

Single-site labeling of lysine in proteins through a metal-free multicomponent approach

Maheshwerreddy Chilamari,^a Neetu Kalra,^b Sanjeev Shukla,^b Vishal Rai^{a*}

Contents

1. General.....	2
2. Methods.....	3
2.1. Procedure for protein labeling.....	3
2.2. Digestion protocol.....	3
2.3. Procedure for late-stage orthogonal tagging of the protein	4
2.4. Procedure for Fab labeling.....	4
2.5. Procedure for synthesis of ADC (Trastuzumab-doxorubicin conjugate)	4
2.6. Procedure for MTT assay (Anti-proliferative activity assay).....	5
3. Synthesis of reagents	5
4. Additional results and discussions	12
4.1. Optimization of reaction conditions	12
4.2. Stability of triethylphosphite in aqueous condition	15
4.3. Plausible mechanism.....	16
5. Effect of labeling on the structure of protein	17
6. Enzymatic assay	18
6.1. Enzymatic assay of Lysozyme C.....	18
6.2. Enzymatic assay of RNase A.....	19
7. MS data of proteins.....	20
7.1. Screening of nucleophiles	20
7.2. Scope of aldehydes	22
7.3. Scope of protein.....	24
8. Late-stage orthogonal tagging of the labeled protein	30
9. Reactions with Fab and monoclonal antibody.....	37
9.1. Labeling of Trastuzumab-Fab.....	37
9.2. Synthesis of antibody-drug conjugate (T-Dox, 13).....	39
9.2.1. Calculation of drug to antibody ratio (DAR).....	39
9.2.2. Anti-proliferative activity test (MTT assay).....	40
10. NMR data of small molecules	41
11. Acknowledgements.....	54
12. References	54

1. General

The reagents, proteins, and enzymes were purchased from Sigma-Aldrich, Alfa Aesar, and Merck Novabiochem. The organic solvents used were reagent grade. Thermo Scientific MaxQ 8000 (350 rpm, 25°C) incubator-shaker was used for conducting the reactions. BUCHI Rotavapor R-210/215 was used to remove organic solvents. CHRiST ALPHA 2-4 LD plus lyophilizer was used for lyophilization of the aqueous samples. UV spectra were recorded on PerkinElmer Lambda 25 UV-Vis spectrometer. Circular Dichroism (CD) measurements were recorded on JASCO J-815 CD spectropolarimeter equipped with Peltier temperature controller. All the spectra were measured in a cuvette of path length 2 cm at 25 °C with a scan speed of 50 nm/min and spectral band width of 1 nm. OriginPro 8 software is used for plotting the UV-Vis and CD spectra. In the enzymatic assay, absorbance was normalized to its values by dividing its absorbance maxima (OriginPro 8). Thin-layer chromatography (TLC) was performed on silica gel coated aluminium TLC plates (Merck, TLC Silica gel 60 F₂₅₄) and visualized using a UV lamp (254 nm). Flash column chromatography was carried out on Combiflash R_f 200 and gravity columns were performed with glass columns using 230-400 mesh silica gel from Merck. ¹H, ¹⁹F, and ¹³C NMR spectra were recorded on Bruker Avance III 400 MHz and Bruker Avance III 500 MHz NMR spectrometer. ¹H NMR spectra were referenced to TMS (0 ppm), CDCl₃ (7.26 ppm) DMSO-d₆ (2.50 ppm), and D₂O (4.79 ppm). ¹³C NMR spectra were referenced to CDCl₃ (77.16 ppm), ¹⁹F NMR spectra were referenced to TFA (-75.45 ppm). Peak multiplicities are designated by the following abbreviations: s, singlet; d, doublet; t, triplet; q, quartet; dd, doublet of doublet; m, multiplet; bs, broad singlet. All the NMR spectra were recorded at 298 K. Agilent Technologies 1200 series HPLC paired to Agilent 6130 mass spectrometer (ESI/APCI) was used to follow the reaction with small molecules and protein. HPLC experiments of compounds were performed on Poroshell 300SB-C18 column (3.0 × 50 mm × 2.7 μ) with the flow rate 0.4 ml/min. Bruker Daltonics MicroTOF-Q-II with electrospray ionization (ESI) was used for the HRMS data. Matrix-assisted laser desorption/ionization time of flight mass spectrometry was performed with on Bruker Daltonics UltrafleXtreme. Sinapic acid matrix was used for Lysozyme C, RNase A, Myoglobin, Ubiquitin, α-Lactalbumin, Cytochrome C and α-Chymotrypsinogen A. On the other hand, α-cyano-4-hydroxycinnamic acid (HCCA) matrix was used for digest of all the proteins. Data analysis was performed using flexAnalysis 3.4. Peptide mass and fragment ion calculator were used for peptide mapping and sequencing (<http://db.systemsbiology.net:8080/teomicsToolkit/FragIonServlet.html>).

2. Methods

2.1. Procedure for protein labeling

In a 1.5 ml eppendorf tube, a protein (**1**, 7.3 nmol) in phosphate buffer (90 μ l, pH 7.8, 0.1 M) was taken. To the protein solution, aldehyde (**2**, 2.19 μ mol) in DMSO (5 μ l) and triethylphosphite (**3**, 2.19 μ mol) in DMSO (5 μ l) were added from freshly prepared stock solutions. The reaction mixture was vortexed (350 rpm) in the incubator-shaker at 25 °C. The progress of reaction was monitored by MALDI-ToF MS using sinapic acid as matrix or ESI-MS. The reaction mixture was treated with hydroxylamine hydrochloride (250 μ g, 3.65 μ mol) in water (5 μ l) for 2 h to quench the excess aldehyde. The excess reagents and salts were removed using spin concentrator (3 kDa MWCO). The reaction mixture was investigated with mass spectrometry. The solvent was removed through lyophilization and the sample was subjected to proteolytic digestion. The peptide mapping of the digest and MS-MS of the labeled peptide(s) was used to confirm the site of labeling.

2.2. Digestion protocol

Procedure for in-solution digestion of RNase A and Lysozyme C, α -Chymotrypsinogen A and α -Lactalbumin

All the solutions were prepared freshly before use in reactions.^[1]

Step 1. Protein (0.1 mg, 7.3 nmol) in 100 mM tris (10 μ l, pH 7.8) with urea (6 M) was taken in 2 ml Eppendorf tube and vortexed for 30 minutes.

Step 2. Disulfide reduction: To reduce the disulfide bonds, reducing agent (1 μ L, 0.2 M DTT in 0.1 M tris) was added to the solution and sample was vortexed for 1 h at 25 °C or 37 °C.

Step 3. Sulfhydryl alkylation: To block the free sulfhydryl groups, alkylating agent (4 μ L, 0.2 M iodoacetamide in 0.1 M tris) was added to the solution and incubated (in the dark) for 1 h at ambient temperature.

Step 4. Quenching alkylating reagent: To quench the unreacted iodoacetamide, reducing agent (DTT, 4 μ L) was added again to the mixture and the sample was vortexed at 25 °C for 1 h. Dilution of the reaction mixture with grade I water reduced the urea concentration to 0.6 M.

Step 5. Enzymatic digestion: To this solution, 10 μ L of α -chymotrypsin or trypsin solution [2 μ g, based on ratio of chymotrypsin or trypsin/protein (1:50); α -chymotrypsin or trypsin in 1 mM HCl was dissolved in 0.4 M tris and 0.01 M CaCl₂] was added and the mixture was incubated at 37 °C for 18 h. Trifluoroacetic acid (0.5 %) was used to adjust the pH of digested solution to < 6 (verified by pH paper). Subsequently, the sample was subjected to peptide mapping by MS and sequencing by MS-MS.

Procedure for in-solution digestion of Ubiquitin, Cytochrome C and Myoglobin

Steps 1 and 5 were used for digestion of Ubiquitin, Subtilisin A, Cytochrome C, and Myoglobin. Steps 2, 3 and 4 are not desired as these proteins do not have disulfide bridges or free sulfhydryl groups.

2.3. Procedure for late-stage orthogonal tagging of the protein

In a 1.5 ml eppendorf tube, α -Chymotrypsinogen A (**1g**, 7.3 nmol) in phosphate buffer (90 μ l, pH 7.8, 0.1 M) was taken. To the protein solution 4-(4-oxopiperidine-1-carbonyl)benzaldehyde (**2j**, 2.19 μ mol) in DMSO (5 μ l) and triethylphosphite (**3f**, 2.19 μ mol) in DMSO (5 μ l) were added from freshly prepared stock solutions. The reaction mixture was vortexed (350 rpm) in the incubator-shaker at 25 °C. The reaction progress was followed by ESI-MS. After 3 h, the excess aldehyde **2j** and triethylphosphite **3h** were removed by spin concentrator (3 kDa MW cut off). The labeled protein sample (**4u**) was lyophilized before subjecting it to the late-stage tagging. The labeled protein (**4u**) was treated with hydroxylamine derivatives (**6**, 250 μ g, 1.46 μ mol in 5 μ l DMSO) in phosphate buffer (pH 7.0, 0.1 M, 95 μ l) for 2 h. Subsequently, the salts and excess hydroxylamine derivative were removed by spin concentrators (10 kDa MW cut off). The sample was used for analysis by MS. The sample was lyophilized and stored at -20 °C.

2.4. Procedure for Fab labeling

In a 1.5 ml eppendorf tube, Trastuzumab Fab (**1h**, 1.25 nmol) in phosphate buffer (45 μ l, pH 7.8, 0.1 M) was taken. Benzaldehyde (**2a**, 0.37 μ mol) in DMSO (2.5 μ l) and triethylphosphite (**3h**, 0.37 μ mol) in DMSO (2.5 μ l) were added from freshly prepared stock solutions. The reaction mixture was vortexed (350 rpm) in the incubator-shaker at 25 °C for 1 h. The progress of reaction was monitored by MALDI-ToF MS using sinapic acid matrix or ESI-MS. The excess aldehyde in the reaction mixture was quenched with hydroxylamine hydrochloride (250 μ g, 0.625 μ mol) in water (5 μ l) for 2 h. The excess reagents and salts were removed by spin concentrators (10 kDa MWCO). The sample was lyophilized and re-suspended in the desired buffer before digestion, peptide mapping, and sequencing by MS-MS.

2.5. Procedure for synthesis of ADC (Trastuzumab-doxorubicin conjugate)

In a 1.5 ml eppendorf tube, Trastuzumab (**1i**, 3.35 nmol) in phosphate buffer (90 μ l, pH 7.8, 0.1 M) was taken. The reagents 4-acetylbenzaldehyde (**2k**, 1.01 μ mol) in DMSO (5 μ l) and triethylphosphite (**3h**, 1.01 μ mol) in DMSO (5 μ l) were added from freshly prepared stock solutions. The reaction mixture was vortexed (350 rpm) in incubator-shaker at 25 °C for 8 h. The progress of the reaction was monitored by MALDI-ToF MS using sinapic acid matrix. The excess reagents and salts were removed by spin concentrators (30 kDa MWCO). The Trastuzumab conjugate (**10**) was lyophilized and re-suspended in phosphate buffer (90 μ l, pH 7.0, 0.1 M). The conjugate solution was treated with O,O'-(propane-1,3-diyl)bis(hydroxylamine) (**11**, 1.67 μ mol) in water (5 μ l). The reaction mixture was vortexed (350 rpm) in incubator-shaker at 25 °C for 2 h. The excess reagent and salts were removed by spin concentrator (30 kDa MWCO). The labeled antibody was lyophilized and re-suspended in phosphate buffer (90 μ l, pH 7.0, 0.1 M). To this solution, doxorubicin (**12**, 0.67 μ mol in 10 μ l water) was added and vortexed for 8 h. Subsequently, the salts and excess doxorubicin (**12**) were removed by using the spin concentrator (30 kDa MWCO) (Note: To ensure the removal of doxorubicin imine adducts from the mAb, hydrazine (3.35 μ M, in 5 μ l water) was added to the reaction mixture and vortexed for 2 h. Subsequently, the excess reagent and salts were removed by the spin concentrator (30 kDa

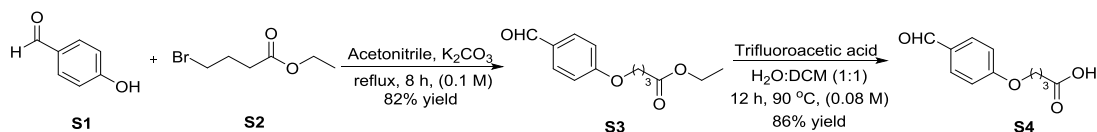
MWCO), and the solvent was removed by lyophilization. The samples were further used for the UV-Vis spectroscopy analysis and MTT assay.

2.6. Procedure for MTT assay (Anti-proliferative activity assay)

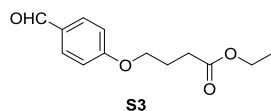
SKBR3 and MDA-MB-231 cells were seeded in 96-well plates (tissue culture grade, flat bottom) at 1×10^4 cells/well in a final volume of 100 μ l of culture medium RPMI and L55 respectively. After seeding for 24 h, cells were treated with various concentrations (0.25-1 μ M) of T-Mab (**1i**), T-Dox (**13**) and Dox (**12**) for next 48 h. All the treatments were given in triplicate. The inhibition of cell proliferation was assessed using the MTT assay from Sigma Aldrich (Sigma Aldrich, Saint Louis, USA). Briefly, 100 μ l of MTT reagent (final concentration 0.5 mg/ml) was added after removing the medium and the plates were incubated at 37 $^{\circ}$ C. After 1 h of incubation, 100 μ l of DMSO was added and absorbance was taken on an ELISA plate reader with a test wavelength of 570 nm and a reference wavelength of 630 nm. Relative growth inhibition rates for the untreated control were calculated and expressed as % inhibition of cell proliferation.

3. Synthesis of reagents

Synthesis of 4-(4-formylphenoxy)butanoate (**S4**)

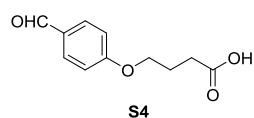


Synthesis of ethyl 4-(4-formylphenoxy)butanoate (**S3**)²



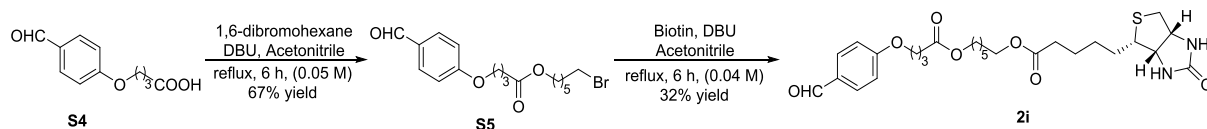
In a 25 ml round bottom flask, *p*-hydroxybenzaldehyde (**S1**, 122.1 mg, 1 mmol) was dissolved in 10 ml acetonitrile. To this aldehyde solution, K_2CO_3 (276.4 mg, 2 mmol) and ethyl 4-bromobutanoate (**S2**, 0.17 ml, 1.2 mmol) were added and reaction mixture was allowed to reflux for 8 h. The reaction progress was monitored by thin layer chromatography (TLC). Upon completion, the reaction mixture was filtered to remove K_2CO_3 . The solution was concentrated under vacuum and the product was purified using flash column chromatography (ethylacetate:*n*-hexane 2:98) to afford ethyl 4-(4-formylphenoxy)butanoate (**S3**) (82% yield). 1H NMR (400 MHz, $CDCl_3$) δ 9.86 (s, 1H), 7.81 (d, $J = 8.6$ Hz, 2H), 6.97 (d, $J = 8.7$ Hz, 2H), 4.14 (q, $J = 7.2$ Hz, 2H), 4.09 (t, $J = 6.2$ Hz, 2H), 2.52 (t, $J = 7.2$ Hz, 2H), 2.14 (m, $J = 6.7$ Hz, 2H), 1.25 (t, $J = 7.1$ Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 190.8, 173.0, 163.9, 132.0, 130.0, 114.8, 67.2, 60.6, 30.6, 24.5, 14.3. MS (ESI) $[MH]^+$ calcd. $C_{13}H_{16}O_4$ 237.1, found 237.1.

Synthesis of 4-(4-formylphenoxy)butanoate (S4)

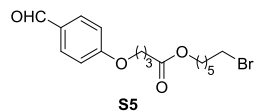


The resulted ester derivative (S3, 194 mg, 0.82 mmol) was dissolved in 10 ml water and DCM mixture (1:1). To this solution, trifluoroacetic acid (4 equiv.) was added and reaction temperature was elevated to 90 °C. The reaction mixture was stirred for another 12 h and the hydrolysis of ester was monitored by TLC. Subsequently, the precipitated crude product was filtered and subjected to silica-gel flash column chromatography (ethylacetate:*n*-hexane 35:65) to afford the pure product S4 (86% yield). ¹H NMR (500 MHz, DMSO-*d*₆) δ 9.86 (s, 1H), 7.86 (d, *J* = 8.7 Hz, 2H), 7.12 (d, *J* = 8.7 Hz, 2H), 4.11 (t, *J* = 6.4 Hz, 2H), 2.39 (t, *J* = 7.3 Hz, 2H), 1.97 (m, 2H). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 191.3, 174.0, 163.5, 131.8, 129.6, 114.9, 67.2, 30.0, 24.1. MS (ESI) [MH]⁺ calcd. C₁₁H₁₃O₄ 209.2, found 209.2.

Synthesis of 6-((4-(4-formylphenoxy)butanoyl)oxy)hexyl 5-((3*a*S,4*S*,6*a*R)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanoate (2i)

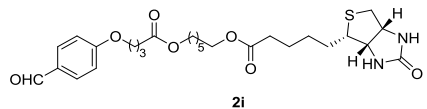


Synthesis of 6-bromohexyl 4-(4-formylphenoxy)butanoate S5



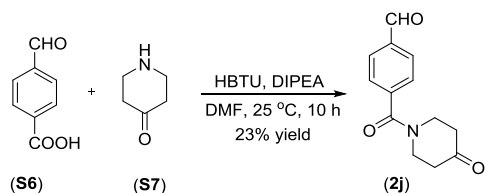
In a 25 ml round bottom flask, ethyl-4-(4-formylphenoxy)butanoate (S4, 105 mg, 0.5 mmol) was dissolved in 10 ml acetonitrile. To this solution, DBU (114 μl, 0.75 mmol) and 1,6-dibromohexane (100 μl, 1 mmol) were added at room temperature and the reaction mixture was allowed to reflux for 6 h. The reaction was monitored by thin layer chromatography (TLC) and upon completion, the solution was concentrated under rotary evaporator. The crude reaction mixture was directly subjected to flash column chromatography (ethylacetate:*n*-hexane 10:90) to afford pure compound 6-bromohexyl 4-(4-formylphenoxy)butanoate S5 (72% yield). ¹H NMR (500 MHz, CDCl₃) δ 9.88 (s, 1H), 7.83 (d, *J* = 8.7 Hz, 2H), 6.99 (d, *J* = 8.7 Hz, 2H), 4.10 (m, *J* = 6.3, 4.0 Hz, 4H), 3.39 (t, *J* = 6.8 Hz, 2H), 2.53 (t, *J* = 7.2 Hz, 2H), 2.19-2.09 (m, 2H), 1.89-1.80 (m, 2H), 1.68-1.58 (m, 2H), 1.50-1.42 (m, 2H), 1.37 (m, *J* = 14.6, 7.3 Hz, 2H). ¹³C NMR (126 MHz, CDCl₃) δ 190.9, 173.1, 163.9, 132.1, 130.1, 114.8, 67.2, 64.6, 33.7, 32.6, 30.7, 28.5, 27.8, 25.2, 24.5. HRMS (ESI) [MNa]⁺ calcd. C₁₇H₂₃BrO₄Na 393.0672, found 393.0684.

Synthesis of 6-((4-(4-formylphenoxy)butanoyl)oxy)hexyl 5-((3a*S*,4*S*,6a*R*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanoate (**2i**)



In a 10 ml round bottom flask, biotin (31 mg, 0.13 mmol) was dissolved in 3 ml acetonitrile. To this solution DBU (27 μ l, 0.19 mmol) and 6-bromohexyl 4-(4-formylphenoxy)butanoate **S5** (100 mg, 0.27 mmol) were added at room temperature and the reaction mixture was allowed to reflux for 6 h. Reaction was monitored by thin layer chromatography (TLC) and upon completion, the solution was concentrated under rotary evaporator. The reaction mixture was re-dissolved in ethyl acetate followed by extraction with NaHCO₃ solution (2 \times 10 ml) and water (2 \times 10 ml). The organic layers were combined and dried with anhydrous Na₂SO₄. The ethyl acetate solution was concentrated under vacuum and the resulted biotinylated product (**2i**) was purified by flash column chromatography (MeOH:DCM 3:97; 63% yield). ¹H NMR (500 MHz, CDCl₃) δ 9.87 (s, 1H), 7.81 (d, *J* = 8.5 Hz, 2H), 6.98 (d, *J* = 8.7 Hz, 2H), 5.81 (bs, 1H), 5.44 (bs, 1H), 4.52-4.46 (m, 1H), 4.34-4.26 (m, 1H), 4.08 (m, 2H), 4.03 (t, *J* = 6.7 Hz, 2H), 3.14 (m, 1H), 2.89 (dd, *J* = 12.8, 4.9 Hz, 1H), 2.72 (d, *J* = 12.8 Hz, 1H), 2.52 (t, *J* = 7.2 Hz, 2H), 2.31 (t, *J* = 7.5 Hz, 2H), 2.18-2.10 (m, 2H), 1.65 (m, 9H), 1.49-1.39 (m, 2H), 1.36 (m, 4H). ¹³C NMR (126 MHz, CDCl₃) δ 190.8, 173.7, 173.0, 163.8, 132.0, 129.9, 114.7, 67.1, 64.5, 64.2, 61.9, 60.1, 55.4, 40.5, 33.9, 30.6, 28.5, 28.4, 28.3, 28.2, 25.5, 24.8, 24.4. HRMS (ESI) [MNa]⁺ calcd. C₂₇H₃₈N₂O₇SNa 557.2292, found 557.2287.

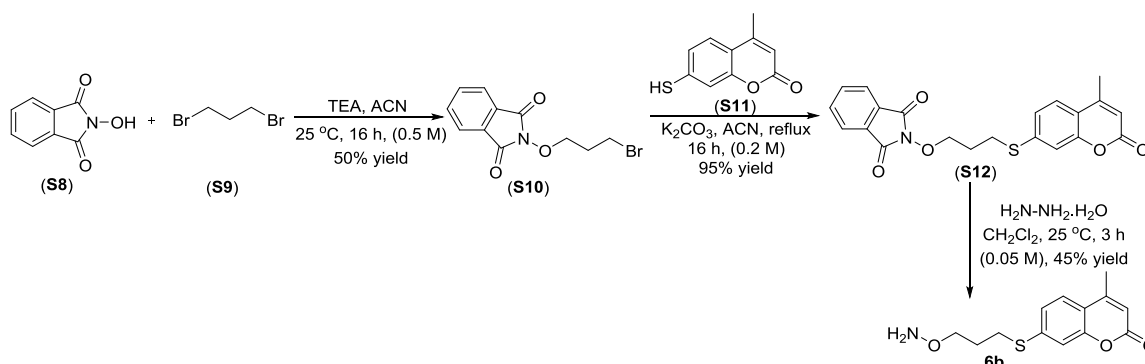
Synthesis of reagent 4-(4-Oxo-piperidine-1-carbonyl)-benzaldehyde (**2j**)³



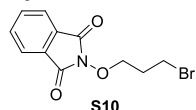
In a 25 mL round bottom flask equipped with a magnetic stirring bar was combined 4-carboxybenzaldehyde (**S6**, 152 mg, 1 mmol), HBTU (574 mg, 1.5 mmol) and dissolved in 5 ml DMF solution. To this solution, DIPEA (254 μ l, 1.5 mmol) was added and stirred for 10 min. To the above reaction mixture, 4-piperidone hydrochloride monohydrate (**S7**, 108 mg, 1 mmol) was added. The reaction was allowed to stir at 25 $^\circ$ C and the reaction progress was monitored by TLC. After completion of the reaction, the DMF was removed under vacuum. The resulted solid was subjected to silica-gel column chromatography. The product **2j** was eluted with 1:1 to 4:1 ethyl acetate/hexanes in 23% yield. ¹H NMR (500 MHz, CDCl₃): δ 10.07 (s, 1H), 7.97 (d, 2H), 7.64 (d, 2H), 4.05 (bs, 2H), 3.70 (bs, 2H), 2.62 (bs, 2H), 2.44 (bs, 2H). ¹³C NMR (126 MHz, CDCl₃): δ 206.1, 191.4, 169.6, 140.8, 137.4, 130.1, 127.6, 41.6-38.9 (-CH₂CH₂- peaks coalesce).³ HRMS (ESI) [MNa]⁺ Calcd. C₁₃H₁₄NO₃Na 254.0788, found 232.0789.

Synthesis of O-hydroxylamine derivatives

Synthesis of 7-((3-(aminooxy)propyl)thio)-4-methyl-2H-chromen-2-one (6b)

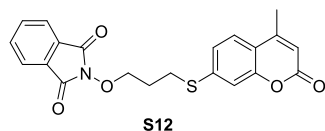


Synthesis of 2-(3-bromopropoxy)isoindoline-1,3-dione (**S10**)⁴



In a 250 ml round bottom flask, N-hydroxyphthalimide **S8** (4894 mg, 30 mmol) and triethyl amine (6.09 ml, 60 mmol) were dissolved in ACN (60 ml). To this solution, 1,3-dibromopropane (**S9**, 8.34 ml, 60 mmol) was added and stirred at 25 °C for 16 h. The reaction mixture was concentrated in vacuo and was added 1 N NaOH solution and ethyl acetate. The organic layer was separated, dried over anhydrous sodium sulfate, filtered, and concentrated in vacuo. The purification of crude mixture was performed by flash column chromatography using ethyl acetate:hexane (3:97) gave **S10** in 50% yield. 1H NMR (400 MHz, $CDCl_3$) δ 7.85-7.81 (m, 2H), 7.75-7.50 (m, 2H), 4.36 (t, $J = 5.8$ Hz, 2H), 3.70 (t, $J = 6.5$ Hz, 2H), 2.27-2.32 (m, $J = 6.2$ Hz, 2H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 163.6, 134.7, 128.9, 123.7, 76.1, 31.6, 29.3 ppm. MS (ESI) $[MH]^+$ calcd. $C_{11}H_{11}Br^{79}NO_3$ 283.9, found 283.9 and calcd. $C_{11}H_{11}Br^{81}NO_3$ 285.9, found 285.9.

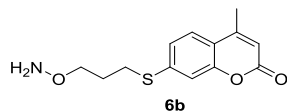
Synthesis of 2-(3-((4-methyl-2-oxo-2H-chromen-7-yl)thio)propoxy)isoindoline-1,3-dione (**S12**)



In a 25 ml round bottom flask, 7-mercapto-4-methylcoumarin **S11** (192 mg, 1 mmol), K_2CO_3 (276 mg, 2 mmol) and 2-(3-bromopropoxy)isoindoline-1,3-dione **S10** (568 mg, 2 mmol) were dissolved in degassed acetonitrile (5 ml) and refluxed for 16 h. The reaction mixture was concentrated in vacuo and purified by silica gel flash column chromatography using ethyl acetate:hexane (7:3) to give **S12** in 95% yield. 1H NMR (400 MHz, $CDCl_3$) δ 7.85-7.7 (m, 2H), 7.81-7.73 (m, 2H), 7.48 (d, $J = 8.2$ Hz, 1H), 7.26-7.20 (m, 2H), 6.21 (d, $J = 0.8$ Hz, 1H), 4.36 (t, $J = 5.8$ Hz, 2H), 3.35 (t, $J = 7.1$ Hz, 2H), 2.40 (d, $J = 0.9$ Hz, 3H), 2.17-2.11 (m, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 163.7, 160.7, 154.0, 152.2, 142.6, 134.7, 128.9, 124.8, 123.3, 123.3, 117.4,

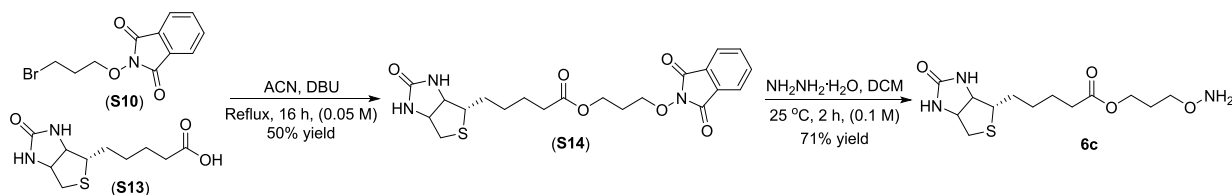
114.7, 114.1, 76.7, 28.6, 27.8, 18.6 ppm. HRMS (ESI) $[MH]^+$ calcd. $C_{21}H_{18}NO_5S$ 396.0906, found 396.0925.

Synthesis of 7-((3-(aminooxy)propyl)thio)-4-methyl-2H-chromen-2-one (6b)

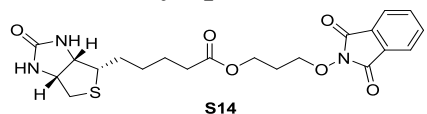


In a 5 ml round bottom flask, 2-(3-((4-methyl-2-oxo-2H-chromen-7-yl)thio)propoxy)isoindoline-1,3-dione **S12** (237 mg, 0.6 mmol) was dissolved in CH_2Cl_2 (12 ml). To this solution, hydrazine monohydrate (80%, 29 μ l, 0.6 mmol) was added and stirred at 25 °C for 3 h. The reaction mixture was filtered and the filtrate was concentrated. The purification of crude mixture was performed by reverse phase preparative HPLC to isolate **6b** (76 mg, 45% yield). 1H NMR (400 MHz, $CDCl_3$) δ 7.44 (d, $J = 8.3$ Hz, 1H), 7.20-7.11 (m, 2H), 6.18 (d, $J = 0.9$ Hz, 1H), 3.77 (t, $J = 5.9$ Hz, 2H), 3.05 (t, $J = 7.3$ Hz, 2H), 2.38 (d, $J = 0.8$ Hz, 3H), 2.02-1.90 (m, 2H) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 160.7, 153.9, 152.2, 143.3, 124.6, 123.0, 117.1, 114.0, 113.8, 73.8, 28.9, 27.8, 18.6 ppm. HRMS (ESI) $[MH]^+$ calcd. $C_{13}H_{16}NO_3S$ 266.0851, found 266.0841.

Synthesis of 3-(aminooxy)propyl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate (6c)



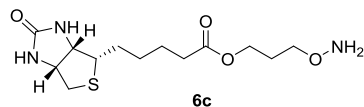
Synthesis of 3-((1,3-dioxoisindolin-2-yl)oxy)propyl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl) pentanoate (S14)



In a 5 ml round bottom flask, biotin **S13** (244 mg, 1 mmol), 2-(3-bromopropoxy)isoindoline-1,3-dione **S10** (568 mg, 2 mmol), and DBU (304 μ l, 2 mmol) were dissolved in acetonitrile (20 ml) to reflux. The progress of the reaction was analyzed by TLC. After 16 h, the reaction mixture was concentrated under vacuum, re-dissolved in ethyl acetate, and subjected to aqueous work up. The combined organic fractions were dried on anhydrous sodium sulfate, filtered, and concentrated under vacuum. The purification of crude reaction mixture was performed by flash column chromatography (MeOH/DCM, 0.5-5%) to result 3-((1,3-dioxoisindolin-2-yl)oxy)propyl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate **S14** (224 mg, 50% yield). 1H NMR (500 MHz, $CDCl_3$) δ 7.87-7.80 (m, 2H), 7.78-7.71 (m, 2H), 5.95 (s, 1H), 5.46 (s, 1H), 4.48 (dd, $J = 15.0, 9.8$ Hz, 1H), 4.35-4.26 (m, $J = 14.0, 6.2$ Hz, 5H), 3.20-3.11 (m, 1H),

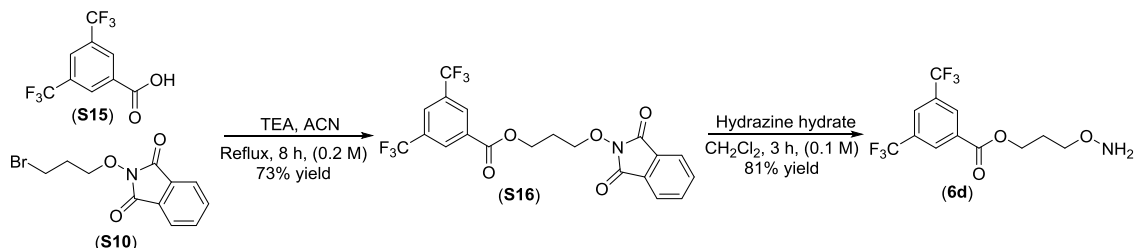
2.89 (dd, $J = 12.8, 5.0$ Hz, 1H), 2.72 (d, $J = 12.8$ Hz, 1H), 2.34 (t, $J = 7.5$ Hz, 2H), 2.12-2.05 (m, 2H), 1.80-1.60 (m, 4H), 1.50-1.40 (m, 2H) ppm. ^{13}C NMR (126 MHz, CDCl_3) δ 173.6, 163.7, 163.7, 134.6, 128.9, 123.7, 75.1, 62.0, 60.7, 60.2, 55.5, 40.6, 33.9, 28.4, 28.3, 27.7, 24.8 ppm. HRMS (ESI) $[\text{MH}]^+$ calcd. $\text{C}_{21}\text{H}_{26}\text{N}_3\text{O}_6\text{S}$ 448.1542, found 448.1548.

Synthesis of 3-(aminooxy)propyl-5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate (**6c**)

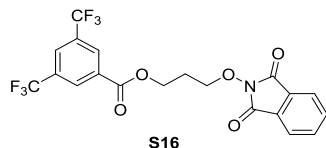


Synthetic protocol of molecule (**6c**) is same as synthetic protocol of **6b**. ^1H NMR (500 MHz, D_2O) δ 4.63 (dd, $J = 7.9, 4.9$ Hz, 1H), 4.45 (dd, $J = 7.9, 4.5$ Hz, 1H), 4.21 (t, $J = 6.3$ Hz, 2H), 3.90 (t, $J = 6.2$ Hz, 2H), 3.42-3.31 (m, 1H), 3.02 (dd, $J = 13.1, 5.0$ Hz, 1H), 2.80 (d, $J = 13.0$ Hz, 1H), 2.44 (t, $J = 7.3$ Hz, 2H), 2.08-1.94 (m, 2H), 1.84-1.55 (m, 4H), 1.53-1.37 (m, 2H) ppm. ^{13}C NMR (126 MHz, D_2O) δ 176.9, 165.3, 72.4, 62.1, 62.0, 60.3, 55.3, 39.7, 33.7, 27.9, 27.6, 26.7, 24.1 ppm. HRMS (ESI) $[\text{MH}]^+$ calcd. $\text{C}_{13}\text{H}_{24}\text{N}_3\text{O}_4\text{S}$ 318.1488, found 318.1467.

Synthesis of 3-(aminooxy)propyl 3,5-bis(trifluoromethyl)benzoate (**6d**)

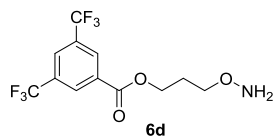


Synthesis of 3-((1,3-dioxoisindolin-2-yl)oxy)propyl 3,5-bis(trifluoromethyl)benzoate (**S16**)



In a 25 ml round bottom flask, 3,5-bis(trifluoromethyl)benzoic acid **S15** (258 mg, 1 mmol), 2-(3-bromopropoxy)isoindoline-1,3-dione **S10** (312 mg, 1.1 mmol) and TEA (418 μl , 3 mmol) were dissolved in acetonitrile (5 ml) to reflux. Progress of the reaction was followed by TLC. After 8 h, reaction mixture was concentrated and purification of crude by flash column chromatography (ethyl acetate:hexane 2:98) gave 3-((1,3-dioxoisindolin-2-yl)oxy)propyl 3,5-bis(trifluoromethyl)benzoate **S16** (335 mg, 73% yield). ^1H NMR (400 MHz, CDCl_3) δ 8.51 (s, 2H), 8.05 (s, 1H), 7.90-7.70 (m, 4H), 4.70 (t, $J = 6.3$ Hz, 2H), 4.39 (t, $J = 6.0$ Hz, 2H), 2.29 (m, $J = 6.1$ Hz, 2H) ppm. ^{13}C NMR (101 MHz, CDCl_3) δ 164.0, 163.7, 134.7, 132.5, 132.3 (q, $J = 34.1$ Hz, 2C), 130.1-129.8 (m, 2C), 129.0, 126.6-126.4 (m, 1C), 124.4, 123.7, 123.0 (q, $J = 272.8$ Hz, 2C), 74.9, 62.7, 27.8 ppm. ^{19}F NMR (376 MHz, CDCl_3) δ -62.94 ppm [Trifluoro acetic acid (TFA) was used as an internal standard, -75.70 ppm]. HRMS (ESI) $[\text{MH}]^+$ calcd. $\text{C}_{20}\text{H}_{14}\text{F}_6\text{NO}_5$ 462.0776, found 462.0775.

Synthesis of 3-(aminooxy)propyl 3,5-bis(trifluoromethyl)benzoate (**6d**)

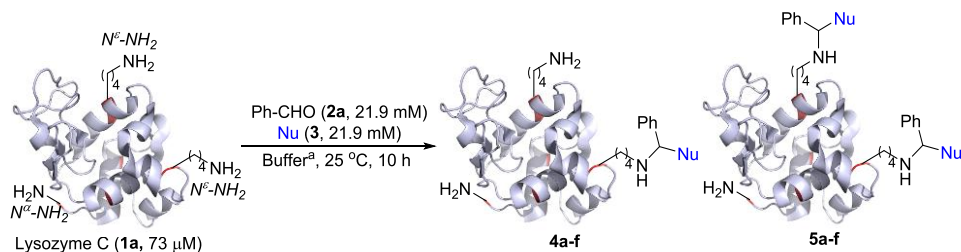


In a 5 ml round bottom flask with 3-((1,3-dioxoisindolin-2-yl)oxy)propyl 3,5-bis(trifluoromethyl)benzoate **S16** (138 mg, 0.3 mmol) in DCM (3 ml) was added hydrazine monohydrate (80%, 37 μ l, 0.75 mmol) and stirred at room temperature. The progress of the reaction was followed by TLC. After 3 h, the reaction mixture was filtered. The filtrate was concentrated in vacuo to result 3-(aminooxy)propyl 3,5-bis(trifluoromethyl)benzoate **6d** (80 mg, 81% yield). ^1H NMR (400 MHz, CDCl_3) δ 8.48 (s, 2H), 8.07 (s, 1H), 5.42 (bs, 2H), 4.50 (t, $J = 6.5$ Hz, 2H), 3.83 (t, $J = 6.1$ Hz, 2H), 2.11 (m, $J = 6.3$ Hz, 2H) ppm. ^{13}C NMR (101 MHz, CDCl_3) δ 164.1, 132.9, 132.6 (q, $J = 33.9$ Hz, 2C), 130.0-129.7 (m, 2C), 126.6-126.3 (m, 1C), 123.02 (q, $J = 273.0$ Hz, 2C), 72.2, 63.6, 27.9 ppm. ^{19}F NMR (376 MHz, CDCl_3) δ -62.54 ppm [TFA was used as an internal standard, -75.70 ppm]. HRMS (ESI) $[\text{MH}]^+$ calcd. $\text{C}_{12}\text{H}_{12}\text{F}_6\text{NO}_3$ 332.0731, found 332.0699.

4. Additional results and discussions

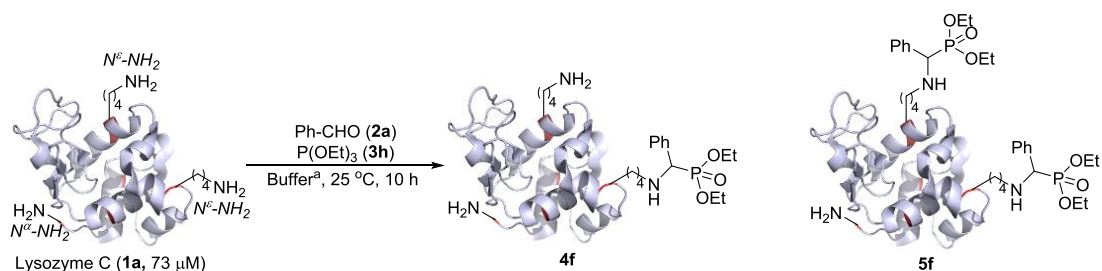
4.1. Optimization of reaction conditions

Table S1. Screening the reactivity of nucleophiles with protein derived imines



Entry	Nu	Observation ^b
1	PhCCH, CuI	- ^c
2	NaBH ₃ CN	Heterogeneous mixture
3	NCCH ₂ CN	<5% (4c), <5% (5c)
4	^t Bu-NC	14% (4d, and 5d)
5	H-P(O)(OCH ₂ CF ₃) ₂	- ^c
6	H-P(O)(O ⁱ Pr) ₂	- ^c
7	H-P(O)(O ^t Bu) ₂	- ^c
8	H-P(O)(OPh) ₂	- ^c
9	P(OPh) ₃	- ^c
10	P(OMe) ₃	40% (4e), 14% (5e)
11	H-P(O)(OMe) ₂	41% (4e), 14% (5e)
12	H-P(O)(OEt) ₂	41% (4f)
13	P(OEt) ₃	42% (4f)

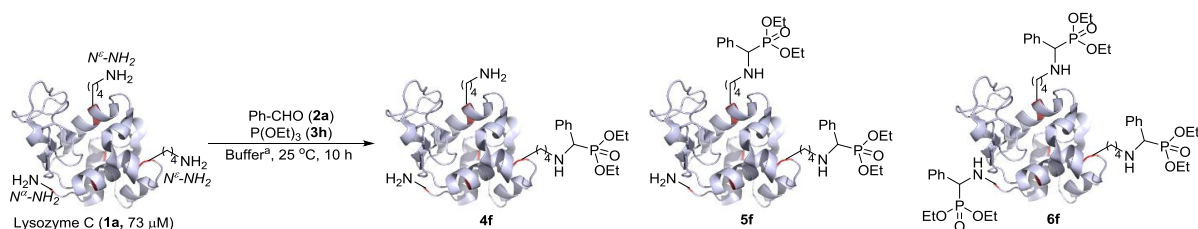
^a Phosphate buffer (pH = 7.8, 0.1 M)/DMSO 9:1. ^b % conversion are based on MALDI-ToF MS data. ^c No reaction was observed.

Table S2. Optimization of stoichiometry of the aldehyde (**2a**) and triethylphosphite (**3h**)

Entry	2a (equiv.)	3h (equiv.)	4f (%conversion) ^b
1	10	300	0
2	25	300	<5
3	50	300	16
4	100	300	22
5	100	100	<5
6	100	200	<5
7	200	300	18
8	300	300	42
9	300	500	30

^a Phosphate buffer (pH = 7.8, 0.1 M)/DMSO 9:1. ^b % conversions are based on ESI-MS data.

We noticed that the stoichiometry of the reagents plays a crucial role in the conversion and selectivity of protein labeling (Table S2). The screening of reaction with the varying stoichiometry of reagents **2a** and **3h** (entries 1-7, Table S2) resulted in low conversions (0-22% conversion). We hypothesized that the protein labeling might be adversely affected due to a competing reaction between aldehyde and triethylphosphite. The conversions improved considerably (42% conversion) with 300 equivalents of reagents **2a** and **3h** (entry 8). Further increase in the stoichiometry of **3h** led to reduced conversions.

Table S3. Effect of pH and buffer salts on the reaction

Entry	Buffer	Conversion ^b
1	Phosphate buffer (pH 7.8, 0.1 M)	4f (42%)
2	Phosphate buffer (pH 7.0, 0.1 M)	4f (39%)
3	Phosphate buffer (pH 6.0, 0.1 M)	4f (0%)
4	Sodium bicarbonate buffer (pH 7.0, 0.1 M)	Complex spectra
5	Tris buffer (pH 7.8, 0.1 M)	4f (44%), 5f (22%), 6f (5%)
6	HEPES buffer (pH 7.8, 0.1 M)	4f (42%), 5f (39%), 6f (6%)
7	PBS buffer (pH 9.0, 0.1 M)	4f (46%), 5f (16%)

^a Buffer (0.1 M)/DMSO 9:1. ^b % Conversion based on ESI-MS and MALDI-ToF-MS data.

Next, we investigated the impact of buffer on the reaction (Table S3). The transformation in the phosphate buffer with pH range 7.8-7.0 renders the mono-labeled product **4f** in good conversion (entries 1 and 2). On the other hand, the protein remains unaltered at pH 6.0 owing to the decreased nucleophilicity of triethylphosphite (entry 3). We also noted an interesting effect of the buffer salt on the reaction (entries 4-6). In particular, high conversions were observed in both tris and HEPES buffer but at the cost of selectivity. We noted a similar trend at pH 9.0 in PBS buffer (entry 7, Table S3). We selected the phosphate buffer (pH 7.8, 0.1 M) conditions for further experiments as it delivers good conversions while retaining excellent site-selectivity.

4.2. Stability of triethylphosphite in aqueous condition

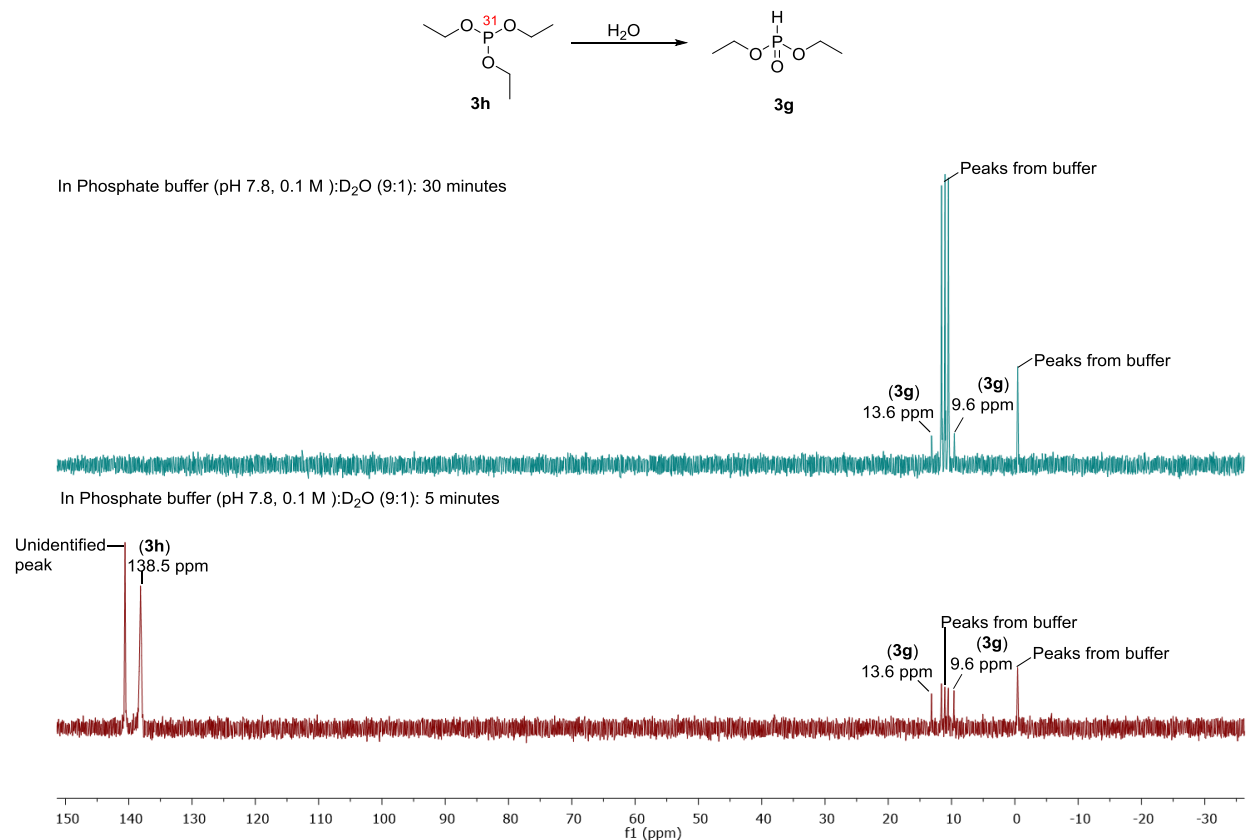
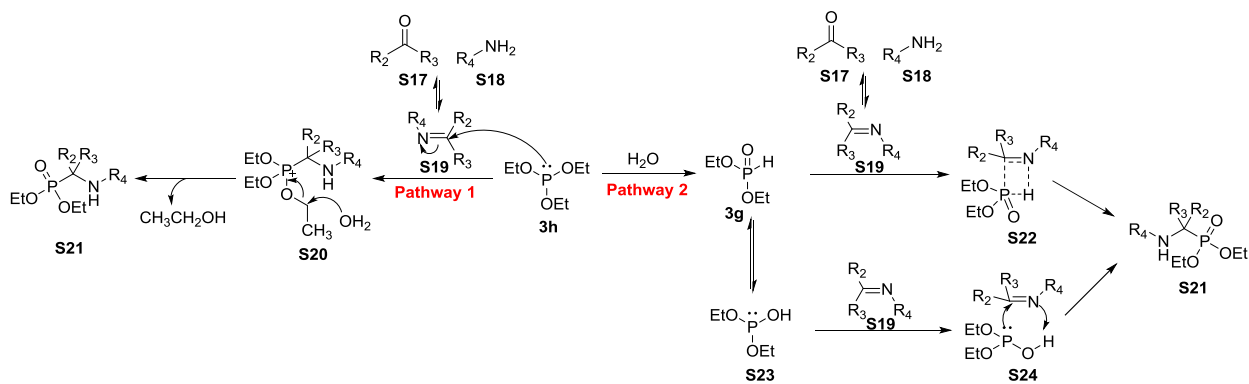


Figure S1. ³¹P-NMR analysis of the triethylphosphite.

³¹P NMR of triethylphosphite (**3h**, δ 137.5 ppm) confirms that it is stable in CDCl₃ for 12 h. However, it converts rapidly into diethylphosphite (**3g**) within 30 minutes in phosphate buffer (pH 7.8, 0.1 M). It indicates that the active nucleophilic species in the reaction is diethylphosphite (**3g**).

4.3. Plausible mechanism



Initially, we anticipated that the phospho-Mannich reaction would involve nucleophilic addition of triethylphosphite (**3h**) to the imine **S19** (pathway 1).⁵ However, the ^{31}P NMR investigations suggest that **3h** converts completely to diethylphosphite (**3g**) in the aqueous buffer. It suggests that the pathway 2 is operative under the given conditions where **3g** reacts with imine and renders the product **S21** through **S22**.⁶ The participation of phosphite (**S23**) is ruled out by ^{31}P NMR.

5. Effect of labeling on the structure of protein

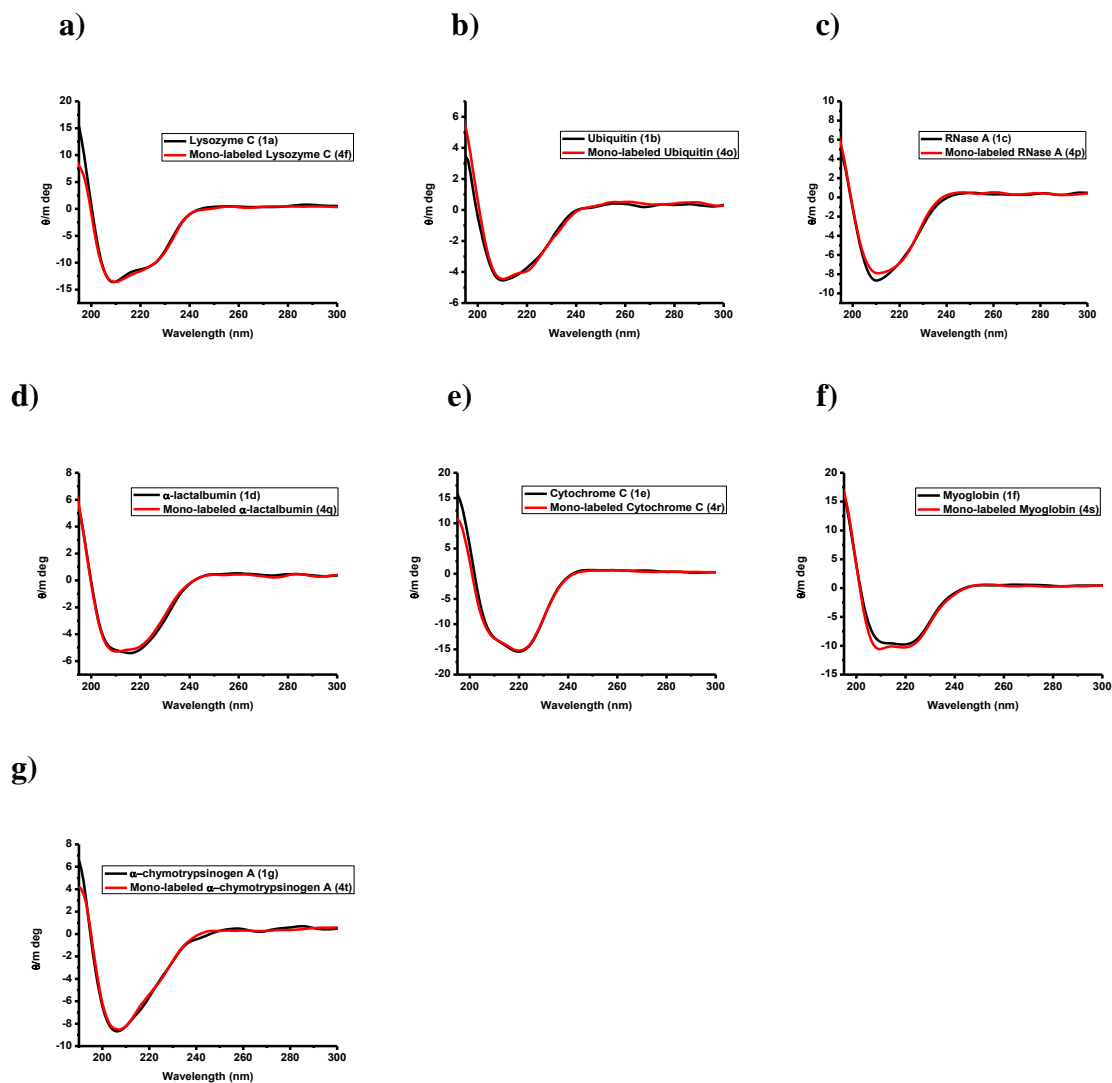


Figure S2: Effect of the reaction conditions on the secondary structure of protein in phosphate buffer (pH 7.8, 0.1 M). (a) Circular Dichroism (CD) spectra of Lysozyme C (**1a**, black line), mono-labeled Lysozyme C (**4f**, red line) at concentration 5 μ M. (b) Circular Dichroism (CD) spectra of Ubiquitin (**1b**, black line), mono-labeled Ubiquitin (**4o**, red line) at concentration 3 μ M. (c) Circular Dichroism (CD) spectra of RNase A (**1c**, black line), mono-labeled RNase A (**4p**, red line) at concentration 5 μ M. (d) Circular Dichroism (CD) spectra of α -lactalbumin (**1d**, black line), mono-labeled α -lactalbumin (**4q**, red line) at concentration 5 μ M. (e) Circular Dichroism (CD) spectra of Cytochrome C (**1e**, black line), mono-labeled Cytochrome C (**4r**, red line) at concentration 5 μ M. (f) Circular Dichroism (CD) spectra of Myoglobin (**1f**, black line), mono-labeled Myoglobin (**4s**, red line) at concentration 5 μ M. (g) Circular Dichroism (CD) spectra of α -chymotrypsinogen A (**1g**, black line), mono-labeled α -chymotrypsinogen A (**4t**, red line) at concentration 5 μ M.

6. Enzymatic assay

6.1. Enzymatic assay of Lysozyme C⁷

Lysozyme C activity before and after the labeling was checked by *Micrococcus lysodeikticus* (ML) lysis at 450 nm (A_{450}) using quartz cuvette (path length, 1 cm at 25 °C). Potassium phosphate buffer (pH 6.2, 0.1 M) was prepared by dissolving potassium phosphate, monobasic in the Millipore Grade I water. The pH was adjusted to 6.2 at 25 °C using 1 M potassium hydroxide (KOH) solution. *Micrococcus lysodeikticus* cell wall suspension [0.01% (w/v)] was prepared in phosphate buffer. The change in absorbance at 450 nm of this suspension versus a buffer blank was in agreement with the literature (0.6-0.7) after adjustment using the appropriate amount of buffer. Freshly prepared Lysozyme C and labeled Lysozyme C solutions (10 µg/1 ml, in phosphate buffer) were used for the assay. 100 µl of Lysozyme C solution and *Micrococcus lysodeikticus* suspension (1 ml) were mixed by inversion. The sample was immediately used for recording the absorbance at 450 nm. The protocol was repeated for measuring the absorption of *Micrococcus lysodeikticus* and labeled Lysozyme C mixture. The enzymatic activity of Lysozyme C remains unperturbed after the chemical modification.

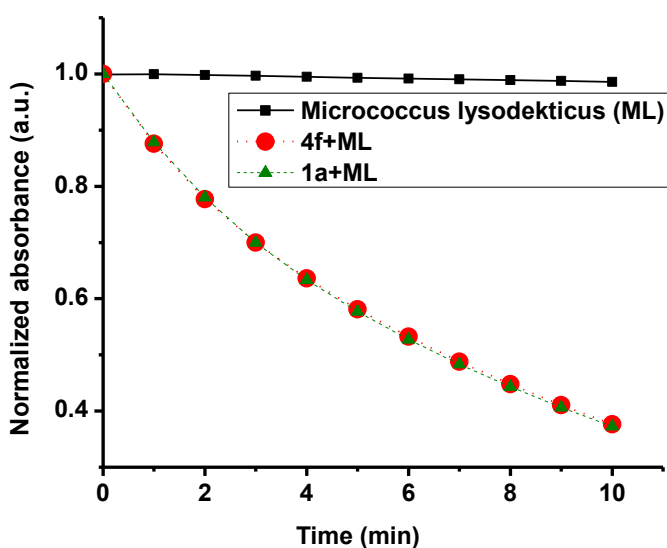


Figure S3: Normalized UV-Vis spectra. Comparison of enzymatic activity between native Lysozyme C (**1a**) and labeled Lysozyme C (**4f**). The enzymatic activity of Lysozyme C remains unperturbed after the chemical modification. Absorbance was normalized to its values by dividing its absorbance maxima (OriginPro 8).

6.2. Enzymatic assay of RNase A⁸

RNase A activity before and after the labeling was checked by Ribonucleic acid (RNA, **16**) hydrolysis at 300 nm (A_{300}) using quartz cuvette (path length, 1 cm at 25 °C). Sodium acetate buffer (pH 5.0, 0.1 M) was prepared by using the Millipore Grade I water (pH was adjusted with 2 M acetic acid). Ribonucleic acid [RNA, 0.1% (w/v), 1 mg/ml] solution was prepared in sodium acetate buffer. Change in absorption at 300 nm of the RNA was performed using RNA solution and blank. Freshly prepared RNase A and labeled RNase A solutions (10 $\mu\text{g}/1\text{ ml}$, in Millipore Grade I water) were used for the assay. RNA (500 μl) and RNase A (500 μl) solutions were mixed by inversion and immediately recorded at 300 nm for 1 h. Similarly, absorption of a mixed solution of RNA and labeled RNase A were performed. The enzymatic activity of RNase A remains unperturbed after the chemical modification.

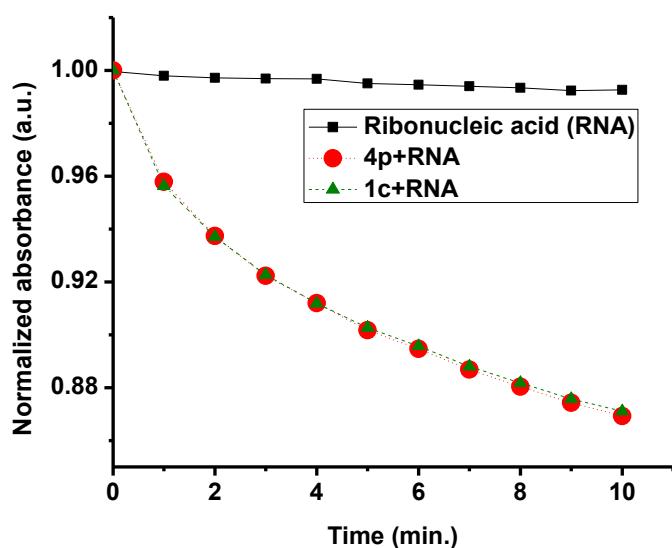


Figure S4. Normalized UV-Vis spectra. Comparison of enzymatic activity between native RNase A (**1c**) and labeled RNase A (**4p**). The enzymatic activity of RNase A (**1c**) remains unperturbed after the chemical modification. Absorbance was normalized to its values by dividing its absorbance maxima (OriginPro 8).

7. MS data of proteins

7.1. Screening of nucleophiles

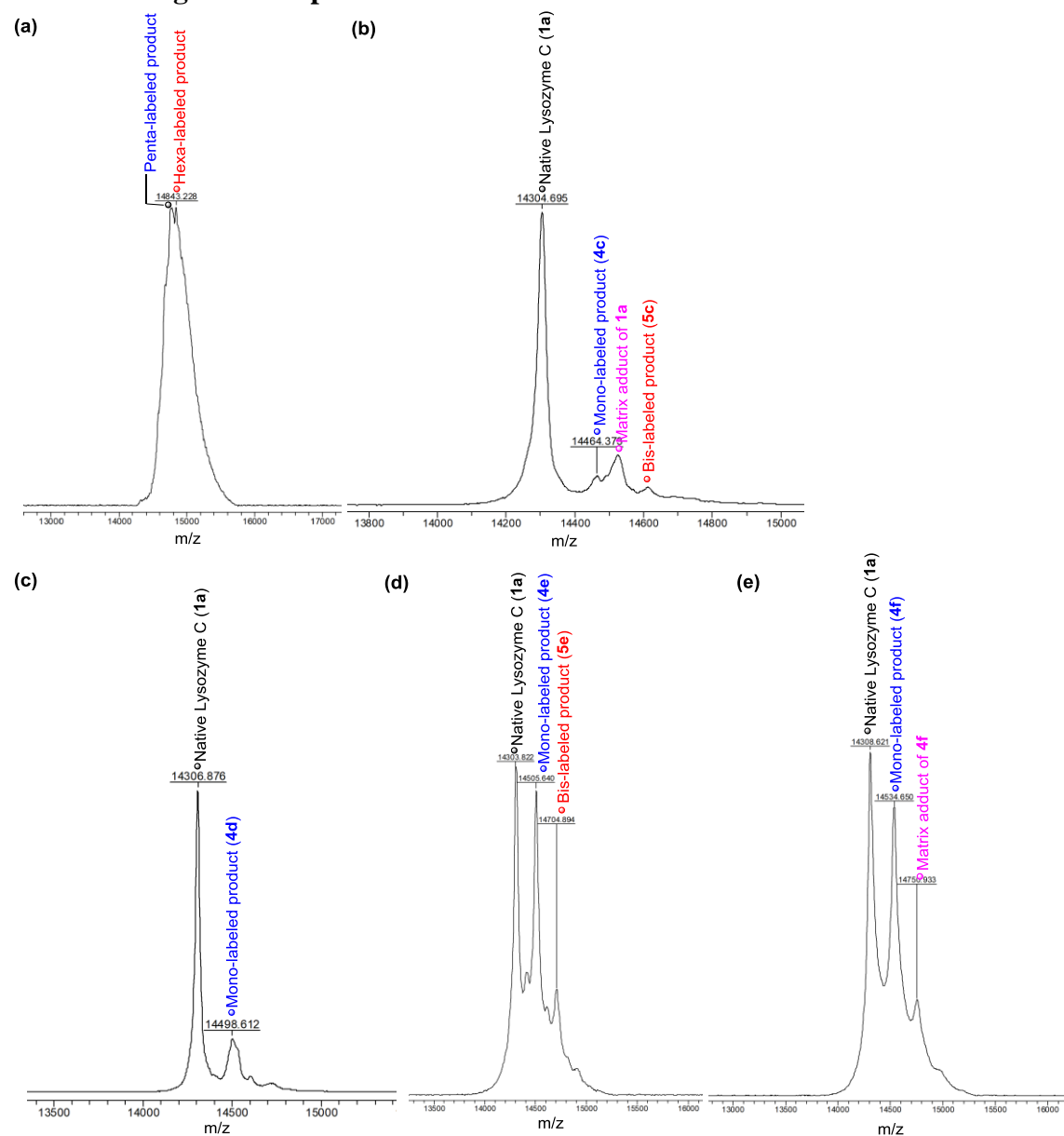


Figure S5. MALDI-ToF MS spectra of various nucleophiles with Lysozyme C (**1a**) and benzaldehyde (**2a**). (a) Reaction with NaBH₃CN (**3a**). (b) Reaction with malononitrile (**3c**). (c) Reaction with ^tBuNC (**3d**). (d) Reaction with trimethylphosphite (**3f**). (e) Reaction with triethylphosphite (**3h**).

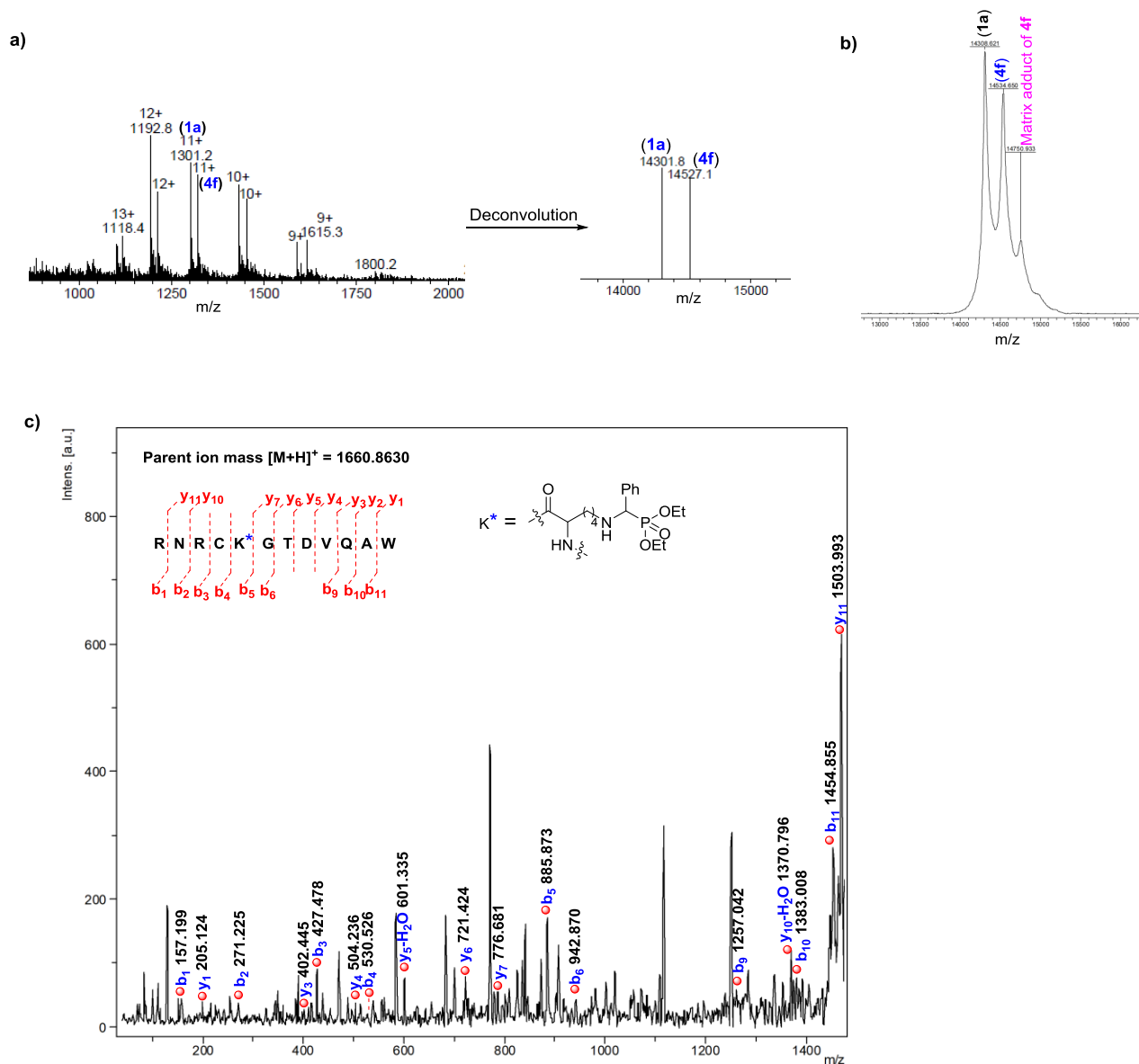


Figure S6. Site-selective native Lysozyme C modification. (a) ESI-MS spectra for Lysozyme C **1a** (1 equiv.) and mono-labeled Lysozyme C **4f**. (b) MALDI-ToF MS spectra of Lysozyme C (**1a**) and mono-labeled Lysozyme C **4f**. (c) MALDI-ToF MS-MS spectrum of labeled RNRCKGTDVQAW (R112-M123). Site of modification is K116 in mono-labeled Lysozyme C **4f**.

7.2. Scope of aldehydes

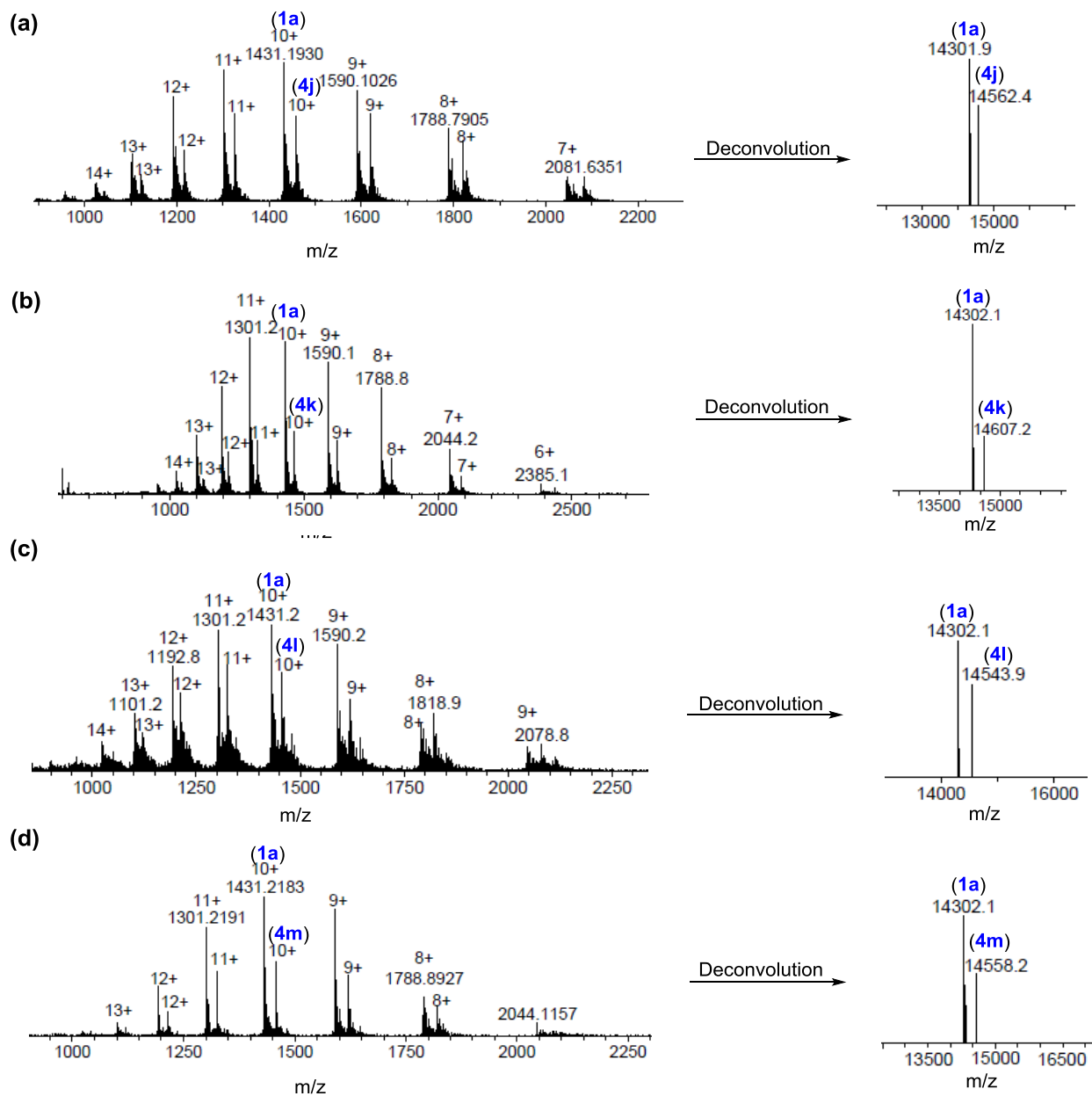


Figure S7. Deconvoluted ESI-MS spectra of the labeled proteins with various aldehydes. (a) Labeling of Lysozyme C (**1a**) with 4-chlorobenzaldehyde (**2e**). (b) Labeling of Lysozyme C (**1a**) with 4-bromobenzaldehyde (**2f**). (c) Labeling of Lysozyme C (**1a**) with 4-hydroxybenzaldehyde (**2g**). (d) Labeling of Lysozyme C (**1a**) with 4-methoxybenzaldehyde (**2h**).

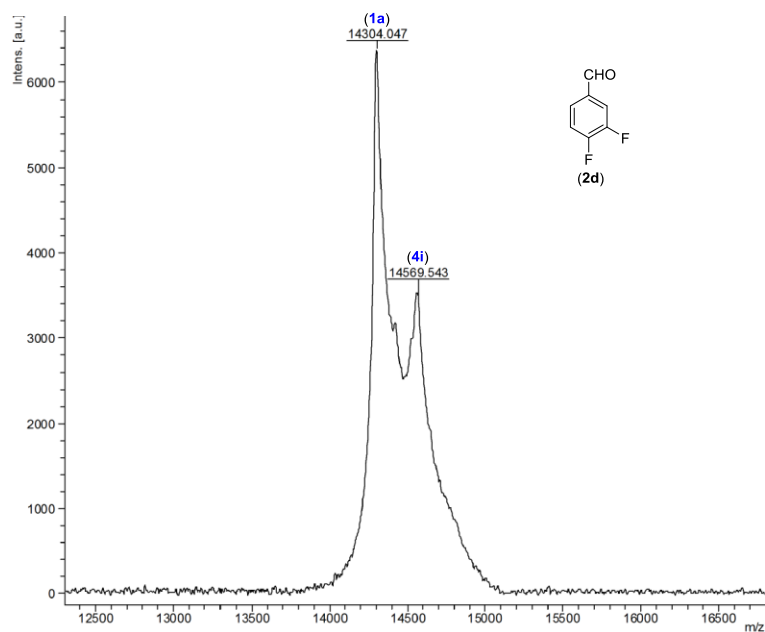


Figure S8. MALDI-ToF MS spectra for site-selective modification of native Lysozyme C (**1a**) with 3,4-difluorobenzaldehyde (**2d**).

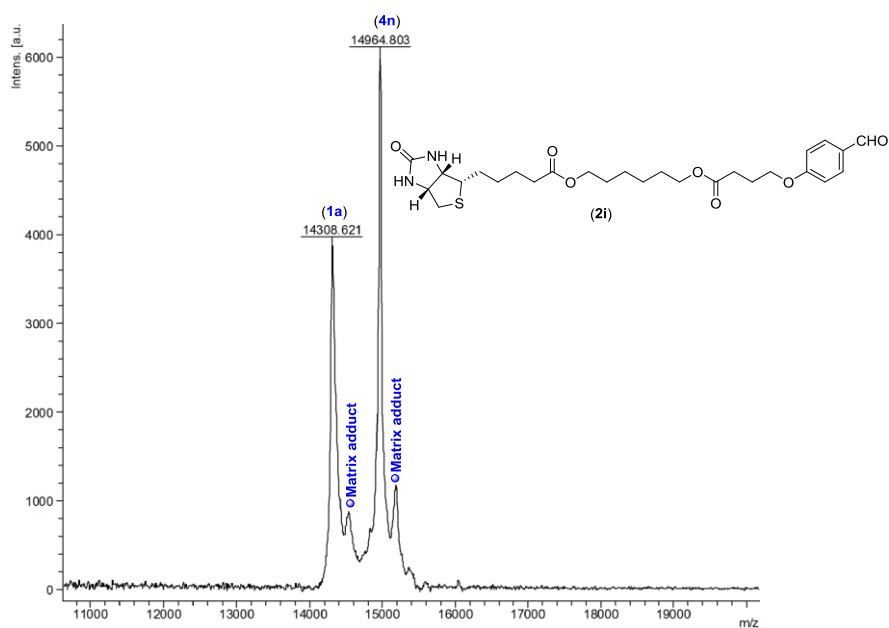


Figure S9. MALDI-ToF MS spectra for site-selective modification of native Lysozyme C (**1a**) with biotinylated aldehyde (**2i**).

7.3. Scope of protein

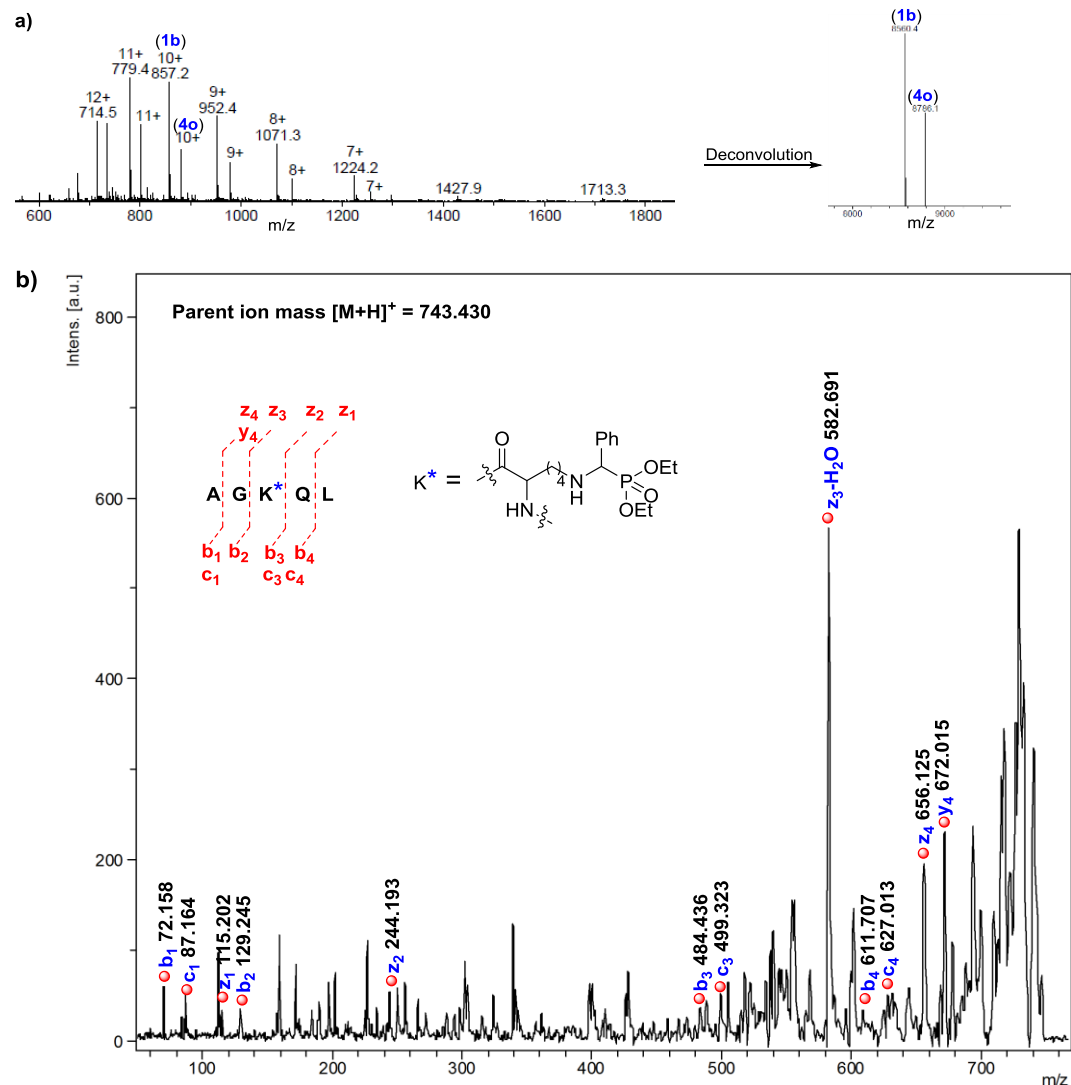


Figure S10. Site-selective native Ubiquitin modification. (a) ESI-MS spectra for Ubiquitin **1b** (1 equiv.) and mono-labeled Ubiquitin **4o**. (b) MALDI-ToF MS-MS spectrum of labeled AGKQL (A46-L50). Site of modification is K48 in mono-labeled Ubiquitin **4o**.

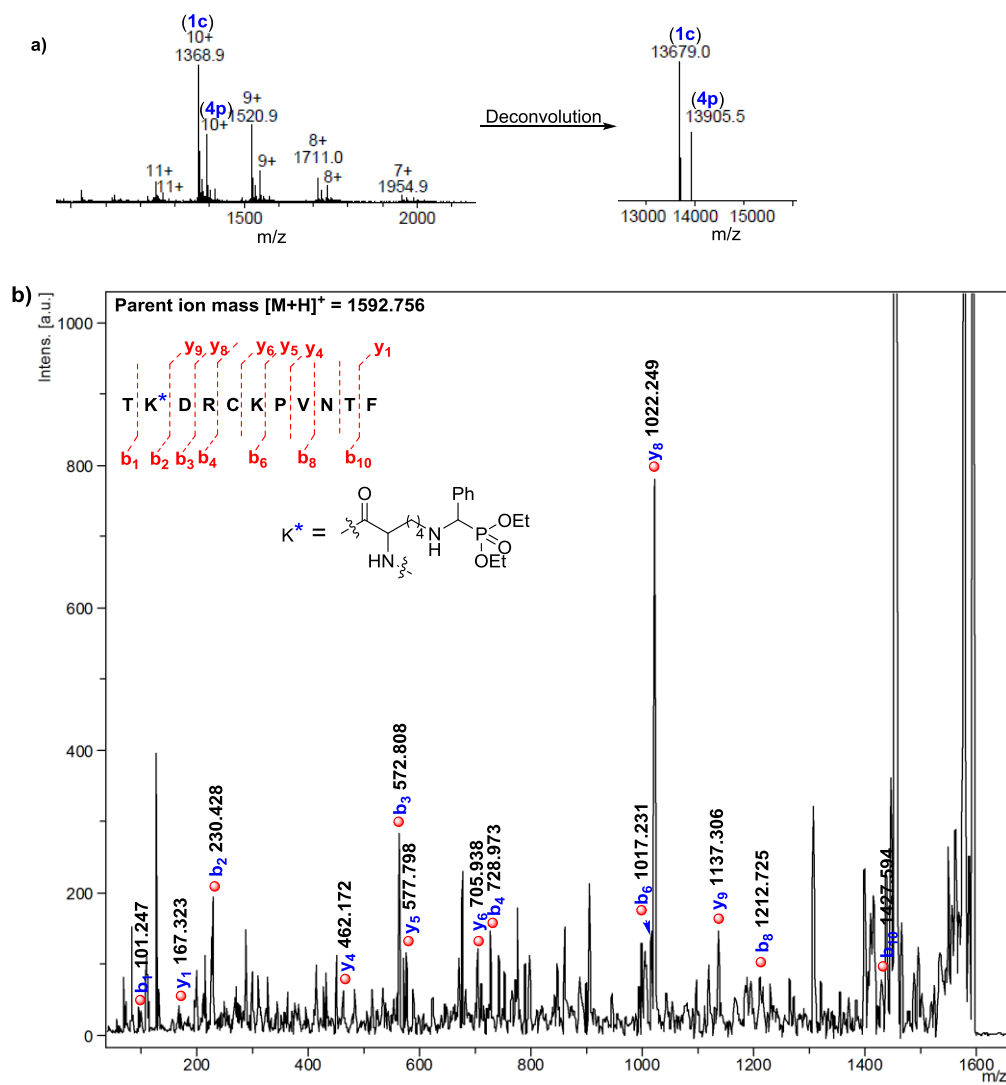


Figure S11. Site-selective native RNase A modification. (a) ESI-MS spectra for RNase A **1c** (1 equiv.) and mono-labeled RNase A **4p**. (b) MALDI-ToF MS-MS spectrum of labeled TKDRCKPVNTF (T36-F46). Site of modification is K37 in mono-labeled RNase A **4p**.

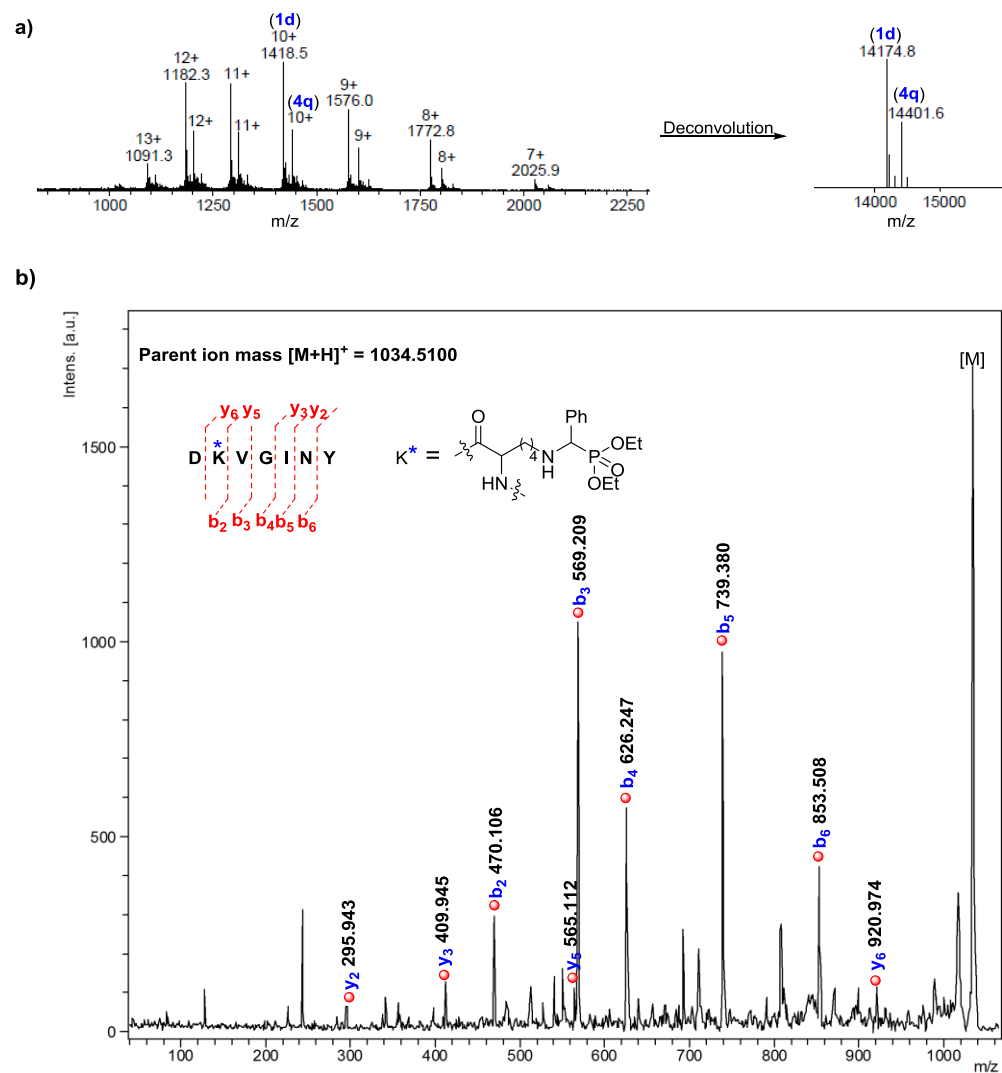


Figure S12. Site-selective native α -lactalbumin modification. (a) ESI-MS spectra for α -lactalbumin **1d** (1 equiv.) and mono-labeled α -lactalbumin **4q**. (b) MALDI-ToF MS-MS spectrum of labeled DKVGINY (D97-Y103). Site of modification is K98 in mono-labeled α -lactalbumin **4q**.

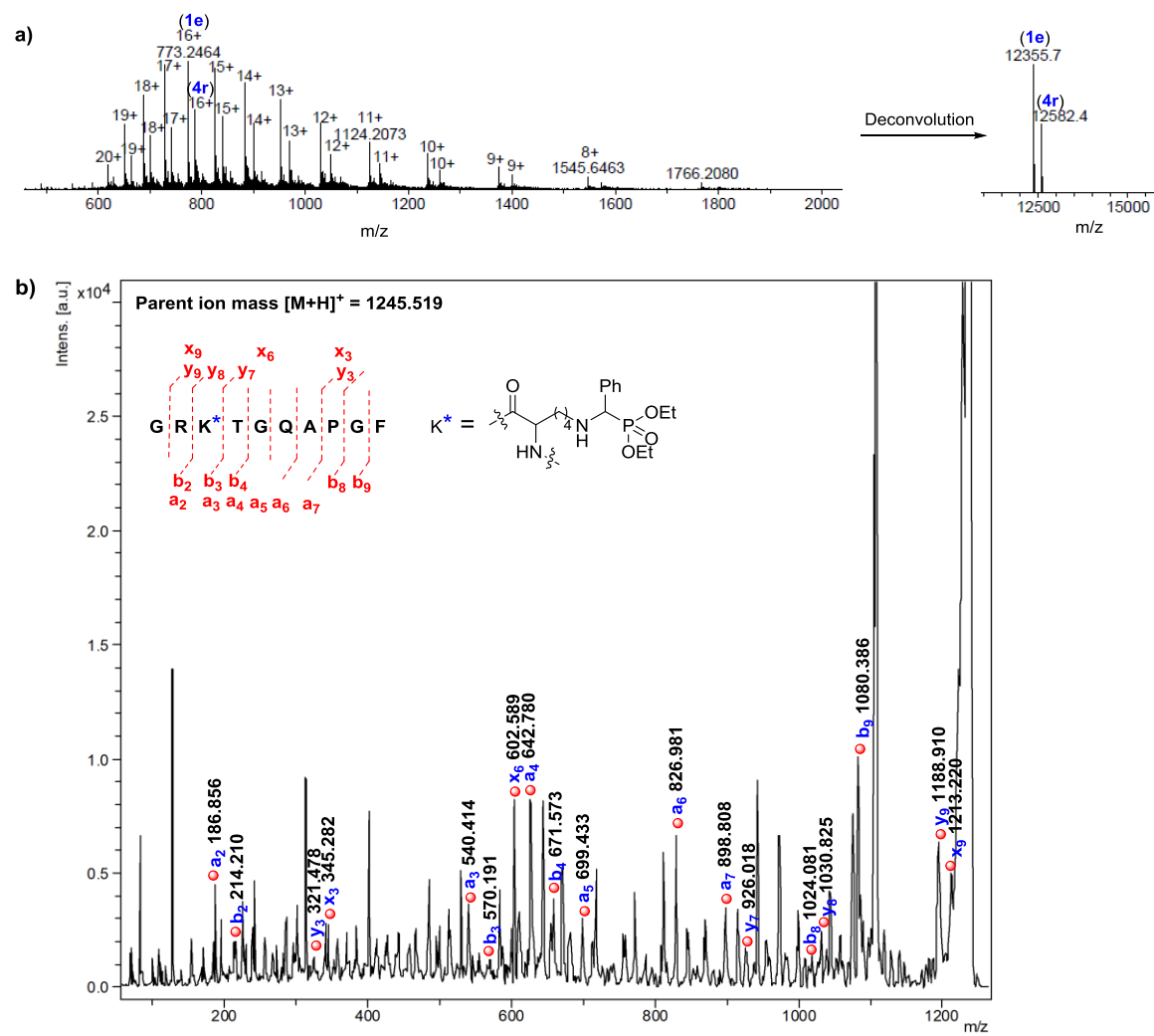


Figure S13. Site-selective native Cytochrome C modification. (a) ESI-MS spectra for Cytochrome C **1e** (1 equiv.) and mono-labeled Cytochrome C **4r**. (b) MALDI-ToF MS-MS spectrum of labeled GRK*⁺TGQAPGF (G37-F46). Site of modification is K39 in mono-labeled Cytochrome C **4r**.

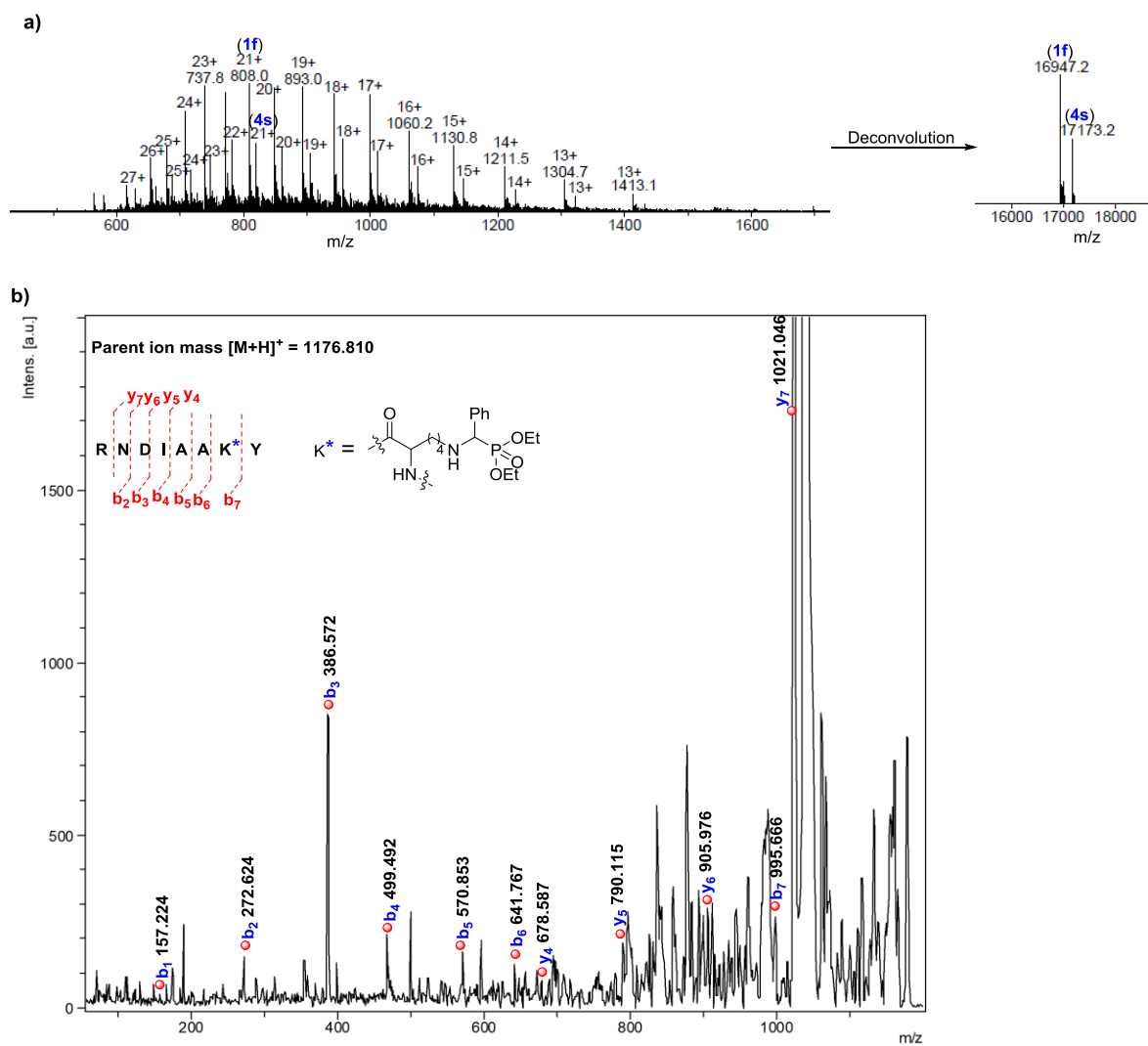


Figure S14. Site-selective native Myoglobin modification. (a) ESI-MS spectra for Myoglobin **1f** and mono-labeled Myoglobin **4s**. (b) MALDI-ToF MS-MS spectrum of labeled RNDIAAKY (R139-Y146). Site of modification is K145 in mono-labeled Myoglobin **4s**.

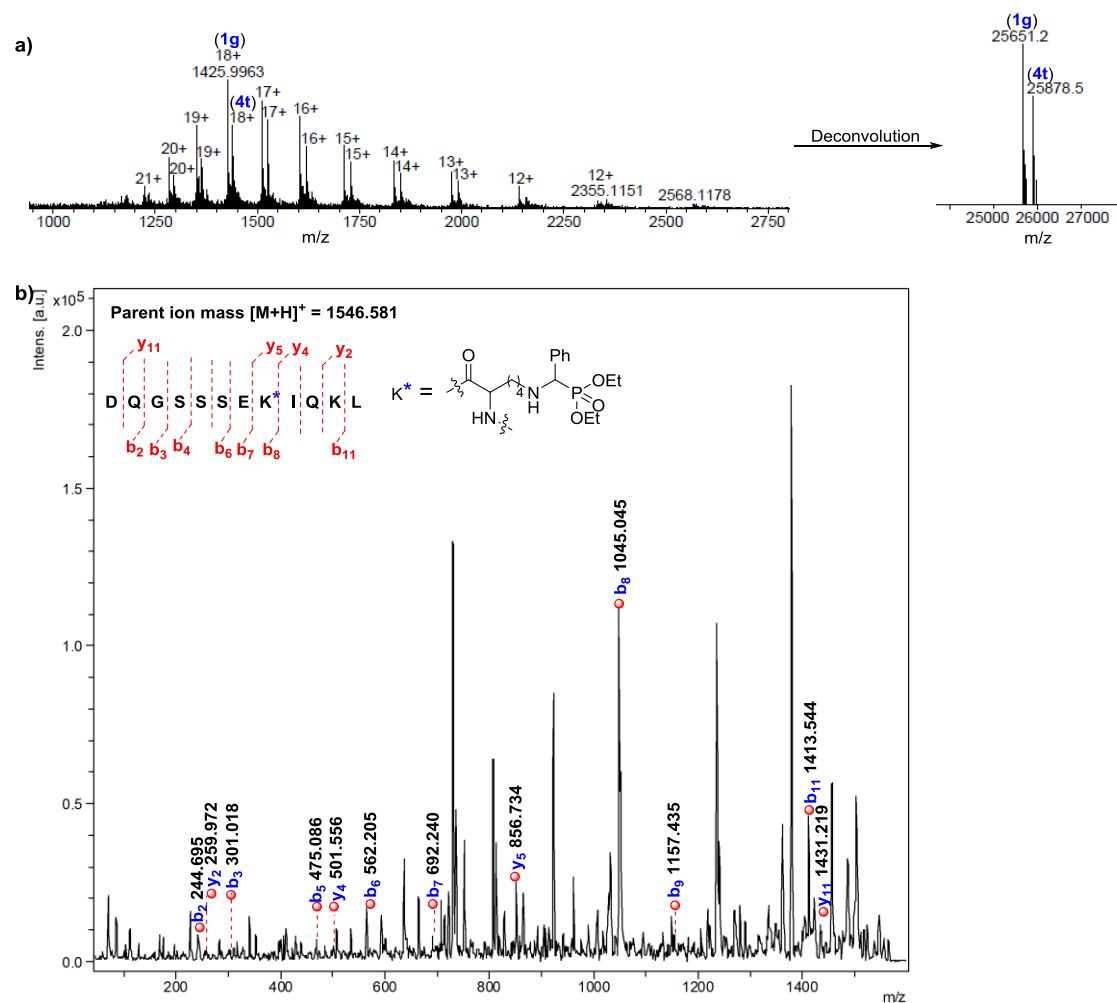
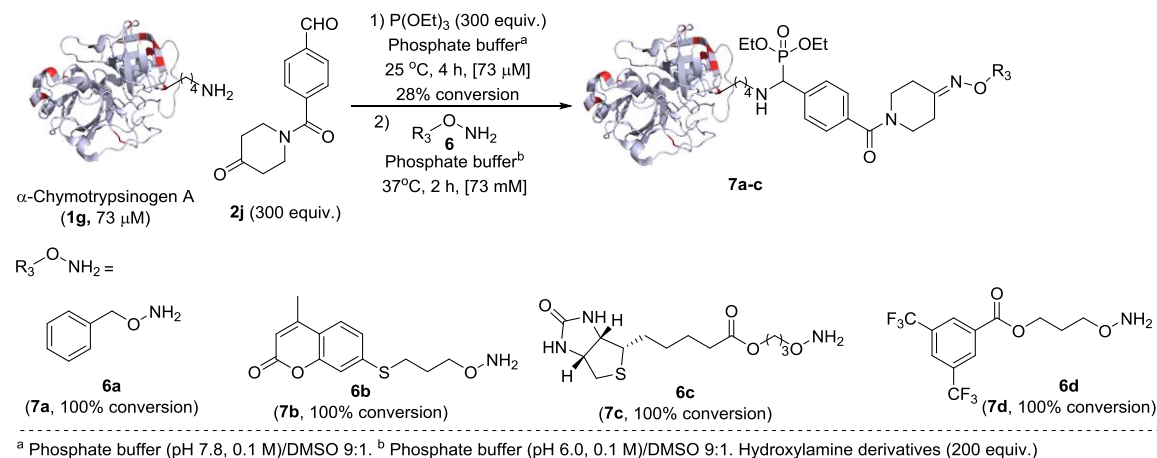


Figure S15. Site-selective native α -chymotrypsinogen A modification. (a) ESI-MS spectra for α -chymotrypsinogen A **1g** (1 equiv.) and mono-labeled α -chymotrypsinogen A **4t**. (b) MALDI-ToF MS-MS spectrum of labeled DQGSSEKIQL (D72-L83). Site of modification is K79 in mono-labeled α -chymotrypsinogen A **4t**.

8. Late-stage orthogonal tagging of the labeled protein



Scheme S1. Late-stage modification of α -chymotrypsinogen A (**1g**) with hydroxylamine derivatives.

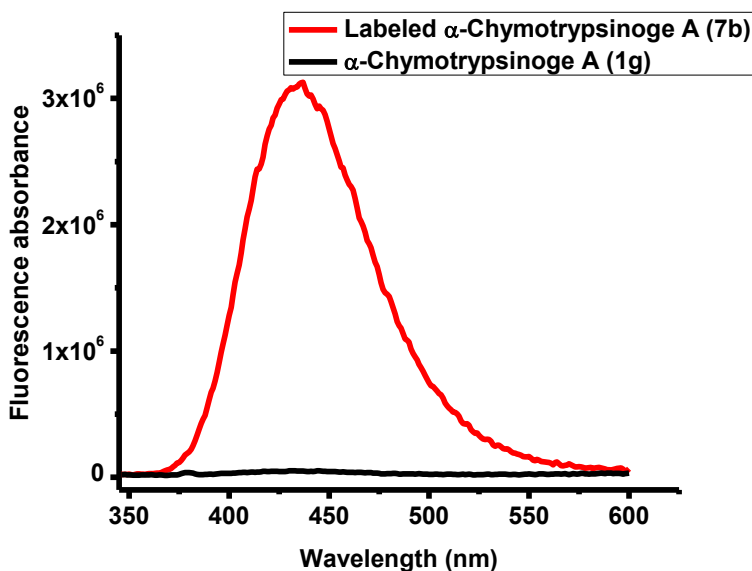


Figure S16. Fluorescence emission spectra of α -chymotrypsinogen A (**1g**, black line), labeled α -chymotrypsinogen A (**7b**, red line) in phosphate buffer (0.1 M, pH 7.0) at concentration 1 mg/ml. The fluorescence spectra of coumarin labeled α -chymotrypsinogen A (**7b**) results emission at 434 nm upon excitation at 336 nm.

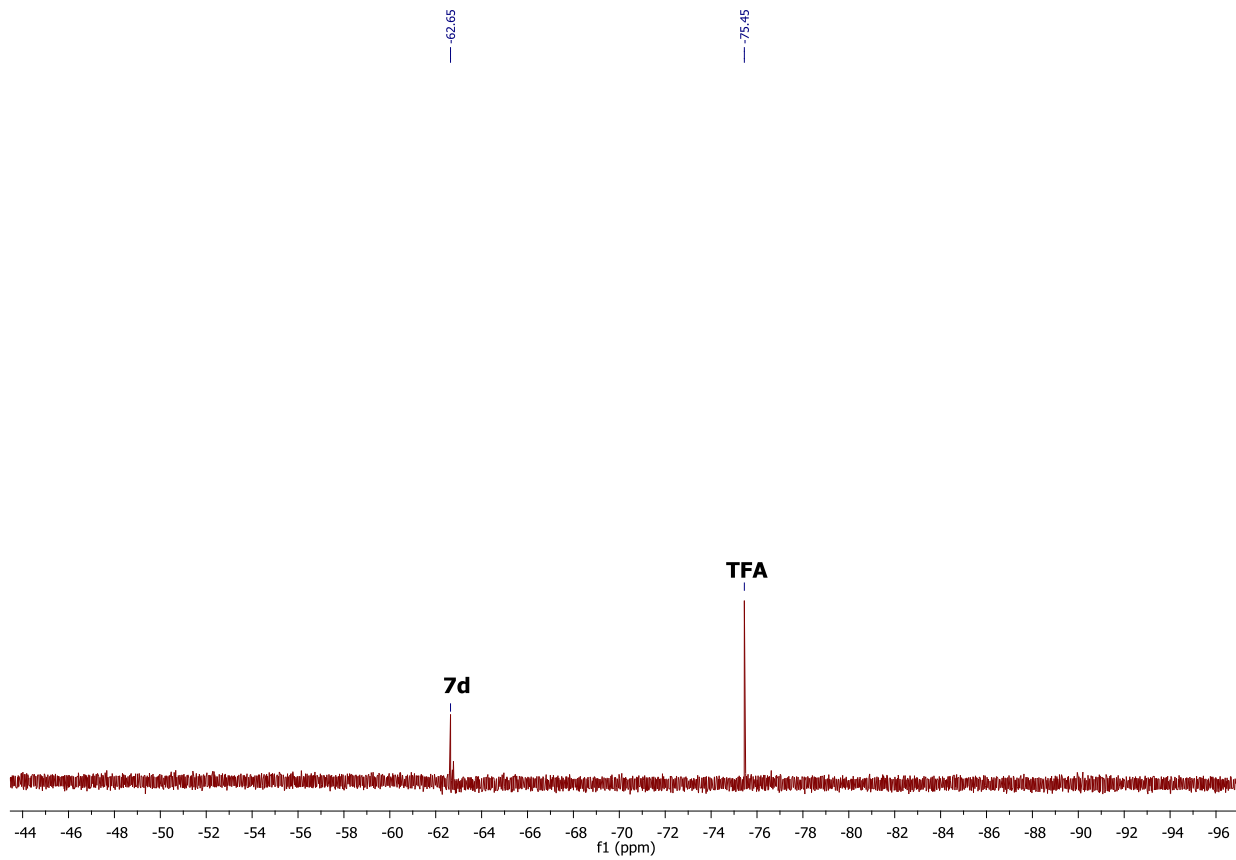


Figure S17. ^{19}F -NMR probe labeled α -chymotrypsinogen A (**7d**), shows a sharp signal at -62.65 ppm by ^{19}F NMR spectroscopy [TFA (0.2 mM) as internal standard at -75.45 ppm, phosphate buffer (0.1 M, pH 7.0): D_2O (9:1)].

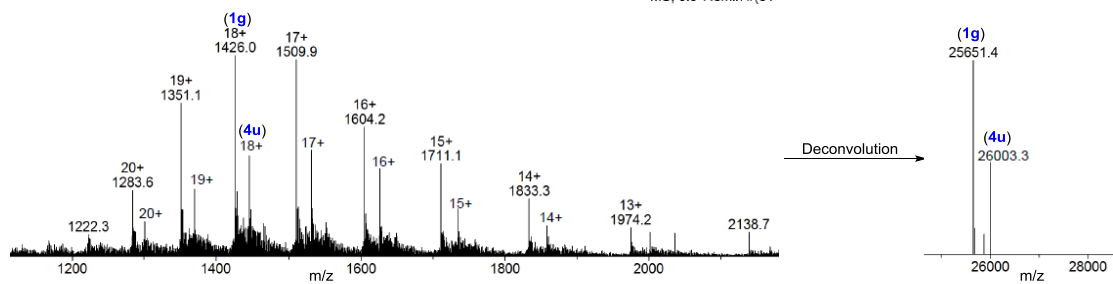


Figure S18. ESI-MS spectra for α -chymotrypsinogen A **1g** (1 equiv.) and mono-labeled α -chymotrypsinogen A **4u**.

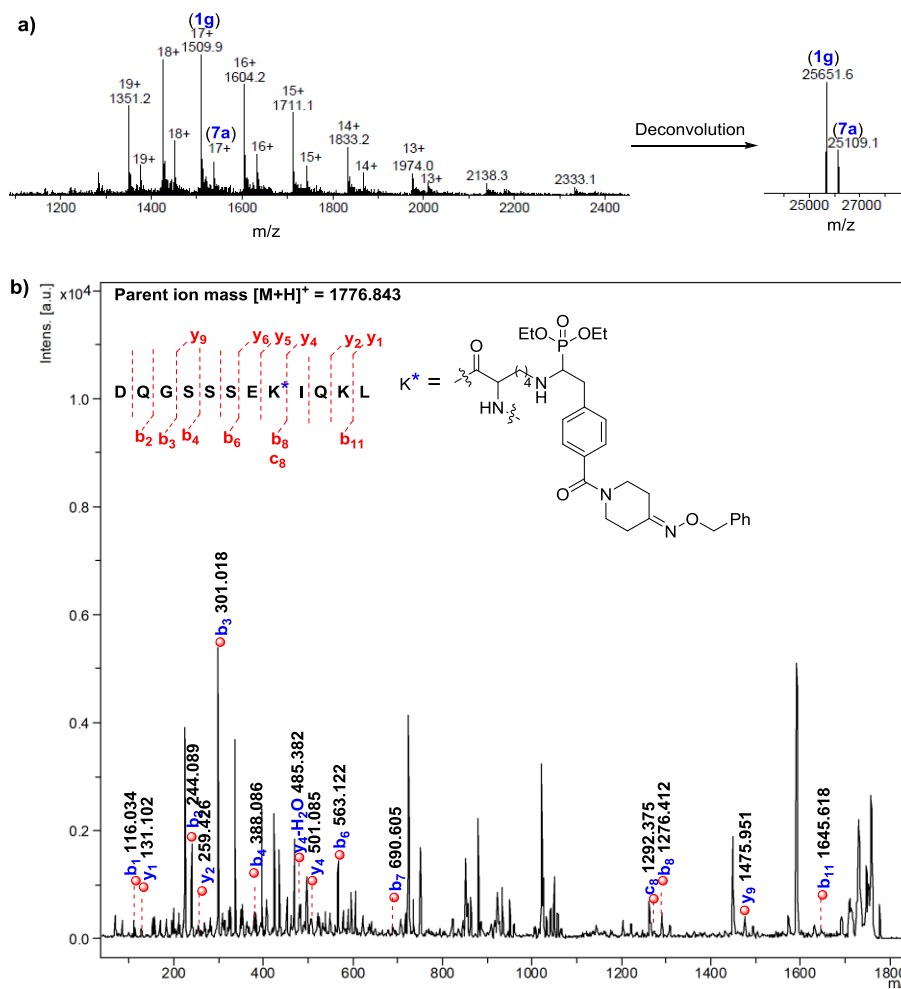


Figure S19. (a) ESI-MS spectra for α -chymotrypsinogen A **1g** (1 equiv.) and mono-labeled α -chymotrypsinogen A **7a**. (b) MALDI-ToF MS-MS spectrum of labeled DQGSSEKIQL (D72-L83). Site of modification is K79 in mono-labeled α -chymotrypsinogen A **7a**.

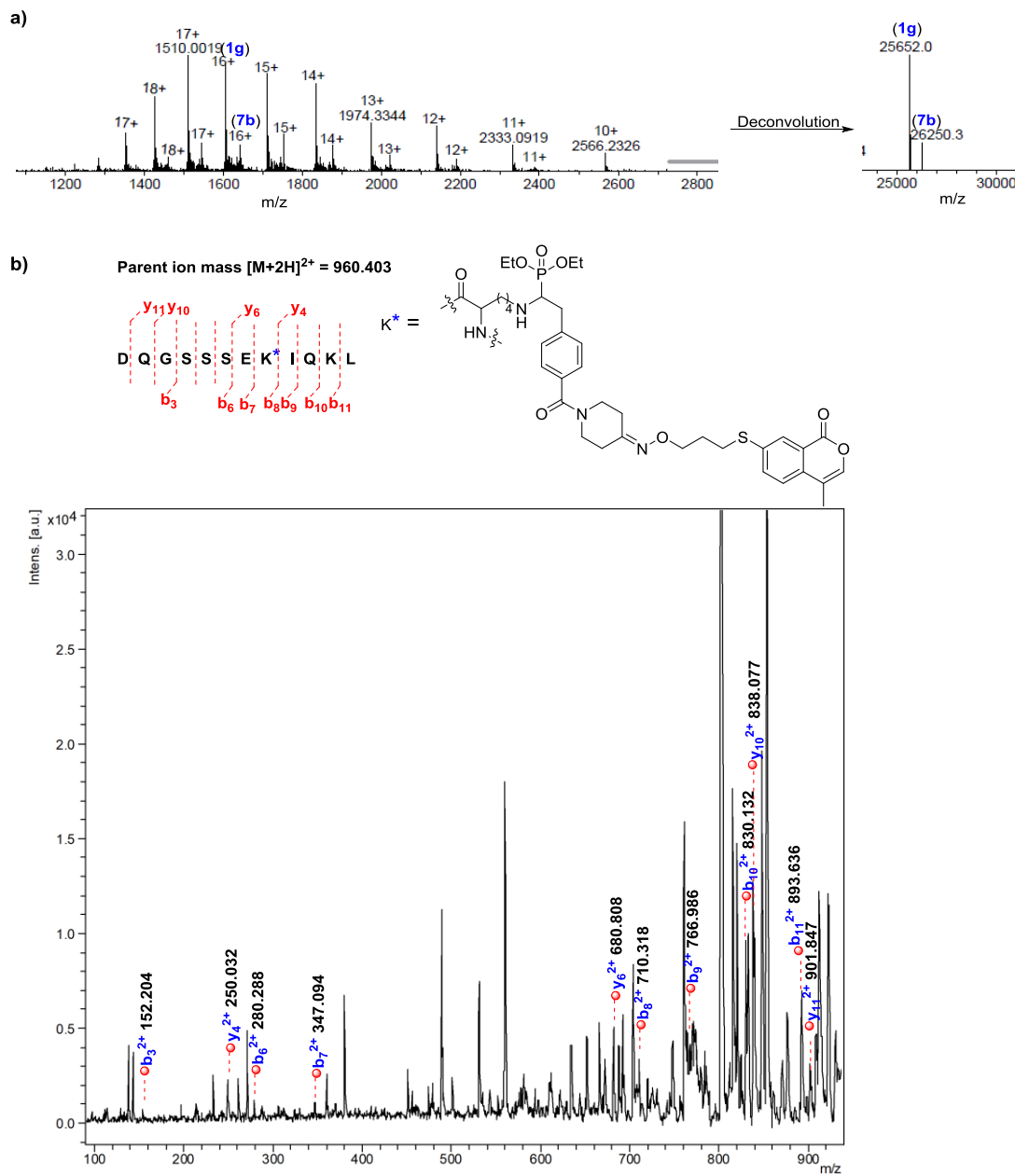
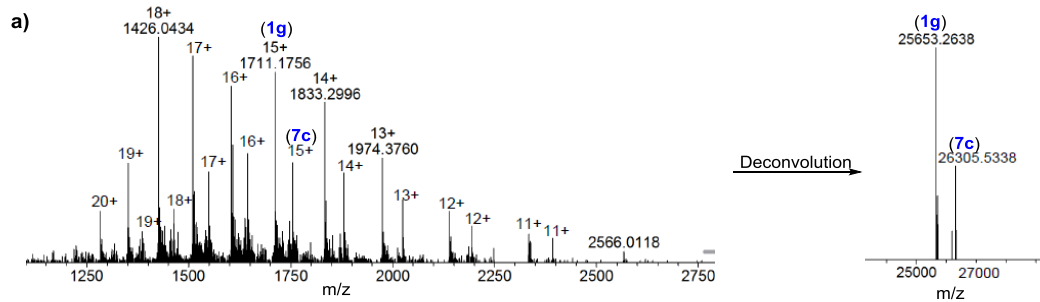
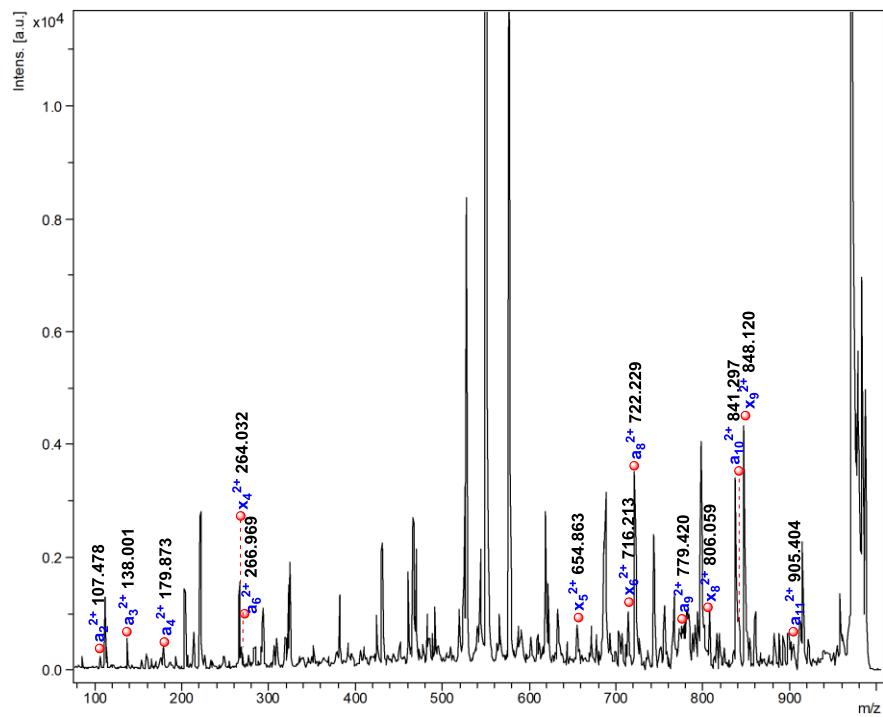
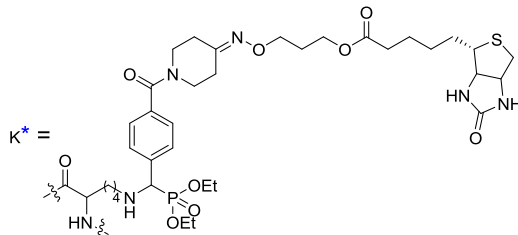
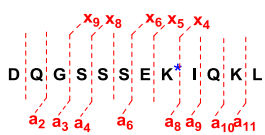


Figure S20. (a) ESI-MS spectra for α -chymotrypsinogen A **1g** (1 equiv.) and mono-labeled α -chymotrypsinogen A **7b**. (b) MALDI-ToF MS-MS spectrum of doubly charged labeled DQGSSEKIQKL (D72-L83). Site of modification is K79 in mono-labeled α -chymotrypsinogen A **7b**.



b) Parent ion mass $[M+2H]^{2+} = 987.896$



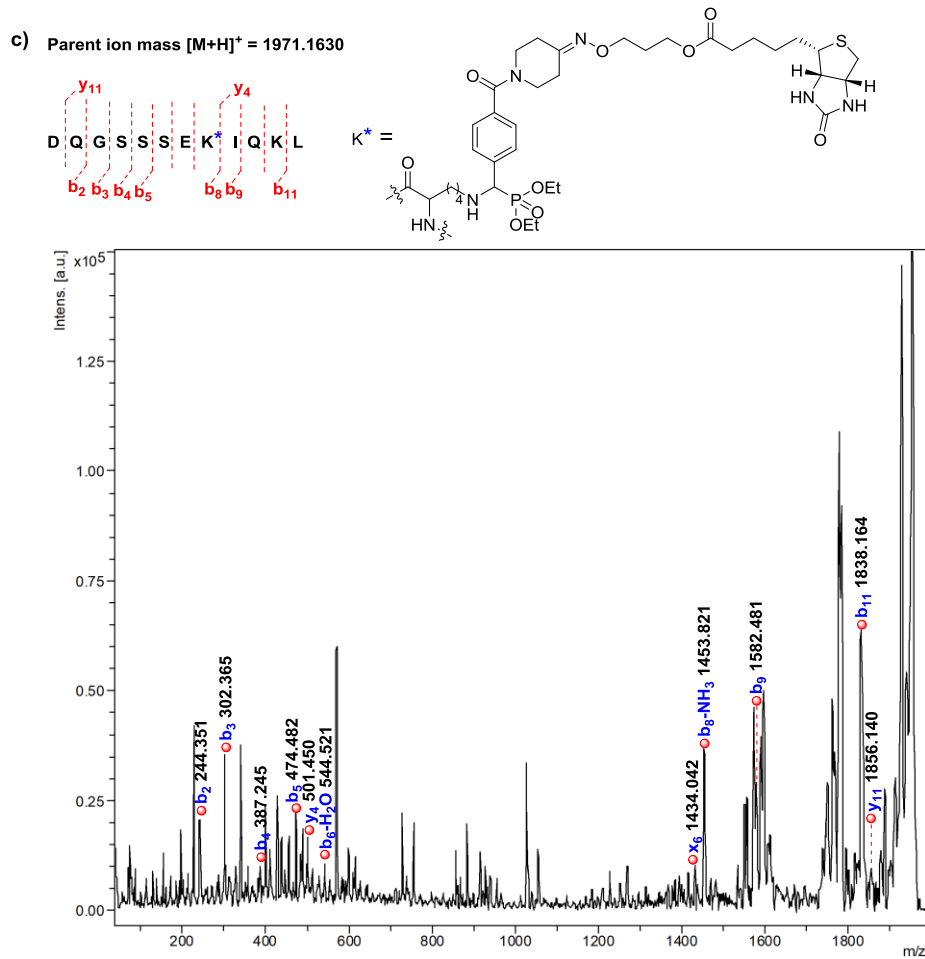


Figure S21. (a) ESI-MS spectra for α -chymotrypsinogen A **1g** (1 equiv.) and mono-labeled α -chymotrypsinogen A **7c**. (b) MALDI-ToF MS-MS spectrum of doubly charged labeled DQGSSSEKIQL (D72-L83). Site of modification is K79 in mono-labeled α -chymotrypsinogen A **7c**. (c) MALDI-ToF MS-MS spectrum of labeled DQGSSSEKIQL (D72-L83). Site of modification is K79 in mono-labeled α -chymotrypsinogen A **7c**.

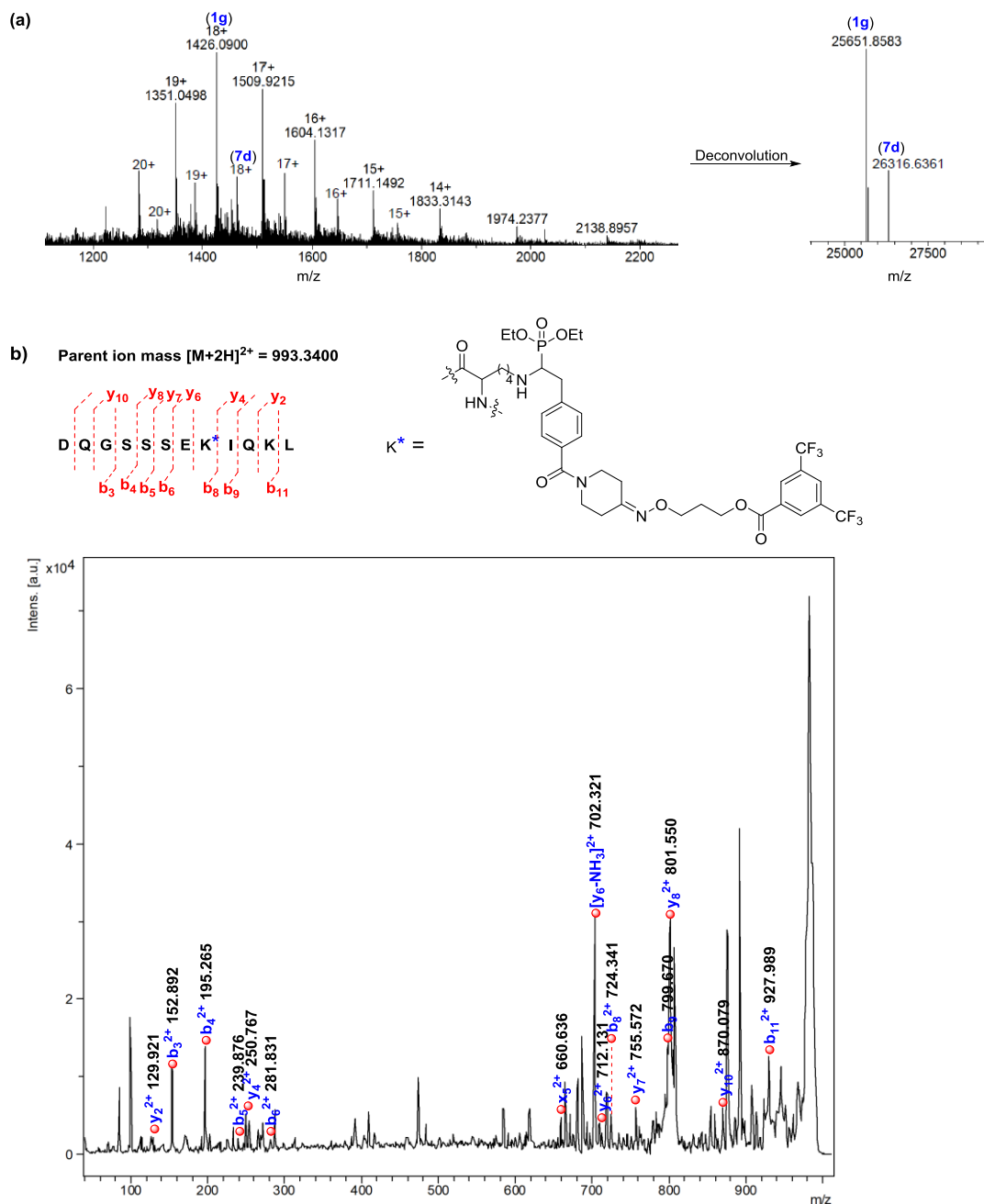
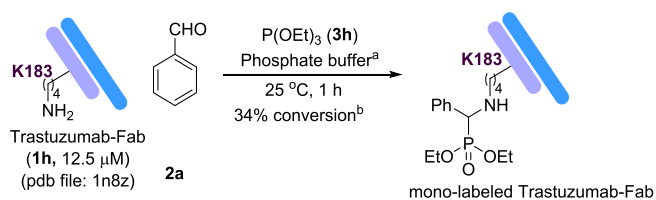


Figure S22. (a) ESI-MS spectra for α -chymotrypsinogen A **1g** (1 equiv.) and mono-labeled α -chymotrypsinogen A **7d**. (b) MALDI-ToF MS-MS spectrum of doubly charged labeled DQGSSSEKIQKL (D72-L83). Site of modification is K79 in mono-labeled α -chymotrypsinogen A **7d**.

9. Reactions with Fab and monoclonal antibody

9.1. Labeling of Trastuzumab-Fab



^a Phosphate buffer (pH=7.8, 0.1 M)/DMSO = 9:1. ^b %conversion were based on the ESI-MS data.

Scheme S2. Labeling of the Trastuzumab-Fab (**1h**) with benzaldehyde (**2a**).

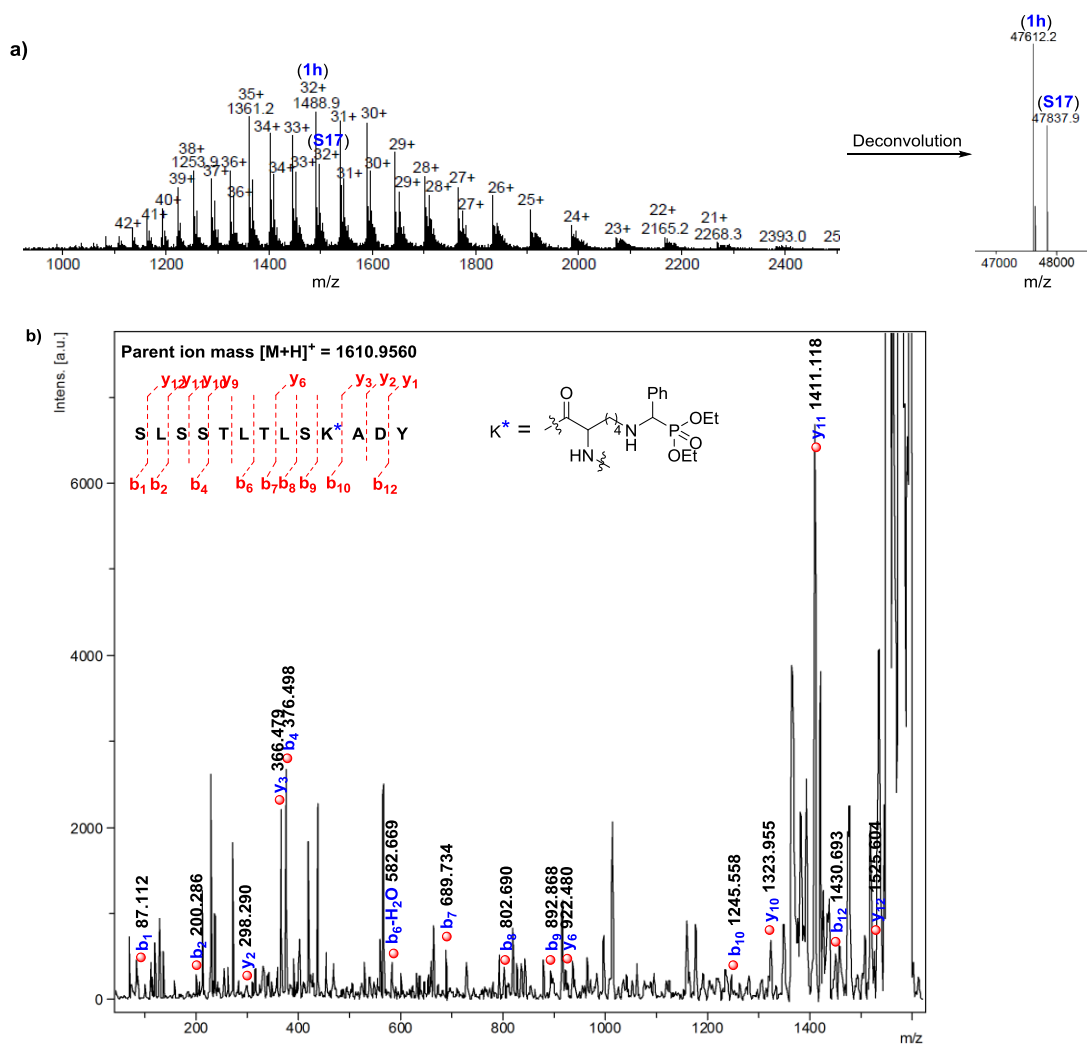
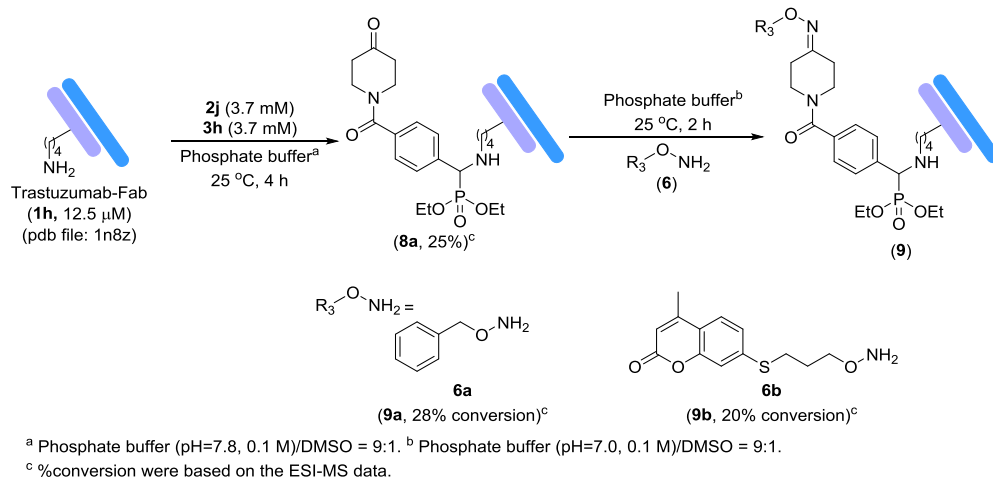


Figure S23. Site-selective native Trastuzumab-Fab modification. (a) ESI-MS spectra for Trastuzumab-Fab **1h** (1 equiv.) and mono-labeled Trastuzumab-Fab **S17**. (b) MALDI-ToF MS-MS spectrum of labeled SLSSTLTLTKADY (S174-Y186) of light chain fragment of Trastuzumab-Fab. Site of modification is K183 in mono-labeled Trastuzumab-Fab **S17**.



Scheme S3. Tagging of Trastuzumab-Fab (**1g**) with hydroxylamine derivatives.

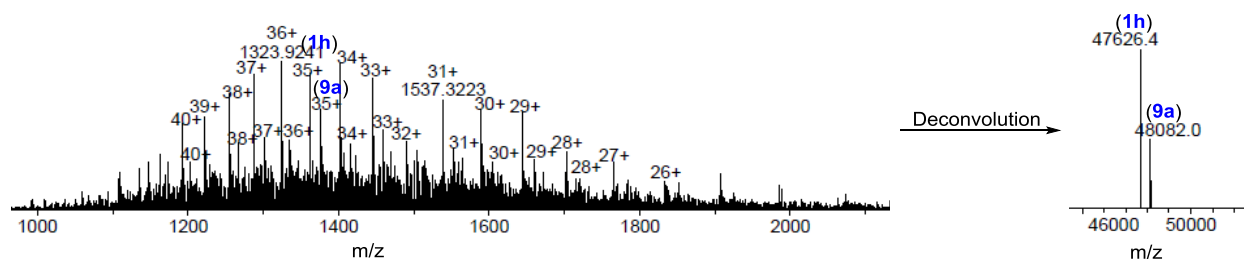


Figure S24. ESI-MS spectra for Trastuzumab-Fab **1h** and mono-labeled Trastuzumab-Fab **9a**

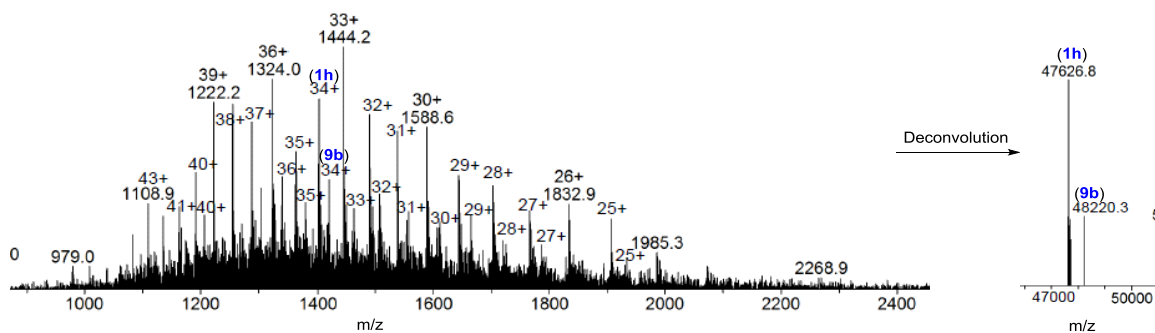
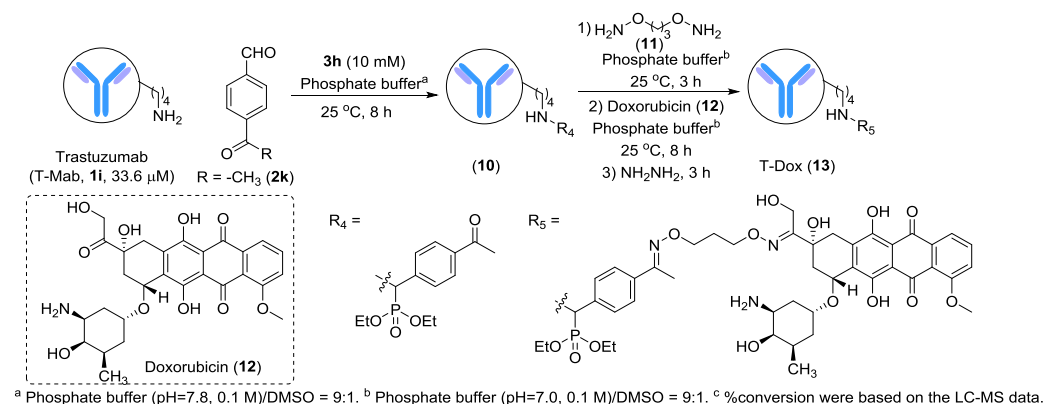


Figure S25. ESI-MS spectra for Trastuzumab-Fab **1h** and mono-labeled Trastuzumab-Fab **9b**

9.2. Synthesis of antibody-drug conjugate (T-Dox, 13)



Scheme S4. Synthesis of trastuzumab-doxorubicin conjugate (ADC, 13).

9.2.1. Calculation of drug to antibody ratio (DAR)⁹

UV spectra were recorded on Agilent Carry 5000 UV-Vis spectrophotometer, operating at 25 °C. Sample buffer (phosphate buffer, pH 7.8, 0.1 M) was used as blank for baseline correction. Calculation of DAR follows the formula below with $\epsilon_{280} = 215380 \text{ M}^{-1} \text{ cm}^{-1}$ for Trastuzumab (T-Mab), $\epsilon_{495} = 8030 \text{ M}^{-1} \text{ cm}^{-1}$ for doxorubicin and 0.006 as a correction factor for DOX absorption at 280 nm. Correction factors are calculated based on the UV-Vis spectroscopy. DAR was calculated as follows:

$$\text{DAR} = \frac{\text{Abs}_{495} / \epsilon_{495}}{(\text{Abs}_{280} - 0.0117 \times \text{Abs}_{495}) / \epsilon_{280}}$$

$$\text{Abs}_{495} = 0.016, \text{Abs}_{280} = 0.43.$$

$$\text{DAR} = 0.92$$

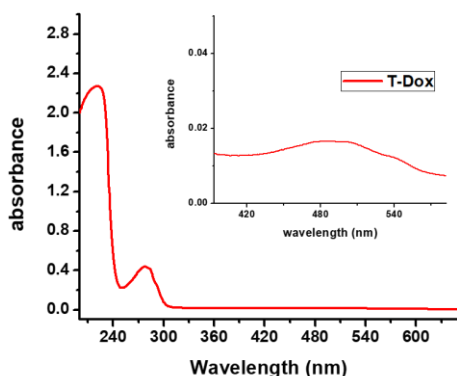


Figure S26. UV-Vis spectra for the determination of DAR for ADC (13) prepared with Dox (12).

9.2.2. Anti-proliferative activity test (MTT assay)

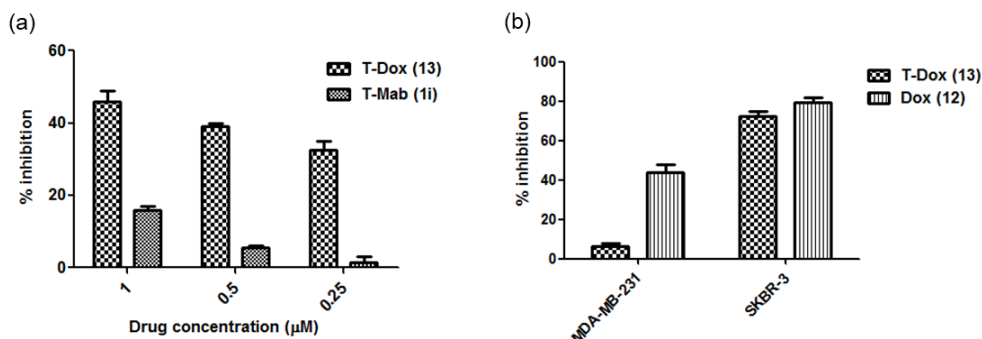


Figure S27: (a) Synthesis of antibody drug conjugate of Trastuzumab (**1i**) through doxorubicin (**12**) conjugation. (b) Inhibition of cell proliferation by T-Dox (**13**) in cancer cell lines SKBR3 (HER2 positive) and MDA-MB-231 (HER2 negative). Cells were treated with various concentrations of T-Dox (**13**) and T-Mab (**1i**) and inhibition of cell proliferation was screened after 48 h using MTT assay. The percentage inhibition was calculated using cells only as control.*shows significant inhibition and the error bars denotes standard error of the mean. (c) Selective inhibition of cancer cell proliferation by T-Dox in comparison with doxorubicin. Inhibition of cell proliferation by T-Dox (**13**) in comparison with equivalent concentrations of doxorubicin in SKBR3 (HER2 positive) and MDA-MB-231(HER2 negative) cancer cells. Cells were treated with T-Dox at 2 µM concentration or equivalent doxorubicin (**12**) for 48 h. Inhibition of cell growth measured by % inhibition of cell proliferation using cell only as control.

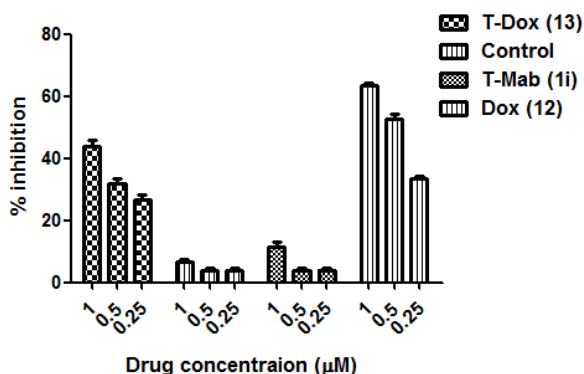
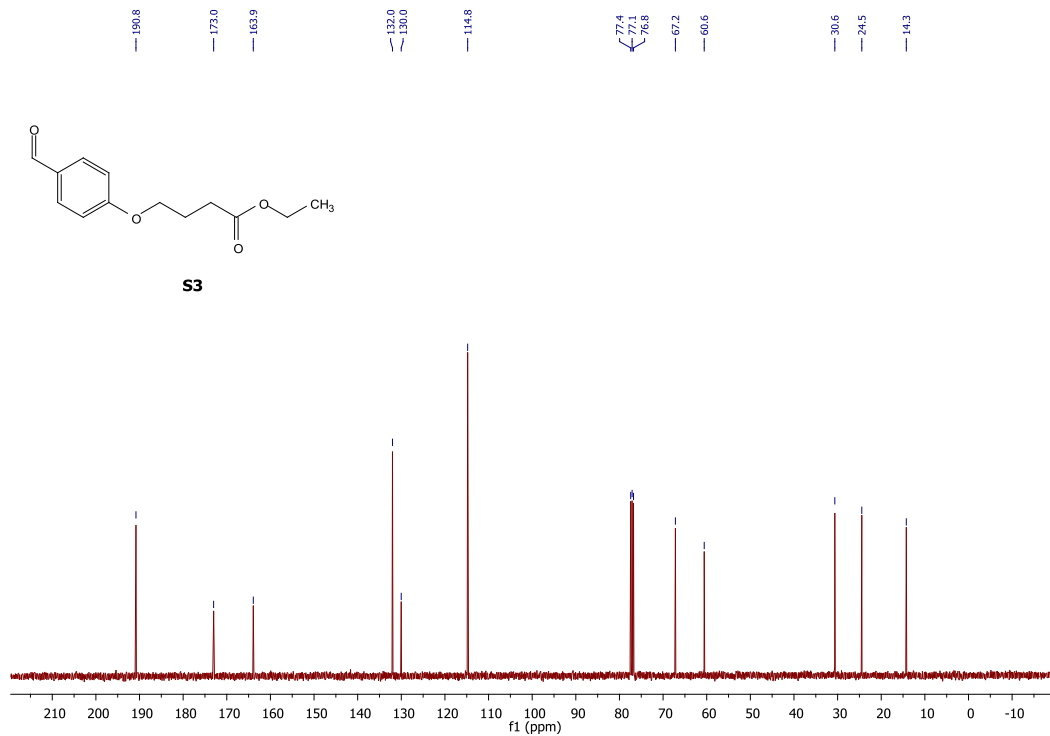
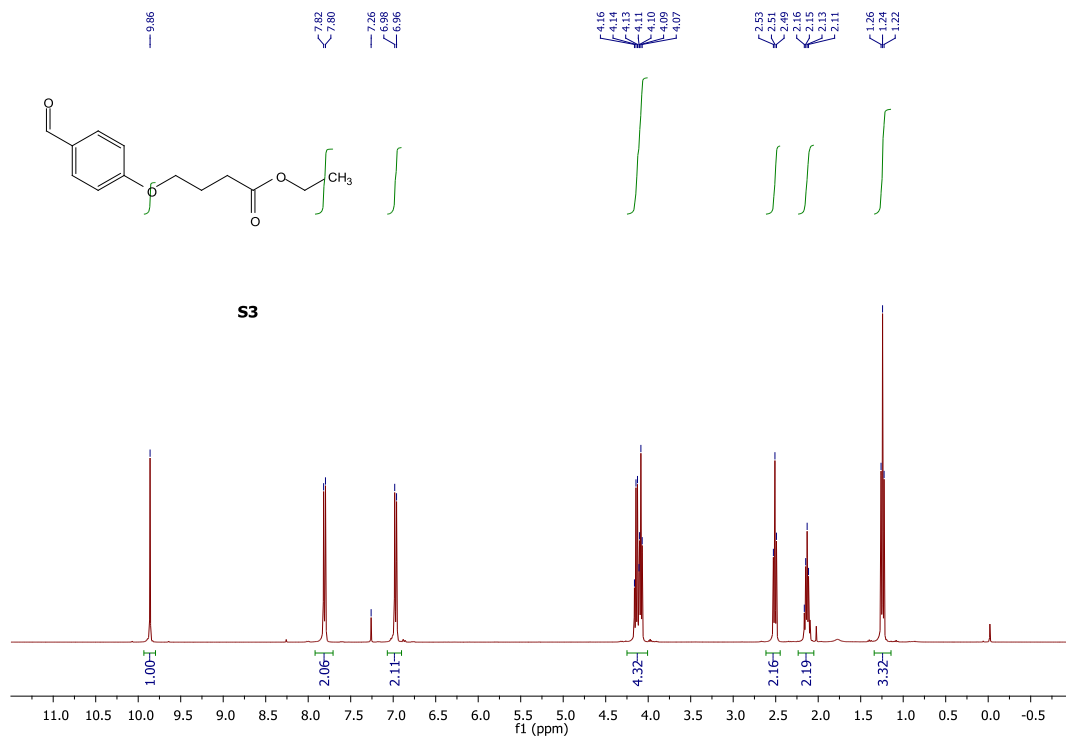
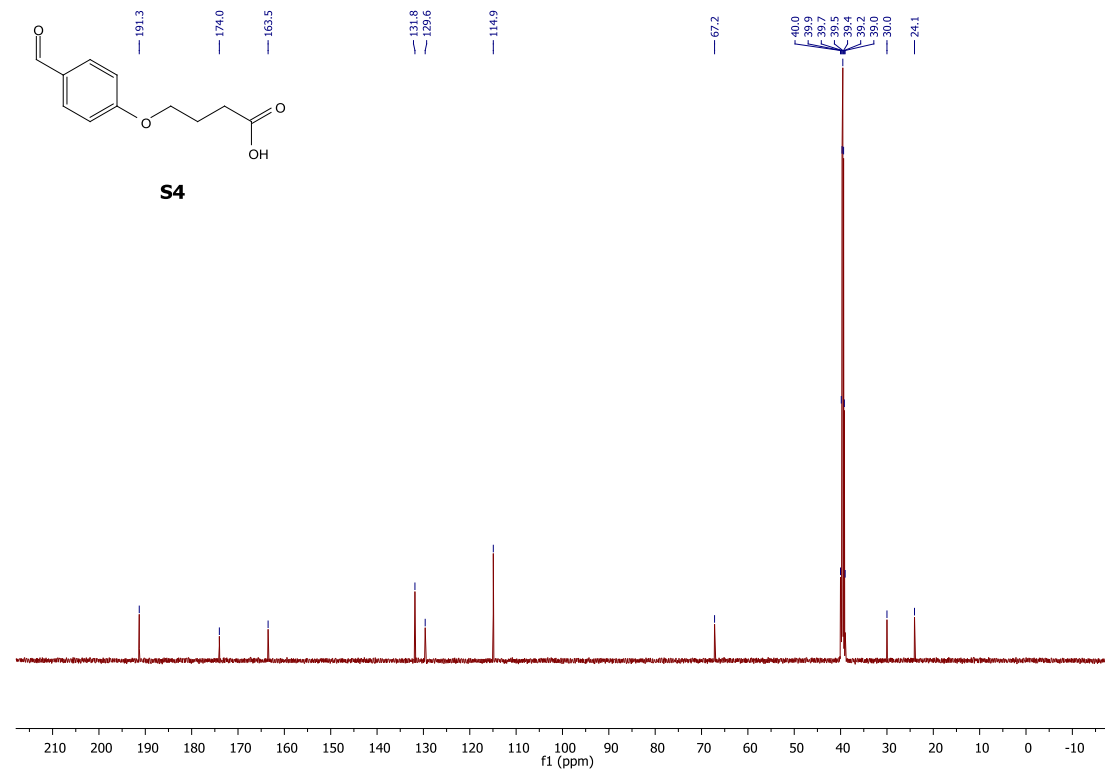
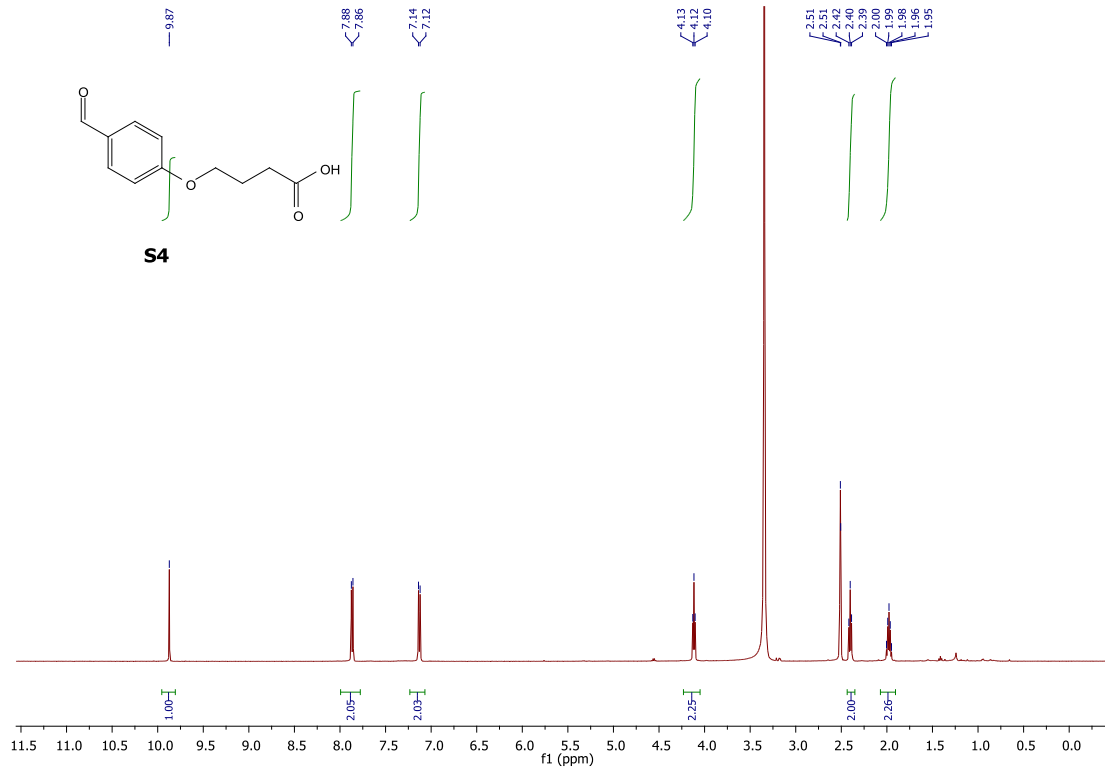


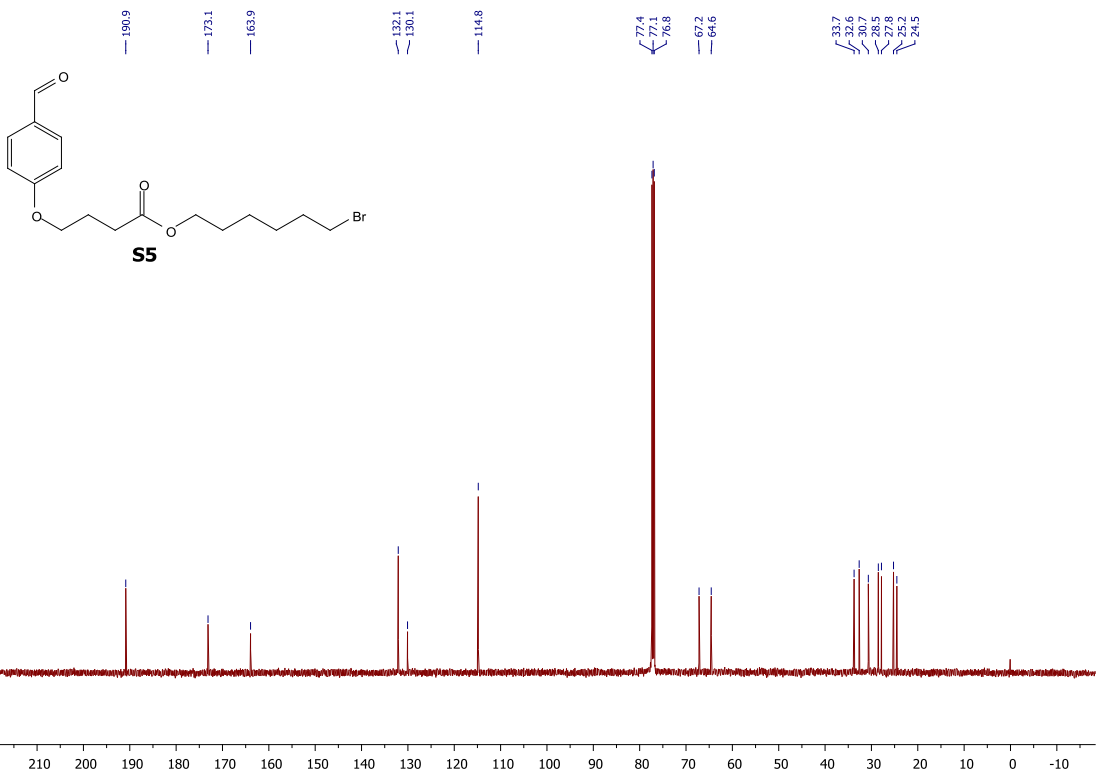
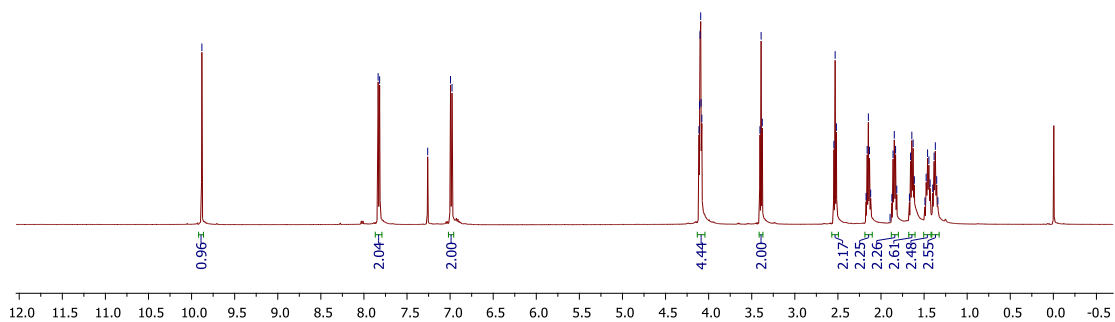
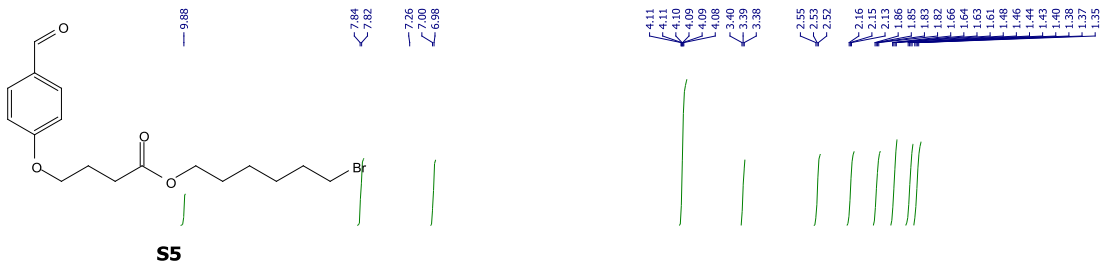
Figure S28. The anti-proliferative activity of the T-Mab (**1i**), T-Dox (**13**) with HER2 positive SKBR3 cells at 48 h.

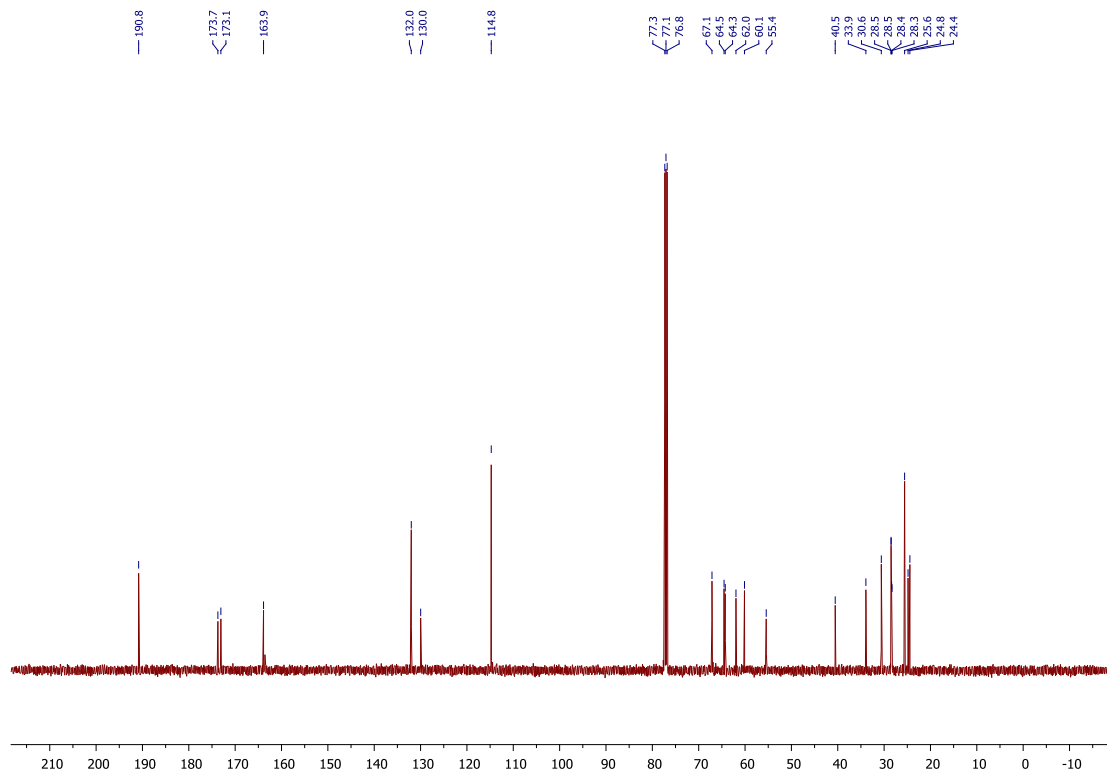
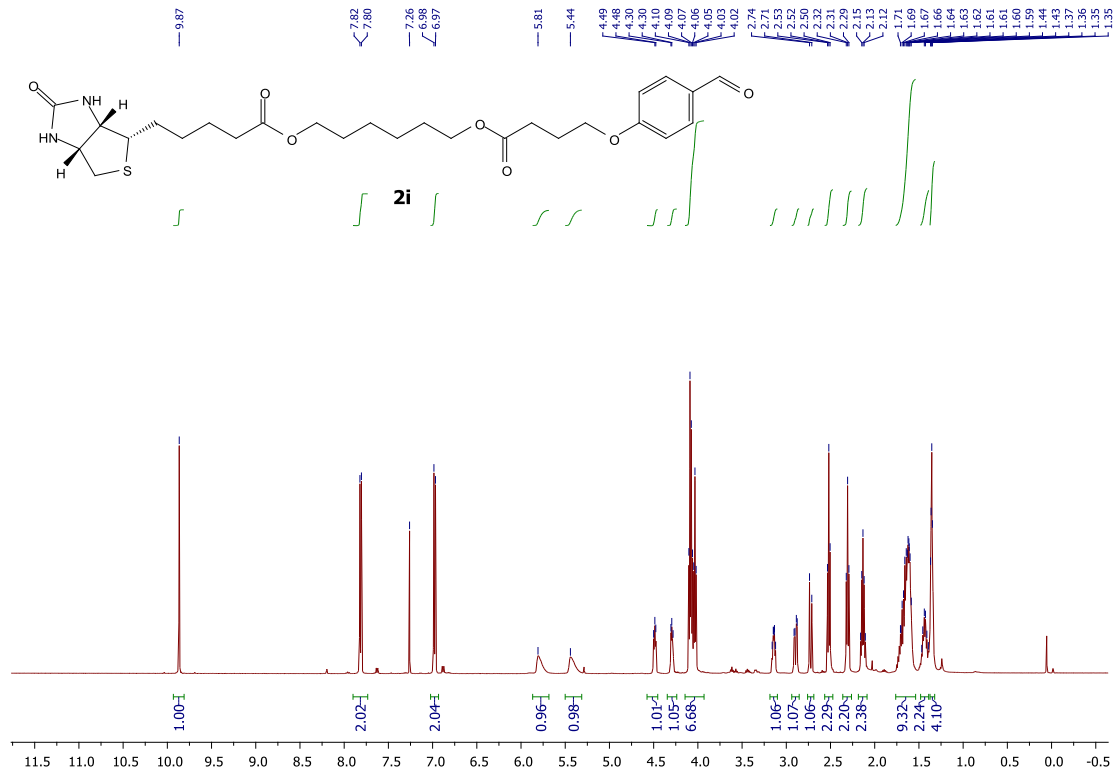
The control experiments for the non-selective binding of Dox (**12**) with antibody performed by mixing 200 equivalents of Dox (**12**) with T-Mab (**1i**) and incubated for the 8 h. The sample was treated with hydrazine for 2 h. The excess Dox (**12**) and hydrazine were washed with water (40 ml) by 30 using spin concentrator (30 kDa MWCO). Further sample (control) was subjected to MTT assay with HER2 positive SKBR3 cells. The MTT suggest that the control samples do not show any effect on the cell proliferation activity.

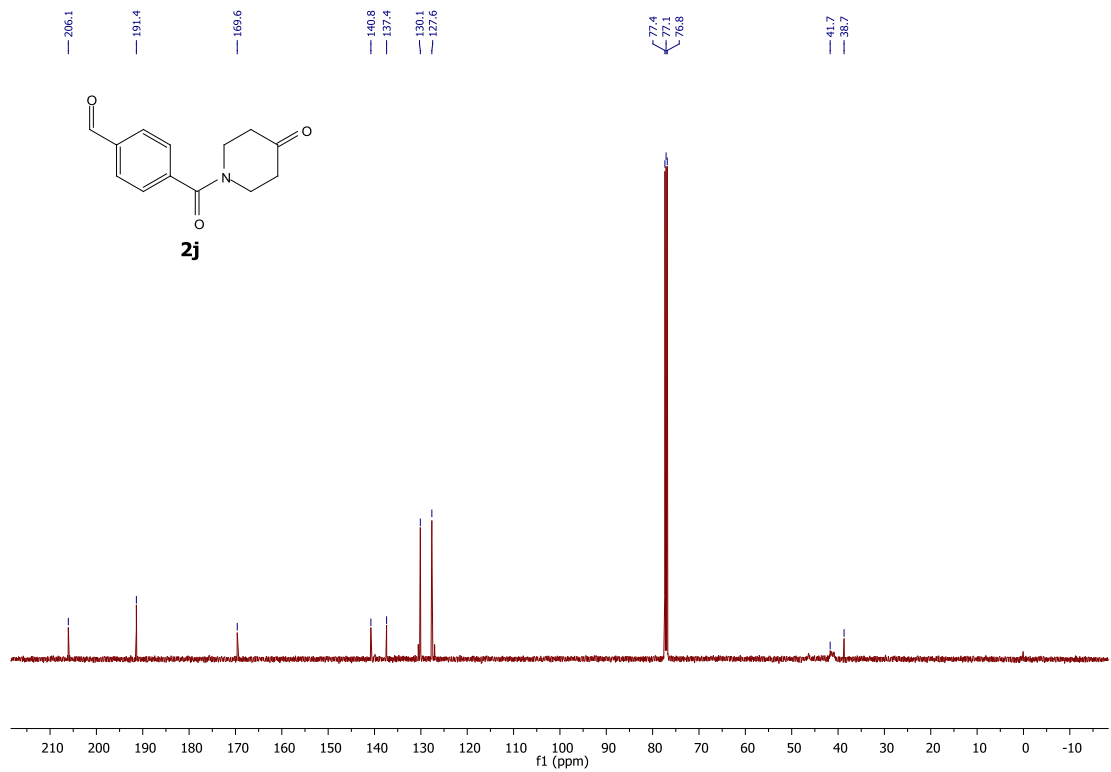
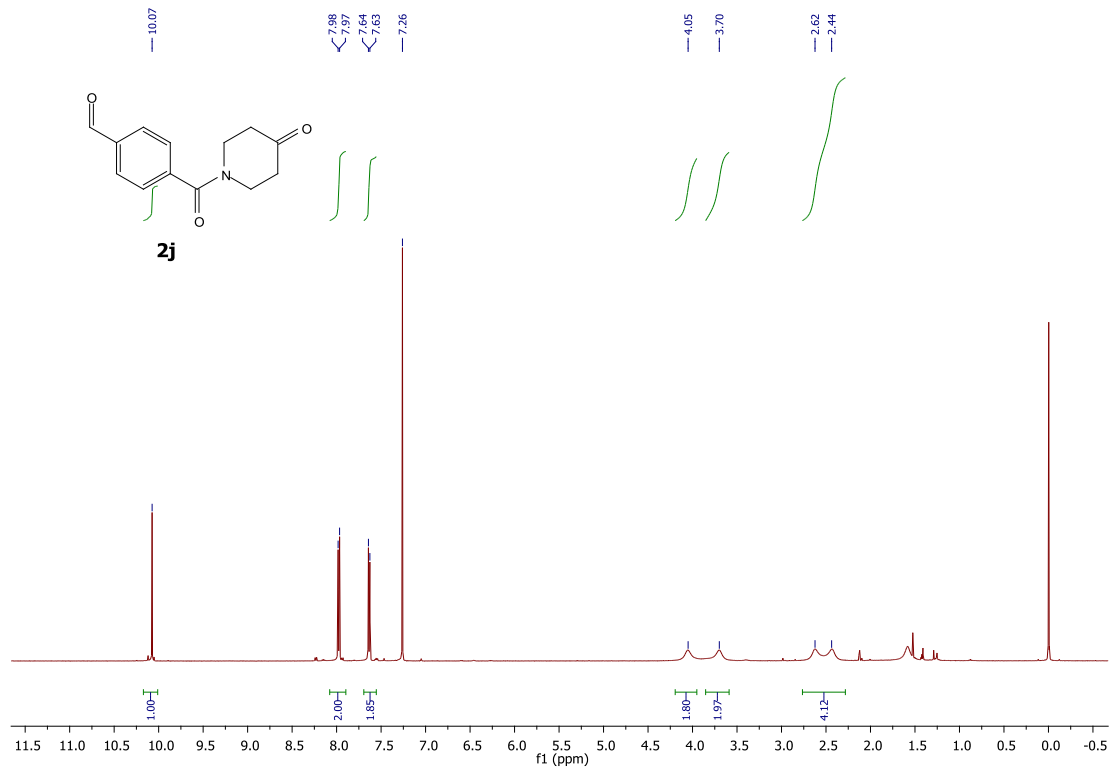
10. NMR data of small molecules

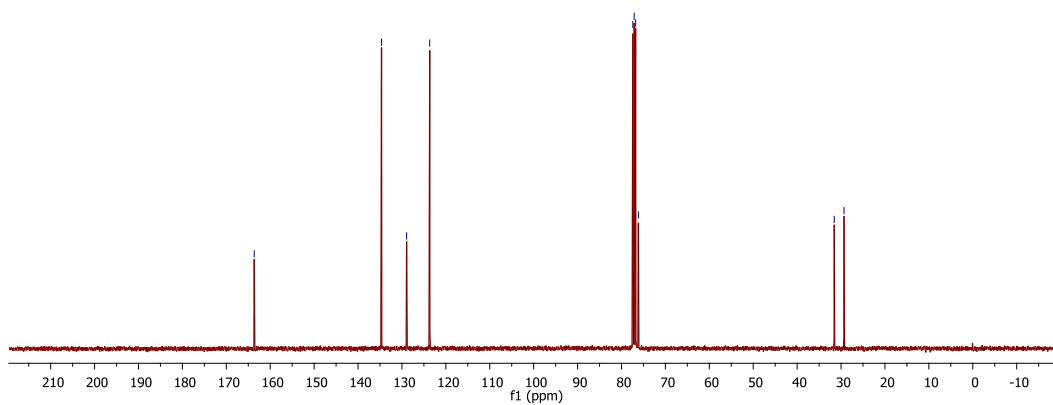
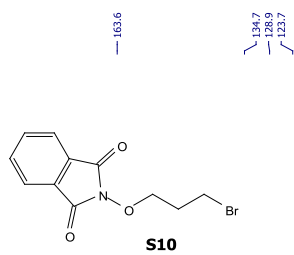
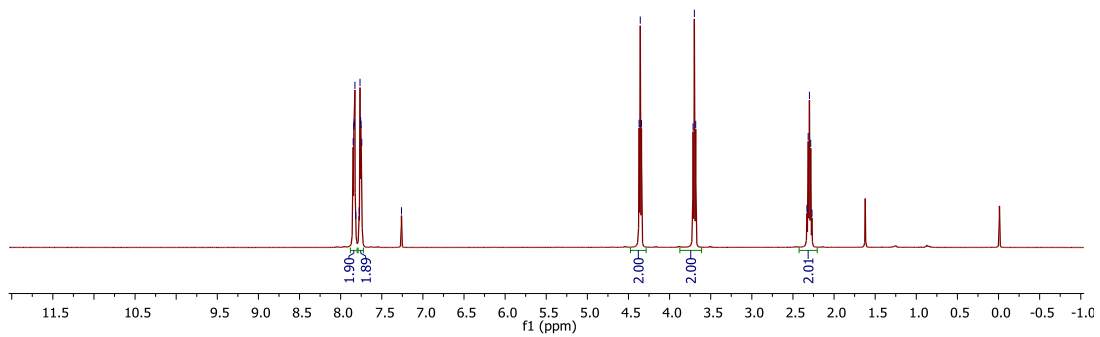
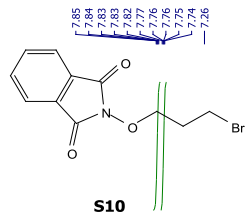


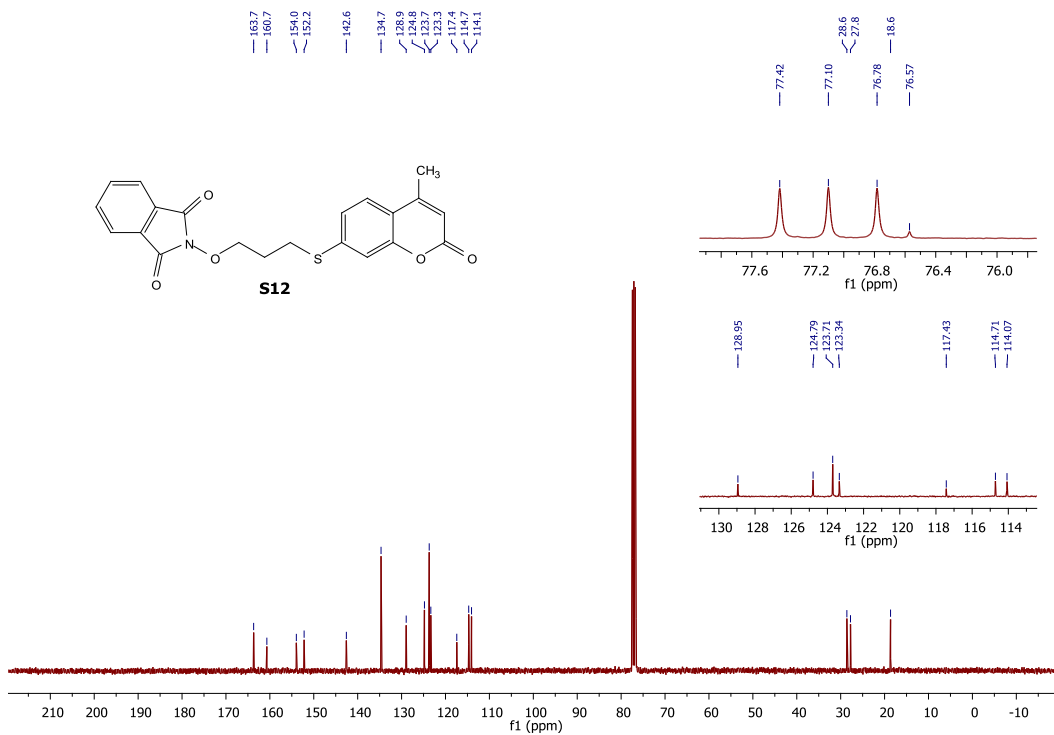
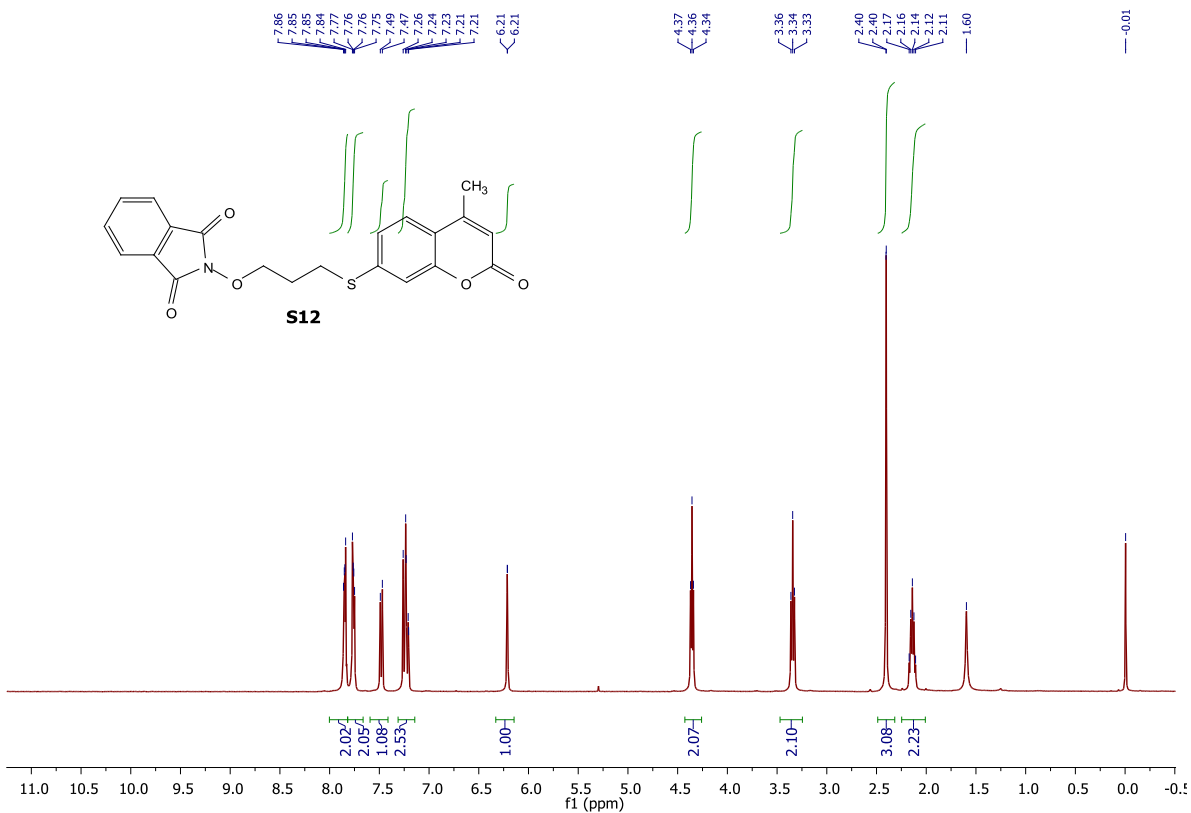


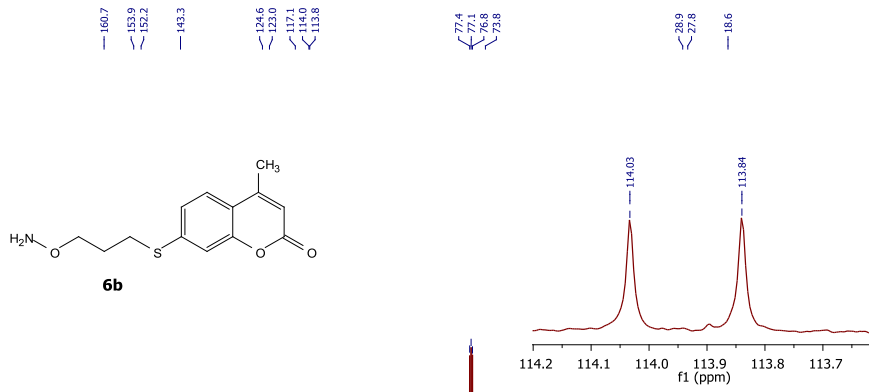
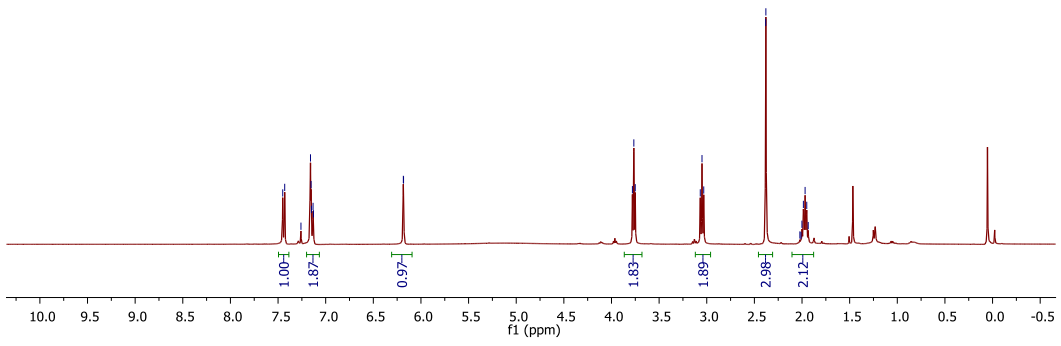
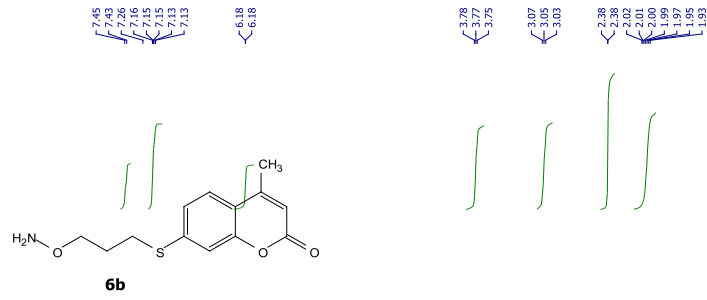


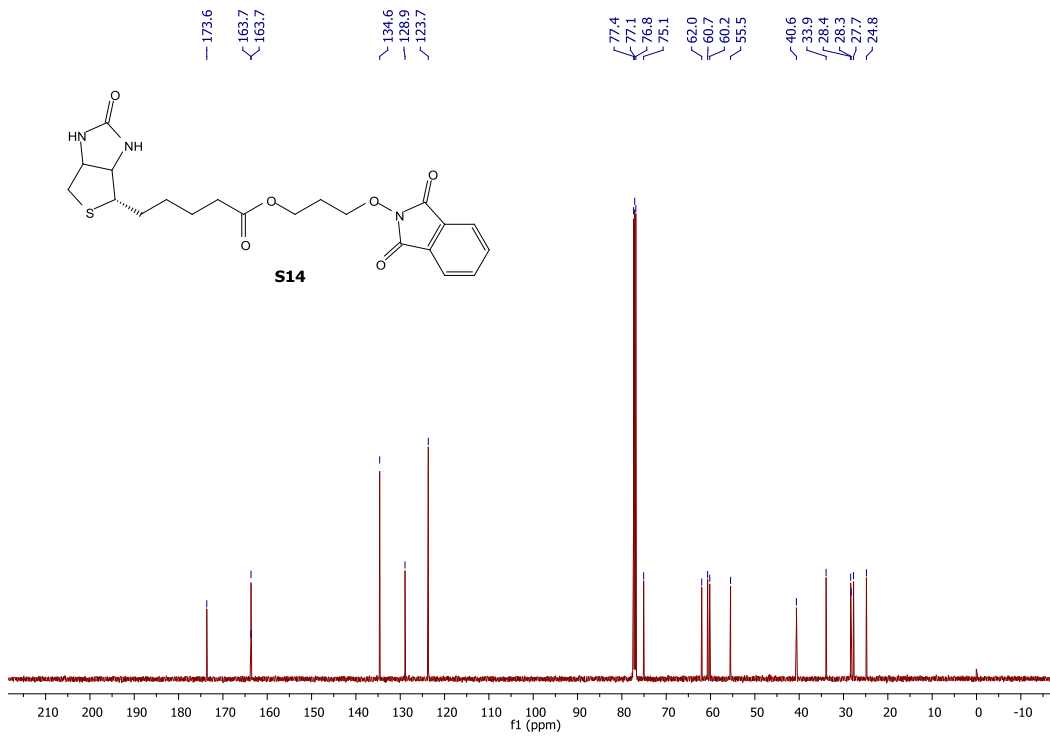
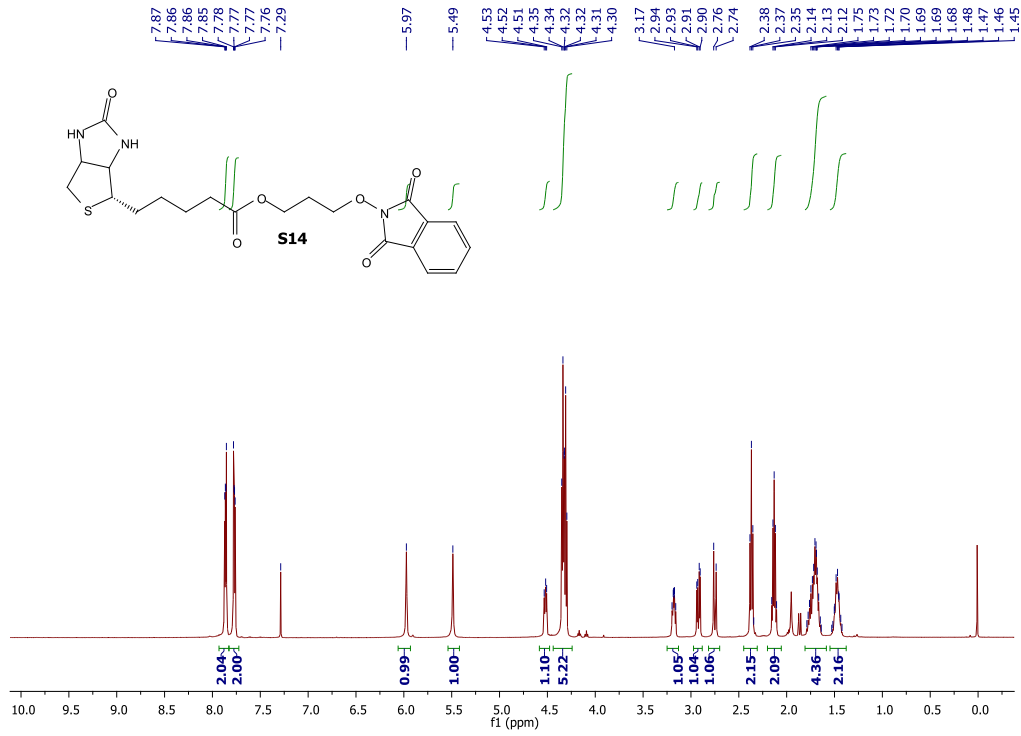


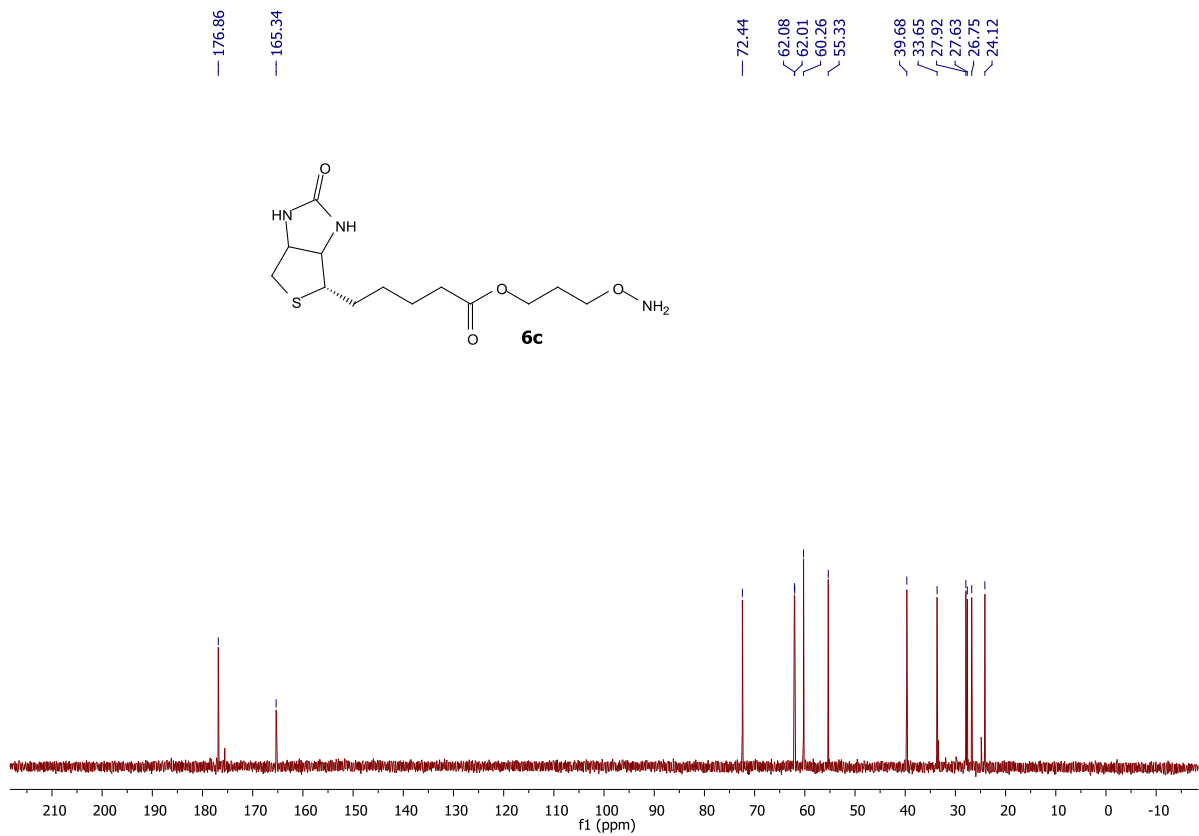
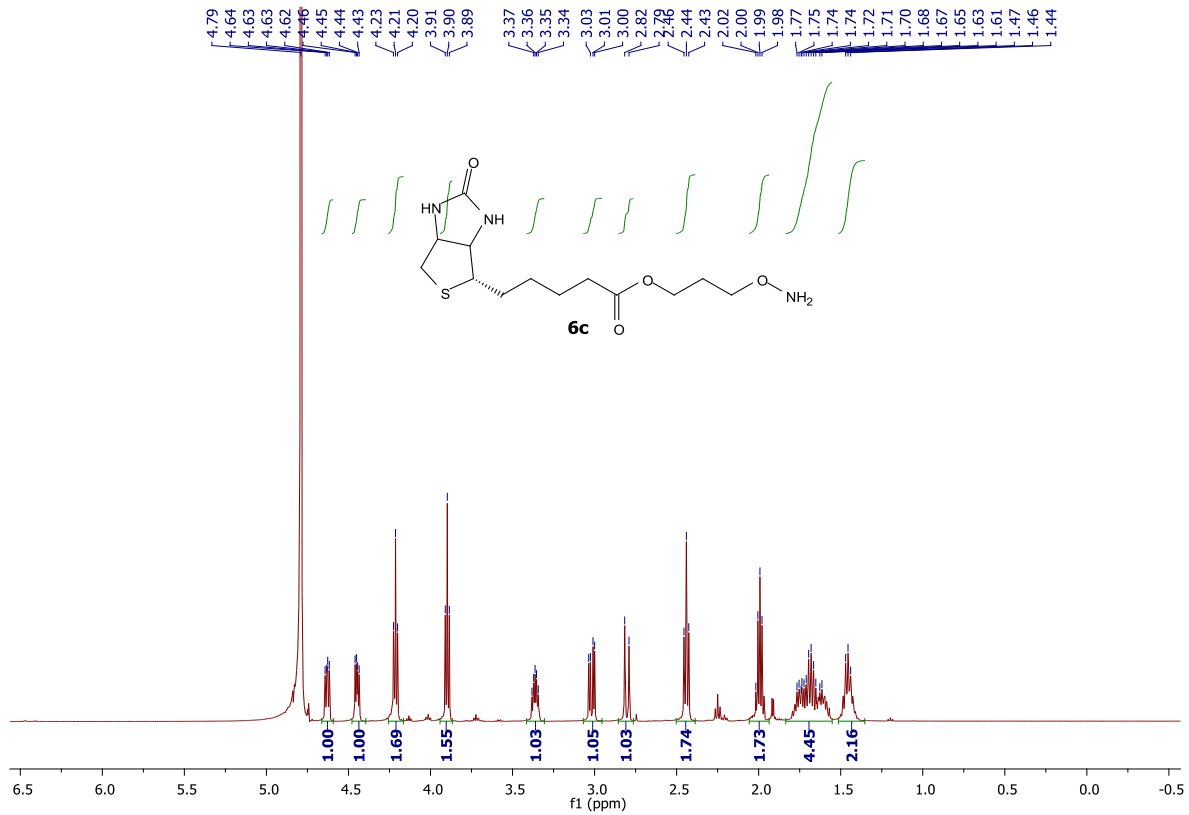


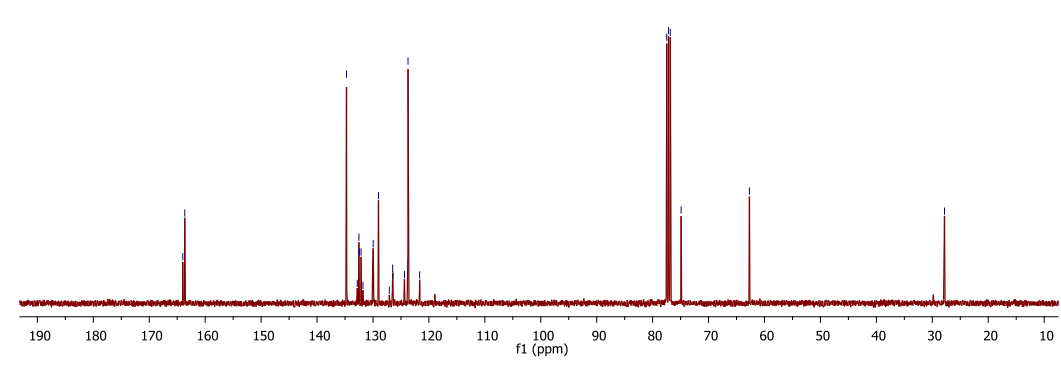
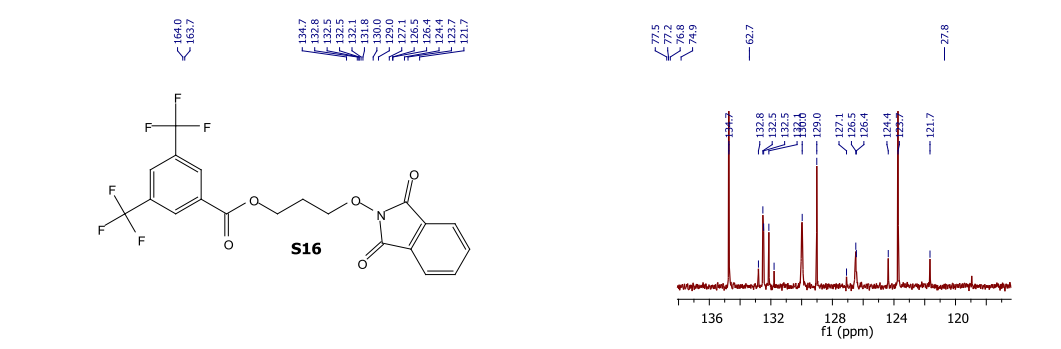
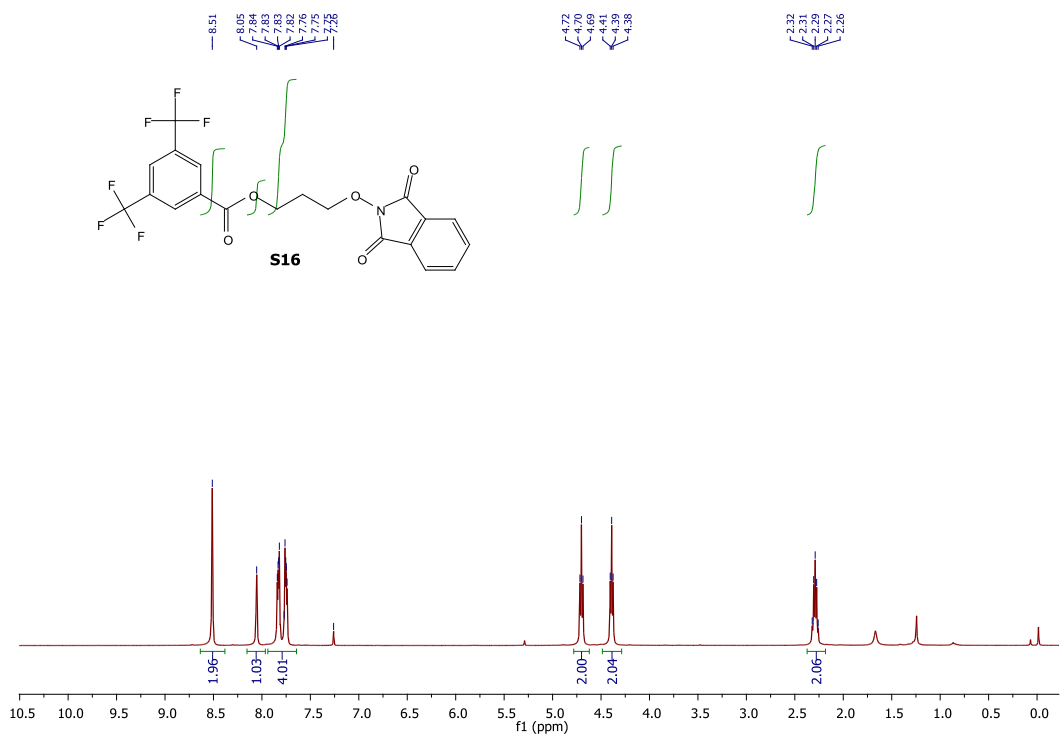


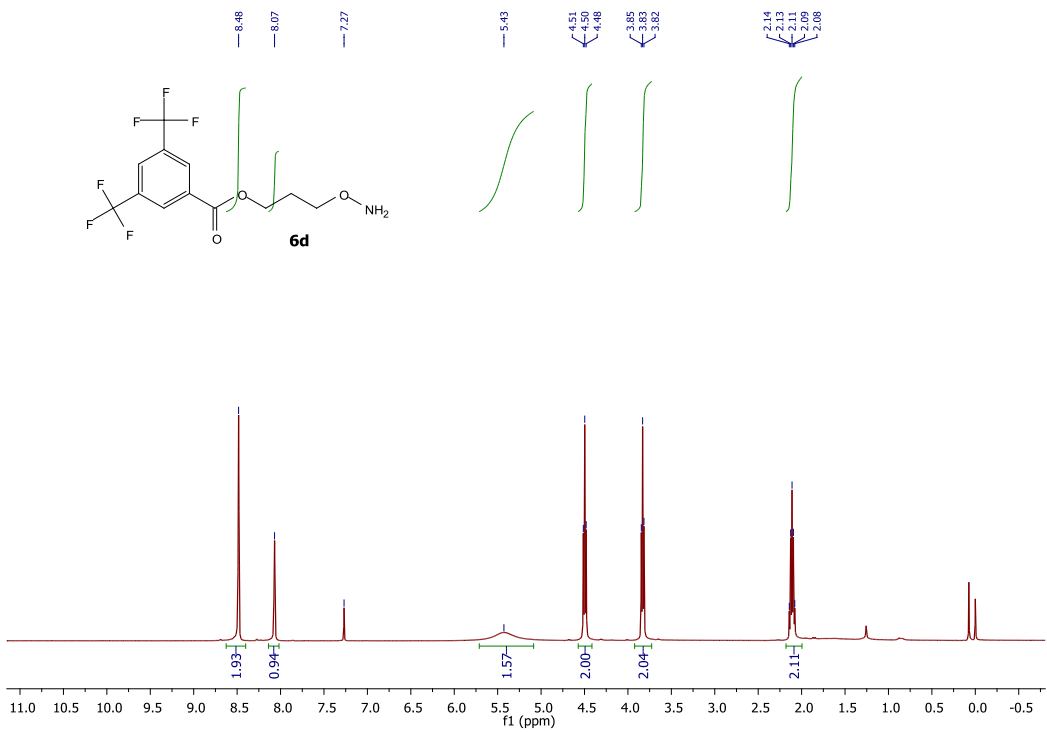
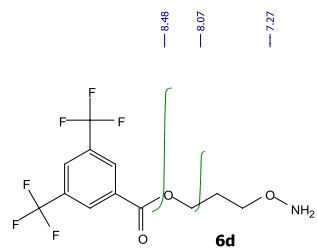
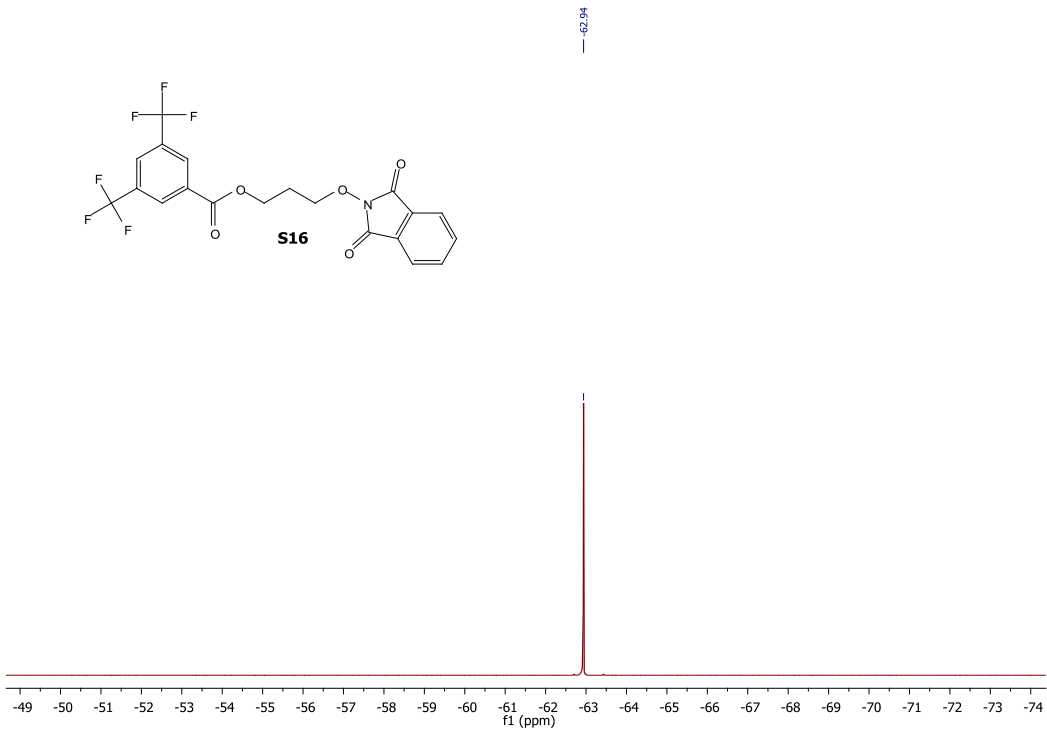
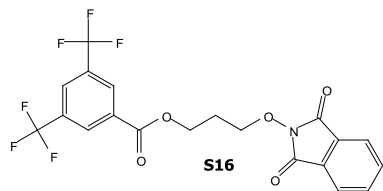


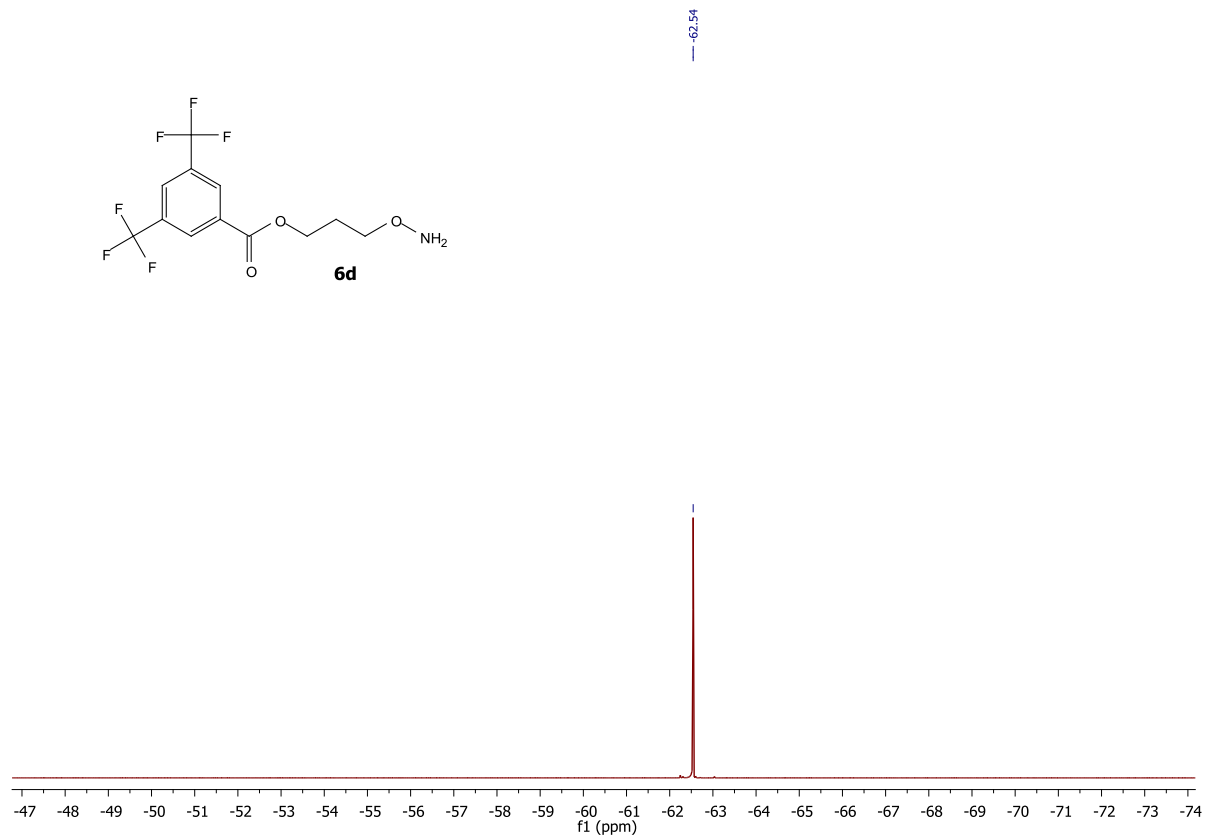
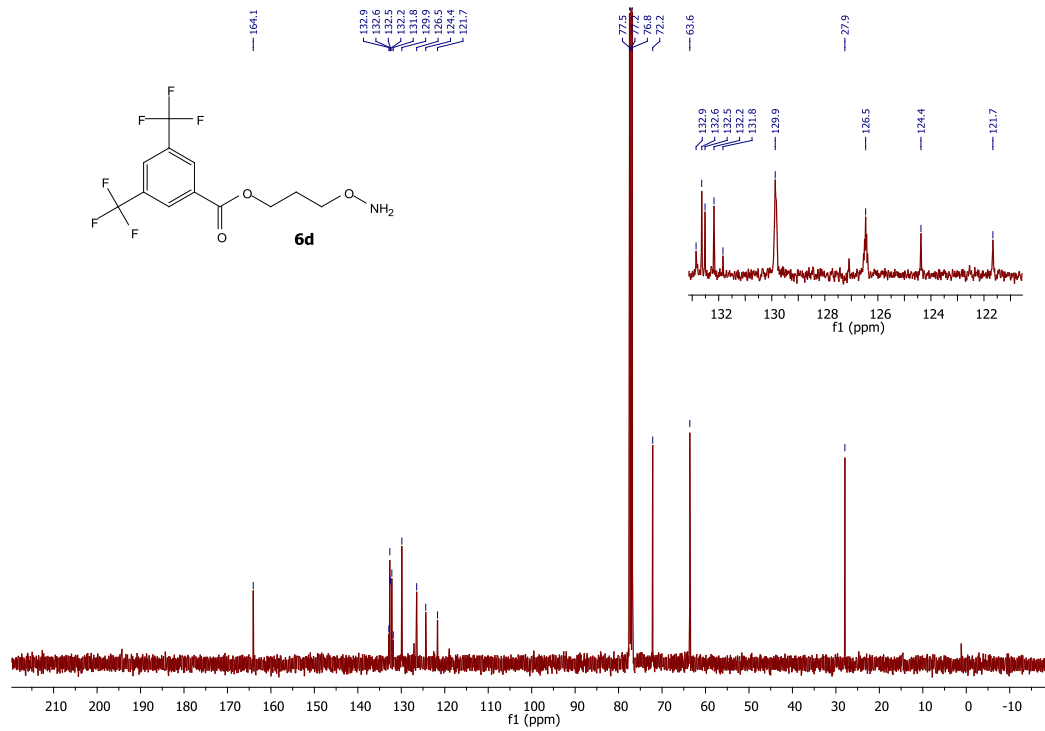












11. Acknowledgements

M.C. is recipient of research fellowship from UGC, India. The authors are thankful to SERB, DAE, DBT, Ramanujan Fellowship, and IISER Bhopal for research grants.

12. References

- ¹ M. Kinter, N. E. Sherman, *Wiley Interscience*, **2000**.
- ² L. Purushottam, S. R. Adusumalli, M. Chilamari, V. Rai, *Chem. Commun.* **2017**, *53*, 959-962
- ³ J. M. McFarland, M. B. Francis, *J. Am. Chem. Soc.* **2005**, *127*, 13490-13491.
- ⁴ F. Ito, S. Ando, M. Iuchi, T. Ukari, M. Takasaki, K. Yamaguchi, *Tetrahedron* **2011**, *67*, 8009-8013.
- ⁵ a) Z. Karimi-Jaberi, M. Amiri, *Heteroatom Chem.* **2010**, *21*, 96-98; b) M. Milen, P. Abranyi-Balogh, A. Dancso, D. Frigyes, L. Pongo, G. Keglevich, *Tetrahedron Lett.* **2013**, *54*, 5430-5433.
- ⁶ R. A. Cherkasov, *et al. Russ. Chem. Rev.* **1998**, *67*, 857-882.
- ⁷ Link for the Lysozyme C enzymatic assay by using the *Micrococcus lysodeikticus*
<http://www.sigmaaldrich.com/technical-documents/protocols/biology/enzymatic-assay-of-lysozyme.html>
- ⁸ Link for the RNase A enzymatic assay by using the Ribonucleic acid
<http://www.sigmaaldrich.com/technical-documents/protocols/biology/enzymatic-assay-of-ribonuclease-a.html>
- ⁹ F. F. Schumacher, J. P. M. Nunes, A. Maruani, C. Vijay, M. E. B. Smith, K. A. Chester, J. R. Baker, S. Caddick, *Org. Biomol. Chem.* **2014**, *12*, 7261-7269.

11.2. Additional references

- ¹ **ADC references:** a) H. Merten, F. Brandl, A. Pluckthun, U. Zangemeister-Wittke, *Bioconjugate Chem.* **2015**, *26*, 2176-2185; b) N. Diamantis, U. Banerji, *Br. J. Cancer*, **2016**, *114*, 362-367; c) J. M. Lambert, R. V. J. Chari, *J. Med. Chem.* **2014**, *57*, 6949-6964; d) N. Zhang, M. E. Klegerman, H. Deng, Y. Shi, E. Golunski, Z. An, *J. Cancer Ther.* **2013**, *4*, 308-322; e) M. R. Gordon, M. Canakci, L. Li, J. Zhuang, B. Osborne and S. Thayumanavan, *Bioconjugate Chem.*, **2015**, *26*, 2198; ; f) P. Agarwal and C. R. Bertozzi, *Bioconjugate Chem.*, **2015**, *26*, 176.; g)) P. M. LoRusso, D. Weiss, E. Guardino, S. Girish and M. X. Sliwkowski, *Clin. Cancer Res.*, **2011**, *17*, 6437.
- ² **Tyrosine modification:** T. L. Schlick, Z. B. Ding, E. W. Kovacs, M. B. Francis, *J. Am. Chem. Soc.* **2005**, *127*, 3718-3723.
- ³ **Selective modification of protein:** a) E. M. Sletten, C. R. Bertozzi, *Angew. Chem. Int. Ed.* **2009**, *48*, 6974-6998; b) C. P. R. Hackenberger and D. Schwarzer, *Angew. Chem. Int. Ed.*, **2008**, *47*, 10030; c) O. Boutureira and G. J. L. Bernardes, *Chem. Rev.*, **2015**, *115*, 2174.
- ⁴ **Chemoenzymatic methods for selective labeling:** a) P. Dennler, A. Chiotellis, E. Fischer, D. Bregeon, C. Belmont, L. Gauthier, F. Lhospice, F. Romagne, R. Schibli, *Bioconjugate Chem.* **2014**, *25*, 569-578; b) S. Jeger, K. Zimmermann, A. Blanc, J. Grnberg, M. Honer, P. Hunziker, H. Struthers, R. Schibli, *Angew. Chem. Int. Ed.* **2010**, *49*, 9995-9997. c) M. Sunbul, J. Yin, *Org. Biomol. Chem.* **2009**, *7*, 3361-3371; d) C. B. Rosen, R. L. Kwant, J. I. MacDonald, M. Rao, M. B. Francis, *Angew. Chem. Int. Ed.* **2016**, *55*, 8585-8589; e) M. J. Appel and C. R. Bertozzi, *ACS Chem. Biol.*, **2015**, *10*, 72-84; f) D. A. Shannon and E. Weerapana, *Curr. Opin. Chem. Biol.*,

-
- 2015, **24**, 18; g) M. J. Appel and C. R. Bertozzi, *ACS Chem. Biol.*, 2015, **10**, 72-84; h) S. C. Larda, D. Pichugin and R. S. Prosser, *Bioconjugate Chem.*, 2015, **26**, 2376.
- ⁵ **Site-directed mutagenesis:** a) D. Shinmi, E. Taguchi, J. Iwano, T. Yamaguchi, K. Masuda, J. Enokizono, Y. Shiraishi, *Bioconjugate Chem.* **2016**, *27*, 1324-1331; b) J. R. Junutula, H. Raab, S. Clark, S. Bhakta, D.D. Leipold, S. Weir, Y. Chen, M. Simpson, S.P. Tsai, M. S. Dennis, Y. Lu, Y. G. Meng, C. Ng, J. Yang, C. C. Lee, E. Duenas, J. Gorrell, V. Katta, A. Kim, K. McDorman, K. Flagella, R. Venook, S. Ross, S. D. Spencer, W. L. Wong, H. B. Lowman, R. Vandlen, M. X. Sliwowski, R. H. Scheller, P. Polakis, W. Mallet, *Nat. Biotechnol.* **2008**, *26*, 925-932; c) C. Zhang, M. Welborn, T. Zhu, N. J. Yang, M. S. Santos, T. V. Voorhis, B. L. Pentelute, *Nat. Chem.* **2016**, *8*, 120-128. d) N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **2004**, *126*, 15046-15047; e) Q. Wang, T. R. Chan, R. Hilgraf, V. V. Fokin, K. B. Sharpless, M. G. Finn, *J. Am. Chem. Soc.* **2003**, *125*, 3192-3193; f) C. Hoppmann, I. Maslennikov, S. Choe, L. Wang, *J. Am. Chem. Soc.* **2015**, *137*, 11218-11221; g)) C. D. Spicer and B. G. Davis, *Nat. Commun.*, 2014, **5**, 4740.
- ⁶ **Nonsense codon suppression:** (a) L. K. Mahal, K. J. Yarema and C. R. Bertozzi, *Science*, 1997, **276**, 1125; (b) A. Dumas, L. Lercher, C. D. Spicer and B. G. Davis, *Chem. Sci.*, 2015, **6**, 50.
- ⁷ **N-terminus modification:** a) A. O.-Y. Chan, C. M. Ho, H.-C. Chong, Y.-C. Leung, J.-S. Huang, M.-K. Wong, C.-M. Che, *J. Am. Chem. Soc.* **2012**, *134*, 2589-2598; b) J. I. MacDonald, H. K. Munch, T. Moore and M. B. Francis. *Nat. Chem. Biol.*, 2015, **11**, 326; c) X. Chen, K. Muthoosamy, A. Pfisterer, B. Neumann and T. Weil, *Bioconjugate Chem.*, 2012, **23**, 500.
- ⁸ ADC references: (a) E. L. Sievers, P. D. Senter, *Annu. Rev. Med.* **2013**, *64*, 15-29; (b) R. V. J. Chari, M. L. Miller, W. C. Widdison, *Angew. Chem. Int. Ed.* **2014**, *53*, 3796-3827; (c) P. Agarwal, C. R. Bertozzi, *Bioconjugate Chem.* **2015**, *26*, 176-192; (d) V. Chudasama, A. Maruani, S. Caddick, *Nat. Chem.* **2016**, *8*, 114-119; (e) A. Beck, L. Goetsch, C. Dumontet, N. Corvaia, *Nat. Rev. Drug Discovery* **2017**, *16*, 315-337; (f) D. Slamon, W. Eiermann, N. Robert, T. Pienkowski, M. Martin, M. Press, J. Mackey, J. Glaspy, A. Chan, M. Pawlicki, T. Pinter, V. Valero, M.-C. Liu, G. Sauter, G. von Minckwitz, F. Visco, V. Bee, M. Buyse, B. Bendahmane, I. Tabah-Fisch, M.-A. Lindsay, A. Riva, J. Crown, *N. Engl. J. Med.* **2011**, *365*, 1273-1283; (g) X. Chen, K. Muthoosamy, A. Pfisterer, B. Neumann and T. Weil, *Bioconjugate Chem.*, 2012, **23**, 500.
- ⁹ **Biologics:** G. Walsh and R. Jefferis, *Nat. Biotechnol.*, 2006, **24**, 1241.
- ¹⁰ **Peptide directed modification:** Z. T. Ball, *Acc. Chem. Res.*, 2013, **46**, 560.
- ¹¹ S. C. Larda, D. Pichugin and R. S. Prosser, *Bioconjugate Chem.*, 2015, **26**, 2376.
- ¹² Methionine oxidation: R. T. Dean, S. Fu, R. Stocker and M. J. Davies, *Biochem. J.*, 1997, **324**, 1.
- ¹³ **Phospha-Mannich reaction:** R. J. Sundberg and Jr. R. H. Smith, *J. Org. Chem.*, 1971, **36**, 295.