Linchpin empowers promiscuous electrophile to enable site-selective modification of histidine and aspartic acid in proteins

Dattatraya Gautam Rawale,a Kalyani Thakur,a Pranav Sreekumar,a Sajeev T. K.,b Ramesh A.,a Srinivasa Rao Adusumalli,a Ram Kumar Mishra,b and Vishal Rai*a

aDepartment of Chemistry and bDepartment of Biological Sciences
Indian Institute of Science Education and Research Bhopal

Contents

1. General information .................................................................................................................. S2
2. General procedures ................................................................................................................... S4
   2a. Protein labeling ..................................................................................................................... S4
   2b. Protein digestion ................................................................................................................... S5
   2c. Insulin bioactivity assay ...................................................................................................... S5
   2d. Docking investigations ....................................................................................................... S6
3. Synthesis and characterization data of reagents ....................................................................... S8
   3a. Alkylating reagents 2a-2f ................................................................................................. S8
   3b. Derivatives of O-hydroxylamine ........................................................................................ S18
4. Protein labeling data ............................................................................................................... S24
   4a. Control reaction of reagent (2a) with native ubiquitin ........................................................ S24
   4b. Control reaction of reagent (2a) with native proteins ........................................................ S26
   4c. Single-site labeling of native proteins .............................................................................. S27
   4d. Tagging of labeled ubiquitin ............................................................................................. S36
   4e. Purification of labeled ubiquitin ........................................................................................ S38
5. Additional data ........................................................................................................................ S38
   5a. Protein labeling ................................................................................................................... S38
   5b. Circular dichroism spectra of protein ................................................................................ S50
   5c. Fluorescence spectra of coumarin tagged ubiquitin .......................................................... S50
6. Spectral data ........................................................................................................................... S51
7. Additional discussion .............................................................................................................. S77
8. Protein sequence ..................................................................................................................... S93
9. References .............................................................................................................................. S94
1. General information

The reagents, proteins, and enzymes were purchased from Sigma-Aldrich, Alfa Aeser, Spectrochem, and Merck. The organic solvents used were reagent grade. Aqueous buffers were prepared freshly using Millipore Grade I water (Resistivity > 5 MΩ cm, Conductivity < 0.2 μS/cm, TOC <30 ppb). Mettler Toledo (FE20) pH meter was used to adjust the final pH. The reaction mixture for the small molecules was stirred (Heidolph, 500-600 rpm). Proteins were either vortexed or incubated in Thermo Scientific MaxQ 8000 incubator shaker (350 rpm, 25-37 °C). Cellulose membrane (MWCO, 6-8 kD) from Spectrum labs was used for dialysis. Amicon® Ultra-0.5 mL 3-10 kDa MWCO centrifugal filters from Merck Millipore was used to remove small molecules from protein mixture, desalting, and buffer exchange. Organic solvents were removed by BUCHI rotavapor R-210/215 whereas aqueous samples were lyophilized by CHRiST ALPHA 2-4 LD plus lyophilizer. Circular Dichroism (CD) measurements were recorded on JASCO J-815 CD spectropolarimeter equipped with peltier temperature controller. All the spectra were measured with a scan speed of 50 nm/min, spectral band width 1 nm using 1 cm path length cuvette at 25 °C. Steady-state fluorescence spectra was carried out in HORIBA JOBIN YVON, FLUOROLOG 3-111. The fluorescence spectra were measured with a quartz cuvette of 1 cm path length.

**Chromatography:** Thin-layer chromatography (TLC) was performed on silica gel coated aluminium TLC plates (Merck, TLC Silica gel 60 F254). The compounds were visualized using a UV lamp (254 nm) and stains such as iodine, ninhydrin, 2,4-dinitrophenylhydrazine. Wherever compounds were purified by chromatography, flash column chromatography was carried out on Combiflash Rf 200, Combiflash NextGen 300+, or gravity columns using silica gel (230-400 or 100-200 mesh) from Merck.

**Nuclear magnetic resonance spectra:** $^1$H, $^{13}$C, $^{19}$F NMR spectra were recorded on Bruker Avance III 400 and 500 MHz NMR spectrometer. $^1$H, $^{13}$C, and $^{19}$F spectra NMR spectra were referenced to TMS (0 ppm), CDCl$_3$ (77.16 ppm), and trifluoroacetic acid (-75.76 ppm) respectively. Peak multiplicities are designated by the following abbreviations: s, singlet; bs, broad singlet; d, doublet; t, triplet; q, quartet; m, multiplet; dd, doublet of doublets; ddd, doublet of doublet of doublets. Spectra were recorded at 298 K.

**Mass spectrometry:** SCIEX X500B QTOF coupled with ExionLC AD UHPLC and Agilent 6130 single quad coupled with Agilent 1200 series HPLC (ESI/APCI) were used for LC-MS and protein sequencing. Poroshell 300 SB-C18 HPLC column (2.1 × 75 mm × 5 μm, flow rate 0.4 ml/min) and XB-C18 UHPLC column (2.5 x 150 mm, 1.7 μm, 100 Å, flow rate 0.3 ml/min) were used for small molecules and protein-derived samples respectively. HRMS data were recorded on Bruker Daltonics MicroTOF-Q-III with electron spray ionization (ESI). Matrix assisted laser desorption/ionisation time of flight mass spectrometry was performed with Bruker Daltonics UltralFleXtreme Software-Flex control version 3.4, using sinapic acid and...
α-cyano-4-hydroxycinnamic acid (HCCA) matrix. Data analysis was performed using SCIEX Bio-pharma view Flex, Flex analysis, and Bruker data analysis software. Peptide mass and fragment ion calculator (http://db.systemsbiology.net:8080/proteomicsToolkit/FragIonServlet.html) were used for peptide mapping and sequencing.

Acetonitrile and H$_2$O were buffered with 0.01% formic acid and used as the mobile phase. Method A was used to record the LC-ESI-MS data for proteins and method B used for peptide mapping and MS/MS.

**Table S1.**

Method A (Column: Agilent, Poroshell 300 Å, SB-C18 5 µm 2.1 × 75 mm, flow rate 0.4 ml/minutes)

| Time (minutes) | H$_2$O (%) | Acetonitrile (%) |
|----------------|------------|------------------|
| 0              | 90         | 10               |
| 1              | 90         | 10               |
| 8              | 40         | 60               |
| 12             | 10         | 90               |
| 15             | 10         | 90               |

**Table S2.**

Method B (Column: Phenomenex, Poroshell 100 Å, 2.5 x 150 mm, 1.7 µm, flow rate 0.3 ml/min)

| Time (minutes) | H$_2$O (%) | Acetonitrile (%) |
|----------------|------------|------------------|
| 0              | 95         | 5                |
| 2              | 95         | 5                |
| 25             | 50         | 50               |
| 26             | 20         | 80               |
| 28             | 95         | 5                |
| 30             | 95         | 5                |

**Reaction conversion determination for protein labeling**

*ESI-MS:* Conversion for protein labeling was calculated based on the relative peak intensity of native protein and labeled protein in the deconvoluted mass spectrum.

\[
\% \text{ Conversion} = \frac{I_{\text{desired product}}}{I_{\text{all relevant species}}} \]

where $I_{\text{desired product}}$ is the peak intensity of labeled protein, and $I_{\text{all relevant species}}$ is the sum of the peak intensities of native protein and labeled protein in the deconvoluted mass spectra.

*MALDI-ToF-MS:* Conversion for protein labeling was calculated based on the relative peak intensity of native protein and labeled protein in the mass spectrum.
2. General procedures

2a. Protein labeling

Site-selective modification of native proteins

Protein 1 (10 nmol) in phosphate buffer (140 µl, 0.1 M, pH 7.8) was taken in a 2 ml microcentrifuge tube. Reagent 2 (1.25 mM) in DMSO (60 µl) from a freshly prepared stock solution was added to it followed by vortexing (350 rpm) at 37 °C. After 4-72 h, the reaction mixture was diluted with acetonitrile:buffer (10:90, 3600 µl). Unreacted reagent and salts were removed by using Amicon® Ultra-4 mL 3-kDa or 10-kDa MWCO centrifugal filters spin concentrator. The sample was further washed with buffer (2 × 4 ml) and concentrated to 190 µl. To this solution, hydroxylamine 4 (5 µmol) in water (10 µl) from a freshly prepared stock solution was added. The reaction mixture was vortexed at 25 °C for 15 minutes to form oxime derivative. The excess of hydroxylamine and salts were removed by using spin concentrator and the sample was collected in water. The modification of protein was analyzed by ESI-MS and MALDI-ToF-MS. The concentrated sample was subjected to digestion, peptide mapping, and sequencing by MS-MS.

Late-stage tagging of labeled ubiquitin

The labeled protein 3a/3b was prepared according to the procedure given above. The labeled protein 3a/3b was concentrated to 160 µl of phosphate buffer (0.1 M, pH 7.0). The O-hydroxylamine derivative 6a/6b/6c (1 µmol) in DMSO (40 µl) was added to convert the mono-labeled ubiquitin 3a/3b into its oxime derivatives (7a/7b/7c). The excess of O-hydroxylamine derivative and salts were removed by the spin concentrator (4 ml, 3-kDa MWCO). The tagged protein (7a-7c) was analyzed by ESI-MS.

Purification of the labeled ubiquitin from reaction mixture

Hydrazide beads 9 (200 µl, hydrazide resin loading: 16 µmol/ml) were taken in a 5 ml fritted polypropylene chromatography column with end tip closures. The beads were re-suspended in phosphate buffer (100 µl) after washing with phosphate buffer (1 M, pH 7.0, 6 x 1 ml). Protein mixture (1a and labeled protein 3a/3b, 200 µM) in phosphate buffer (200 µl, 1 M, pH 7.0) and p-phenylenediamine (p-PDA, 100 mM) in phosphate buffer (100 µl, 1 M, pH 7.0) were added to the beads followed by end-to-end rotation (30 rpm, rotary mixer) at 25 °C for 4 h. The immobilization of the labeled protein on hydrazide resin was monitored by UV-absorbance of the supernatant (complete loading of labeled protein was also confirmed by ESI-MS of supernatant). After collecting the supernatant, the beads were washed with phosphate buffer (0.3 M, pH 7.3, 5 x 1 ml) and KCl (1 M, 4 x 1 ml) to remove the adsorbed protein from resin. The beads were further washed with buffer (0.3 M, pH 7.0, 5 x 1 ml) and re-suspended in phosphate buffer (200 µl, 0.3 M, pH 7.0).
The labeled protein from its immobilized derivative was released by adding O-hydroxylamine hydrochloride (50 μl, 1 M in buffer, 0.3 M, pH 7.0) (end-to-end rotation at 25 °C for 2 h). The supernatant was collected and then the salts, p-PDA and O-hydroxylamine were removed using the spin concentrator (3 kDa or 10-kDa MWCO). The purity of the labeled protein 5a was confirmed by ESI-MS.

2b. Protein digestion

All solutions were made freshly prior to use.1

**Procedure for in-solution digestion of insulin and α-lactalbumin**

In a 1.5 ml microcentrifuge tube, protein (0.1 mg) in 100 mM tris (10 μl, pH 7.8) with urea (6 M) was incubated for 30 minutes at 37 °C. To this solution, reducing agent (1 μl, 0.2 M DTT in 0.1 M tris) was added and sample was incubated for 1 h at 37 °C. Alkylating agent (4 μl, 0.2 M iodoacetamide in 0.1 M tris) was added to the solution and incubated (in dark) for 1 h at 25 °C for blocking the free sulphhydryl groups. The unreacted iodoacetamide was quenched with reducing agent (4 μl, 0.2 M DTT in 0.1 M tris) for 1 h at 25 °C. The sample was desalted using Amicon® Ultra-1.5 mL 3-kDa spin concentrator. The sample was collected in 100 μl grade I water. To this solution, 5 μl of enzyme (α-chymotrypsin/trypsin) solution [5 μg, enzyme/protein (1:20); enzyme was dissolved in grade I water] was added and the mixture was incubated at 37 °C for 12 h. Subsequently, the sample was used for peptide mapping by MS and sequencing by MS-MS.

**Procedure for in-solution digestion of ubiquitin and myoglobin**

Protein (0.1 mg) in 100 mM tris (10 μl, pH 7.8) with urea (6 M) was taken in a 1.5 ml microcentrifuge tube. To this solution, tert-butanol (10 μl) was added and incubated for 3 h at 37 °C. The sample was desalted using Amicon® Ultra-1.5 mL 3-kDa spin concentrator. The sample was collected in 100 μl grade I water. To this solution, 5 μl of enzyme (α-chymotrypsin/trypsin) solution [5 μg, enzyme/protein (1:20); enzyme was dissolved in grade I water] was added and the mixture was incubated at 37 °C for 12 h. Subsequently, the sample was used for peptide mapping by MS and sequencing by MS-MS.

2c. Insulin bioactivity assay

**Immunofluorescence**

HEK293T cells were grown on coverslips in a six-well plate format in Dulbecco’s modified Eagle’s medium (DMEM) and serum-starved for 24 h in 1% serum-containing DMEM medium. Subsequently, cells were treated with 50 μM LY294002 for 30 min. After LY294002 treatment, cells were washed twice
with PBS and treated with either native insulin or purified insulin (3 μg each) in 1 ml of 1% FBS containing DMEM media for 30 min. Post-treatment, cells were again washed twice with PBS and fixed using 100% chilled methanol for 15 min at 4 °C. The cells were then rehydrated and permeabilized with rehydration buffer (10 mM Tris, 150 mM NaCl, 0.1% Triton X-100) for 10 min. Cells were blocked with 5% Normal Goat Serum (NGS) for 1 h at 4 °C after rehydration. The cells were stained for 2 h with pAkt antibodies (1:600, anti-pAKT-S473, CST #4060) at 4 °C. After primary antibody incubation, cells were washed three times with PBS-T (5 min each) and incubated with Alexa Fluor-568 conjugated goat anti-rabbit IgG (1:1000, Life Technologies) for 1 h. Cells were washed thrice with PBS-T and mounted on slides using DAPI containing mounting medium (SIGMA F6057). Fluorescence signals were captured on Zeiss LSM 780 confocal microscope, and images were analyzed using Image J software.

**Western blotting**

Cells were cultured and treated as mentioned in the immunofluorescence section. Post-treatment, cells were washed twice with PBS and lysed directly in the 1X Laemmli buffer (containing 1 mM Sodium orthovanadate and 1 mM PMSF) by boiling at 100 °C for 10 min. The cleared lysates were obtained by centrifugation and analyzed on 8% SDS-PAGE. Following wet transfer protocols, proteins were transferred onto the methanol-activated 0.2 μm PVDF membrane (MERCK, Cat #ISEQ85R) using 1X transfer buffer (2.5 mM Tris-HCl pH 7.5, 19.2 mM Glycine). The membrane was blocked for 1 h in 5% BSA. Further, the membrane was incubated overnight with anti-pAkt-S473 (1:2000, CST #4060) and anti-Actin (1:8000, BD #612656) antibodies. The membrane was later washed three times for 10 min each with TBS-T buffer (20 mM Tris HCl pH 7.5, 150 mM NaCl, 0.1% Tween-20). The membrane was incubated with Alexa fluor Plus 680 secondary antibodies (Invitrogen #A32734) for 1 h. Later, it was washed with TBS-T buffer (three times for 10 min each). Images were taken using LI-COR IR system. Anti-Actin (1:8000, BD #612656) antibodies were used for loading control detection.

**2d. Docking investigations**

**Coordinate Preparation:** Docking of reagent 2c to ubiquitin's surface was performed by using AutoDock suite version 4.2. The coordinates were obtained from protein data bank (www.rcsb.org) with PDB ID: 1UBQ. The ligand coordinate file was prepared after structural optimization with semi-empirical calculations. For the PDB file, we deleted the water molecules and added hydrogen atoms. The protonation state and valency of heteroatoms in both the protein and ligand PDB files were verified. Using AutoDockTools (ADT) GUI, the coordinates were read, the Gasteiger-Marsili charges of the molecules were computed, non-polar hydrogen atoms were merged, and atom types were assigned. The edited protein
molecule was saved as a PDBQT file. After customizing the torsional degrees of freedom of the ligand (12 rotatable bonds and 17 non-rotatable bonds), an output file in the PDBQT format was generated.

**AutoGrid:** The PDBQT files for both protein and ligand were read by ADT. The searching space for docking was defined by generating a grid box that covers the entire protein surface. The map types were determined by running AutoGrid.

**AutoDock:** The ligand and macromolecule for docking simulation were specified by selecting the corresponding PDBQT files. The Genetic Algorithm (GA) search method was employed with 1000 GA runs. The docking parameter file for GA was taken as a DPF output file. Docking simulation was performed using this file, and the results were analyzed using ADT. The optimum energy conformation was written as a PDB file and visualized in PyMOL along with the protein. The image files were rendered using PyMOL.
3. Synthesis and characterization data of reagents

3a. Alkylating reagents 2a-2f

\[ \text{Scheme S1. Synthesis of 6-phenoxyhexyl 4-methylbenzenesulfonate (2a)} \]

\[ \text{Synthesis of ((6-bromohexyl)oxy)benzene (S13)} \]

\[ \text{In a 25 ml round bottom flask, phenol S11 (188 mg, 2 mmol), potassium carbonate (276 mg, 2 mmol), sodium iodide (298 mg, 2 mmol), and 1,6-dibromohexane S12 (303 µl, 2 mmol) were dissolved in 10 ml acetone. The reaction mixture was refluxed and the progress of reaction was monitored by TLC. After 16 h, reaction mixture was filtered, concentrated, and purified by silica gel flash column chromatography using n-hexane to isolate S13 (382 mg, 55% yield; Rf 0.69, ethyl acetate:n-hexane 10:90; colorless liquid).} \]

\[ \text{\textsuperscript{1}H NMR (500 MHz, CDCl}_3\text{) } \delta \text{ 7.33-7.23 (m, 2H), 6.96-6.91 (m, 1H), 6.96-6.86 (m, 2H), 3.96 (t, J = 6.4 Hz, 2H), 3.42 (t, J = 6.8 Hz, 2H), 1.94-1.86 (m, 2H), 1.85-1.76 (m, 2H), 1.56-1.46 (m, 4H) ppm.} \]

\[ \text{\textsuperscript{13}C NMR (126 MHz, CDCl}_3\text{) } \delta \text{ 159.1, 129.5, 120.6, 114.6, 67.7, 33.9, 32.8, 29.2, 28.0, 25.4 ppm. LRMS (ESI) [M+H]	extsuperscript{+} calcd. C}_{12}\text{H}_{18}\text{BrO 257.1, found 257.1.} \]

\[ \text{Synthesis of 6-phenoxyhexyl 4-methylbenzenesulfonate (2a)} \]

\[ \text{In a 10 ml round bottom flask, ((6-bromohexyl)oxy) benzene S13 (280 mg, 1.1 mmol), silver tosylate S14 (337 mg, 1.21 mmol) were taken and dissolved in 5 ml acetone to reflux. The progress of reaction was monitored by TLC. After 12 h, the reaction mixture was filtered, concentrated, and purified by silica gel flash column chromatography (ethyl acetate:n-hexane, 6:94) to isolate 2a (340 mg, 89% yield; Rf 0.25, ethyl acetate:n-hexane 10:90; White solid).} \]

\[ \text{\textsuperscript{1}H NMR (500 MHz, CDCl}_3\text{) } \delta \text{ 7.79 (d, J = 8.3 Hz, 2H), 7.33} \]
(d, J = 8.3 Hz, 2H), 7.30-7.24 (m, 2H), 6.96-6.90 (m, 1H), 6.90-6.84 (m, 2H), 4.04 (t, J = 6.5 Hz, 2H), 3.91 (t, J = 6.4 Hz, 2H), 2.44 (s, 3H), 1.78-1.62 (m, 4H), 1.47-1.33 (m, 4H) ppm. **¹³C NMR** (126 MHz, CDCl₃) δ 159.1, 144.8, 133.3, 129.9, 129.6, 128.1, 120.7, 114.6, 70.6, 67.6, 29.2, 28.9, 25.6, 25.3, 21.7 ppm. **HRMS** (ESI) [M+Na]+ calcd. C₁₉H₂₄O₄SNa 371.1293, found 371.1301.

**Scheme S2.** 3-(4-formyl-3-hydroxyphenoxy)propyl 4-methylbenzenesulfonate (2b)

**Synthesis of 4-(3-bromopropoxy)-2-hydroxybenzaldehyde (S17)**

In a 25 ml round bottom flask, 2, 4 dihydroxy phenol S15 (500 mg, 3.6 mmol), potassium carbonate (500 mg, 3.6 mmol), sodium iodide (536 mg, 3.6 mmol), and 1,3-dibromohexane S16 (377 µl, 3.6 mmol) were dissolved in 16 ml acetone. The reaction mixture was refluxed and the progress of reaction was monitored by TLC. After 16 h, reaction mixture was filtered, concentrated, and purified by silica gel flash column chromatography using n-hexane ethyl acetate to isolate S17 (472 mg, 51% yield; Rf 0.53, ethyl acetate:n-hexane 12:88; white solid). **¹H NMR** (400 MHz, CDCl₃) δ 11.46 (s, 1H), 9.72 (s, 1H), 7.44 (d, J = 8.7 Hz, 1H), 6.54 (dd, J = 8.7, 2.3 Hz, 1H), 6.44 (d, J = 2.3 Hz, 1H), 4.17 (t, J = 5.8 Hz, 2H), 3.59 (t, J = 6.4 Hz, 2H), 2.40-2.27 (m, 2H) ppm. **¹³C NMR** (101 MHz, CDCl₃) δ 194.5, 165.9, 164.6, 135.4, 115.4, 108.6, 101.4, 65.9, 32.0, 29.5 ppm. **LRMS** (ESI) [M+H]+ calcd. C₁₀H₁₁BrO₃ 258.9, found 258.3.

**Synthesis of 3-(4-formyl-3-hydroxyphenoxy)propyl 4-methylbenzenesulfonate (2b)**

In a 25 ml round bottom flask, 4-(3-bromopropoxy)-2-hydroxybenzaldehyde S17 (283 mg, 1.1 mmol), silver tosylate S14 (337 mg, 1.21 mmol) were taken and dissolved in 11 ml acetonitrile to reflux. The progress of reaction was monitored by TLC. After 12 h, the reaction mixture was filtered, concentrated,
and purified by silica gel flash column chromatography (ethyl acetate:n-hexane, 6:94) to isolate 2b (309 mg, 80% yield; Rf 0.35, ethyl acetate:n-hexane 15:85; White solid). 1H NMR (400 MHz, CDCl3) δ 11.44 (s, 1H), 9.72 (s, 1H), 7.76 (d, J = 8.2 Hz, 2H), 7.41 (d, J = 8.6 Hz, 1H), 7.27 (d, J = 8.2 Hz, 2H), 6.42 (dd, J = 8.6, 2.2 Hz, 1H), 6.27 (d, J = 2.2 Hz, 1H), 4.24 (t, J = 5.9 Hz, 2H), 3.99 (t, J = 5.8 Hz, 2H), 2.38 (s, 3H), 2.20-2.07 (m, 2H) ppm. 13C NMR (101 MHz, CDCl3) δ 194.5, 165.7, 164.4, 145.0, 135.3, 132.8, 129.9, 127.9, 115.4, 108.4, 101.4, 66.6, 63.7, 28.6, 21.7 ppm. HRMS (ESI) [M+Na]+ calcd. C17H18O6SNa 373.0722, found 373.0739.

Scheme S3. Synthesis of 6-(4-formyl-3-hydroxyphenoxy)hexyl 4-methylbenzenesulfonate (2c)

4-((6-bromoheptyl)oxy)-2-hydroxybenzaldehyde (S18)

In a 25 ml round bottom flask, 2,4-dihydroxybenzaldehyde S15 (276 mg, 2 mmol), potassium carbonate (276 mg, 2 mmol), sodium iodide (298 mg, 2 mmol), and 1,6-dibromohexane S12 (606 µl, 4 mmol) were dissolved in 10 ml acetone to reflux. The progress of reaction was monitored by TLC. After 16 h, the reaction mixture was filtered, concentrated and purified by silica gel flash column chromatography (ethyl acetate:n-hexane, 2:98) to isolate S18 (270 mg, 45% yield; Rf 0.51, white solid). 1H NMR (500 MHz, CDCl3) δ 11.48 (s, 1H), 9.70 (s, 1H), 7.42 (d, J = 8.7 Hz, 1H), 6.52 (dd, J = 8.5, 2.3 Hz, 1H), 6.40 (d, J = 2.3 Hz, 1H), 4.01 (t, J = 6.4 Hz, 2H), 3.42 (t, J = 6.8 Hz, 2H), 1.96-1.86 (m, 2H), 1.85-1.77 (m, 2H), 1.58-1.43 (m, 4H) ppm. 13C NMR (126 MHz, CDCl3) δ 194.4, 166.4, 164.6, 135.3, 115.2, 108.8, 101.2, 68.4, 33.8, 32.7, 28.9, 27.9, 25.3 ppm. HRMS (ESI) [M+H]+ calcd. C13H15BrO3 301.0439, found 301.0453.

6-(4-formyl-3-hydroxyphenoxy)hexyl 4-methylbenzenesulfonate (2c)
In a 5 ml round bottom flask of 4-((6-bromohexyl)oxy)-2-hydroxybenzaldehyde \textbf{S18} (360 mg, 1.2 mmol) and silver tosylate \textbf{S14} (363 mg, 1.3 mmol) were dissolved in 5 ml acetone. The reaction mixture was refluxed and the progress of reaction was monitored by TLC. After 12 h, the reaction mixture was filtered, concentrated, and purified by silica gel flash column chromatography (ethyl acetate:n-hexane, 15:85) to isolate \textbf{2c} (442 mg, 94% yield; Rf 0.57, white solid). \textbf{\textsuperscript{1}H NMR} (500 MHz, CDCl\textsubscript{3}) \(\delta\) 11.47 (s, 1H), 9.71 (s, 1H), 7.78 (d, \(J = 8.3\) Hz, 2H), 7.42 (d, \(J = 8.7\) Hz, 1H), 7.34 (d, \(J = 8.0\) Hz, 2H), 6.50 (dd, \(J = 8.7, 2.3\) Hz, 1H), 6.38 (d, \(J = 2.3\) Hz, 1H), 4.04 (t, \(J = 6.4\) Hz, 2H), 3.96 (t, \(J = 6.4\) Hz, 2H), 2.44 (s, 3H), 1.79-1.71 (m, 2H), 1.71-1.64 (m, 2H), 1.49-1.32 (m, 4H) ppm. \textbf{\textsuperscript{13}C NMR} (126 MHz, CDCl\textsubscript{3}) \(\delta\) 194.4, 166.3, 164.4, 144.8, 135.3, 133.1, 129.9, 127.8, 115.0, 108.6, 101.1, 70.5, 68.3, 28.7, 28.7, 25.3, 25.1, 21.6 ppm. \textbf{HRMS} (ESI) [M+H]\textsuperscript+ calcd. C\textsubscript{20}H\textsubscript{25}O\textsubscript{6}S 393.1372, found 393.1390.

\textbf{Scheme S4}. Synthesis of 6-(tosyloxy)hexyl 2-(2-(4-formyl-3-hydroxyphenoxy)acetamido)benzoate (2d)

**Synthesis of 6-bromohexyl 2-aminobenzoate (S20)**

In a 50 ml round bottom flask, anthranilic acid \textbf{S19} (411 mg, 3 mmol) was dissolved in acetone (15 ml) at room temperature followed by addition of 1,6-dibromohexane \textbf{S12} (686 µl, 4.5 mmol), sodium iodide (89 mg, 0.6 mmol), and potassium carbonate (828 mg, 6 mmol). The reaction mixture was heated to reflux for 24 h. The progress of reaction was monitored by TLC. Finally, the reaction mixture was cooled to room temperature, subjected to aqueous workup, and purification by silica gel column chromatography (ethyl acetate:n-hexane, 2:98) to isolate \textbf{S20} (585 mg, 64% yield; Rf 0.62, off-white solid). \textbf{\textsuperscript{1}H NMR} (500 MHz,
CDCl$_3$ δ 7.92 – 7.85 (m, 1H), 7.34 – 7.24 (m, 1H), 6.71 – 6.62 (m, 2H), 5.73 (s, 2H), 4.29 (t, $J = 6.6$ Hz, 2H), 3.43 (t, $J = 6.8$ Hz, 2H), 1.95 – 1.87 (m, 2H), 1.84 – 1.75 (m, 2H), 1.63 – 1.42 (m, 4H) ppm. $^{13}$C NMR (126 MHz, CDCl$_3$) δ 168.2, 150.6, 134.1, 131.2, 116.3, 111.0, 64.2, 33.8, 32.7, 28.7, 27.9, 25.4 ppm. HRMS (ESI) [M+Na]$^+$ calcd. C$_{13}$H$_{18}$BrNO$_3$ 322.0419, found 322.0393.

**Synthesis of 6-bromohexyl 2-(2-bromoacetamido)benzoate (S22)**

In a 10 ml round bottom flask, a stirred solution of 6-bromohexyl 2-aminobenzoate S20 (585 mg, 1.9 mmol) and potassium carbonate (538 mg, 3.9 mmol) in DCM:H$_2$O (4 ml) bromoacetyl bromide S21 (341 µl) were added slowly by pressure equalizer funnel at 0 °C and stirred for 2 h. The reaction mixture was brought to room temperature and stirred for another 2 h. The progress of reaction was monitored by TLC. Later, a saturated Na$_2$CO$_3$ solution was added to the reaction mixture followed by extraction by dichloromethane (50 ml, 3 times). Next, the organic layer was washed with brine solution and dried with Na$_2$SO$_4$. Subsequently, it was concentrated on a rotary evaporator and purified by silica gel flash column chromatography (ethyl acetate:hexane, 1:99) to render white solid product S22 (750 mg, 93%). $^1$H NMR (500 MHz, CDCl$_3$) δ 11.74 (s, 1H), 8.71 – 8.63 (m, 1H), 8.10 – 8.01 (m, 1H), 7.60 – 7.52 (m, 1H), 7.19 – 7.09 (m, 1H), 4.35 (t, $J = 6.6$ Hz, 2H), 4.01 (s, 2H), 3.41 (t, $J = 6.7$ Hz, 2H), 1.93 – 1.85 (m, 2H), 1.84 – 1.77 (m, 2H), 1.60 – 1.43 (m, 4H) ppm. $^{13}$C NMR (126 MHz, CDCl$_3$) δ 168.0, 164.9, 140.7, 134.6, 130.9, 123.4, 120.4, 116.0, 65.5, 33.7, 32.6, 29.7, 28.4, 27.8, 25.3 ppm. HRMS (ESI) [M+H]$^+$ calcd. C$_{15}$H$_{19}$BrNO$_3$ 419.9810, found 419.9812.

**Synthesis of 6-bromohexyl 2-(2-(4-formyl-3-hydroxyphenoxy)acetamido)benzoate (S23)**

In a 25 ml round bottom flask, 6-bromohexyl 2-(2-bromoacetamido) benzoate S22 (750 mg, 1.7 mmol), 2,4-dihydroxybenzaldehyde S15 (234 mg, 1.7 mmol), and potassium carbonate (234 mg, 1.7 mmol) were dissolved in acetone (8.5 ml) and heated to reflux for 8 h. The progress of reaction was monitored by TLC. The reaction mixture was concentrated on a rotary evaporator and purified by silica gel flash column chromatography (ethyl acetate:hexane, 4:96) to give compound S23 (410 mg, 50%). $^1$H NMR (500 MHz,
CDCl$_3$) δ 12.11 (s, 1H), 11.42 (s, 1H), 9.73 (s, 1H), 8.81 – 8.70 (m, 1H), 8.10 – 7.97 (m, 1H), 7.57 – 7.52 (m, 1H), 7.49 (d, $J = 8.6$ Hz, 1H), 7.16 – 7.09 (m, 1H), 6.74 (dd, $J = 8.6$, 2.3 Hz, 1H), 6.57 (d, $J = 2.3$ Hz, 1H), 4.65 (s, 2H), 4.34 (t, $J = 6.6$ Hz, 2H), 3.41 (t, $J = 6.7$ Hz, 2H), 1.95 – 1.84 (m, 2H), 1.84 – 1.72 (m, 2H), 1.58 – 1.41 (m, 4H) ppm. **$^{13}$C NMR** (126 MHz, CDCl$_3$) δ 194.7, 167.8, 165.9, 164.2, 164.0, 140.3, 135.7, 134.5, 130.8, 123.3, 120.4, 116.1, 115.9, 108.2, 102.4, 67.5, 65.3, 33.7, 32.6, 28.4, 27.8, 25.3 ppm.

**HRMS** (ESI) [M+Na]$^+$ calcd. C$_{22}$H$_{24}$BrNO$_6$ 500.0685, found 500.0706.

**Synthesis of 6-(tosyloxy)hexyl 2-(2-(4-formyl-3-hydroxyphenoxy)acetamido)benzoate (2d)**

In a 10 ml round bottom flask, 6-bromohexyl 2-(2-(4-formyl-3-hydroxyphenoxy)acetamido) benzoate S23 (239 mg, 0.5 mmol) and silver tosylate S14 (167 mg, 0.6 mmol) were added in acetonitrile (5 ml). The resultant solution was heated for reflux for 12 h. After completion, the reaction mixture was concentrated on a rotary evaporator followed by aqueous workup and extraction with dichloromethane. The organic fraction was dried and the crude reaction mixture was purified by flash column chromatography (ethyl acetate: n-hexane, 30:70) to give 2d (145 mg, 51%). **$^1$H NMR** (500 MHz, CDCl$_3$) δ 12.12 (s, 1H), 11.43 (s, 1H), 9.75 (s, 1H), 8.76 (d, $J = 8.5$ Hz, 1H), 8.08 – 7.99 (m, 1H), 7.77 (d, $J = 8.2$ Hz, 2H), 7.59 – 7.54 (m, 1H), 7.52 (d, $J = 8.6$ Hz, 1H), 7.33 (d, $J = 8.3$ Hz, 2H), 7.18 – 7.10 (m, 1H), 6.75 (dd, $J = 8.6$, 2.3 Hz, 1H), 6.58 (d, $J = 2.2$ Hz, 1H), 4.67 (s, 2H), 4.31 (t, $J = 6.6$ Hz, 2H), 4.03 (t, $J = 6.4$ Hz, 2H), 2.43 (s, 3H), 1.80 – 1.62 (m, 4H), 1.46 – 1.37 (m, 4H) ppm. **$^{13}$C NMR** (126 MHz, CDCl$_3$) δ 194.8, 167.9, 166.0, 164.3, 164.1, 144.8, 140.3, 135.8, 134.6, 133.2, 130.9, 129.9, 127.9, 123.4, 120.5, 116.2, 116.0, 108.3, 102.4, 70.4, 67.6, 65.3, 28.8, 28.4, 25.5, 25.2, 21.7 ppm. **HRMS** (ESI) [M+Na]$^+$ calcd. C$_{29}$H$_{31}$NO$_5$SNa 592.1617, found 592.1628.
Scheme S5. Synthesis of 6-(tosyloxy)hexyl 3-(2-(4-formyl-3-hydroxyphenoxy)acetamido)benzoate (2e)

The procedure for synthesis of LDM reagent 2e is similar to the LDM reagent 2d.

**Synthesis of 6-bromohexyl 3-aminobenzoate (S25)**

Yield 62%; Rf 0.48, white solid. $^1$H NMR (500 MHz, CDCl$_3$) δ 7.42 (d, $J$ = 7.7 Hz, 1H), 7.37 – 7.31 (m, 1H), 7.24 – 7.18 (m, 1H), 6.90 – 6.82 (m, 1H), 4.29 (t, $J$ = 6.6 Hz, 2H), 3.59 (s, $J$ = 12.9 Hz, 2H), 3.52 (s, $J$ = 6.8 Hz, 2H), 0.57, off-white solid. $^1$H NMR (500 MHz, CDCl$_3$) δ 8.34 (s, 1H), 8.07 – 8.01 (m, 1H), 7.96 – 7.87 (m, 1H), 7.87 – 7.80 (m, 1H), 7.49 – 7.40 (m, 1H), 4.33 (t, $J$ = 6.6 Hz, 2H), 4.05 (s, 2H), 3.42 (t, $J$ = 6.8 Hz, 2H), 1.94 – 1.83 (m, 2H), 1.83 – 1.73 (m, 2H), 1.62 – 1.39 (m, 4H) ppm. $^{13}$C NMR (126 MHz, CDCl$_3$) δ 168.5, 150.3, 134.5, 131.9, 116.5, 116.5, 111.3, 64.1, 33.6, 32.8, 28.4, 27.4, 25.2 ppm. HRMS (ESI) [M+H]$^+$ calcd. C$_{13}$HNO$_2$ 300.0599, found 300.0575.

**Synthesis of 6-bromohexyl 3-(2-bromoacetamido)benzoate (S26)**

Yield 85%; Rf 0.57, off-white solid. $^1$H NMR (500 MHz, CDCl$_3$) δ 8.34 (s, 1H), 8.07 – 8.01 (m, 1H), 7.96 – 7.87 (m, 1H), 7.87 – 7.80 (m, 1H), 7.49 – 7.40 (m, 1H), 4.33 (t, $J$ = 6.6 Hz, 2H), 4.05 (s, 2H), 3.42 (t, $J$ = 6.8 Hz, 2H), 1.94 – 1.83 (m, 2H), 1.83 – 1.73 (m, 2H), 1.62 – 1.39 (m, 4H) ppm. $^{13}$C NMR (126 MHz, CDCl$_3$) δ 166.2, 164.1, 137.2, 131.4, 129.4, 126.3, 124.7, 121.1, 65.3, 33.8, 32.7, 29.4, 28.6, 27.9, 25.3 ppm. HRMS (ESI) [M+H]$^+$ calcd. C$_{15}$H$_{10}$Br$_2$NO$_3$ 419.9810, found 419.9813.
Synthesis of 6-bromohexyl 3-(2-(4-formyl-3-hydroxyphenoxy)acetamido)benzoate (S27)

\[
\text{\includegraphics[width=0.2\textwidth]{image.png}}
\]

Yield 43%, off-white solid. \(^1\)H NMR (500 MHz, CDCl\(_3\)) \(\delta\) 11.44 (s, 1H), 9.78 (s, 1H), 8.26 (s, 1H), 8.09 – 8.03 (m, 1H), 8.03 – 7.96 (m, 1H), 7.87 – 7.81 (m, 1H), 7.54 (d, \(J = 8.6\) Hz, 1H), 7.49 – 7.41 (m, 1H), 6.67 (dd, \(J = 8.6, 2.4\) Hz, 1H), 6.54 (d, \(J = 2.4\) Hz, 1H), 4.68 (s, 2H), 4.33 (t, \(J = 6.6\) Hz, 2H), 3.42 (t, \(J = 6.8\) Hz, 2H), 1.95 – 1.86 (m, 2H), 1.84 – 1.74 (m, 2H), 1.58 – 1.42 (m, 4H) ppm. \(^{13}\)C NMR (126 MHz, CDCl\(_3\)) \(\delta\) 194.6, 166.0, 165.1, 164.3, 163.4, 136.8, 135.8, 131.3, 129.3, 126.1, 124.6, 121.0, 116.4, 107.9, 102.3, 67.3, 65.1, 33.7, 32.6, 28.5, 27.8, 25.2 ppm. HRMS (ESI) [M+H]\(^+\) calcd. C\(_{22}\)H\(_{24}\)BrNO\(_6\) 478.0865, found 478.0849.

Synthesis of 6-(tosyloxy)hexyl 3-(2-(4-formyl-3-hydroxyphenoxy)acetamido)benzoate (2e)

\[
\text{\includegraphics[width=0.2\textwidth]{image.png}}
\]

Yield 67%; \(R_t\) 0.32, off-white solid. \(^1\)H NMR (500 MHz, CDCl\(_3\)) \(\delta\) 11.42 (s, 1H), 9.77 (s, 1H), 8.35 (s, 1H), 8.07 – 8.00 (m, 2H), 7.86 – 7.81 (m, 1H), 7.76 (d, \(J = 8.3\) Hz, 2H), 7.52 (d, \(J = 8.6\) Hz, 1H), 7.48 – 7.42 (m, 1H), 7.32 (d, \(J = 8.0\) Hz, 2H), 6.66 (dd, \(J = 8.6, 2.4\) Hz, 1H), 6.51 (d, \(J = 2.3\) Hz, 1H), 4.66 (s, 2H), 4.29 (t, \(J = 6.5\) Hz, 2H), 4.04 (t, \(J = 6.3\) Hz, 2H), 2.43 (s, 3H), 1.80 – 1.62 (m, 4H), 1.49 – 1.35 (m, 4H) ppm. \(^{13}\)C NMR (126 MHz, CDCl\(_3\)) \(\delta\) 194.8, 166.1, 165.3, 164.4, 163.6, 144.9, 137.0, 135.9, 133.1, 131.4, 129.9, 129.4, 127.9, 126.2, 124.9, 121.1, 116.4, 108.1, 102.4, 70.5, 67.4, 65.1, 28.7, 28.5, 25.5, 25.2, 21.7 ppm. HRMS (ESI) [M+Na]\(^+\) calcd. C\(_{29}\)H\(_{31}\)NO\(_6\)SNa 592.1617, found 592.1633.
Scheme S6. Synthesis of 6-bromohexyl 4-(2-(4-formyl-3-hydroxyphenoxy)acetamido)benzoate (2f)

The procedure for synthesis of LDM reagent 2f is similar to the LDM reagent 2d.

**Synthesis of 6-bromohexyl 4-aminobenzoate (S29)**

Yield 58%; Rf 0.45, white solid. $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 7.84 (d, $J = 8.6$ Hz, 2H), 6.63 (d, $J = 8.6$ Hz, 2H), 4.25 (t, $J = 6.6$ Hz, 2H), 4.09 (s, 2H), 3.40 (t, $J = 6.8$ Hz, 2H), 1.93 – 1.82 (m, 2H), 1.80 – 1.71 (m, 2H), 1.56 – 1.40 (m, 4H). $^{13}$C NMR (176 MHz, CDCl$_3$) $\delta$ 166.8, 150.9, 131.7, 120.2, 113.9, 64.4, 33.9, 32.8, 28.8, 28.0, 25.5. HRMS (ESI) [M+H]$^+$ calcd. C$_{13}$H$_{18}$BrNO$_2$ 300.0599, found 300.0587.

**Synthesis of 6-bromohexyl 4-(2-bromoacetamido) benzoate (S30)**

Yield 86%; Rf 0.49, white solid. $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 8.26 (s, 1H), 8.09 – 8.00 (m, 2H), 7.70 – 7.58 (m, 2H), 4.31 (t, $J = 6.6$ Hz, 2H), 4.04 (s, 2H), 3.42 (t, $J = 6.8$ Hz, 2H), 1.96 – 1.84 (m, 2H), 1.84 –
1.70 (m, 2H), 1.61 – 1.40 (m, 4H) ppm. \(^{13}\text{C NMR}\) (126 MHz, CDCl\(_3\)) \(\delta\) 166.1, 163.6, 141.0, 130.9, 127.0, 119.2, 65.0, 33.8, 32.7, 29.4, 28.7, 27.9, 25.4 ppm. \(\text{HRMS (ESI)}\) [M+Na]\(^+\) calcd. C\(_{13}\)H\(_{19}\)Br\(_2\)NO\(_3\) 441.9629, found 441.9641.

**Synthesis of 6-bromohexyl 4-(2-(4-formyl-3-hydroxyphenoxy)acetamido)benzoate (S31)**

Yield 24%; \(R_f\) 0.37, white solid. \(^1\text{H NMR}\) (500 MHz, CDCl\(_3\)) \(\delta\) 11.43 (s, 1H), 9.79 (s, 1H), 8.28 (s, 1H), 8.11 – 7.98 (m, 2H), 7.73 – 7.64 (m, 2H), 7.55 (d, \(J = 8.6\) Hz, 1H), 6.67 (dd, \(J = 8.6, 2.4\) Hz, 1H), 6.54 (d, \(J = 2.4\) Hz, 1H), 4.68 (s, 2H), 4.31 (t, \(J = 6.6\) Hz, 2H), 3.42 (t, \(J = 6.8\) Hz, 2H), 1.97 – 1.83 (m, 2H), 1.84 – 1.74 (m, 2H), 1.58 – 1.43 (m, 4H) ppm. \(^{13}\text{C NMR}\) (126 MHz, CDCl\(_3\)) \(\delta\) 194.8, 166.0, 165.2, 164.4, 163.4, 140.7, 136.0, 131.0, 126.9, 119.4, 116.6, 107.9, 102.5, 67.5, 65.0, 33.8, 32.7, 28.7, 28.2, 25.4 ppm. \(\text{HRMS (ESI)}\) [M+H]\(^+\) calcd. C\(_{22}\)H\(_{24}\)BrNO\(_6\) 478.0865 found 478.0848.

**Synthesis of 6-bromohexyl 4-(2-(4-formyl-3-hydroxyphenoxy)acetamido)benzoate (2f)**

Yield 50%; \(R_f\) 0.34, white solid. \(^1\text{H NMR}\) (500 MHz, CDCl\(_3\)) \(\delta\) 11.43 (s, 1H), 9.78 (s, \(J = 13.5\) Hz, 1H), 8.31 (s, 1H), 8.03 (d, \(J = 8.7\) Hz, 2H), 7.78 (d, \(J = 8.2\) Hz, 2H), 7.68 (d, \(J = 8.7\) Hz, 2H), 7.54 (d, \(J = 8.6\) Hz, 1H), 7.33 (d, \(J = 8.0\) Hz, 2H), 6.66 (dd, \(J = 8.6, 2.4\) Hz, 1H), 6.54 (d, \(J = 2.3\) Hz, 1H), 4.67 (s, 2H), 4.26 (t, \(J = 6.6\) Hz, 2H), 4.03 (t, \(J = 6.4\) Hz, 2H), 2.44 (s, 3H), 1.76 – 1.64 (m, 4H), 1.47 – 1.34 (m, 4H) ppm. \(^{13}\text{C NMR}\) (126 MHz, CDCl\(_3\)) \(\delta\) 194.8, 166.0, 165.2, 164.4, 163.4, 144.8, 140.7, 136.0, 133.2, 131.0, 129.9, 128.0, 126.8, 119.4, 116.5, 107.9, 102.5, 70.5, 67.4, 64.9, 28.8, 28.6, 25.5, 25.2, 21.7 ppm. \(\text{HRMS (ESI)}\) [M+Na]\(^+\) calcd. C\(_{29}\)H\(_{31}\)NO\(_9\)SNa 592.1617, found 592.1618.
3b. Derivatives of O-hydroxylamine

![Scheme S7](image)

**Scheme S7.** N-(1-(aminoxy)-2-oxo-7,10,13-trioxa-3-azahexadecan-16-yl)-3,5-bis(trifluoromethyl)benzamide (6a)

**Synthesis of N-(1-bromo-2-oxo-7,10,13-trioxa-3-azahexadecan-16-yl)-3,5-bis(trifluoromethyl)benzamide (S36)**

In a 10 ml round bottom flask, N-(3-(2-(2-(3-aminopropoxy)ethoxy)ethoxy)propyl)-3,5-bis(trifluoromethyl)benzamide S34 (200 mg, 0.43 mmol) and potassium carbonate (237 mg, 1.72 mmol) in DCM:H₂O (1:1, 4.3 ml) were stirred at 0 °C for 5 minutes. To this solution, 2-bromoacetyl bromide S35 (162 μl, 1.29 mmol) was added slowly over a period of 1 hour. The reaction mixture was brought to room temperature and stirred for another 3 h. The progress of the reaction was monitored by TLC. The reaction mixture was concentrated using rotary evaporator. The purification of crude reaction mixture was performed by silica gel flash column chromatography using ethyl acetate:n-hexane (70:30) to isolate S36 (225 mg, 90% yield; viscous liquid). **¹H NMR** (500 MHz, CD₃OD) δ 8.45 (s, 2H), 8.17 (s, 1H), 3.84 (s, 2H), 3.69 – 3.51 (m, 14H), 3.35 – 3.32 (m, 2H), 1.98 – 1.89 (m, 2H), 1.83 – 1.74 (m, 2H). **¹³C NMR** (126
MHz, CD$_3$OD) δ 169.3, 166.5, 138.2, 132.9, 128.9, 125.8, 123.5, 71.5, 71.2, 70.0, 69.8, 38.9, 38.8, 38.5, 30.3, 30.0, 28.8. HRMS (ESI) [M+Na]$^+$ calcd. C$_{21}$H$_{30}$BrF$_6$N$_2$O$_5$Na 603.0905, found 603.0911.

**Synthesis of tert-butyl ((1-(3,5-bis(trifluoromethyl)phenyl)-1,17-dioxo-6,9,12-trioxa-2,16-diazaoctadecan-18-yl)oxy)carbamate (S38)**

In a 5 ml round bottom flask, N-(1-bromo-2-oxo-7,10,13-trioxa-3-azahexadecan-16-yl)-3,5-bis(trifluoromethyl)benzamide S36 (200 mg, 0.34 mmol), tert-butyl hydroxycarbamate S37 (91 mg, 0.69 mmol), and triethyl amine (103 µl, 1.0 mmol) were dissolved in acetonitrile (3.4 ml) and heated to reflux for 12 h. The progress of reaction was monitored by TLC. The reaction mixture was concentrated on a rotary evaporator and purified by silica gel flash column chromatography (ethyl acetate:hexane, 60:40) to give compound S38 (146 mg, 68%). $^1$H NMR (500 MHz, CD$_3$OD) δ 8.43 (s, 2H), 8.16 (s, 1H), 4.23 (s, 2H), 3.68 – 3.50 (m, 14H), 3.34 (t, $J$ = 6.8 Hz, 2H), 1.95 – 1.88 (m, 2H), 1.82 – 1.75 (m, 2H), 1.47 (s, 9H).

$^{13}$C NMR (126 MHz, CD$_3$OD) δ 171.3, 166.5, 159.7, 138.2, 132.9, 128.9, 125.6, 123.5, 82.9, 76.4, 71.5, 71.2, 71.2, 70.0, 69.8, 38.9, 37.4, 30.32, 30.2 (2C), 28.4 (3C). HRMS (ESI) [M+Na]$^+$ calcd. C$_{21}$H$_{30}$BrF$_6$N$_2$O$_5$Na 656.2383, found 656.2400.

**N-(1-(aminoxy)-2-oxo-7,10,13-trioxa-3-azahexadecan-16-yl)-3,5-bis(trifluoromethyl) benzamide (6a)**

In a 5 ml round bottom flask, tert-butyl ((1-(3,5-bis(trifluoromethyl)phenyl)-1,17-dioxo-6,9,12-trioxa-2,16-diazaoctadecan-18-yl)oxy)carbamate S38 (130 mg, 0.20 mmol), was dissolved in ACN:H$_2$O (2 ml). To this solution, trifluoroacetic acid (92 µl, 0.8 mmol) was added slowly over 15 minutes and stirred at 25 °C for 12 h. The reaction mixture was concentrate and dried on rotary evaporator, and triturate in diethyl ether, concentrated in vacuo to give 6a (105 mg, yield 99%; pale yellow liquid. $^1$H NMR (700 MHz, CD$_3$OD) δ 8.44 (s, 2H), 8.16 (s, 1H), 4.56 (s, 2H), 3.72 – 3.46 (m, 14H), 3.35 – 3.32 (m, 2H), 1.97 – 1.86 (m, 2H), 1.83 – 1.73 (m, 2H). $^{13}$C NMR (176 MHz, CD$_3$OD) δ 168.3, 165.2, 136.8, 127.6, 124.5, 123.9, 122.4, 70.9, 70.1, 69.8, 68.5, 38.9, 37.5, 36.4, 33.1, 28.9, 28.8 (2). HRMS (ESI) [M+H]$^+$ calcd. C$_{26}$H$_{38}$F$_6$N$_3$O$_5$ 534.2039, found 534.2006.
Scheme S8. Synthesis of 3-(aminoxy)propyl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate (6b)

Synthesis of 2-(3-bromopropoxy)isoindoline-1,3-dione (S40)²

In a 250 ml round bottom flask, N-hydroxyphthalimide S39 (4.894 g, 30 mmol) and triethyl amine (6.09 ml, 60 mmol) were dissolved in acetonitrile (60 ml). To this solution, 1,3-dibromopropane S16 (8.34 ml, 60 mmol) was added and stirred at 25 °C for 16 h. The reaction mixture was concentrated in vacuo followed by addition of 1 N NaOH solution and ethyl acetate. The organic layer was separated, dried over anh. sodium sulfate, filtered, and concentrated in vacuo. The purification of the crude reaction mixture was performed by silica gel flash column chromatography using ethyl acetate:hexane (3:97) to isolate S40 (4.259 g, 50% yield; Rf 0.57, ethyl acetate:n-hexane 30:70; white solid). ¹H NMR (400 MHz, CDCl₃) δ 7.89-7.81 (m, 2H), 7.80-7.73 (m, 2H), 4.37 (t, J = 5.8 Hz, 2H), 3.71 (t, J = 6.5 Hz, 2H), 2.36-2.26 (m, 2H) ppm. ¹³C NMR (101 MHz, CDCl₃) δ 163.7, 134.7, 129.0, 123.7, 76.2, 31.6, 29.4 ppm. HRMS (ESI) [M+Na]⁺ calcd. C₁₁H₁₁BrNO₃Na 305.9742 found 305.9739.

Synthesis of 3-((1,3-dioxoisindolin-2-yl)oxy)propyl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate (S42)
In a 5 ml round bottom flask, biotin S41 (244 mg, 1 mmol), 2-(3-bromopropoxy)isoindoline-1,3-dione S40 (568 mg, 2 mmol), and DBU (304 µl, 2 mmol) were dissolved in acetonitrile (20 ml). The reaction mixture was refluxed and the progress of the reaction was monitored by TLC. After 16 h, the reaction mixture was concentrated in vacuum. This was followed by solvent-solvent extraction using ethyl acetate and water. The organic fractions were combined, dried on anh. sodium sulfate, filtered, and concentrated on rotary evaporator. The purification of crude reaction mixture was performed by silica gel flash column chromatography (MeOH:DCM, 0.5-5%) to isolate 3-((1,3-dioxoisooindolin-2-yl)oxy)propyl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate S42 (224 mg, 50% yield; Rf 0.33, MeOH:DCM 05:95; white solid).

\[^{1}H\] NMR (500 MHz, CDCl\(_3\)) \(\delta\) 7.89-7.82 (m, 2H), 7.80-7.73 (m, 2H), 5.96 (s, 1H), 5.48 (s, 1H), 4.56-4.47 (m, 1H), 4.38-4.26 (m, 5H), 3.23-3.13 (m, 1H), 2.92 (dd, \(J = 12.8, 5.0\) Hz, 1H), 2.74 (d, \(J = 12.8\) Hz, 1H), 2.36 (t, \(J = 7.4\) Hz, 2H), 2.16-2.09 (m, 2H), 1.80-1.60 (m, 4H), 1.54-1.39 (m, 2H) ppm.

\[^{13}C\] NMR (126 MHz, CDCl\(_3\)) \(\delta\) 173.7, 163.8, 163.7, 134.7, 129.0, 123.7, 75.1, 62.0, 60.7, 60.2, 55.5, 40.7, 34.0, 28.4, 28.3, 27.8, 24.9 ppm.

HRMS (ESI) [M+H]+ calcd. C\(_{21}\)H\(_{26}\)N\(_3\)O\(_6\)S 448.1542, found 448.1548.

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

In a 5 ml round bottom flask, 3-((1,3-dioxoisooindolin-2-yl)oxy)propyl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate S42 (134 mg, 0.3 mmol) in DCM (3 ml) and hydrazine monohydrate (80%, 37 µl, 0.75 mmol) were stirred at room temperature. The progress of the reaction was followed by TLC. After 2 h, the reaction mixture was filtered and concentration of the filtrate in vacuo led to the isolation of 3-(aminooxy)propyl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate 6b (67 mg, 71% yield; Rf 0.21, MeOH:DCM 05:95; white solid. \[^{1}H\] NMR (500 MHz, D\(_2\)O) \(\delta\) 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), 3.45-3.27 (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\(_2\)O) \(\delta\) 176.9, 163.8, 163.7, 134.7, 129.0, 123.7, 75.1, 62.0, 60.7, 60.2, 55.3, 40.7, 34.0, 28.4, 28.3, 27.8, 24.9 ppm. HRMS (ESI) [M+H]+ calcd. C\(_{19}\)H\(_{24}\)N\(_3\)O\(_4\)S 318.1488, found 318.1467.
Scheme S9. Synthesis of 7-((3-(aminooxy)propyl)thio)-4-methyl-2H-chromen-2-one (6c)

Synthesis of 2-(3-((4-methyl-2-oxo-2H-chromen-7-yl)thio)propoxy)isoindoline-1,3-dione (S44)

In a 25 ml round bottom flask, 7-mercapto-4-methylcoumarin S43 (192 mg, 1 mmol), K$_2$CO$_3$ (276 mg, 2 mmol), and 2-(3-bromopropoxy)isoindoline-1,3-dione S40 (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 S44 (375 mg, 95% yield; R$_f$ 0.37, ethyl acetate:n-hexane 50:50; white solid). $^1$H NMR (400 MHz, CDCl$_3$) $\delta$ 7.90-7.81 (m, 2H), 7.80-7.73 (m, 2H), 7.48 (d, $J$ = 8.2 Hz, 1H), 7.26-7.20 (m, 2H), 6.22 (d, $J$ = 0.8 Hz, 1H), 4.36 (t, $J$ = 5.8 Hz, 2H), 3.35 (t, $J$ = 7.1 Hz, 2H), 2.41 (d, $J$ = 0.9 Hz, 3H), 2.23-2.08 (m, 2H) ppm. $^{13}$C NMR (101 MHz, CDCl$_3$) $\delta$ 163.8, 160.7, 154.0, 152.6, 142.6, 134.7, 129.0, 124.9, 123.8, 123.4, 117.5, 114.8, 114.1, 76.6, 28.7, 27.8, 18.7 ppm. HRMS (ESI) [M+H]$^+$ calcd. C$_{21}$H$_{18}$NO$_5$S 396.0906, found 396.0925.

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

2-(3-((4-methyl-2-oxo-2H-chromen-7-yl)thio)propoxy)isoindoline-1,3-dione S44 (237 mg, 0.6 mmol) was dissolved in CH$_2$Cl$_2$ (12 ml) in a 50 ml round bottom flask. To this solution, hydrazine monohydrate (80%, 29 µl, 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 reaction mixture was performed by reverse phase preparative HPLC to isolate 6c (76 mg, 45% yield; R$_f$ 0.6, ethyl acetate:n-hexane 50:50; pale green viscous liquid). $^1$H NMR (400 MHz, CDCl$_3$) $\delta$ 7.46 (d, $J$ = 8.3 Hz, 1H), 7.23-7.13 (m, 2H), 6.18 (d, $J$ = 0.9 Hz, 1H), 3.79 (t, $J$ = 5.9 Hz, 2H), 3.07 (t, $J$ = 7.3 Hz, 2H), 2.40 (d, $J$ = 0.8 Hz, 3H), 2.02-1.90 (m, 2H) ppm. $^{13}$C NMR (101
MHz, CDCl$_3$) δ 160.7, 154.0, 152.3, 143.3, 124.7, 123.1, 117.2, 114.1, 113.9, 73.9, 29.0, 27.8, 18.6 ppm.

**HRMS (ESI) [M+H]$^+$** calcd. C$_{13}$H$_{16}$NO$_3$S 266.0851, found 266.0841.
4. Protein labeling data

4a. Table S3 Control reaction of reagent (2a) with native ubiquitin

| Sr. No. | Equivalent (2a) | Temp (°C) | % Conversiona |
|---------|-----------------|-----------|---------------|
| 1       | 25              | 25        | 0             |
| 2       | 25              | 37        | 0             |
| 3       | 50              | 25        | 0             |
| 4       | 50              | 37        | 0             |
| 5       | 100             | 25        | 0             |
| 6       | 100             | 37        | 0             |

a % Conversion was determined by ESI-MS.

Figure S1. ESI-MS spectrum for labeled ubiquitin (3a)

Figure S2. ESI-MS spectrum for labeled ubiquitin (3a)
Figure S3. ESI-MS spectrum for labeled ubiquitin (3a)

Figure S4. ESI-MS spectrum for labeled ubiquitin (3a)

Figure S5. ESI-MS spectrum for labeled ubiquitin (3a)
**Figure S6.** ESI-MS spectrum for labeled ubiquitin (3a)

4b. Table S4 Control reaction of reagent (2a) with native proteins

| Entry | Proteins                  | % Conversion<sup>a</sup> (3) |
|-------|---------------------------|------------------------------|
| 1     | Insulin (1b)              | 0                            |
| 2     | α-Lactalbumin (1c)        | 4                            |
| 3     | Myoglobin (1d)            | 5                            |

<sup>a</sup> % Conversion was determined by ESI-MS.

**Figure S7.** ESI-MS spectrum for labeled insulin (3b)
Figure S8. ESI-MS spectrum for labeled α-lactalbumin (3c)

Figure S9. ESI-MS spectrum for labeled myoglobin (3d)

4c. Single-site labeling of native proteins
Figure S10. (a) Site-selective labeling of ubiquitin 1a enabled by LDM reagent 2c (25 equiv.). (b) ESI-MS spectrum for mono-labeled ubiquitin (5a) after oxime formation. (c) MS-MS spectrum of labeled ubiquitin after the digestion of 5a with trypsin. The site of modification is D32.
**Figure S11.** (a) Site-selective labeling of insulin 1b enabled by LDM reagent 2c (25 equiv.). (b) ESI-MS spectrum for mono-labeled insulin (5b) after oxime formation. (c) MS-MS spectrum of labeled insulin after the reduced the 5b with DTT. The site of modification is H10.
a) 

\[ \text{α-Lactalbumin} \]

\[ 1c \]

\[ + \]

\[ 2c \]

1. PB (pH 7.8, 0.1 M)

\[ 50 \mu M, 37 \degree C, 4 \text{ h} \]

\[ \text{α-Lactalbumin} \]

\[ 4a \]

\[ \rightarrow \]

\[ 5c \] (44% Conversion)

b) 

Deconvoluted m/z values:

- 1091.0
- 1181.8
- 1299.2
- 1418.0
- 1575.5
- 1772.3
- 2025.5

Deconvoluted Mass (Da):

- 14170.0
- 14495.1

m/z range:

1000 to 2200
Figure S12. (a) Site-selective labeling of α-lactalbumin 1c enabled by LDM reagent 2c (25 equiv.). (b) ESI-MS spectrum for mono-labeled α-lactalbumin (5c) after oxime formation. (c) MS-MS spectrum of labeled α-lactalbumin after the digestion of 5c with α-chymotrypsin. The site of modification is H107.
Figure S13. (a) Site-selective labeling of myoglobin 1d enabled by LDM reagent 2c (25 equiv.). (b) ESI-MS spectrum for mono-labeled myoglobin (5d) after oxime formation. (c) MS-MS spectrum of labeled myoglobin after the digestion of 5d with α-chymotrypsin. The site of modification is H116.
4d. Tagging of labeled ubiquitin

a) Site-selective labeling of ubiquitin 1a enabled by LDM reagent 2c (25 equiv.).

b) ESI-MS spectrum of 19F NMR probe tagged ubiquitin (7a).

Figure S14. (a) Site-selective labeling of ubiquitin 1a enabled by LDM reagent 2c (25 equiv.). (b) ESI-MS spectrum of 19F NMR probe tagged ubiquitin (7a).
**Figure S15.** (a) Site-selective labeling of ubiquitin **1a** enabled by LDM reagent **2c** (25 equiv.). (b) ESI-MS spectrum of biotin probe tagged ubiquitin (**7b**).

**Figure S16.** (a) Site-selective labeling of ubiquitin **1a** enabled by LDM reagent **2c** (25 equiv.). (b) ESI-MS spectrum of coumarin probe tagged ubiquitin (**7c**).
4e. Purification of labeled ubiquitin

Figure S17. ESI-MS spectrum of purified mono-labeled ubiquitin (10).

5. Additional data

5a. Protein labeling

a) Site-selective labeling of ubiquitin 1a enabled by LDM reagent 2c (25 equiv.).

b) ESI-MS spectrum of labeled ubiquitin (32).

Figure S18. (a) Site-selective labeling of ubiquitin 1a enabled by LDM reagent 2c (25 equiv.). (b) ESI-MS spectrum of labeled ubiquitin (32).
Figure S19. (a) Site-selective labeling of ubiquitin 1a (1 equiv.) enabled by LDM reagent 2c (25 equiv.). (b) ESI-MS spectrum of labeled ubiquitin (33).
Figure S20. (a) Site-selective labeling of ubiquitin 1a enabled by LDM reagent 2f (25 equiv.). (b) ESI-MS spectrum labeled ubiquitin (34).

Figure S21. (a) Site-selective labeling of ubiquitin 1a enabled by LDM reagent 2b (25 equiv.). (b) ESI-MS spectrum of labeled ubiquitin (35).
**Figure S22.** (a) Site-selective labeling of insulin 1b enabled by LDM reagent 2d (25 equiv.). (b) ESI-MS spectrum of labeled insulin (36).
Figure S23. (a) Site-selective labeling of insulin 1b enabled by LDM reagent 2e (25 equiv.). (b) ESI-MS spectrum of labeled insulin (37).

Figure S24. (a) Site-selective labeling of insulin 1b enabled by LDM reagent 2f (25 equiv.). (b) ESI-MS spectrum of labeled insulin (38).
Figure S25. (a) Site-selective labeling of insulin 1b enabled by LDM reagent 2b (25 equiv.). (b) ESI-MS spectrum of labeled insulin (39).
Figure S26. (a) Site-selective labeling of α-lactalbumin 1c enabled by LDM reagent 2d (25 equiv.). (b) ESI-MS spectrum of labeled α-lactalbumin (40).

Figure S27. (a) Site-selective labeling of α-lactalbumin 1c enabled by LDM reagent 2e (25 equiv.). (b) ESI-MS spectrum of labeled α-lactalbumin (41).
Figure S28. (a) Site-selective labeling of α-lactalbumin 1c enabled by LDM reagent 2f (25 equiv.). (b) ESI-MS spectrum of labeled α-lactalbumin (42).
Figure S29. (a) Site-selective labeling of α-lactalbumin 1c enabled by LDM reagent 2b (25 equiv.). (b) ESI-MS spectrum of labeled α-lactalbumin (43).

Figure S30. (a) Site-selective labeling of myoglobin 1d enabled by LDM reagent 2b (25 equiv.). (b) ESI-MS spectrum of labeled myoglobin (44).
Figure S31. (a) Site-selective labeling of myoglobin 1d enabled by LDM reagent 2d (25 equiv.). (b) MALDI-ToF-MS spectrum of labeled myoglobin (45).
a) Myoglobin

1d

\[
\text{HN} \quad \text{O} \\
\text{O} \quad \text{O} \\
\text{O} \quad \text{O} \\
\text{CH} \quad \text{O} \quad \text{SO}_3^+ \\
\text{HO} \quad \text{NH}_2
\]

1. PB (pH 7.8, 0.1 M) 50 μM, 37 °C, 24 h
2. 4b, 15 min.

Myoglobin

46 (13% Conversion)

1d

\[
\text{HN} \quad \text{O} \\
\text{O} \quad \text{O} \\
\text{O} \quad \text{O} \\
\text{CH} \quad \text{O} \quad \text{SO}_3^+ \\
\text{HO} \quad \text{NH}_2
\]

46

\[
\text{HN} \quad \text{O} \\
\text{O} \quad \text{O} \\
\text{O} \quad \text{O} \\
\text{CH} \quad \text{O} \quad \text{SO}_3^+ \\
\text{HO} \quad \text{NH}_2
\]

Matrix

m/z

15000 20000
Figure S32. (a) Site-selective labeling of myoglobin 1d enabled by LDM reagent 2e (25 equiv.). (b) MALDI-ToF-MS spectrum of labeled myoglobin (46).
**Figure S33.** (a) Site-selective labeling of myoglobin 1d enabled by LDM reagent 2f (25 equiv.). (b) MALDI-ToF-MS spectrum of labeled myoglobin (47).

5b. **Circular dichroism spectra of protein**

![Circular Dichroism (CD) spectra of ubiquitin (1a) and labeled ubiquitin (5a) in phosphate buffer (0.1 M, pH 7.0) at concentration 0.1 mg/ml.](image)

**Figure S34.** Circular Dichroism (CD) spectra of ubiquitin (1a) and labeled ubiquitin (5a) in phosphate buffer (0.1 M, pH 7.0) at concentration 0.1 mg/ml.

5c. **Fluorescence spectra of coumarin tagged ubiquitin**

![Fluorescence spectra of coumarin tagged ubiquitin](image)
**Figure S35.** Steady-state fluorescence spectra of coumarin tagged ubiquitin (6c) in phosphate buffer (0.1 M, pH 7.0). 6c exhibits emission band peaked at 428 nm (excitation at 333 nm). For MS data, see Scheme 4.

6. Spectral data

![S13 NMR spectrum](image)

**Figure S36.** $^1$H-NMR spectrum of compound S13.
Figure S37. $^{13}$C-NMR spectrum of compound S13.

Figure S38. $^1$H-NMR spectrum of compound 2a.
Figure S39. $^{13}$C-NMR spectrum of compound 2a.

Figure S40. $^1$H-NMR spectrum of compound S17.
Figure S41. $^{13}$C-NMR spectrum of compound S17.

Figure S42. $^1$H-NMR spectrum of compound 2b.
Figure S43. $^{13}$C-NMR spectrum of compound 2b.

Figure S44. $^1$H-NMR spectrum of compound S18.
Figure S45. $^{13}$C-NMR spectrum of compound S18.

Figure S46. $^1$H-NMR spectrum of compound 2c.
Figure S47. $^{13}$C-NMR spectrum of compound 2c.

Figure S48. $^1$H-NMR spectrum of compound S20.
Figure S49. $^{13}$C-NMR spectrum of compound S20.

Figure S50. $^1$H-NMR spectrum of compound S22.
Figure S51. $^{13}$C-NMR spectrum of compound S22.

Figure S52. $^1$H-NMR spectrum of compound S23.
Figure S53. $^{13}$C-NMR spectrum of compound S23.

Figure S54. $^1$H-NMR spectrum of compound 2d.
Figure S55. $^{13}$C-NMR spectrum of compound 2d.

Figure S56. $^1$H-NMR spectrum of compound S25
Figure S57. $^{13}$C-NMR spectrum of compound S25.

Figure S58. $^1$H-NMR spectrum of compound S26.
Figure S59. $^{13}$C-NMR spectrum of compound S26.

Figure S60. $^1$H-NMR spectrum of compound S27.
Figure S61. $^{13}$C-NMR spectrum of compound S27.

Figure S62. $^1$H-NMR spectrum of compound 2e.
Figure S63. $^{13}$C-NMR spectrum of compound 2e.

Figure S64. $^1$H-NMR spectrum of compound S29.
Figure S65. $^{13}$C-NMR spectrum of compound S29.

Figure S66. $^1$H-NMR spectrum of compound S30.
Figure S67. $^{13}$C-NMR spectrum of compound S30.

Figure S68. $^1$H-NMR spectrum of compound S31.
Figure S69. $^{13}$C-NMR spectrum of compound S31.

Figure S70. $^1$H-NMR spectrum of compound 2f.
Figure S71. $^{13}$C-NMR spectrum of compound 2f.

Figure S72. $^1$H-NMR spectrum of compound S36.
Figure S73. $^{13}$C-NMR spectrum of compound S36.

Figure S74. $^1$H-NMR spectrum of compound S38.
Figure S75. $^{13}$C-NMR spectrum of compound S38.

Figure S76. $^1$H-NMR spectrum of compound 6a.
Figure S77. $^{13}$C-NMR spectrum of compound 6a.

Figure S78. $^1$H-NMR spectrum of compound S40.
Figure S79. $^{13}$C-NMR spectrum of compound S40.

Figure S80. $^1$H-NMR spectrum of compound S42.
Figure S8. $^{13}$C-NMR spectrum of compound S42.

Figure S82. $^1$H-NMR spectrum of compound 6b.
Figure S83. $^{13}$C-NMR spectrum of compound 6b.

Figure S84. $^1$H-NMR spectrum of compound S44.
Figure S85. $^{13}$C-NMR spectrum of compound S44.

Figure S86. $^1$H-NMR spectrum of compound 6c.
7. Additional discussion

[1] Reaction of proteins with higher concentration of alkylating reagent 2a.

The reactions were performed with a higher concentration of electrophile (2a, 1000-5000 equivalents) with different proteins. In the case of α-lactalbumin and myoglobin, we noted ~5% conversion with 5000 equivalents of alkylating reagent 2a. For α-lactalbumin, E1 is labeled preferentially in contrast to H107 by LDM reagent 2c. For myoglobin, E18 is labeled for myoglobin (H116 with 2c), validating the reagent’s promiscuity. These observations are also supported by the reaction between ubiquitin and reagent 2c that renders the modification of glutamic acid, aspartic acid, and histidine (entry 4, Table S7).
Figure S88. (a) α-Lactalbumin labeling with 2a: MS and MS-MS. (b) Myoglobin labeling with 2a: MS and MS-MS.

[2] Reaction of ubiquitin (1a) with LDM reagent (2c)

Table S5. Effect of LDM reagent concentration on protein bioconjugation.
We treated ubiquitin with higher equivalents of LDM reagent (2c, Table S5). The experiments in this series were performed from a common stock solution of protein and reagent. At first, 25 equivalents of reagent 2c resulted in 37% labeling of D32. The increase in relative stoichiometry (50 equivalents) led to increased conversions (46%) while retaining the single-site D32 labeling. A further increase to 100 and 200 equivalents resulted in 5-6% bis-labeled product along with the mono-labeled bioconjugates (43% and 35%, respectively). In both cases, D58 was labeled in addition to D32. Interestingly, 500 and 1000 equivalents of reagent led to a considerable drop in conversions (25% and 17%, respectively) while regaining the single-site selectivity.

| S. No. | Equivalents (2c) | % Conversiona | Site(s) of modification |
|--------|------------------|---------------|-------------------------|
| 1      | 25               | 37b           | 0                       |
| 2      | 50               | 46            | 0                       |
| 3      | 100              | 43            | 6                       |
| 4      | 250              | 35            | 5                       |
| 5      | 500              | 25            | 0                       |
| 6      | 1000             | 17            | 0                       |

*a% Conversion was determined by ESI-MS. bIn another trial; 42% conversion was observed under these conditions.*
+H5S/M3 (140-200) from VR539PS39KT3, sample 1 - VR539PS39KT3, Experiment 2 @ 9.71 min, Precursor: 368.7888 Da.

+H4/M3 (140-200) from VR5507530KT3, sample 1 - VR5507530KT3, Experiment 2 @ 10.76 min, Precursor: 403.0044 Da.
Figure S89. Reaction of ubiquitin with LDM reagent 2c (a) 25 equivalents; (b) 50 equivalents; (c) 100 equivalents; (d) 250 equivalents; (e) 500 equivalents; (f) 1000 equivalents.

[3] Effect of protein concentration and pH

The concentration has a role to play in the regulation of reactivity and selectivity. The ubiquitin at 5 µM results in 18% D32-labeled bioconjugate (entry 1, Table S6). The increase of concentration to 50 µM (entries 2-4, Table S6) results in increased conversions without compromise in chemoselectivity and site-selectivity. However, a further increase in concentration (75 and 100 µM) shows 4-8% bis-labeling (D58) along with mono-labeling (D32).
**Table S6.** Effect of protein concentration on bioconjugation.

| S. No. | Concentration (x µM) | % Conversion* | Site(s) of modification |
|--------|----------------------|---------------|------------------------|
|        | Mono-labeled | Bis-labeled |                          |
| 1      | 5         | 18           | 0          | D32                |
| 2      | 10        | 20           | 0          | D32                |
| 3      | 20        | 28           | 0          | D32                |
| 4      | 50        | 37*          | 0          | D32                |
| 5      | 75        | 38           | 8          | D32, D58           |
| 6      | 100       | 30           | 4          | D32, D58           |

*% Conversion was determined by ESI-MS. *In another trial, 42% conversion was observed under these conditions.

![Chemical structures and reaction scheme](image)

**Chemical structures and reaction scheme:**

- **Ubiquitin**
- **1a**
- **2c (25 eq)**
- **1. PB (pH 7.8, 0.1 M) x µM, 37°C, 72 h**
- **2. 1 h**
- **HO-NH2**
- **4b**
- **Ubiquitin**
- **5a**

**Chemical structures and reactions:**

- **1a**
- **2c (25 eq)**
- **5a**

**Mass spectra:**

- **Deconvoluted mass spectra:**
  - **1a** 855.6 Da
  - **5a** 6782.7 Da

**Mass spectra details:**

- **+MS/MS (m/z 2000-3500) from VR540P540<Tr1>**
- **+MS/MS (sample1) from VR540P540<Tr1>, Experiment2 @ 9.73 min. Precursor 380.7056 Da.**
b) 

\[ +[\text{HS/IS([14]2000)}]_{\text{m/z}} \text{ sample 2} \to \text{VR540 P540 KT5, Experiment 2 @ 9.73 min, Precursor: 369.705 Da.} \]

\[ +[\text{HS/IS(14)} - 2000)_{\text{m/z}} \text{ sample 2} \to \text{VR540 P540 KT3, Experiment 2 @ 9.83 min, Precursor: 369.798 Da.} \]

---

S85
(a) +HS/HS(84–2000)from VRS33/PS53/KT1.wiff2 (sample 1) - VRS33/PS53/KT1, Experiment 2 @ 9.70 min, Precursor: 388706 Da.
Figure S90. Reaction of ubiquitin with concentration (a) 5 µM; (b) 10 µM; (c) 20 µM; (d) 50 µM; (e) 75 µM; (f) 100 µM.

Next, we investigated the effect of the aqueous buffer’s pH (Table S7). While the reaction ceased altogether at pH 6, 10% single-site labeled ubiquitin was observed at pH 7. Further increase of pH to 7.8 allowed the reaction to deliver 37% D32-labeled ubiquitin. Interestingly, the promiscuity of electrophile overpowers at pH 9 and results in the labeling of D32, D58, E64, and H68.
Table S7. Effect of reaction medium’s pH on bioconjugation

| S. No. | pH | % Conversion<sup>a</sup> | Site of modification |
|--------|----|--------------------------|----------------------|
| 1      | 6.0| 0                        | -                    |
| 2      | 7.0| 10 mono-labeled          | D32                  |
| 3      | 7.8| 37<sup>b</sup> mono-labeled | D32                  |
| 4      | 9.0| 39 mono-, 20 bis-, 8 tris-, 5 tetra-labeled | D32, D58, E64, H68 |

<sup>a</sup> % Conversion was determined by ESI-MS. <sup>b</sup>In another trial; 42% conversion was observed under these conditions.
Figure S91. Reaction of ubiquitin with pH (a) 6.0; (b) 7.0; (c) 7.8; (d) 9.0.

[4] Insulin labeling and bioactivity assay

Insulin has three primary amines with discrete placement in space (Figure S92a). The structural evaluation indicates that the N\(^{\alpha}\)-NH\(_2\) of N-Phe is suitably positioned (Figure S92b) to guide the alkylating electrophile in the proximity of H10 to enable its irreversible modification.
Figure S92. (a) Amine-H10 inter-residue distance (Å) in insulin. (b) N-Phe-derived linchpin places the alkylationating electrophile in proximity of H10. (Insulin, PDB ID: 3I40)

![Image of Figure S92](image)

Figure S93. Labeling of insulin with reagent 2c (a) ESI-MS spectra of purified insulin. (b) CD spectra of purified insulin.

The insulin was treated with the LDM reagent to render analytically pure H10-labeled insulin (11; Figure S93a). The circular dichroism data confirmed that the labeling and enrichment protocol does not alter the insulin structure (Figure S93b). Next, we took this forward to test the consequences of bioconjugation on its binding to the insulin receptor and its effect on the downstream signalling pathway (Figure 6).

[5] Docking investigations
Figure S94. Ubiquitin (PDB ID: 1UBQ) and LDM reagent 2c (a) Model highlighting K29-derived linchpin directed placement of electrophile near D32. (b) Docking investigations.

The structural model suggests that the imine formation at K29 with reagent 2c can potentially place the alkylating electrophile in the proximity of D32 residue (Figure S94a). On the other hand, the docking results indicate that the reagent has a binding preference that is not appropriate for the linchpin formation at K29 or irreversible covalent modification of D32 (Figure S94b). Overall, the data indicate that the ligand effect is unlikely to contribute while the linchpin-directed modification regulates the site-of-conjugation.

8. Protein sequence

1. Ubiquitin from bovine erythrocytes
   PDB ID: 1UBQ
   amino acid sequence:
   MQIFVKTLTGKTTITLEVEPSDIIIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTL HLVRLRGG

2. Insulin from human recombinant
   PDB ID: 3I40
   amino acid sequence:
   Chain A: GIVEQCTSICSLYQLENYN
   Chain B: FVNQHLCGSHLVEALYLVCGERGFFYTPKT

3. α-Lactalbumin from bovine milk
   amino acid sequence:3
   EQLTKCEVFRELLDKLGYGGVSLPEWVCTTFHTSGYDTQAIVQNNDSTEYGFLQINNKIDCKND QDPHSSNINSCDKFLNNDLTNNIMCVKKILDKVGNYWLAHKALCSEKLDQWLCEKL
4. Myoglobin from equine skeletal muscle
PDB ID: 1WLA

*amino acid sequence:*
GLSDGEWQQVLNVWGKVEADIAHGQEVLRFLGHTPELEKFDKFKHLKTEAEMKASEDLKK
HGTVVLTALGGILKKHHEAILKPLAQSHATKHKIPIKYLEFISDAIIHVLHSKHPDGADAQ
GAMTKALELRNDIAAKYKELGFQG

9. References

1. M. Kinter and N. E. Sherman, *Wiley Interscience*, **2000**.
2. F. Ito, S. Ando, M. Luchi, T. Ukari, M. Takasaki and K. Yamaguchi, *Tetrahedron* 2011, **67**, 8009-8013.
3. K. Brew, F. J. Castellino, T. C. Vanaman and R. L. Hill, *J. Biol. Chem.* 1970, **245**, 4570-4582.