 8.12 – 7.89 (m, 2H), 7.86 – 7.70 (m, 1H), 7.25 (d, *J* = 7.5 Hz, 2H), 6.97 (d, *J* = 8.4 Hz, 1H), 2.41 (s, 3H); ^{13}C NMR (151 MHz, cdcl$_3$) δ 154.13, 150.89, 148.53, 137.55, 136.32, 133.61, 129.87, 124.97, 122.72, 118.22, 77.21, 77.00, 76.79, 20.28; HRMS (ESI-TOF-MS) (m/z): $[M+H]^+$ Calculated for $C_{13}H_{12}N_3O_3$ 258.0873, found 258.0873.

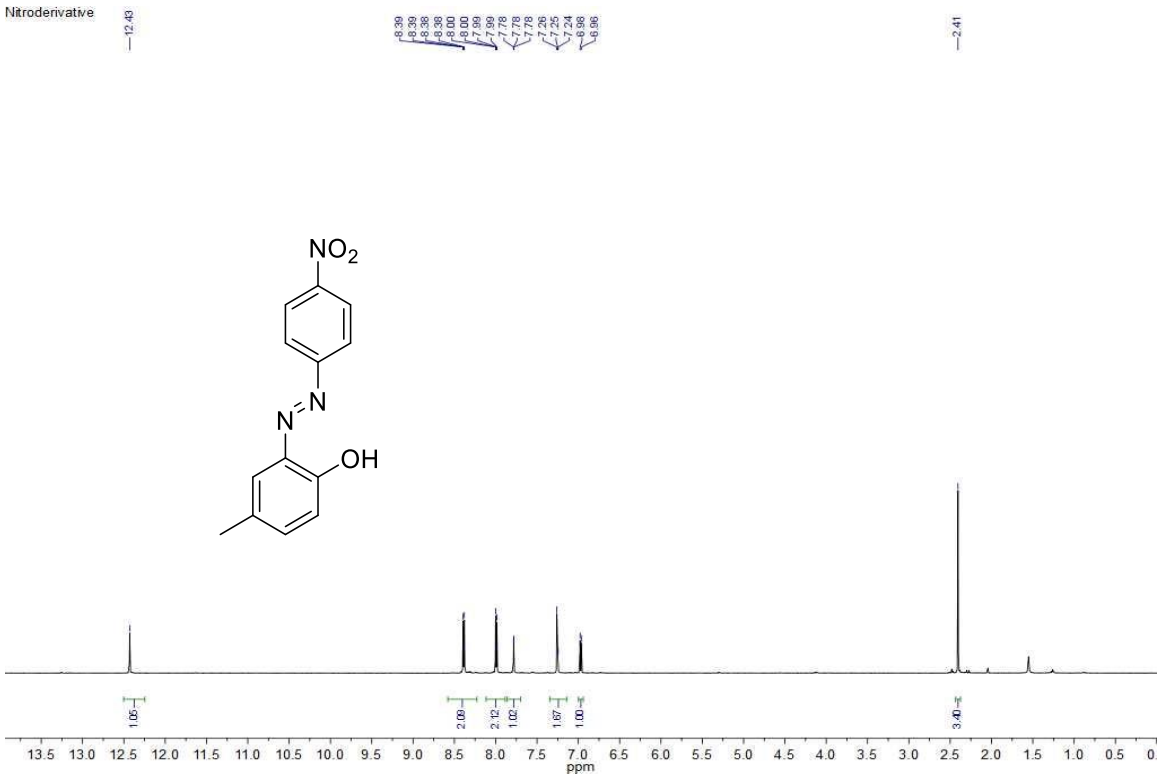
NMR spectra:

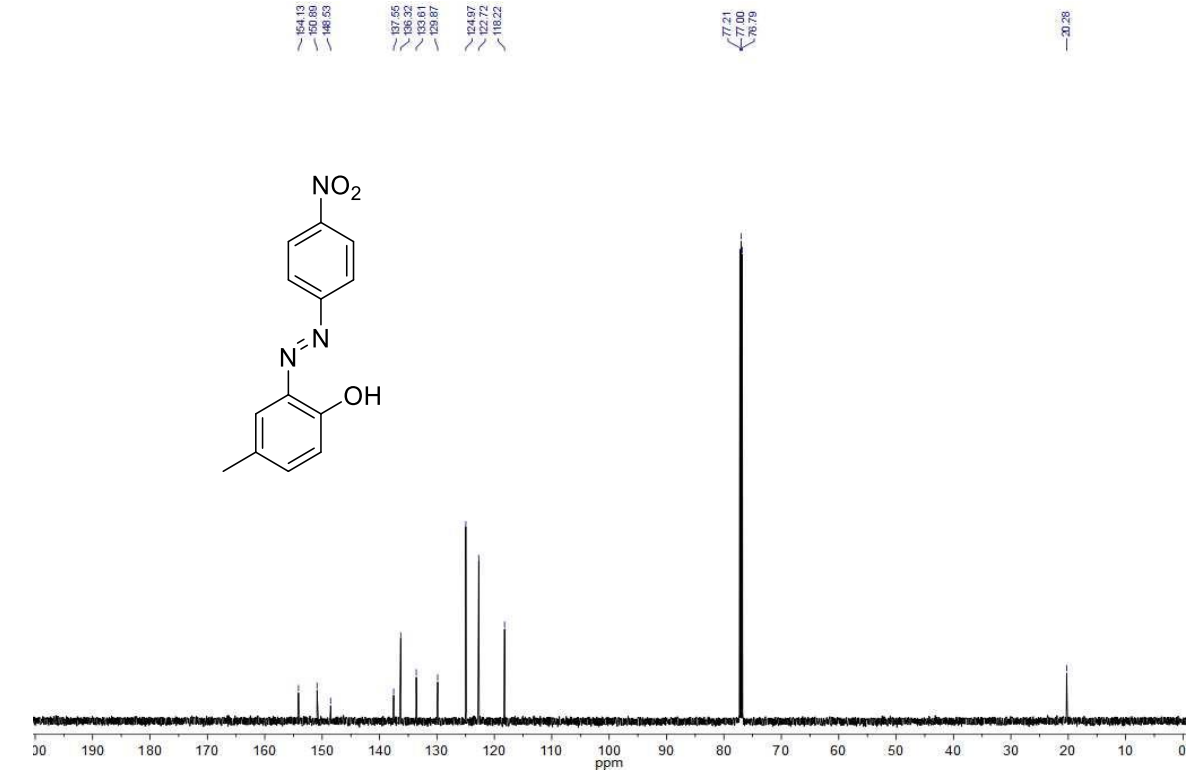




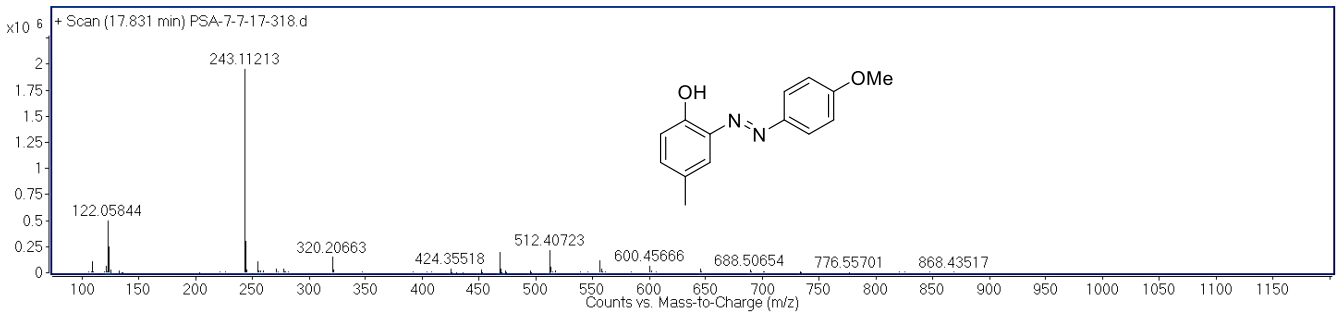

12.60
8.20
8.20
8.20
8.19
8.19
8.18
8.18
7.92
7.91
7.91
7.90
7.90
7.77
7.77
7.77
7.26
7.22
7.21
7.21
7.20
7.20
6.96
6.94
3.96
3.96
2.40
COOMe
N
N
OH
0.89
2.12
1.06
1.00
0.95
0.90
3.14
3.04
13.5 13.0 12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.
ppm

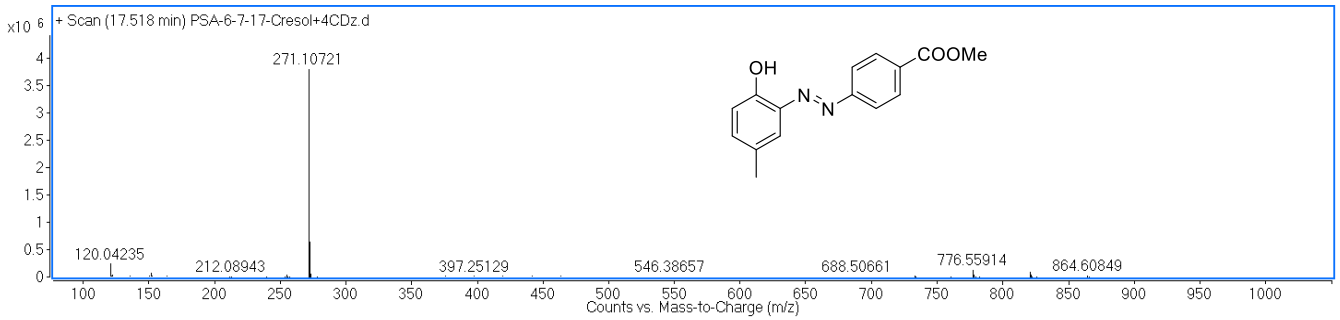

166.33
153.31
150.70
137.34
136.30
133.36
131.79
130.22
129.50
121.99
118.04
77.21
77.00
76.79
52.38
20.29
COOMe
N
N
OH
230 220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
ppm

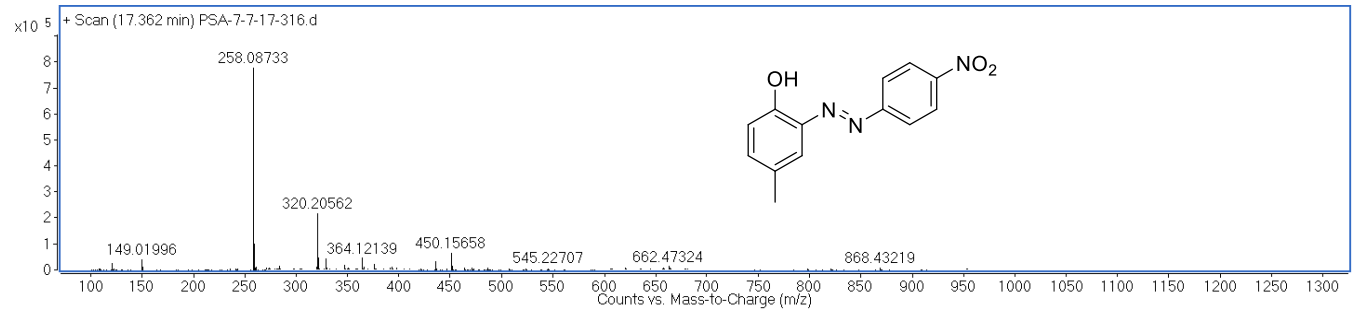

Nitroderivative
12.43
8.39
8.39
8.38
8.38
8.00
8.00
7.99
7.99
7.78
7.78
7.26
7.25
7.24
6.98
6.96
2.41
NO2
N
N
OH
1.05
2.09
2.12
1.02
1.67
1.00
3.40
ppm


154.13
150.89
148.53
137.55
136.32
133.61
129.87
124.97
122.72
118.22
77.21
77.00
76.79
20.28
NO2
N
N
OH
ppm

Mass spectra:







References:

1. Italia, J. S.; Addy, P. S.; Wrobel, C. J.; Crawford, L. A.; Lajoie, M. J.; Zheng, Y.; Chatterjee, A. *Nat. Chem. Biol.* **2017,** *13* (4), 446-450.
2. (a) Hutchins, B. M.; Kazane, S. A.; Staflin, K.; Forsyth, J. S.; Felding-Habermann, B.; Schultz, P. G.; Smider, V. V. *J. Mol. Biol.* **2011,** *406* (4), 595-603; (b) Hutchins, B. M.; Kazane, S. A.; Staflin, K.; Forsyth, J. S.; Felding-Habermann, B.; Smider, V. V.; Schultz, P. G. *Chem. Biol.* **2011,** *18* (3), 299-303.
3. Gavrilyuk, J.; Ban, H.; Nagano, M.; Hakamata, W.; Barbas, C. F., 3rd. *Bioconjugate Chem.* **2012,** *23* (12), 2321-8.
4. Jensen, S. M.; Kimani, F. W.; Jewett, J. C. *Chembiochem* **2016,** *17* (23), 2216-2219.