Supporting Information

Heterobifunctional cross-linker with dinitroimidazole and azide modules for protein and oligonucleotide functionalization

Qunfeng Luo[a,*], Shuli Liu[a], Yaoguang Hua[a], Chunqiu Long[a], Sijia lv[a], Juncheng Li[a], Yuzhi Zhang[a]

[a] Schoolof Basic Medical Sciences, Jiangxi Medical College, NanchangUniversity, Nanchang, Jiangxi 330006, People's Republic of China;
*Correspondence E−mail: luoqunfeng@ncu.edu.cn

Table of Contents

1. General information

Solvents were obtained from Energy Chemical, Xilong Science and Bidepharm companies and used directly without further purification unless indicated. Amino acids and peptides were obtained from commercial sources. Protein 3CL used in this work is a variant of the main protease (M^{pro}) of SARS-COV-2 with Ser46 changed to Phe (S46F), and a purification tag with 33 amino acids is added at its N terminal. RhoA* in this work is a fusion protein that consists of GST and RhoA. GST is a 218-amino acid tag used for protein purification, while RhoA is a small GTPase with 193 residues, involved in regulating various cellular processes. Bovine serum albumin (BSA) was purchased from Solarbio Life Sciences and used without further purification. Analytical thin layer chromatography was performed on 0.25 mm silica gel 60-F254. Visualization was carried out with UV light. ^{1}H NMR spectra were recorded on Bruker AVANCE III HD 400 MHz. The following abbreviations (or combinations thereof) were used to explain multiplicities: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad. Coupling constants, J, were reported in Hertz unit (Hz). ^{13}C NMR spectra were recorded on Bruker AVANCE III HD 400 MHz instrument (100 MHz) and were fully decoupled by broad band proton decoupling. High-resolution mass spectra (HRMS) were recorded on an Agilent Mass spectrometer using ESI-TOF (electrospray ionization-time of flight). HPLC data were obtained on Agilent 1260 infinity II system using commercially available columns.

Sequence of 3CL

3CL used in this work is composed of 339 amino acids and contains twelve free cysteine residues and eleven lysine residues, while only three cysteine residues are exposed to solvent [1]. The amino acid sequence marked in orange is the purification tag. Calculated M.W.(average mass): 37269.44 Da.

GSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSSGFRKMAFPSGKVEGCM
VQVTCGTTTLNGLWLDDVVYCPRHVICTFEDMLNPNYEDLLIRKSNHNFLVQ
AGNVQLRVIGHSMQNCVLKLKVDTANPKTPKYKFVRIQPGQTFSVLACYNG
SPSGVYQCAMRPNFTIKGSFLNGSCGSVGFNIDYDCVSFCYMHHMELPTGVH
AGTDLEGNFYGPFVDRQTAQAAGTDTTITVNVLAWLYAAVINGDRWFLNRF
TTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCASLKELLQNG
MNGRTILGSALLEDEFTPFDVVRQCSGVTFQ

Sequence of RhoA

RhoA is composed of 422 amino acids and contains 10 cysteine residues and 38 lysine residues. Calculated M. W.(average): 48434.10 Da.

MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEF
PNLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYG
VSRIAYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYD
ALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQAT
FGGGDHPPKSDLVPRGSPEFMAAIRKKLVIVGDGACGKTCLLIVFSKDQFPEV
YVPTVFENYVADIEVDGKQVELALWDTAGQEDYDRLRPLSYPDTDVILMCFS

IDSPDSLENIPEKWTPEVKHFCPNVPIILVGNKKDLRNDEHTRRELAKMKQEP
VKPEEGRDMANRIGAFGYMECSAKTKDGVREVFEMATRAALQARRGKKKS
GCLVL

Sequence of bovine serum albumin (BSA)

BSA is composed of 583 amino acids and contains one free cysteine residue Cys34 (35 Cys residues in total, of which 34 Cys residues are in the form of disulfide bond) and 59 lysines. Calculated M.W. (average): 66429.98 Da

DTHKSEIAHRFKDLGEEHFKGLVLIAFSQYLQQC^{34}PFDEHVKLVNELTEFAKT
CVADESHAGCEKSLHTLFGDELCKVASLRETYGDMADCCEKQEPERNECFLS
HKDDSPDLPKLKPDPNTLCDEFKADEKKFWGKYLYEIARRHPYFYAPELLYY
ANKYNGVFQECCQAEDKGACLLPKIETMREKVLTSSARQRLRCASIQKFGER
ALKAWSVARLSQKFPKAEFVEVTKLVTDLTKVHKECCHGDLLECADDRADL
AKYICDNQDTISSKLKECCDKPLLEKSHCIAEVEKDAIPENLPPLTADFAEDKD
VCKNYQEAKDAFLGSFLYEYSRRHPEYAVSVLLRLAKEYEATLEECCAKDDP
HACYSTVFDKLKHLVDEPQNLIKQNCDQFEKLGEYGFQNALIVRYTRKVPQV
STPTLVEVSRSLGKVGTRCCTKPESERMPCTEDYLSLILNRLCVLHEKTPVSEK
VTKCCTESLVNRRPCFSALTPDETYVPKAFDEKLFTFHADICTLPDTEKQIKKQ
TALVELLKHKPKATEEQLKTVMENFVAFVDKCCAADDKEACFAVEGPKLVV
STQTALA

Sequence of oligonucleotide o1

The oligonucleotide **o1** was ordered from Takara Biotechnology (Dalian) Co., Ltd. Its 5' terminal was attached a click tag, dibenzoazacyclooctyne (DBCO) group. It contains 22 bases (5'-GGTTCAGGAGGCTCAATTTGGT-3') and exists in a single strand. So it is termed as 5'-(DBCON)GGTTCAGGAGGCTCAATTTGGT-3' with calculated M.W. (average): 7316 Da. It was confirmed by HPLC and ESI-MS analysis by the company.

2. Experimental procedures

A. Methods for RP-HPLC analysis

Analytical HPLC analysis was performed using Phenomenex C18 (Luna 5 μm C18(2) 100 Å, 250 x 4.6 mm) analytical columns with mobile phase of water-methanol (0.1% formic acid or 0.1% acetic acid or neutral) at a flow rate of 1.0 mL/min. Mobile phase A: 100 % v/v water, or with 0.1 % v/v formic acid or acetic acid; mobile phase B: 100 % v/v methanol or with 0.1 % v/v formic acid. Gradient used：isocratic 0% methanol for 5 min, then 0% to 100% methanol in 25 min, then 100% to 0% methanol in 5 min, then isocratic 0% methanol for 3 min.

B. Modification of peptides and proteins by DNIm-N_3

Typically, peptide (1 mM) was incubated with DNIm-$\mathbf{N_3}$ (1 mM) in HEPES buffer (50 or 100 mM) at indicated pH at 25 °C for indicated time before analysed by HPLC and MS.

For protein modification, functionalized agents bearing DNIm warhead were first prepared from DNIm-N_3 via SPAAC reaction. Then approximately 30 μM of protein (BSA) in phosphate buffer was incubated with each functionalized agent at pH 7.0 at r.t. for 1-2 h. Alternatively, incubation of protein samples with cross-linker DNIm-N_3 under neutral buffer was conducted to attach an azide group onto proteins. Then excess cross-linkers were removed upon desalting through PD MiniTrap™ G-25 columns (Cytiva). The azido-bearing proteins were then functionalized with various BCN or DBCO derivatives via SPAAC reaction. All the reaction mixture was quenched by addition of 10 μl acetic acid and desalted through PD MiniTrap™ G-25 columns. Through these two approaches, the proteins were attached with various functional groups, including biotin or fluorescent group. For biotin group, it was demonstrated by Western Blot analysis with the key reagent HRP-Streptavidin and LC-MS. For fluorescent group, it was analysed by SDS-PAGE and irradiated by UV light (302 nm) or blue light (440-485 nm) to visualize the desired bands and LC-MS. The protein gels were also stained by Coomassie Brilliant Blue (CBB).

C. Modification of oligonucleotide by DNIm-N_3

20 μM of oligonucleotide **o1** was incubated with equal amount of DNIm-N_3 in neutral HEPES buffer for 2-4 h followed by the addition of various sulfhydryl containing agents including PEG (80 μM), peptide **c2** (80 μM) and protein BSA (20 μM). The mixture was kept at room temperature for 1.5-2 h, and then subjected to Urea-PAGE or SDS-PAGE analysis. Second, incubation of 100 μM of BSA with excess cross-linker DNIm-N_3 under neutral buffer was performed to attach an azide group onto BSA. Then excess cross-linkers were removed upon ultrafiltration through a membrane (MWCO 30 kDa). The azido-bearing BSA (20 μM) was then functionalized with equal amount of **o1** via SPAAC reaction. Third, we also conducted a one-pot reaction in which **o1** (20 μM) and BSA (20 μM) were mixed with the cross-linker DNIm-N_3 (20 μM) at the same time in HEPES buffer at pH 7.0.

These products were analyzed by Urea-PAGE or SDS-PAGE. For oligonucleotide-PEG and oligonucleotide-peptide conjugates, they were analysed by Urea-PAGE. For oligonucleotide-protein conjugate, it was analysed by SDS-PAGE.

D. Western blot analysis of biotinylated proteins

After the chemical modification was completed, equal amounts of proteins from each sample were mixed with loading buffer and subjected to 10% or 12% SDS-PAGE followed by transferring onto PVDF membranes (Millipore, IPVH00010). The membrane was then blocked with 5% skim milk (Solarbio, D8340) for 1 h at room temperature and incubated with horseradish peroxidase-labeled streptavidin (SA-HRP) (Diamond, BN10063-0001) for 2 h at room temperature. Subsequently the membrane was washed with 1×TBST (20 mM Tris-HCl, 150 mM NaCl, 0.05% Tween-20) for three times, 10 min each time. Protein bands were visualized by taking advantage of an enhanced chemiluminescence (ECL) kit (UElandy Inc, S6008M) and a digital gel image analysis system (Tanon 5500).

E. Urea-PAGE analysis of oligonucleotide conjugates

After the chemical modification was completed, an equal amount of each oligonucleotide conjugate sample was mixed with 2X TBE-Urea sample buffer (Sangon Biotech, C506046), heated at 70 ˚C for 3 ~ 5 min and analyzed with 20% (w/v) Urea-PAGE. We first prepared a polyacrylamide gel comprised of separating gel and concentrating gel. A 10 ml 20% (w/v) separating gel was configured by mixing together 6.7 ml 30% (w/v) acrylamide / bisacrylamide 29:1 solution (Solarbio, A1010), 4.8 g urea (3A Materials, A60507), 1 ml 10×TBE buffer (890 mM Tris-Boric acid, 100 mM EDTA), 80 μl 10% ammonium persulfate (Aamas-Beta, P1384950), 10 μl N', N', N', N'-tetraethylmethylenediamine (Aamas-Beta, P1511906) and 2.3 ml double distilled water. A 3 ml 4% (w/v) concentrating gel was configured by mixing together 0.4 ml 30% (w/v) acrylamide / bisacrylamide 29:1 solution (Solarbio, A1010), 1.44 g urea (3A Materials, A60507), 0.3 ml 10×TBE buffer (890 mM Tris-Boric acid, 100 mM EDTA), 24 μl 10% ammonium persulfate (Aamas-Beta, P1384950), 3 μl N', N', N', N'-tetraethylmethylenediamine (Aamas-Beta, P1511906) and 2.3 ml double distilled water.

Next, the prepared polyacrylamide gel was pretreated with a constant 200 volts voltage for 30 minutes without sample loaded. Then the samples were loaded and subjected to electrophoresis for 60 minutes at 200 volts. Subsequently the gel was dyed with nucleic acid gel dye TS-GelRed (Tsingke, TSJ003) according to the instructions. Finally, the bands were visualized by ultraviolet analyzer (Qingke, ZF-2) at 302nm.

F. Mass spectrometric measurement of the protein adducts

F1. Capillary electrophoresis (CE)

Protein samples (in dd H_2O, 0.2-1 mg/ml) were analyzed using a 908 Devices (Boston, MA) ZipChip system connected in-line to a Thermo Scientific Orbitrap Ascend Tribrid. Protein samples were separated on HS chips using the "Intact Antibodies Peptide Background Electrolyte" premixed solution obtained from 908 Devices. Samples were injected into the CE separation channel by an injection volume of 2 nL. The electrical field strength of 500 V/cm was used during CE separation, resulting in a total run time of 6 min.

F2. Mass spectrometry

A Thermo Scientific Orbitrap Ascend Tribrid™ mass spectrometer was used in this study in Intact Protein mode for the analysis of intact Protein samples. The mass range was set to normal and the scan range mode was defined by the m/z range. In this work, a mass range between *m/z* 400 and 2000 was sufficient to detect the Protein samples. For Full MS scans, automatic gain control (AGC) was set to 5.0E5, and maximum injection time was set to 500 ms. For data-dependent analysis, MS2 were acquired at 60000 resolution using collision energy type EThcD with two microscans. Peptides with charge states 10-50 were selected for MS2 analysis.

F3. Data analysis

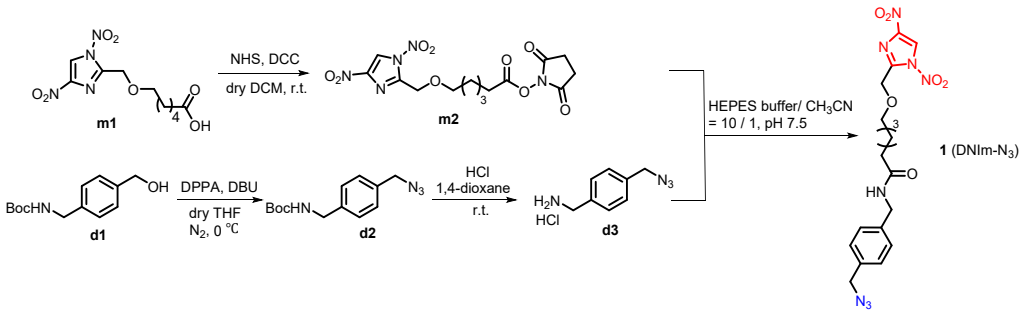
The raw data from the Thermo Scientific Orbitrap Ascend were analyzed with Thermo Scientific BioPharmaFinder™ 5.3 software (P/N OPTON-31029). For the

characterization of the total mass of the Protein samples, BSA protein sequence of 583 amino acids and 3Cl protein sequence of 306 amino acids were created in the protein sequence manager. For the study of Protein samples, the modification was considered as the variable modification. The average mass of Protein samples was calculated with the modified default ReSpect deconvolution method using the ReSpect™ algorithm and the identification automatically aligned with the structural information in protein sequence manager. And the conversion rate was calculated based on Fractional Abundance Algorithm of this software (BioPharmaFinder™) as per the equation: modified protein / (modified protein + unmodified protein).

F4. Experimental conditions of the CE ZipChip, Orbitrap Ascend and BioPharma Finder2.0 software.

CE ZipChip	
Chip:	HS
CE Electrical Field Strength:	500 V/cm
Pressure Assistance:	ON
Background Electrolyte Solution (BGE):	Intact Antibodies
Sample Injection:	2 nL
Orbitrap Ascend MS	
Mode:	Normal
Detector Type:	Orbitrap
m/z Range:	400-2000
Fragmentation:	EThcD
Resolution:	60000
Sheath Gas (Arb):	4
Capillary Temp:	300℃
S-lens RF Level:	60
BioPharma Finder 2.0 Software	
Deconvolution Algorithm:	ReSpect™

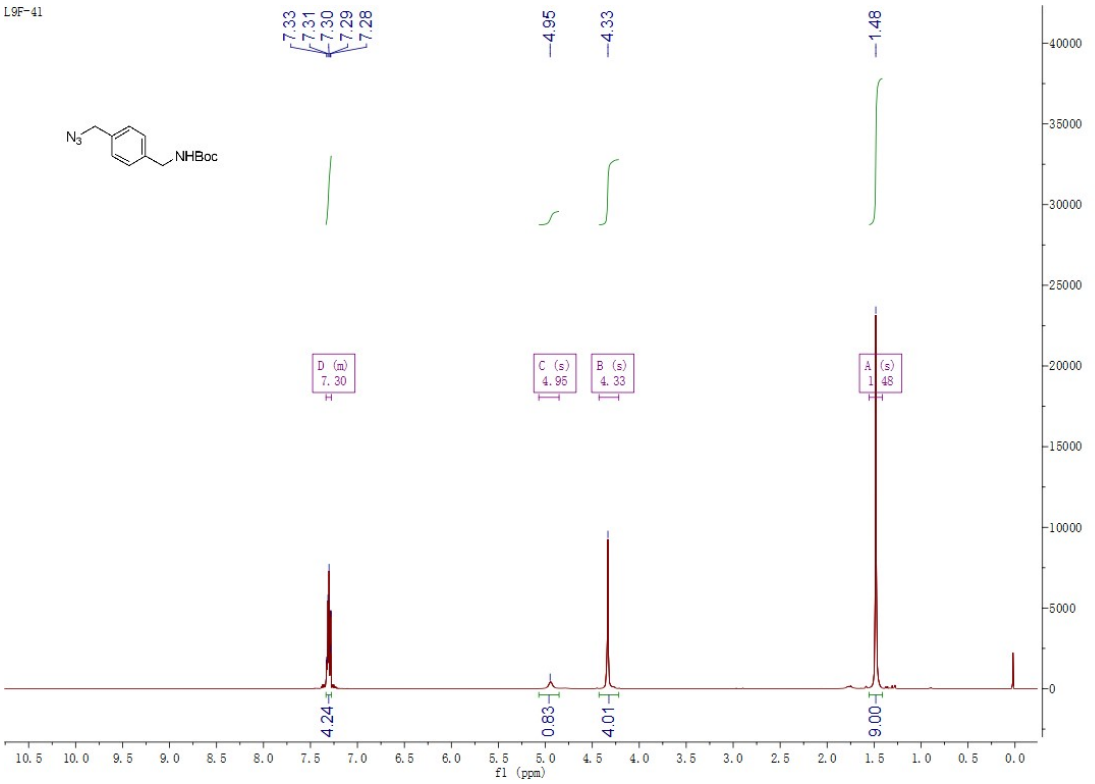
3. Synthesis of cross-linker 1 (DNIm-N_3).



Supplementary figure 1. Synthesis of cross-linker **1** (DNIm-N_3). To obtain the starting material **m1**, we proceeded the synthetic route referring to the reported method [2] with a slight modification. The procedure for preparing **1** starting from **m1**

was detailed as follows.

Synthesis of compound *tert*-butyl (4-(azidomethyl)benzyl)carbamate (d2): Under N_2 atmosphere and ice bath, to a two-neck flask were added successively compound **d1** (240 mg, 1.0 mmol) in dry THF (2 ml), and DPPA (0.3 ml, 1.66 mmol, 1.66 equiv) and DBU (0.15 ml, 1.15 mmol, 1.15 equiv) in dry THF (8 mL) via a syringe. The mixture was stirred for 18 h with the temperature gradually increasing to r.t. TLC analysis showed no starting material remained and the mixture was concentrated in vacuum. The resulting residue was subjected to flash silica gel chromatography (PE/EA = 20/1, v/v, then PE/EA = 10/1) to provide the title compound **d2** as a white solid (214 mg, 82% yield). ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.33 – 7.28 (m, 4H), 4.95 (s, 1H), 4.33 (s, 4H), 1.48 (s, 9H). The ^{1}H NMR spectrum was in accordance with that reported by Christian and his co-authors [3].



Supplementary figure 2. ^{1}H NMR spectrum of compound **d2**.

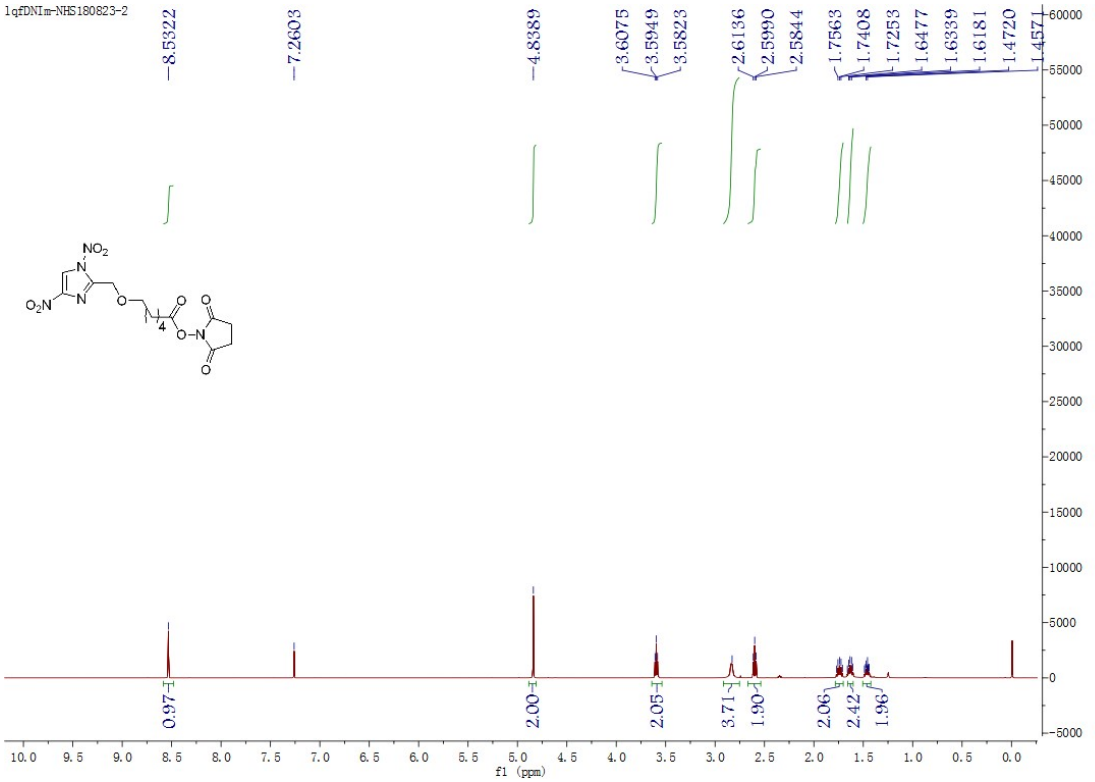
Synthesis of compound (4-(azidomethyl)phenyl)methanamine hydrochloride (d3): To a solution of precursor **d2** (16.5 mg, 63 μmol) in 1,4-dioxane (2 ml), 0.55 ml of HCl/1,4-dioxane solution (4 M HCL in 1,4-dioxane) was added slowly under 0 ˚C. And the mixture was stirred overnight until TLC analysis showed no starting material remained. Then it was concentrated in vacuum to give hydrochloride of compound **d3** as white solid and subjected to the next step.

Synthesis of compound 2, 5-dioxopyrrolidin-1-yl 6-((1,4-dinitro-1H-imidazol-2-yl) methoxy)hexanoate (m2): To a single-neck flask was added successively compound **m1** (130 mg, 0.43 mmol), DCC (97 mg, 0.47 mmol, 1.1 equiv), NHS (54 mg, 0.47 mmol, 1.1 equiv) and dry DCM (10 mL). The mixture was stirred at r.t. for 15.5 h before concentrated in vacuum. The resulting residue was subjected to flash silica gel

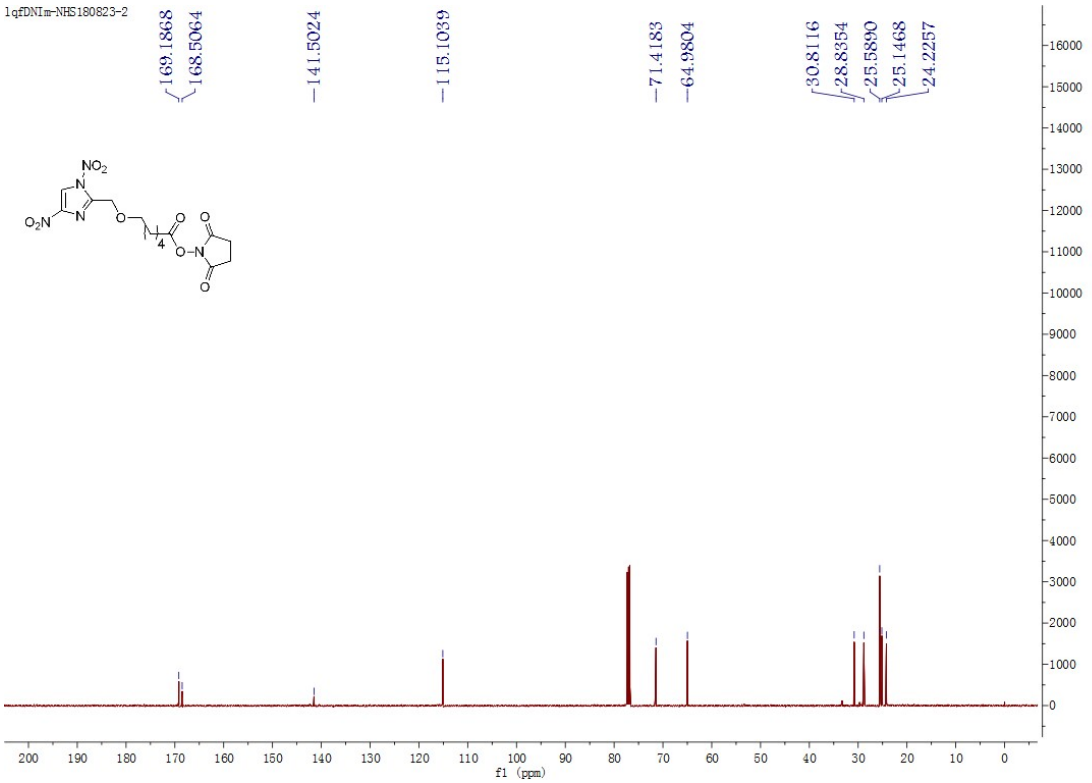
chromatography (PE/EA = 10/6, v/v) to provide the title compound **m2** as a white solid (127 mg, 74%). It was analyzed by NMR dissolved in $CDCl_3$. Its NMR spectra were in accordance with that reported in previous work [4].

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.53 (s, 1H), 4.84 (s, 2H), 3.59 (t, J = 6.3 Hz, 2H), 2.83 (s, 4H), 2.60 (t, J = 7.3 Hz, 2H), 1.78 – 1.70 (m, 2H), 1.63 (m, 2H), 1.46 (m, 2H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 169.19, 168.51, 141.50, 115.10, 71.42, 64.98, 30.81, 28.84, 25.59, 25.15, 24.23.



Supplementary figure 3. ^{1}H NMR spectrum of compound **m2**.



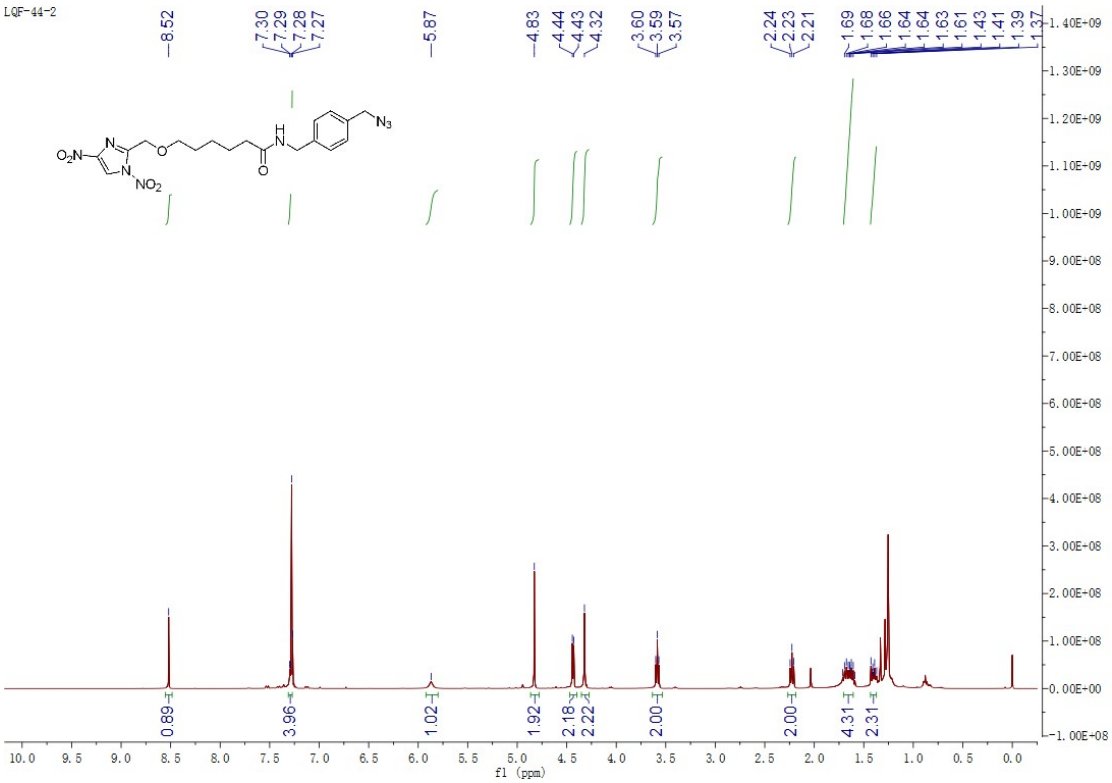
Supplementary figure 4. ^{13}C NMR spectrum of compound **m2**.

Synthesis of compound *N*-(4-(azidomethyl)benzyl)-6-((1,4-dinitro-1*H*-imidazol-2-yl)methoxy)hexanamide (1): The crude compound **d3** (0.063 mmol) was dissolved in 10 ml HEPES buffer (pH 7.5, 0.2 M) in a single-neck flask and stirred vigorously. Then compound **m2** (0.063 mmol) in 1 ml CH_3CN was added. The resulting mixture was stirred at r.t. for 1 h before extracted with DCM (60 ml + 20 ml). The organic layer was washed with brine (3 ml X 3), dried with anhydrous Na_2SO_4, filtered and concentrated. The resulting residue was subjected to flash silica gel chromatography (PE/EA/MeOH = 100/50/0.5, to 100/50/2, V/V/V) to afford the title compound **1** as a white solid (25 mg, 89% yield).

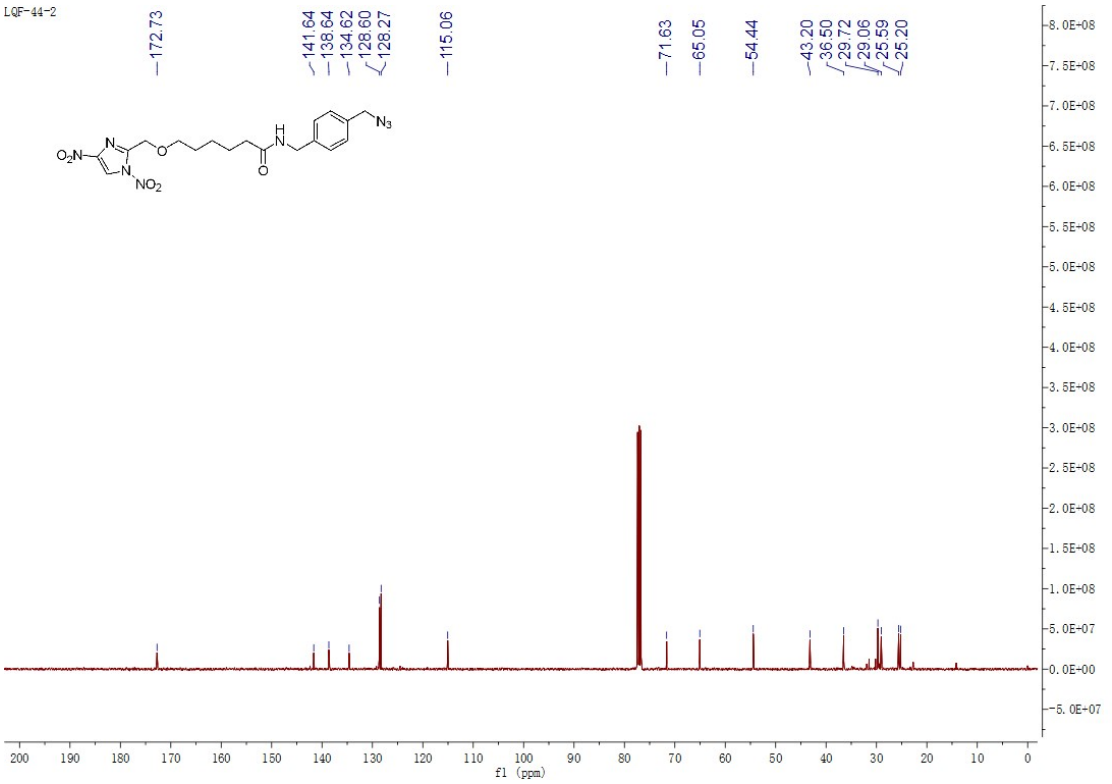
^{1}H NMR (400 MHz, $CDCl_3$) δ 8.52 (s, 1H), 7.31 – 7.27 (m, 4H), 5.87 (s, 1H), 4.83 (s, 2H), 4.44 (d, *J* = 5.8 Hz, 2H), 4.32 (s, 2H), 3.59 (t, *J* = 6.3 Hz, 2H), 2.23 (t, *J* = 7.5 Hz, 2H), 1.70 – 1.61 (m, 4H), 1.43 – 1.37 (m, 2H).

^{13}C NMR (100 MHz, $CDCl_3$) δ 172.73, 141.64, 138.64, 134.62, 128.60, 128.27, 115.06, 71.63, 65.05, 54.44, 43.20, 36.50, 29.72, 29.06, 25.59, 25.20.

HRMS (ESI) m/z calcd. for $C_{18}H_{23}N_8O_6^+$ $[M+H]^+$ 447.1735, found 447.1732.



Supplementary figure 5. ^{1}H NMR spectrum of cross-linker **1**.



Supplementary figure 6. ^{13}C NMR spectrum of cross-linker **1**.

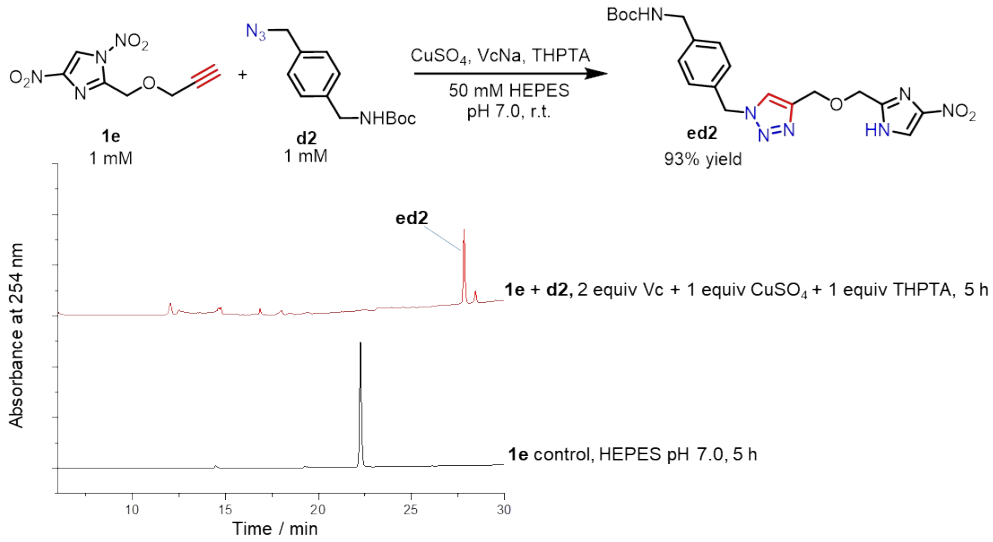
4. CuAAC reaction between cross-linker 1e and azide d2.

Cross-linker **1e** was synthesized according to previous method [2]. **1e** (1 mM) and azide **d2** (1 mM) was dissolved in neutral HEPES buffer (50 mM). Then $CuSO_4$ (1 equiv), THPTA (1 equiv) and VcNa (2 equiv) was added successively. The mixture was stirred for 5 h and then subjected to HPLC analysis (supplementary figure 7, red HPLC trace). A second sample with cross-linker **1e** (1 mM) incubated in neutral HEPES buffer (50 mM) for 5 h was also prepared as a control (black HPLC trace).

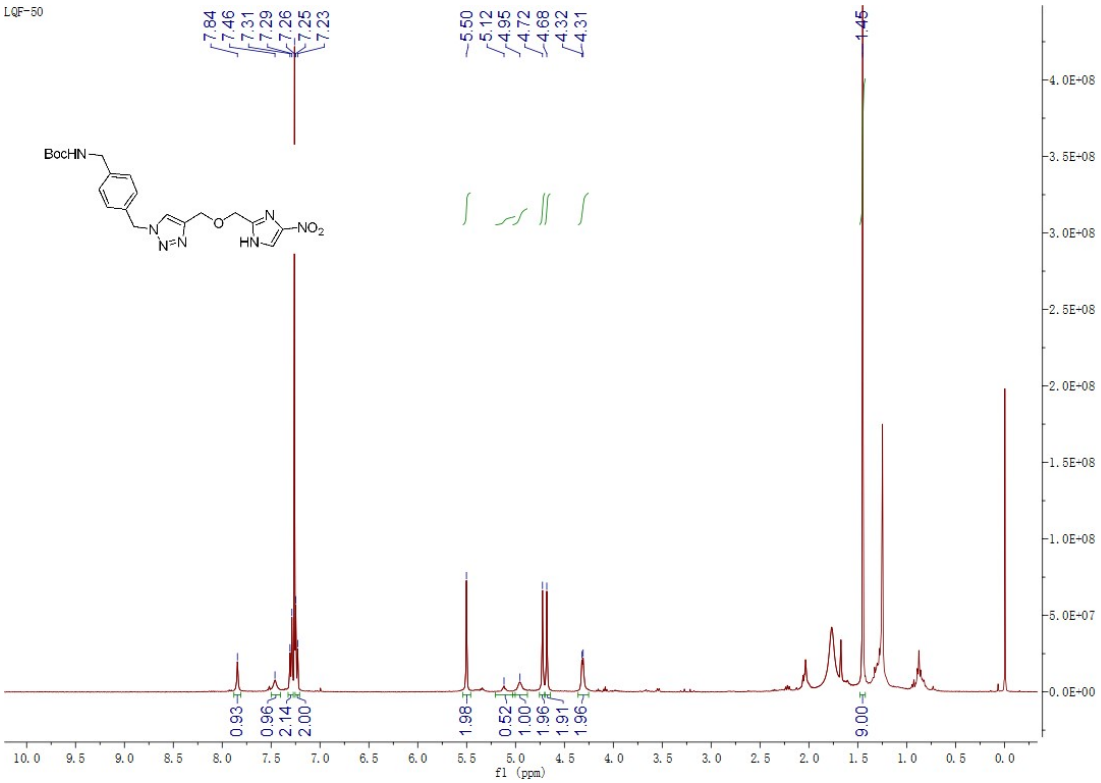
To obtain enough **ed2** for structure elucidation, we amplified the reaction of red HPLC trace for tens of times. After completion of the reaction showed by TLC analysis, the mixture was extracted with ethyl acetate. The organic layer was washed with brine, dried with anhydrous Na_2SO_4 and concentrated in vacuum. The residue was subjected to flash silica gel chromatography (DCM/MeOH = 60/1, to 40/1, V/V) to afford compound **ed2** as a brown solid (8 mg, 93% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.84 (s, 1H), 7.46 (s, 1H), 7.30 (d, J = 8.0 Hz, 2H), 7.24 (d, J = 8.1 Hz, 2H), 5.50 (s, 2H), 5.12 (s, 1H), 4.95 (s, 1H), 4.72 (s, 2H), 4.68 (s, 2H), 4.31 (d, J = 5.3 Hz, 2H), 1.45 (s, 9H). That the peak of H atom on imidazole ring is at 7.84 ppm denotes -NO_2 group at N1 position on the imidazole ring disappears, compared to that at 8.5 ppm in cross-linker **1**.

HRMS (ESI) m/z calcd. for $C_{20}H_{25}N_7NaO_5^+$[M+Na]$^+$ 466.1809, found 466.1808.

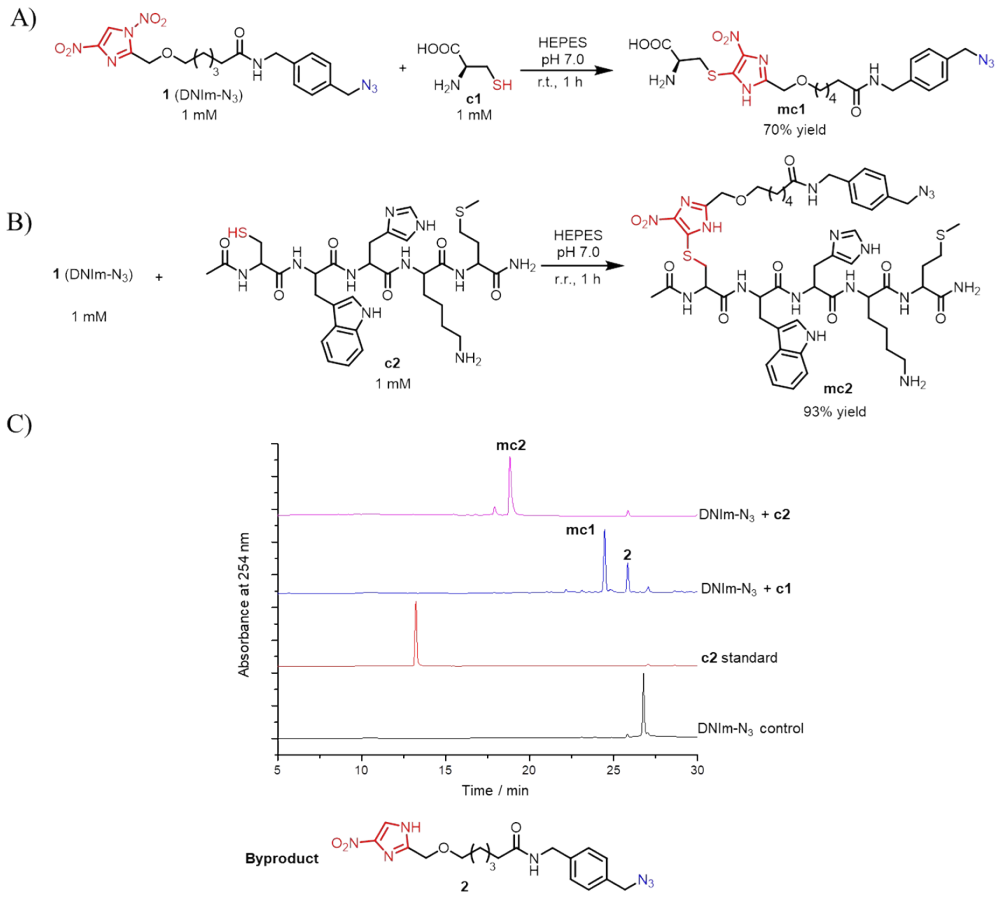


Supplementary figure 7. CuAAC reaction between compound **1e** and **d2** generated compound **ed2** as main product.

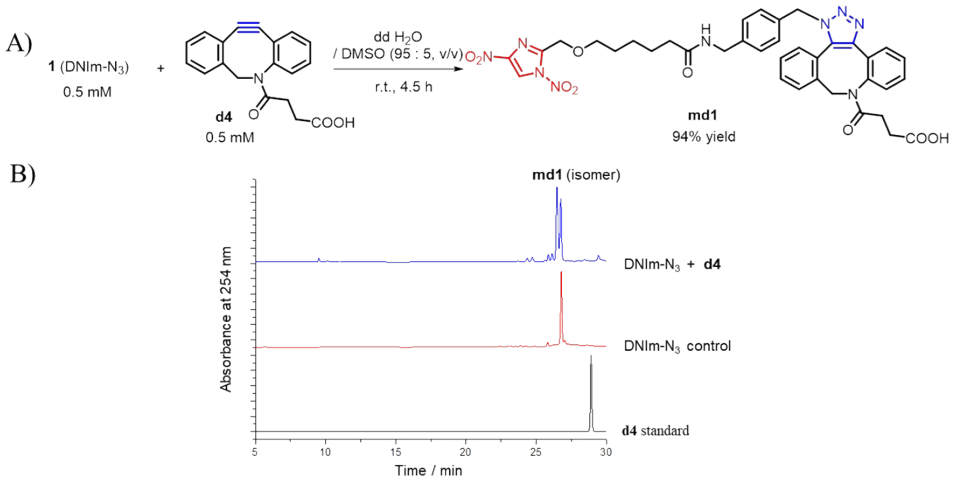


Supplementary figure 8. ^{1}H NMR spectrum of compound **ed2**.

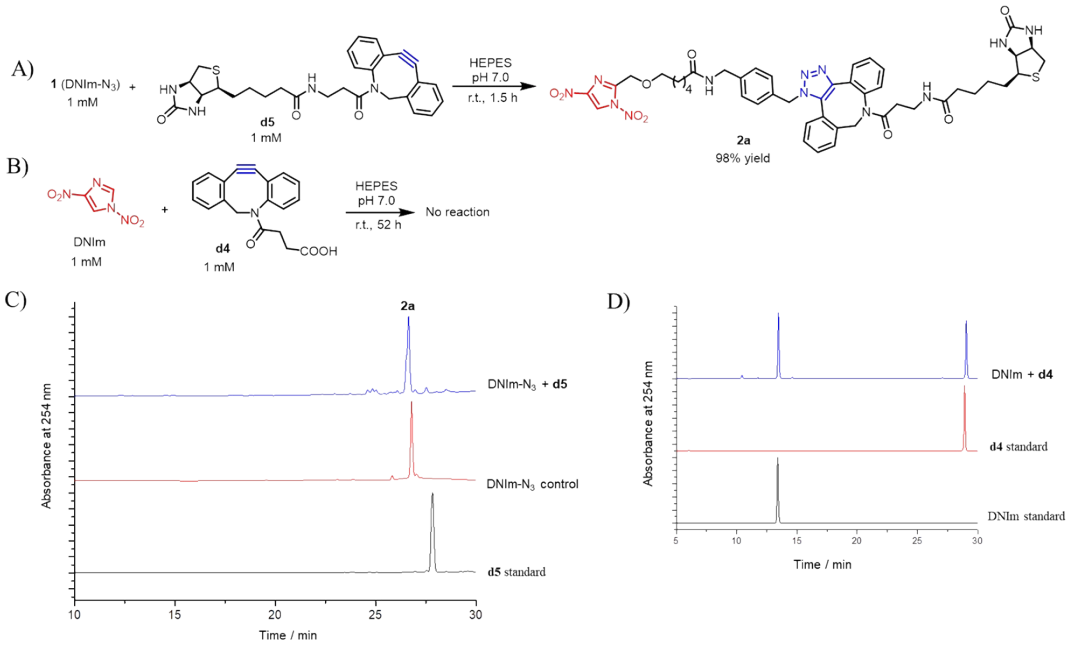
5. Orthogonal reaction of DNIm and azide motifs in DNIm-N_3 with Cys and BCN or DBCO derivatives.



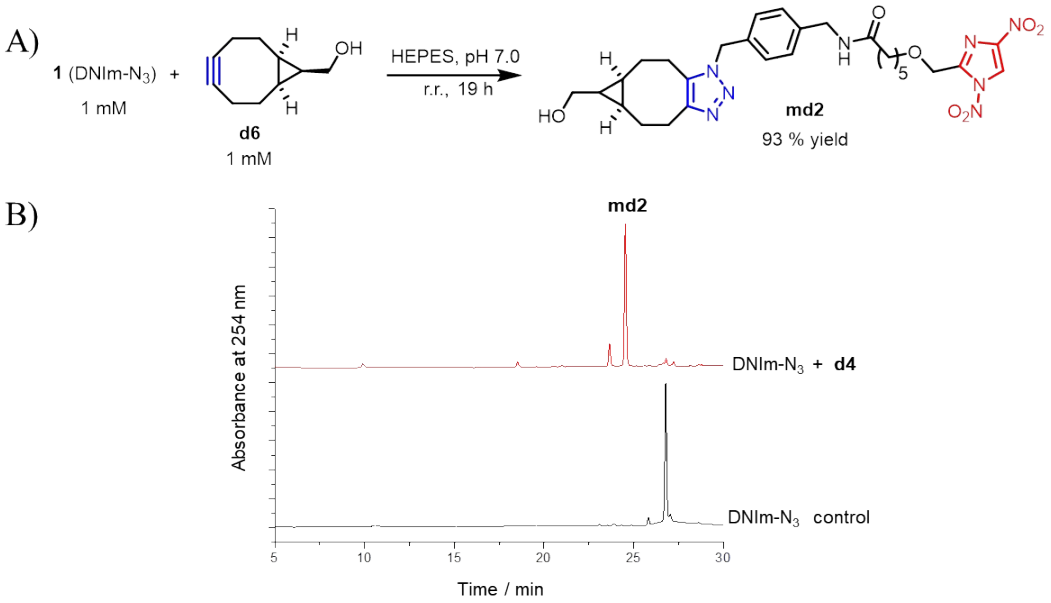
Supplementary figure 9. Reaction of DNIm motif in cross-linker **1** with Cys. A) Reaction of DNIm motif with L-Cysteine **c1**. B) Reaction of DNIm motif with cysteine-containing peptide **c2** (AcNH-CWHKM-$CONH_2$). C) HPLC spectrum of the reaction in A and B. Analytical data for **mc1**: HRMS (ESI) m/z calcd. for $C_{21}H_{29}N_8O_6S^+$ $[M+H]^+$ 521.1925, found 521.1929. **mc2**: HRMS (ESI) m/z calcd. for $C_{51}H_{71}N_{17}O_{10}S_2^{2+}$ $[M+2H]^{2+}$ 572.7500, found 572.7497. Analytical data for **2**: HRMS (ESI) m/z calcd. for $C_{18}H_{24}N_7O_4^+$ $[M+H]^+$ 402.1884, found 402.1881. Yield for **2**: 19%. Peptide **c2** was purchased from Jiangsu Shenlang Biotech CO,.LTD.



Supplementary figure 10. Reaction of azide motif in cross-linker **1** with compound **d4** bearing DBCO group. A) The reaction of azide motif with **d4**. B) HPLC spectrum of the reaction. Analytical data for **md1**: HRMS (ESI) m/z calcd. for $C_{37}H_{37}N_9NaO_9^+$ $[M+Na]^+$ 774.2606, found 774.2607.

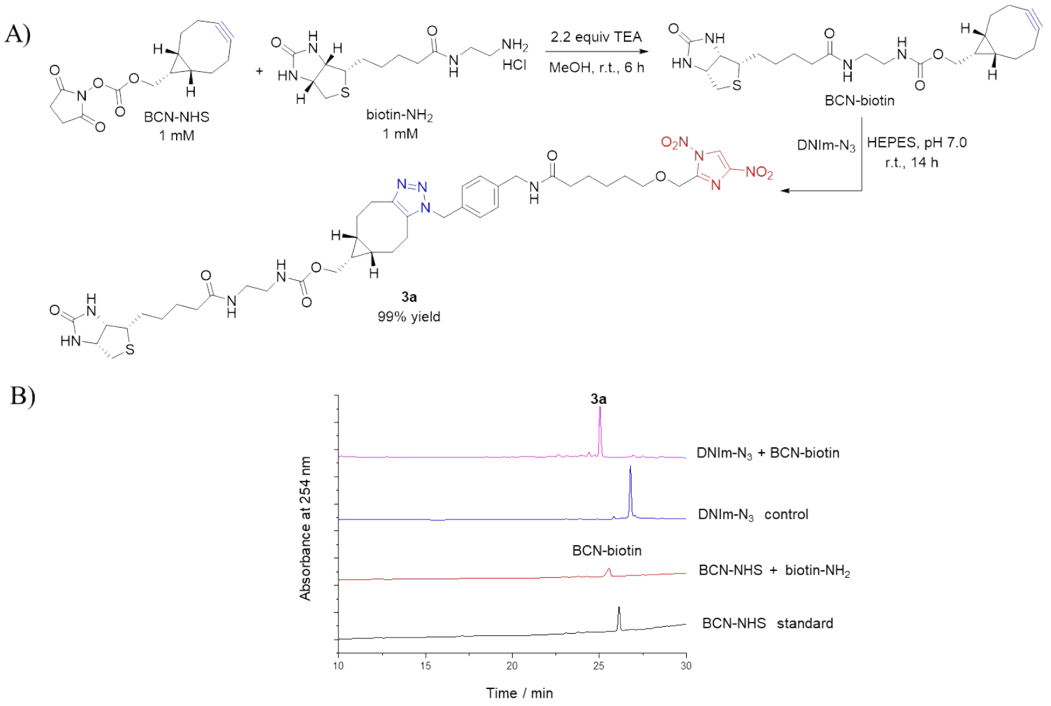


Supplementary figure 11. Reaction of azide motif in cross-linker DNIm-N_3 with **d5** (bearing DBCO group). A) The reaction of azide motif in cross-linker **1** with **d5**. B) The reaction of DNIm with **d4**. C) HPLC spectrum of the reaction between **1** and **d5.** D) HPLC spectrum of the reaction between DNIm and **d4.** Analytical data for **2a**: HRMS (ESI) m/z calcd. for $C_{46}H_{53}N_{12}O_9S^+$ $[M+H]^+$ 949.3774, found 949.3784.



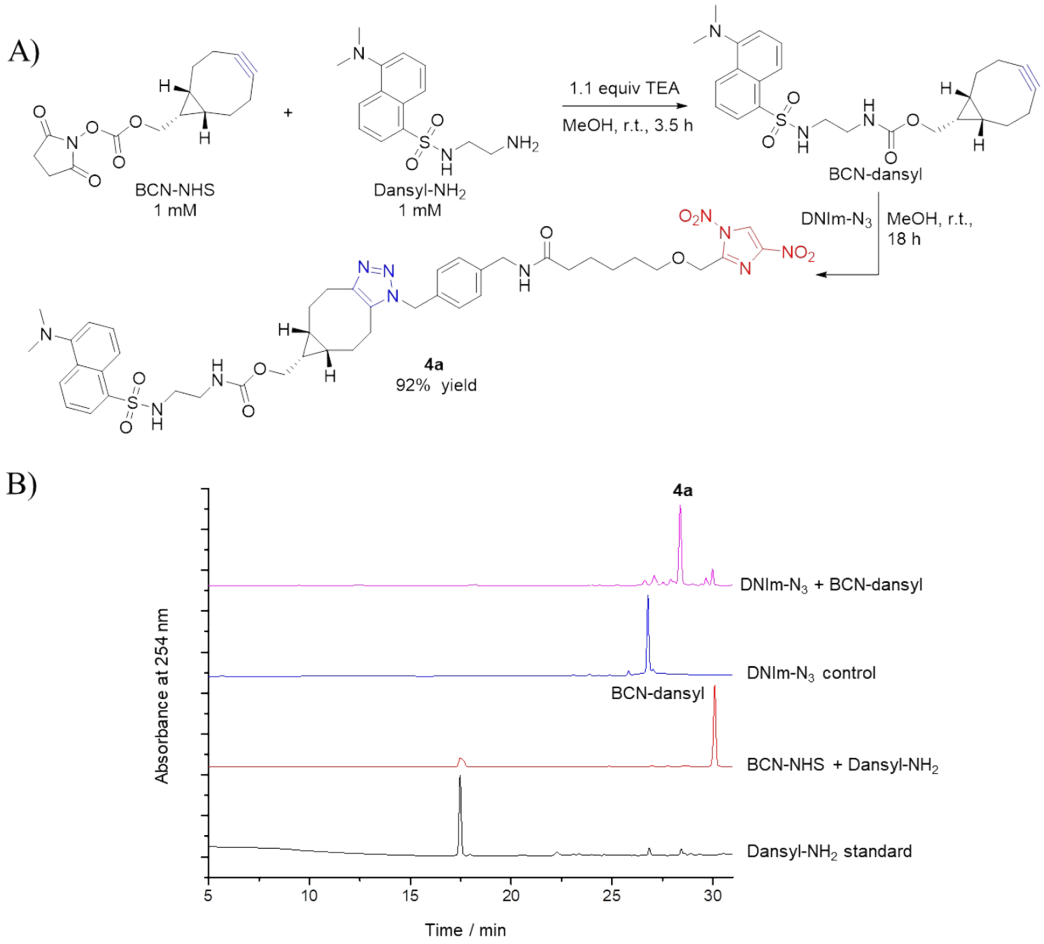
Supplementary figure 12. A) The reaction of azide motif in cross-linker **1** with **d6** bearing BCN ring. B) HPLC spectrum of the reaction between **1** and **d6**. Analytical data for **md2**: HRMS (ESI) m/z calcd. for $C_{28}H_{36}N_8NaO_7^+$ $[M+Na]^+$ 619.2599, found 619.2601.

6. Preparation of functional agents bearing DNIm warhead with cross-linker 1 via SPAAC reaction.

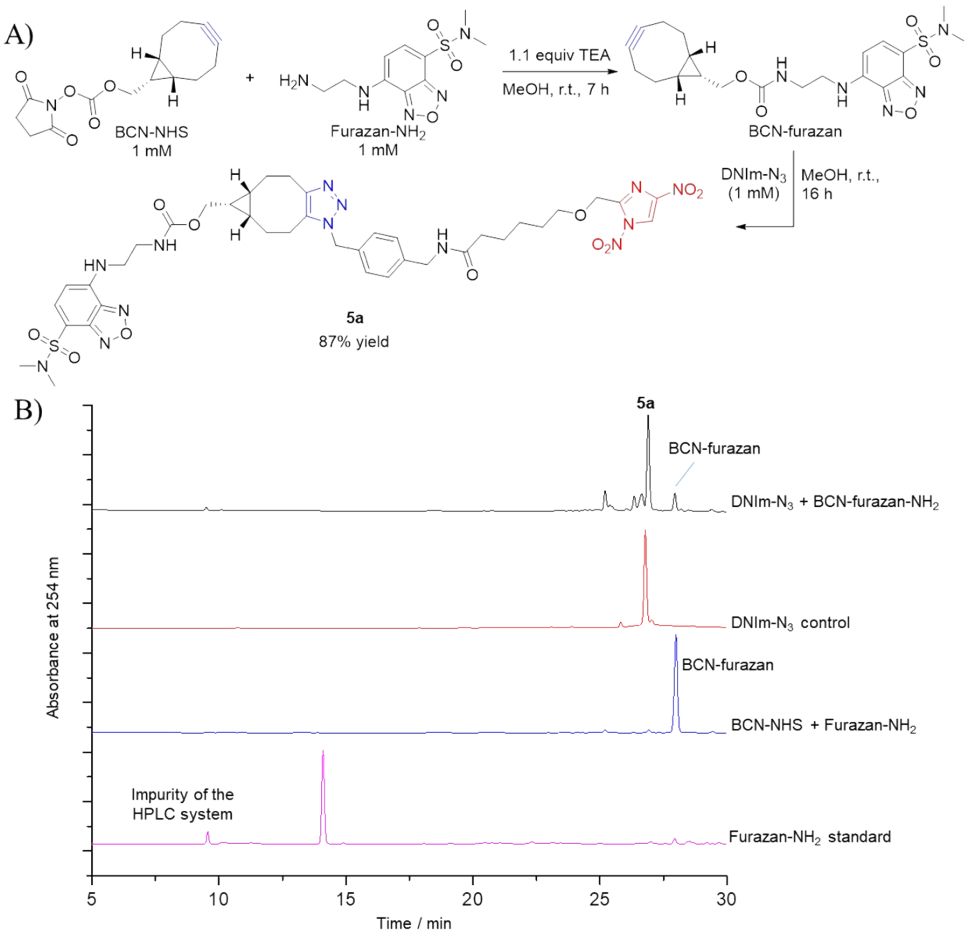


Supplementary figure 13. A) Synthesis of the biotin-based functional agent **3a** via cross-linker **1**. B) HPLC spectrum of the two steps of the synthetic route. Analytical data for BCN-biotin: HRMS (ESI) m/z calcd. for $C_{23}H_{34}N_4NaO_4S^+$ $[M+Na]^+$ 485.2193, found 485.2198. Analytical data for **3a**: HRMS (ESI) m/z calcd. for

$C_{41}H_{56}N_{12}NaO_{10}S^+$ $[M+Na]^+$ 931.3855, found 931.3861.



Supplementary figure 14. A) Synthesis of the dansyl-based functional agent **4a** via cross-linker **1**. B) HPLC spectrum of the two steps of the synthetic route. Analytical data for BCN-dansyl: HRMS (ESI) m/z calcd. for $C_{25}H_{31}N_3NaO_4S^+$ $[M+Na]^+$ 492.1927, found 492.1929. Analytical data for **4a**: HRMS (ESI) m/z calcd. for $C_{43}H_{55}N_{11}O_{10}S^{2+}$ $[M+2H]^{2+}$ 458.6922, found 458.6920.

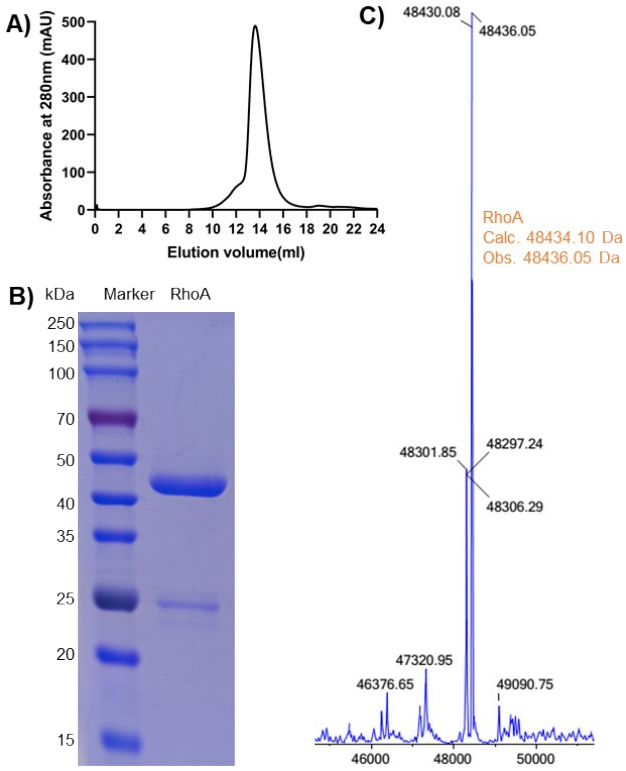


Supplementary figure 15. A) Synthesis of furazan-based functional agent **5a** via cross-linker **1**. B) HPLC spectrum of the two steps of the synthetic route. Analytical data for BCN-furazan: HRMS (ESI) m/z calcd. for $C_{21}H_{28}N_5O_5S^+$ $[M+H]^+$ 462.1806, found 462.1812. Analytical data for **5a**: HRMS (ESI) m/z calcd. for $C_{39}H_{50}N_{13}O_{11}S^+$ $[M+H]^+$ 908.3468, found 908.3472.

7. Preparation of the protein RhoA

The plasmid (P57370) encoding the RhoA was obtained from MiaoLingBio, China. Subsequently the plasmid was transformed into Escherichia coli BL21 (DE3) cells. The transformed bacteria were plated on Luria-Bertani (LB) agar containing 100 μg/mL ampicillin and incubated at 37°C. A single colony was picked from the plate and inoculated into 10 mL of LB medium containing 100 μg/mL ampicillin. This culture was incubated at 37°C with shaking at 200 rpm for 12 hours to generate the seed culture, which was subsequently diluted 1:100 into 1000 mL of LB medium. When the optical density at 600 nm reached 0.6, 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce RhoA protein expression. After 12 hours of growth at 18°C, the cells were harvested by centrifugation, resuspended, and lysed. The supernatant containing the target protein was collected by centrifugation and incubated with BeyoGold™ GST-tag Purification Resin (Beyotime Biotechnology, China) for 1 hour at 4°C, followed by loading onto a gravity column.

The target protein was eluted using elute buffer containing 10mM L-Glutathione reduced (Solarbio, China), and the eluent was further purified using a Superdex™ 200 Increase 10/300 GL column (Cytiva) via size exclusion chromatography on Protein Purification System (AKTA pure 25). The purified protein was analyzed by SDS-PAGE and confirmed by mass spectrometry.

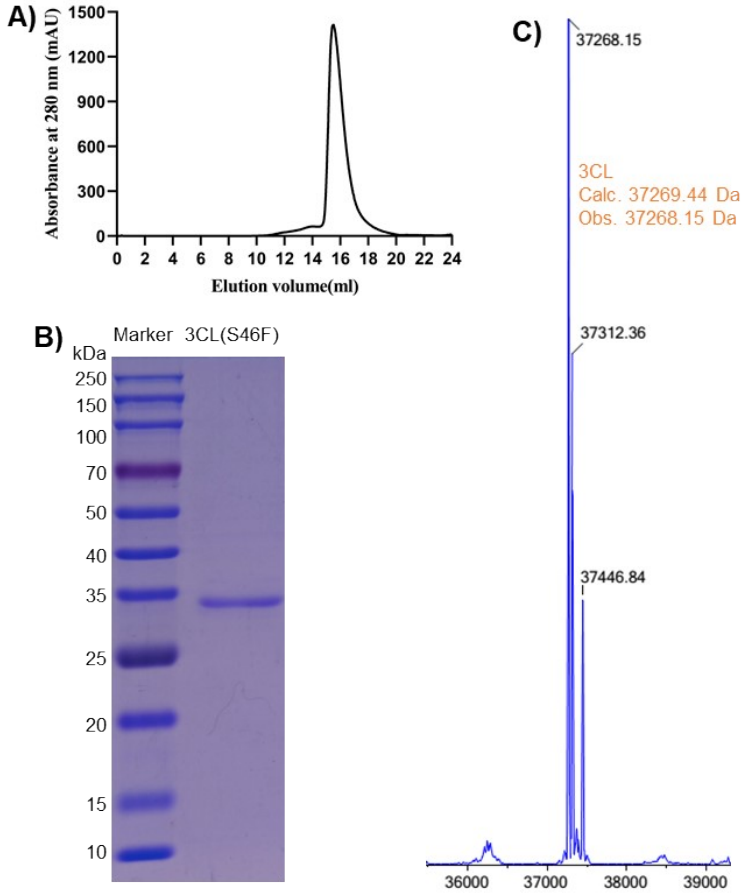


Supplementary figure 16. A) Spectrum of purification of the protein RhoA by size exclusion chromatography. B) Result for SDS-PAGE analysis of the obtained protein RhoA. C) Deconvoluted mass spectrum of RhoA.

8. Preparation of the protein 3CL

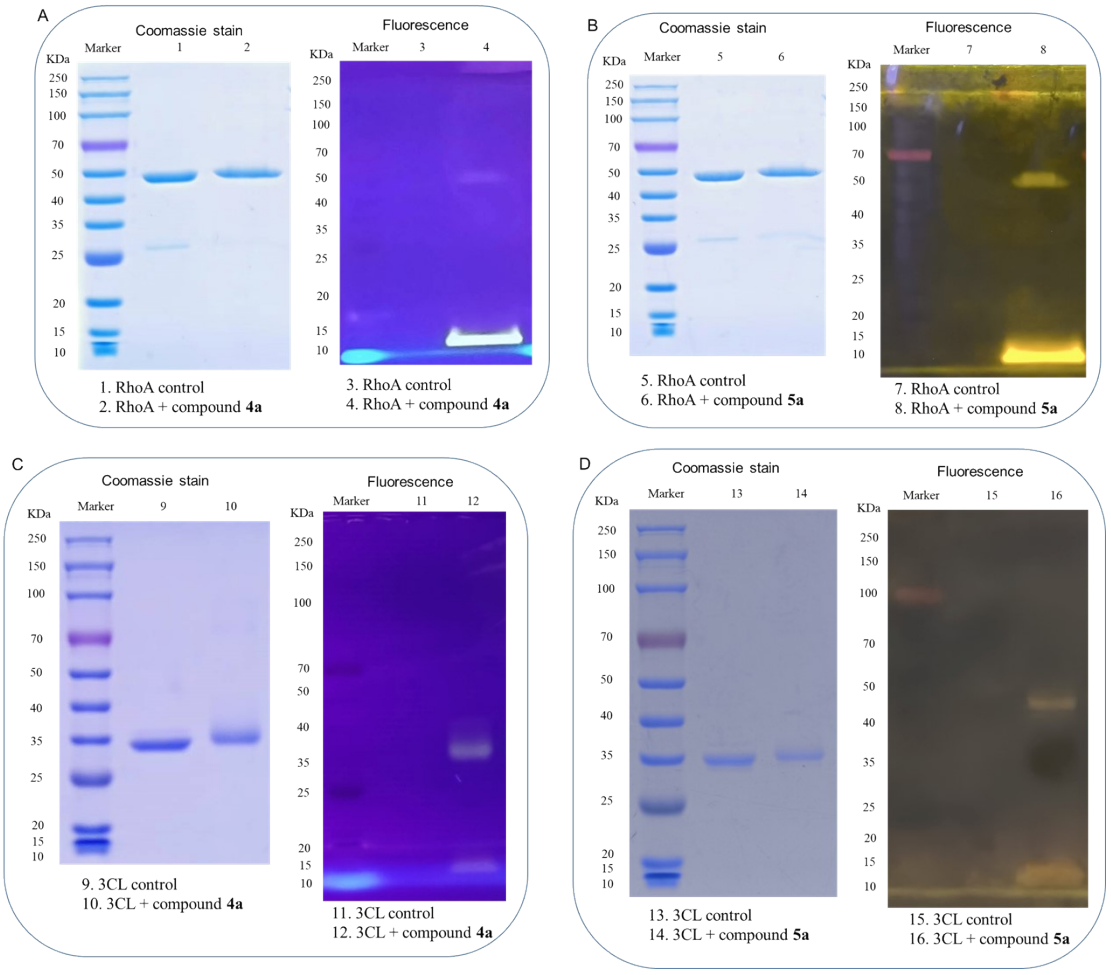
The full-length gene encoding the 3CL(S46F) variant of SARS-CoV-2 was codon-optimized, synthesized, and cloned into a pET-28a vector with a 6-histidine tag at the N-terminus and a stop codon at the C-terminus. The recombinant plasmid was then transformed into Escherichia coli BL21 (DE3) cells, which were plated on Luria-Bertani (LB) agar containing 50 μg/mL kanamycin and incubated at 37°C. A single colony was picked from the plate and inoculated into 5 mL of LB medium containing 50 μg/mL kanamycin. This culture was incubated at 37°C with shaking at 200 rpm for 12 hours to generate the seed culture, which was subsequently diluted 1:100 into 200 mL of LB medium. When the optical density at 600 nm reached 0.6, 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce 3CL protein expression. After 12 hours of growth at 18°C, the cells were harvested by centrifugation, resuspended, and lysed. The supernatant containing the target protein was collected by centrifugation and incubated with Ni-NTA resin (GE Healthcare, USA) for 1 hour at 4°C, followed by loading onto a gravity column. The target

protein was eluted using an imidazole gradient, and the eluent was further purified using a Superdex™ 200 Increase 10/300 GL column (Cytiva) via size exclusion chromatography on Protein Purification System (AKTA pure 25). The purified protein was analyzed by SDS-PAGE and confirmed by mass spectrometry.



Supplementary figure 17. A) Spectrum of purification of the protein 3CL by size exclusion chromatography. B) Result for SDS-PAGE analysis of the obtained protein 3CL. C) Deconvoluted mass spectrum of 3CL.

9. The full gel spectra for the SDS-PAGE analysis of fluorescently labeled proteins

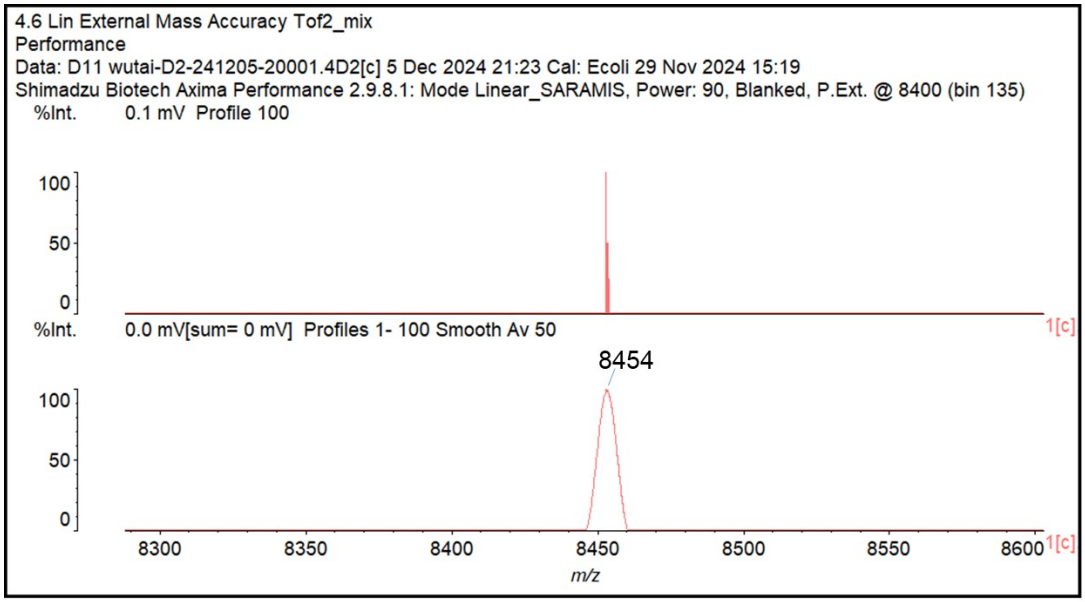


Supplementary figure 18. The full gel spectra for the SDS-PAGE analysis of proteins modified by compound **4a** and **5a**. On the left is the full Coomassie staining image, and on the right is the full fluorescence image. A) The full gel spectra for the SDS-PAGE analysis of RhoA modified by compound **4a**. The Coomassie staining and fluorescence images in Fig. 6B in the main text were cut from these full gel spectra. B) The full gel spectra for the SDS-PAGE analysis of RhoA modified by compound **5a**. The Coomassie staining and fluorescence images in Fig. 6D in the main text were cut from these full gel spectra. C) The full gel spectra for the SDS-PAGE analysis of 3CL modified by compound **4a**. The Coomassie staining and fluorescence images in Fig. 6C in the main text were cut from these full gel spectra. D) The full gel spectra for the SDS-PAGE analysis of 3CL modified by compound **5a**. The Coomassie staining and fluorescence images in Fig. 6E in the main text were cut from these full gel spectra.

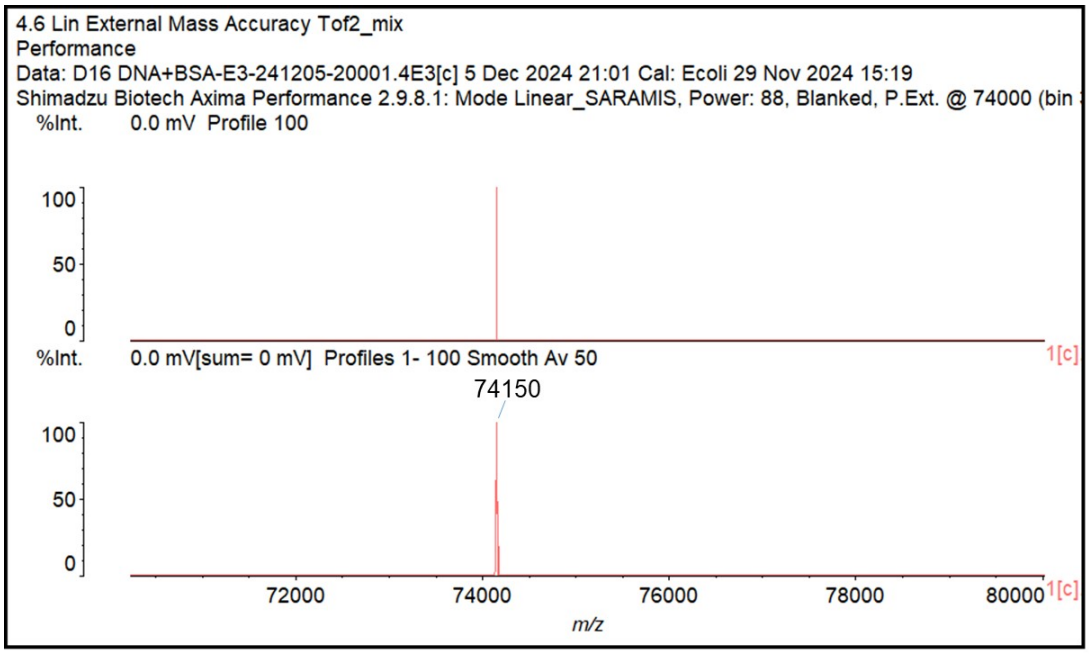
10. MALDI-TOF analysis of the oligonucleotide conjugates

The oligonucleotide conjugates were analyzed by MALDI-TOF mass. The test was

performed on SHIMADZU iD plus Performance. Matrix: α-Cyano-4-hydroxycinnamic acid (CHCA); Linear mode (positive or negtive). However, we only obtained the mass signals of oligonucleotide-peptide and oligonucleotide-BSA, and that of oligonucleotide-PEG was not detected, regardless of in positive or negative linear mode.

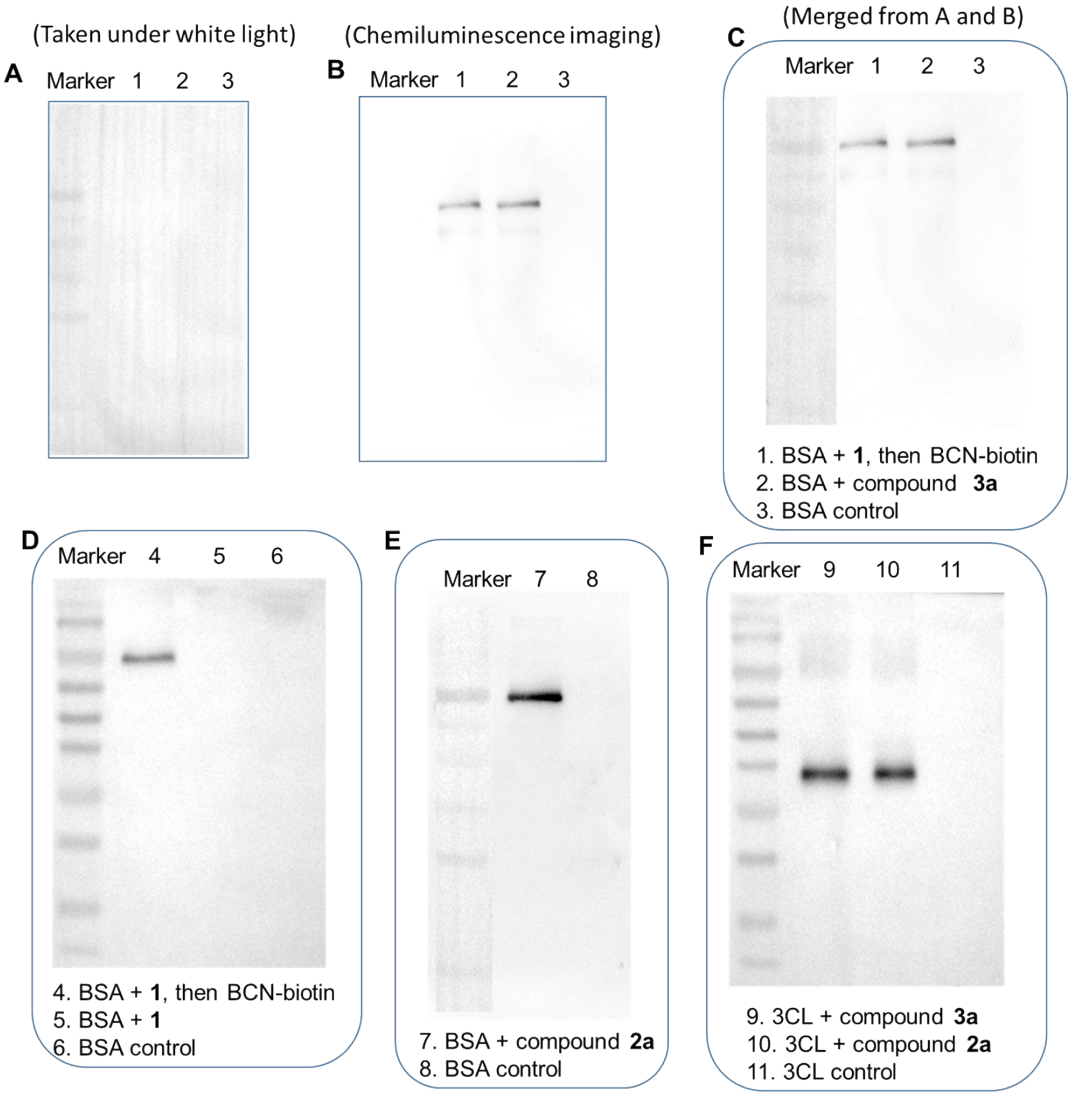


Supplementary figure 19. MALDI-TOF mass of the oligonucleotide-peptide conjugate in lane 4 of Fig. 7B in the main text. Mass for the conjugate oligonucleotide-peptide (Da): calculated 8460, found 8454.

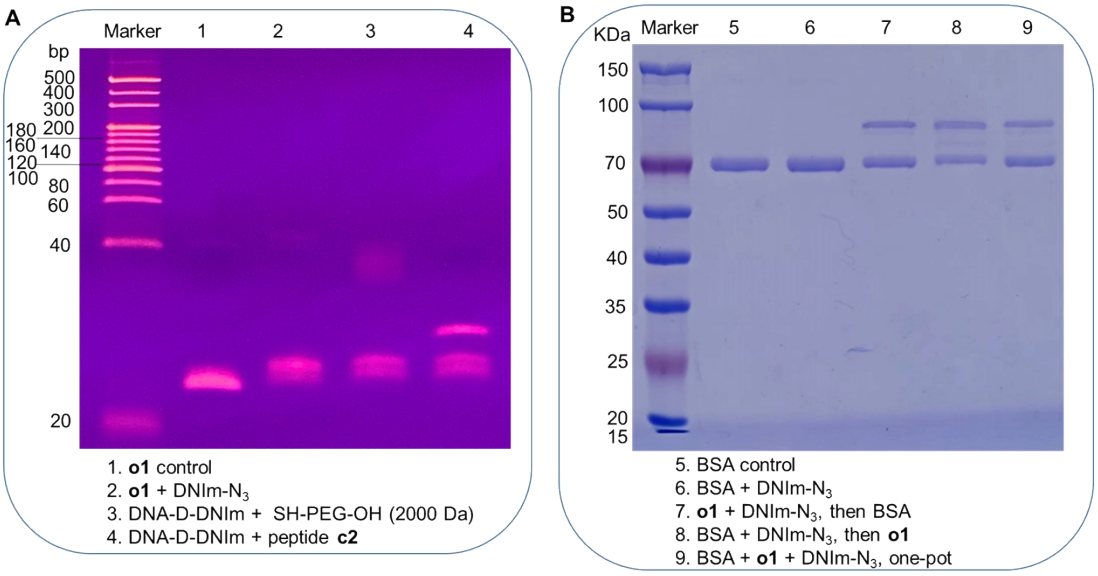


Supplementary figure 20. MALDI-TOF mass of the oligonucleotide-BSA conjugate in lane 9 of Fig. 7C in the main text. The three samples in lane 7-9 of Fig. 7C all showed the same peak around 74150 Da, while only one was presented here. Mass for the conjugate oligonucleotide-BSA (Da): calculated 74143, found 74150.

11. The original figures for the biotinylated proteins and oligonucleotide bioconjugates analysis.



Supplementary figure 21. The original figures for western blot analysis of the biotinylated proteins. A) Figure for western blot analysis of biotinylated BSA modified by compound **3a** taken under white light. It was used to see the marker; B) Figure for western blot analysis of biotinylated BSA taken by chemiluminescence imaging. It was the uncropped version of Fig. 4C; C) A merged version from A and B of Supplementary figure 21. It was also the uncropped version of Fig. 4C; D) Uncropped version for Fig. 4D; E) Uncropped version for Fig. 4E ; F) Uncropped version for Fig. 4F.



Supplementary figure 22. The original figures for oligonucleotide bioconjugates analysis. A) Uncropped version for Fig. 7B; B) Uncropped version for Fig. 7C.

References

1. Hu, Q.; Xiong, Y.; Zhu, G. H.; Zhang, Y. N.; Zhang, Y. W.; Huang, P.; Ge, G. B., The SARS-CoV-2 main protease (M(pro)): Structure, function, and emerging therapies for COVID-19. *MedComm* **2020,** *3* (3).

2. Luo, Q.; Tao, Y.; Sheng, W.; Lu, J.; Wang, H., Dinitroimidazoles as bifunctional bioconjugation reagents for protein functionalization and peptide macrocyclization. *Nat Commun* **2019,** *10* (1), 142.

3. Büll, C.; Heise, T.; van Hilten, N.; Pijnenborg, J. F.; Bloemendal, V. R.; Gerrits, L.; Kers-Rebel, E. D.; Ritschel, T.; den Brok, M. H.; Adema, G. J.; Boltje, T. J., Steering Siglec-Sialic Acid Interactions on Living Cells using Bioorthogonal Chemistry. *Angew Chem Int Ed Engl* **2017,** *56* (12), 3309-3313.

4. Hua, Y.; Liu, S.; Xie, S. S.; Shi, L.; Li, J.; Luo, Q., Heterobifunctional Cross-Linker with Dinitroimidazole and N-Hydroxysuccinimide Ester Motifs for Protein Functionalization and Cysteine-Lysine Peptide Stapling. *Org Lett* **2023,** *25* (49), 8792-8796.