Supporting information

Chemoselective Dual Labelling of Native and Recombinant Proteins

Bikram Keshari Agrawalla[‡,†] Tao Wang,[†,§] Andreas Riegger,[‡,†] Matthias P. Domogalla,[¥,‡] Kerstin Steinbrink,[¥] Thilo Dörfler,[‼] Xi Chen,[†] Felix Boldt,[†] Markus Lamla,[†] Jens Michaelis,[‼] Seah Ling Kuan,*[,‡,†] and Tanja Weil*[,‡,†]

[‡]Max-Planck-Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany.

[†]Institute of Inorganic Chemistry I, Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany.

[§]Key Laboratory of Advanced Technologies of Materials, Ministry of Education, School of Materials Science and Engineering, Southwest Jiaotong University, Chengdu 610031, P.R. China,

[¥]Department of Dermatology, University Medical Center Mainz, Johannes Gutenberg-University Mainz, Mainz D-55099, Germany

[‼]Institute of Biophysics, Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany

Corresponding Authors email address.

*weil@mpip-mainz.mpg.de

*kuan@mpip-mainz.mpg.de

1. General Experimental

Unless noted, all operations were performed without taking precautions to exclude air and moisture. All solvents and reagents were purchased from commercial sources and were used without further purification. The sulfo-cyanine3-maleimide (Cy3-MI) and cyanine5-N-Hydroxysuccinimide-ester (Cy5-NHS) were purchased from Lumiprobe. 7-Diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM) and Alexa fluor 647 C2 maleimide (Alexa647-MI) were purchased from Thermo Fisher Life Technologies. Atto550-maleimide (Atto550-MI) was purchased from ATTO-TEC GmbH. The peptide (CATSICGGGC, 95% purity) was purchased from GL Biochem China. Recombinant Human Interleukin-2 was received from Cell Science, MA. MALDI-TOF-MS spectra were acquired on a Bruker Reflex III. HR-MALDI-MS was recorded on a Solarix (Bruker) FTICR-MS. The absorbance and emission measurements were recorded using TECAN M1000 microplate reader.

2. Sequential modification of Interleukin-2

2.1 Structure of sulfo-cyanine3-maleimide and allyl-sulfone-cyanine5

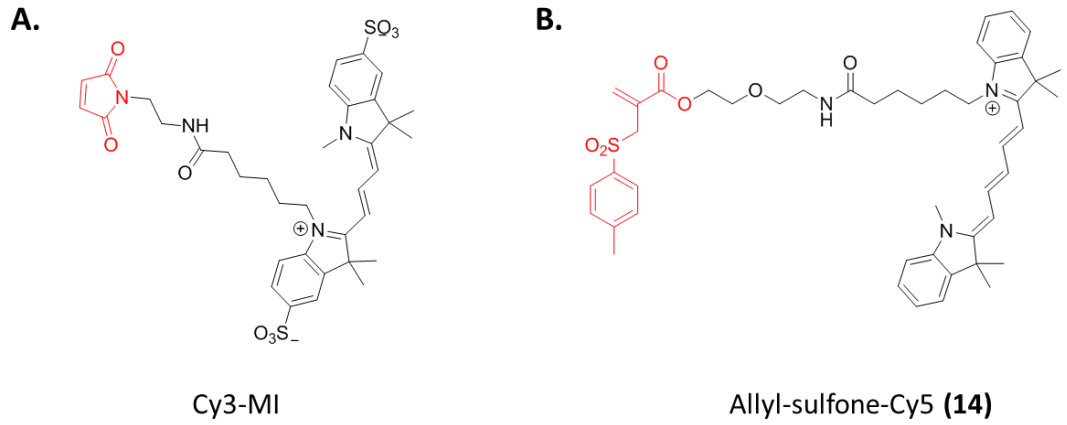
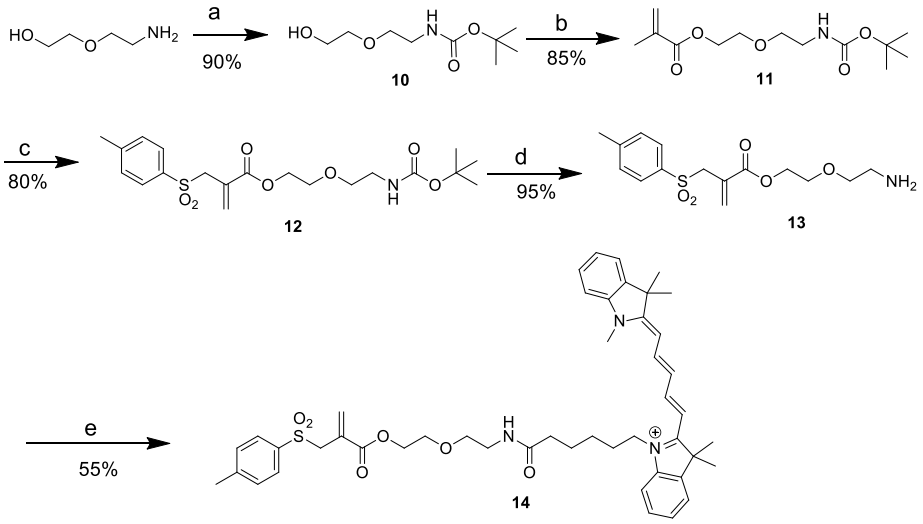


Figure S1. Chemical structure of (A) sulfo-cyanine3-maleimide (Cy3-MI) and (B) allyl-sulfone-cyanine5 (Allyl-sulfone-Cy5 **14**)

2.2 Synthesis and characterization of 3,3-dimethyl-1-(6-oxo-6-((2-(2-((2-(tosylmethyl)acryloyl)oxy)ethoxy)ethyl)amino)hexyl)-2-((1E,3E)-5-((Z)-1,3,3-trimethylindolin-2-ylidene)penta-1,3-dien-1-yl)-3H-indol-1-ium (Allyl-sulfone-Cy5, 14).



Scheme S1. The synthetic scheme of Allyl-sulfone-Cy5 **14**. (a) Triethylamine (TEA), di-tert-butyl dicarbonate (Boc), dry dichloromethane (DCM), overnight; (b) TEA, methacryloyl chloride, dry DCM, overnight; (c) 1. Sodium p-toluenesulfinate, Iodine, DCM, 3 days, 2. TEA, DCM, overnight, 3. TEA, Ethyl acetate (EA), 95 °C, overnight; (d) Trifluoroacetic acid (TFA), DCM; (e) Cy5-NHS, TEA, dry dimethylformamide (DMF), 24 h.

The Allyl-sulfone-Cy5 **14** was synthesized following the synthetic procedure reported earlier.[1]

Synthesis of tert-butyl (2-(2-hydroxyethoxy)ethyl)carbamate 10.

In a 250 mL round bottom flask, 2-(2-aminoethoxy)ethanol (3.15 g, 30 mmol) was dissolved in 80 mL of dichloromethane (DCM) and Triethylamine (TEA, 5 mL, 36 mmol, 1.2 equiv.) was added. Thereafter, di-tert-butyl dicarbonate (Boc, 7.85 g, 36 mmol, 1.2 equiv.) dissolved in 10 mL of DCM and dropped slowly to the solution. The reaction mixture was stirred overnight at room temperature and the solvent was removed under vacuum. The residue was purified by column chromatography with ethyl acetate (EA)/Hexane (2:1) to afford 5.8 g of the product **10** as colorless oil in 94% yield. ^{1}H NMR (300 MHz, Chloroform-*d*) δ 5.15 (s, 1H), 3.67 (dd, *J* = 5.3, 3.7 Hz, 2H), 3.55 – 3.43 (m, 4H), 3.26 (q, *J* = 5.4 Hz, 2H), 1.38 (s, 9H). ^{13}C NMR (75 MHz, Chloroform-*d*) δ 156.24, 79.39, 72.35, 70.37, 61.69, 40.45, 28.47. LC-MS (ESI): m/z = 228 $[M+Na]^+$ (calcd. mass: 205.13, formula: $C_9H_{19}NO_4$).

Synthesis of 2-(2-((tert-butoxycarbonyl)amino)ethoxy)ethyl methacrylate 11.

In a 250 mL Schlenk flask, tert-butyl (2-(2-hydroxyethoxy)ethyl)carbamate **10** (1.85 g, 9 mmol) and TEA (2.22 mL,16 mmol, 1.8 equiv.) were dissolved in 60 mL of dry DCM and cooled to 0 °C. Then methacryloyl chloride (1.62 mL, 16 mmol,

1.8 equiv.) was added dropwise to the solution. After stirring overnight at RT, the reaction mixture was filtered and the solvent was evaporated by rotary evaporator. The residue was purified by column chromatography with EA/Hexane (1:4) to afford 2.1 g of the ester **11** as colorless oil in 85% yield. ^{1}H NMR (300 MHz, Chloroform-*d*) δ 6.13 (dq, *J* = 2.0, 1.0 Hz, 1H), 5.58 (p, *J* = 1.6 Hz, 1H), 4.33 – 4.24 (m, 2H), 3.74 – 3.65 (m, 2H), 3.54 (dd, *J* = 5.6, 4.8 Hz, 2H), 3.30 (q, *J* = 4.9 Hz, 2H), 1.95 (dd, *J* = 1.6, 1.0 Hz, 3H), 1.43 (s, 9H). ^{13}C NMR (75 MHz, Chloroform-*d*) δ 168.64, 158.28, 137.60, 126.31, 80.00, 70.96, 69.85, 64.99, 49.85, 49.57, 49.28, 49.00, 48.72, 48.43, 48.15, 41.19, 28.75, 18.45. LC-MS (ESI): m/z = 296 $[M+Na]^+$ (calcd. mass: 273.16, formula: $C_{13}H_{23}NO_5$).

Synthesis of 2-(2-((tert-butoxycarbonyl)amino)ethoxy)ethyl 2-(tosylmethyl)acrylate 12.

Compound **11** (1.4 g, 5.12 mmol) was dissolved in 50 mL of DCM. Sodium p-toluenesulfinate (1.37 g, 7.69 mmol, 1.5 equiv.) and Iodine (968 mg, 7.69 mmol, 1.5 equiv.) were added sequentially. The suspension was stirred for three days at room temperature before TEA (1.4 mL, 10.2 mmol, 2 equiv.) was injected slowly at 0 °C. After stirring overnight, the organic layer was washed with H_2O, $NaHCO_3$, $Na_2S_2O_3$ and brine solution. The organic layer was dried over anhydrous $MgSO_4$ and the solvent was evaporated. The residue was dissolved in 20 mL of EA and TEA (1.4 mL, 10.2 mmol, 2 equiv.) was added. The mixture was then refluxed overnight at 95 °C and the solvent was removed under vacuum. The residue was purified by column chromatography with EA/Hexane (1:1.5) to afford 1.75 g of the product **12** as yellowish oil in 80% yield. ^{1}H NMR (300 MHz, Chloroform-*d*) δ 7.77 – 7.67 (m, 2H), 7.31 (d, *J* = 8.0 Hz, 2H), 6.49 (s, 1H), 5.84 (s, 1H), 4.20 – 4.03 (m, 5H), 3.64 – 3.55 (m, 2H), 3.50 (t, *J* = 5.2 Hz, 2H), 3.29 (q, *J* = 5.3 Hz, 2H), 2.42 (s, 3H), 1.41 (s, 9H). ^{13}C NMR (75 MHz, Chloroform-*d*) δ 164.84, 155.97, 144.91, 135.40, 133.57, 129.70, 128.91, 128.73, 77.49, 77.06, 76.64, 70.15, 68.54, 64.34, 60.39, 57.63, 28.38, 21.65. LC-MS (ESI): m/z = 450 $[M+Na]^+$, 328 $[M-Boc+H]^+$ (calcd. mass: 427.17, formula: $C_{20}H_{29}NO_7S$).

Synthesis of 2-(2-aminoethoxy)ethyl 2-(tosylmethyl)acrylate 13.

In a 100 mL round bottom flask, compound **12** (1 g, 2.3 mmol) was dissolved in 20 mL of DCM and trifluoroacetic acid (TFA, 3.86 mL, 46 mmol, 20 equiv.) was added. The resulting mixture was stirred overnight at RT. The DCM and TFA were removed under vacuum to obtain 750 mg of the allyl-sulfone amine **13** as yellow-brown oil in 98% yield. ^{1}H NMR (300 MHz, Methanol-d_4) δ 7.75 – 7.67 (m, 2H), 7.42 (d, *J* = 8.0 Hz, 2H), 6.40 (s, 1H), 5.68 (d, *J* = 1.2 Hz, 1H), 4.32 – 4.18 (m, 4H), 3.80 – 3.62 (m, 5H), 3.12 (t, *J* = 5.0 Hz, 2H), 2.43 (s, 3H). ^{13}C NMR (75 MHz, Methanol-*d4*) δ 166.60, 146.75, 136.54, 134.13, 130.92, 130.87, 129.75, 70.11, 67.69, 65.25, 58.51, 48.98, 40.64, 21.56. LC-MS (ESI): m/z = 328 $[M+H]^+$, 655 $[2M+H]^+$ (calcd. mass: 327.11, formula: $C_{15}H_{21}NO_5S$).

Synthesis of allyl-sulfone-Cy5,14.

Under argon atmosphere, Cy5-NHS (4 mg, 6.5 µmol) and **13** (4 mg, 13 µmol, 2 equiv.) were dissolved in 1 mL of dry dimethylformamide (DMF) and then TEA(1.7 µL, 26 µmol, 4 equiv.) was added. After stirring for 24 h, at RT, the solvent was evaporated in vacuum. The crude product was purified by column chromatography with 5% methanol in DCM to afford 2.8 mg of the Allyl-sulfone-Cy5 as blue solid in 54% yield. ^{1}H NMR (300 MHz, Methanol-d_4) δ 8.30 (t, *J* = 13.1 Hz, 1H), 7.98 (s, 1H), 7.88 (d, *J* = 7.7 Hz, 2H), 7.68 (t, *J* = 7.3 Hz, 1H), 7.41 (d, *J* = 8.1 Hz, 1H), 7.33 (dd, *J* = 8.7, 3.1 Hz, 1H), 6.74 – 6.58 (m, 1H), 6.37 – 6.26 (m, 1H), 4.28 – 4.15 (m, 2H), 4.10 (s, 1H), 3.65 (d, *J* = 7.8 Hz, 2H), 3.54 (t, *J* = 5.4 Hz, 1H), 3.36 (d, *J* = 4.9 Hz, 3H), 3.28 (d, *J* = 7.2 Hz, 2H), 3.25 – 3.13 (m, 1H), 3.00 (s, 1H), 2.86 (s, 1H), 2.67 (s, 1H), 2.44 (d, *J* = 6.8 Hz, 1H), 2.20 (q, *J* = 9.8, 8.6 Hz, 1H), 1.94 (s, 9H), 1.78 (d, *J* = 16.8 Hz, 2H), 1.74 (s, 4H), 1.67 (d, *J* = 7.3 Hz, 1H), 1.36 – 1.19 (m, 9H), 0.89 (d, *J* = 7.1 Hz, 1H). LC-MS (ESI): m/z = 792 $[M]^+$ (calcd. mass: 792.40, formula: $C_{47}H_{58}N_3O_6S^+$), HPLC purity 93 %.

2.3 Sequential modification of Interleukin-2 (IL-2)

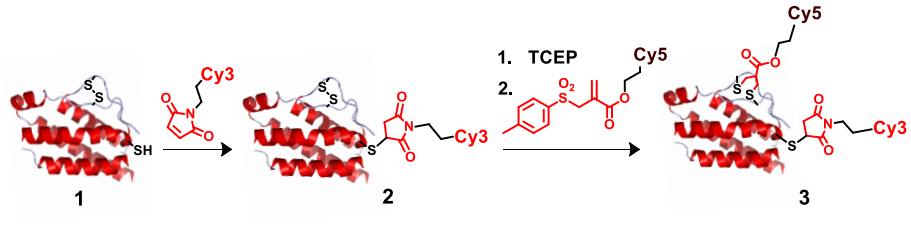


Figure S2 Sequential dual labeling of IL-2 (**1**) with Cy3-MI and Allyl-sulfone-Cy5 (**14**) respectively.

IL-2 **1** (100 µg, 0.0064 µmol) and Cy3-MI (10 µg, 0.013 µmol, 2 equiv.) were stirred in 50 mM phosphate buffer (PB) at pH 7.4 at 15 °C for overnight under argon atmosphere. The reaction mixture was purified by size exclusion chromatography using a Zeba Spin Desalting Columns (MWCO 7000) to obtain the conjugate IL2-Cy3 **2**. Optical spectra were recorded and the concentration of the construct was estimated *via* gel densitometry (please see the calculation below). For the second labeling step, IL2-Cy3 **2** (50 µg, 3.1 nmol) was first reacted with 2 equivalents of tris(2-carboxyethyl)phosphine (TCEP, 1.5 µg, 6.25 nmol) in 50 mM Phosphate buffer at pH 7.8 at 15 °C, stirring for 30 minutes under argon atmosphere. Then, Allyl-sulfone-Cy5 **14** (4.92 µg, 6.25 nmol, 2 equiv.) was added and the reaction was stirred at 15 °C for 24 h. The reaction mixture was purified by size exclusion chromatography using Zeba Spin desalting columns (MWCO 7000) to yield the conjugate IL2-Cy3-Cy5 **3**. The optical spectra were collected and the concentration of the construct was estimated *via* gel densitometry (please see the calculation below).

2.4 Determination of IL-2 labelling efficiency.

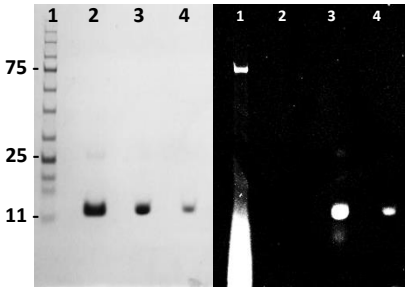


Figure S3 SDS-PAGE gel image with (left) or without (right) coomassie blue staining. lane 1: Protein Marker VI, AppliChem (from top to the bottom: 245, 180, 135, 100, 75, 63, 48, 35, 25, 20, 17, 11 kD), lane 2: IL-2 **1**, lane 3: IL2-Cy3 **2**, lane 4: IL2-Cy3-Cy5 **3**.

The absorbance of IL2-Cy3 **2** at 554 nm is 0.0102, which was fitted in the calibration curve (Figure S4A) and 10.76 µg/mL (14.62 µM) of Cy3-MI was calculated to be attached on the IL-2. In addition, the concentration of IL-2 is 300 µg/mL (19.5 µM) from the gel densitometry. Therefore, the labeling efficiency of Cy3-MI is 75% (=14.62/19.5*100%). Similarly, the absorbance of Cy5 in IL2-Cy3-Cy5 **3** at 650 nm is 0.0194, which was fitted in the calibration curve (Figure S4B) and 0.00163 mg/mL (1.43 µM) of Allyl-sulfone-Cy5 was calculated to be attached on the IL2-Cy3 **2**. From the gel densitometry, the concentration of IL-2 in IL2-Cy3-Cy5 **3** is 40 µg/mL (2.6 µM). Therefore, the labeling efficiency of Allyl-sulfone-Cy5 is 55% (=1.43/2.6*100%).

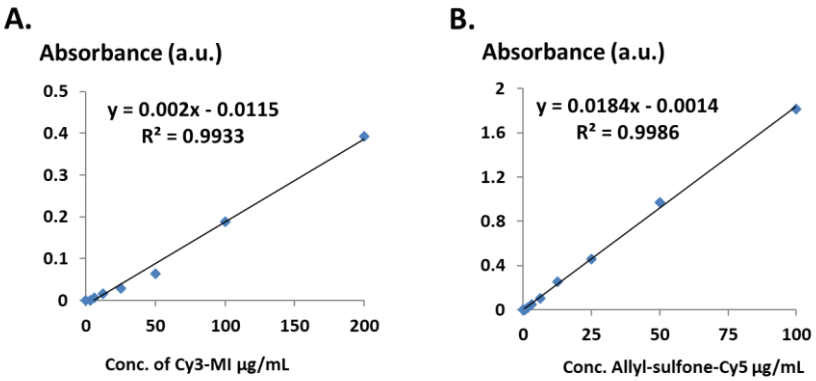


Figure S4 Calibration curve for the (A) Cy3-MI and (B) Allyl-sulfone-Cy5 based on absorbance at 550 nm and 650 nm respectively in phosphate buffer at pH 7.4.

2.5 Bio-activity study of IL-2 1, IL2-Cy3 2 and IL2-Cy3-Cy5 3

To determine the biological activity of the modified IL-2, the IL-2-dependent murine cell line: CTLL-2 was used. Proliferation experiments were performed in RPMI1640 medium, supplemented with 10% FCS, 2 mM L-glutamine and 50

µM β-mercaptoethanol (CTLL-2 medium). CTLL-2 cells were seeded at a density of $3x10^3$ in 200 µl in a 96 well plate and cultivated for 48 h. The proliferation of the cells was determined by adding [^{3}H] thymidine for 16 h. Subsequently, incorporated ^{3}H thymidine was measured by a ß-scintillation counter in counts per minute. The IL2-Cy3 **2** and IL2-Cy3-Cy5 **3** were titrated and compared to native IL-2 **1** in the following concentrations 200, 100, 10, 1, 0.5, 0.25, 0.1, 0.01 U/mL. For culture maintenance the CTLL-2 cells were incubated in CTLL-2 medium supplemented with 50 U/mL of IL-2.

2.5.1 Bio-activity study of disulfide reduced Interleukin-2.

To study the effect of disulfide bond of IL-2 in its bio-activity we synthesized a derivatives in which, the disulfide bond of IL-2 was reduced and labelled by N-(2-aminoethyl)maleimide. The IL-2 **1** (100 µg, 6.2 nmol) was first reacted with 10 equivalents of (tris(2-carboxyethyl)phosphine) (62 nmol) in 50 mM Phosphate buffer at pH 7.4 at 15 °C, stirring for 30 minutes under argon atmosphere. Then, the reaction mixture was divided into two portions and one portion containing was used for subsequent reaction. To 50 ug of reduced IL-2 (3.2 nmol) was added excess N-(2-aminoethyl)maleimide (20 equiv., 62 nmol) and the reaction was stirred at 15 °C for 24 h. Thereafter, the protein was purified by size exclusion chromatography using Zeba Spin desalting columns (MWCO 7000) to yield **2'**. IL-2 received from Cell Science was used as a control for evaluation of the biological activity. Prior to evaluation of the bioactivity, the concentration of IL-2 in each of the three samples was normalized by absorbance measurement at 280 nm. The bio-activity was measured by CTLL-2 cells assay as described above.

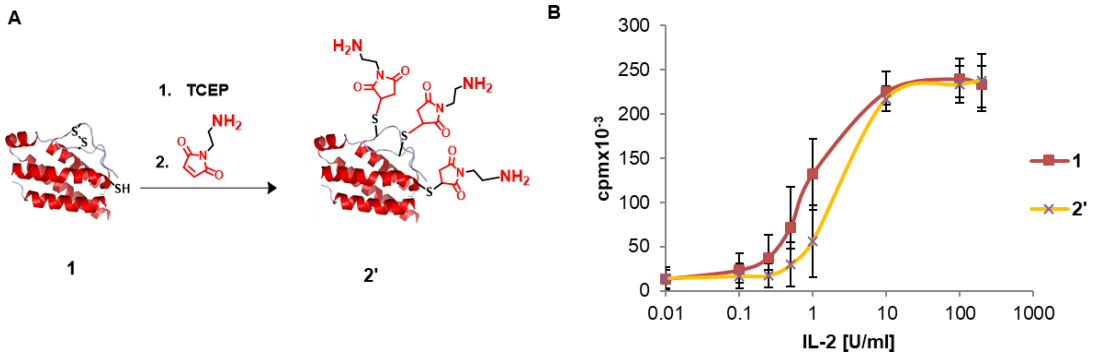


Figure S5 Cell proliferation assay for disulfide reduced IL-2. (A) Reduction of disulfide bond of IL-2 with TCEP followed by labelling with excess N-(2-aminoethyl)maleimide to yield **2'**. (B) T-Cell proliferation assay for IL-2 **1** (red) and **2'** (yellow). The graph shows no significant reduction of cell proliferation efficiency for **2'** in comparison to native IL-2 **1**.

3. Sequential modification of model peptide 4

3.1 Structure of 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM)

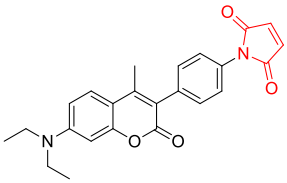


Figure S6 Chemical structure of 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM)

3.2 Sequential modification of Dis-Tag

3.2.1 Synthesis of mono-CPM labelled model peptide 5

The model peptide **4** (CATSICGGGC, 11 mg, 0.01266 mmol) and 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM, 7.6 mg, 0.01899 mmol) were mixed in 14 mL of 55 % Acetonitrile (ACN) in 100 mM Phosphate buffer (pH 7.4). The resulting mixture was stirred at RT for 3 h. The product was purified by preparative HPLC using an Atlantis Prep OBD T3 column (19 x 100 mm, 5 µm) with the mobile phase starting from 100 % solvent A (0.1 % TFA in water) and 0 % solvent B (0.1 % Trifluoroacetic acid in ACN) (0-5 min) to 40 % solvent B in 5 min, 40% B for 9 min and finally reaching 100% solvent B in 35 min with a flow rate of 10 mL/min. The absorbance was monitored at 365 nm and 254 nm. **5** has been isolated with an overall yield of 61%. Characterization via HR MALDI-TOF MS using Sinapinic acid as matrix afforded the signal at m/z 1271.45193 $[M+H]^+$, 1293.43544 $[M+Na]^+$, 1309.40974 $[M+K]^+$, corresponding to the calculated molecular weight 1270.45 and additional cations.

3.2.2 Synthesis of tri-CPM labelled model peptide 6

The mono-labeled model peptide **5** (0.5 mg, 0.39 µmol) was dissolved in 1 mL of 100 mM Phosphate buffer. TCEP (0.2 mg, 0.7 µmol) was added and the solution was incubated for 30 min. CPM (0.6 mg, 1.49 µmol) was dissolved in 1 mL ACN and added slowly to the above mixture. The reaction solution was stirred for 3 h at RT. The product was purified by preparative HPLC using an Atlantis Prep OBD T3 column (19 x 100 mm, 5 µm) with the mobile phase starting from 100% solvent A (0.1 % TFA in water) and 0% solvent B (0.1 % TFA in ACN) (0-5 min) to 43% solvent B in 5 min, raising to 65% solvent B in 10 min and finally reaching 100% solvent B in 35 min with a flow rate of 10 mL/min. The absorbance was monitored at 365 nm and 254 nm. **6** has been isolated in 72% yield and characterized via HR MALDI-TOF MS using α-cyano-4-hydroxycinnamic acid (CHCA) as matrix afforded the signal at m/z 2100.2224 $[M+Na]^+$, 2116.2168 $[M+K]^+$ corresponding to the calculated molecular weight 2076.78.

4. Sequential modification of Dis-tag-EYFP

4.1 Expression of EYFP

4.1.1 Plasmid preparation

The method add-on-PCR was employed to attach the tag sequence at the N-terminus of a model protein EYFP. Plasmid PURE-YFP was used as a template and the primers (ordered from Eurofins MWG Operon, Ebersberg, Germany) comprised 25 bases homologous to the template and an additional sequence coding for the add-on tag. The sequence of the forward primer was AAA AAAAAACAT ATG TGC GCG ACC TCT ATT TGC GTG AGC AAG GCG GAG GAG CTG TTTC and the sequence of the reverse primer TTT TTTTTT AAG CTT CTT GTA GAG CTC GTC CAT GCC GAGA. The PCR product as well as the cloning plasmid pET22b(+) (Novagen, Madison, USA) were treated with the restriction enzymes NdeI and HindIII (both from Fermentas, St. Leon-Roth, Germany). After ligation the novel plasmid pET22b(+)-Y3 was transformed into E.coli DH5-alpha bacterial cells and extracted from bacterial culture using the GenElute™ Plasmid MiniPrep Kit from Sigma-Aldrich. Successful addition of the tag sequence and ligation into the expression vector were confirmed by sequencing which was performed by Eurofins MWG Operon, Ebersberg, Germany. The PCR primers were chosen so that the stop codon of the template was excluded from amplification. Therefore the His-tag coding region situated after the HindIII restriction site on the cloning plasmid was included in transcription and the recombinant protein carried a C-terminal His-tag to allow for easy purification.

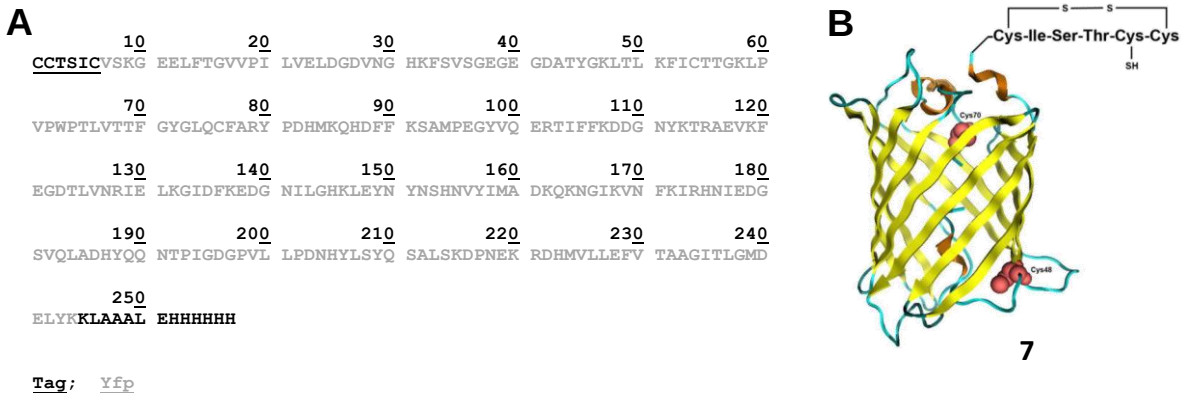


Figure S7 (A) Amino acid sequence of the Dis-tag-EYFP **7** containing a His-tag. (B) Visualization of Dis-tag-EYFP.

4.1.2 Protein purification

The pET22b(+)-Y3 plasmid was transferred into E.*coli* BL21 (DE3) bacteria *via* electroporation. 200 mL LB medium (Alfa Aesar, 20 g/L) supplemented with 100 µg/mL ampicillin (Sigma-Aldrich) was inoculated and the expression culture grown at 37 °C 170 rpm until OD_{578} reached ~0.6. Protein production was initiated by the addition of 0.01 mM IPTG (AppliChem,

final concentration) and was allowed to proceed at 30 °C, 150 rpm for 16 h. Subsequently the bacterial cells were collected *via* centrifugation and re-suspended into 2x Phosphate-buffered saline (PBS, 274 mM NaCl, 5.4 mM KCl, 20 mM Na_2HPO_4, 4 mM KH_2PO_4) pH 7.4 supplemented with 1x protease inhibitor cocktail (Calbiochem, Protease Inhibitor Cocktail Set V, EDTA free). The lyophilized cocktail was reconstituted with dH_2O to obtain a 100x stock solution.) and 6 µl (0.79 µl/10 mL lysis buffer) benzonase (Merck, >90% purity) normalizing to OD_{578} 7.

The bacterial cells were disrupted *via* ultra-sonication. The lysate was centrifuged to separate the soluble portion which was filtered through a 0.2 µm syringe filter prior loading onto a Ni^{2+} charged and 2x PBS equilibrated 1 mLHiTrap™ FF IMAC column from GE Healthcare. The column was washed with ten column volumes (cv) of 2x PBS, 10 mM imidazole and the protein eluted with three cv of 2x PBS, 500 mM imidazole. The protein was transferred into fresh 2x PBS pH 7.4 using a Vivaspin 6 (Sartorius Group, MWCO 10k) centrifugal concentrator.

4.2 Structure of Alexa647-Maleimide

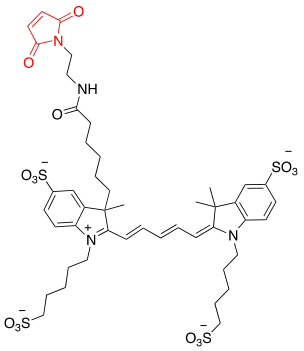


Figure S8 Chemical Structure of Alexa 647 C2 Maleimide (Alexa647-MI). Chemical Structure of Atto550-maleimide (Atto550-MI) is not available.

4.3 Sequential dual-labelling of Dis-tag-EYFP

The Dis-tag-EYFP **7** (1 mg, 0.034 µmol) and 5 equivalents of Atto550-MI (0.14 mg, 0.172 µmol) were stirred in PBS buffer (2x, pH 7.4) for 3 h under argon atmosphere. The reaction mixture was purified by size exclusion chromatography using a Sephadex G-25 matrix to obtain the Atto-EYFP **8**. Optical spectra were recorded and the concentration of the construct was estimated via gel densitometry. In order to cap all the unreacted free thiols on Dis-tag-EYFP, Atto-EYFP **8** was further reacted with *N*-(2-Aminoethyl)maleimide trifluoroacetate salt (0.9 mg, 3.4 µmol) in PBS buffer (2x, pH 7.4) for 3 h. The reaction mixture was purified by size exclusion chromatography using Sephadex G-25 matrix to isolate the amine capped Atto-EYFP. For the second labeling step, Atto-EYFP **8** or the amine capped Atto-EYFP (0.725 mg, 0.025 µmol) was first reacted with an excess of TCEP (19.2 µg, 0.068 µmol) by stirring for 1 h under argon atmosphere. Then, 3 equivalents of

Alexa647-MI (100 µg, 0,077 µmol) were added and the reaction allowed proceeding for 3 h. The reaction mixture was purified by size exclusion chromatography using Sephadex G-25 matrix to yield the Atto-Alexa-EYFP **9a**. The optical spectra were recorded and the concentration of the construct was estimated via gel densitometry.

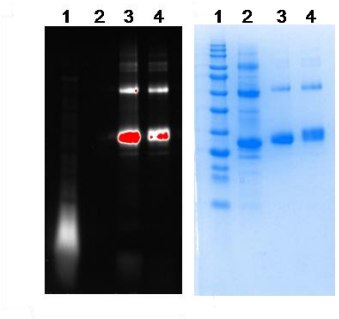


Figure S9 SDS-PAGE gel image with (right) or without (left) coomassie blue staining. lane 1: Protein Marker VI, AppliChem (From top to the bottom: 245, 180, 135, 100, 75, 63, 48, 35, 25, 20, 17, 11kD), lane 2: Dis-tag-EYFP **7**, lane 3: Atto-EYFP **8**, lane 4: Atto-Alexa-EYFP **9a**.

4.3.1 Determination of labelling efficiency

The absorbance of Atto-EYFP **8a** at 554 nm is 0.3187, which was fitted in the calibration curve (Figure S10A) and 0.005724 mg/mL (7.02 µM) of Atto550-MI was calculated to be attached on the EYFP. In addition, the concentration of Atto-EYFP **8a** is 0.338 mg/mL (11.66 µM) from the gel densitometry. Therefore, the labeling efficiency of Atto550-MI is 60% (=7.015/11.66*100%). Similarly, the absorbance of Atto-Alexa-EYFP **9a** at 652 nm is 0.104, which was fitted in the calibration curve (Figure S10B) and 0.00163 mg/mL (1.25 µM) of Alexa647-MI was calculated to be attached on the EYFP. From the gel densitometry, the concentration of Atto-Alexa-EYFP **9a** is 0.0875 mg/mL (3.02 µM). Therefore, the labeling efficiency of Alexa647 is 42% (=1.254/3.017*100%).

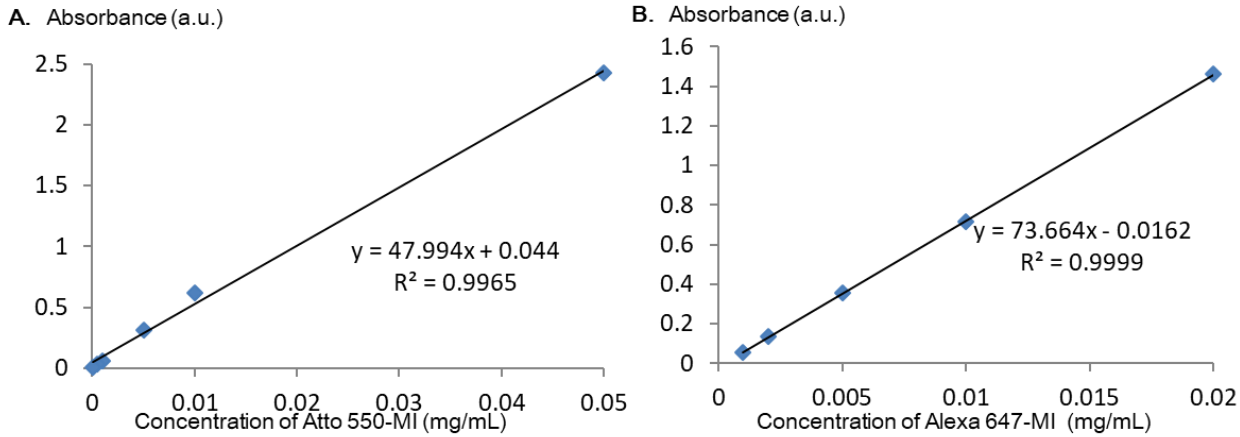


Figure S10 The calibration curve of the concentration of (A) Atto-550 and (B) Alexa-647 based on the absorbance at 554 nm and 652 nm respectively measured in phosphate buffer at pH 7.4.

Dual color labeling on Dis-tag-EYFP for single-molecule FRET.

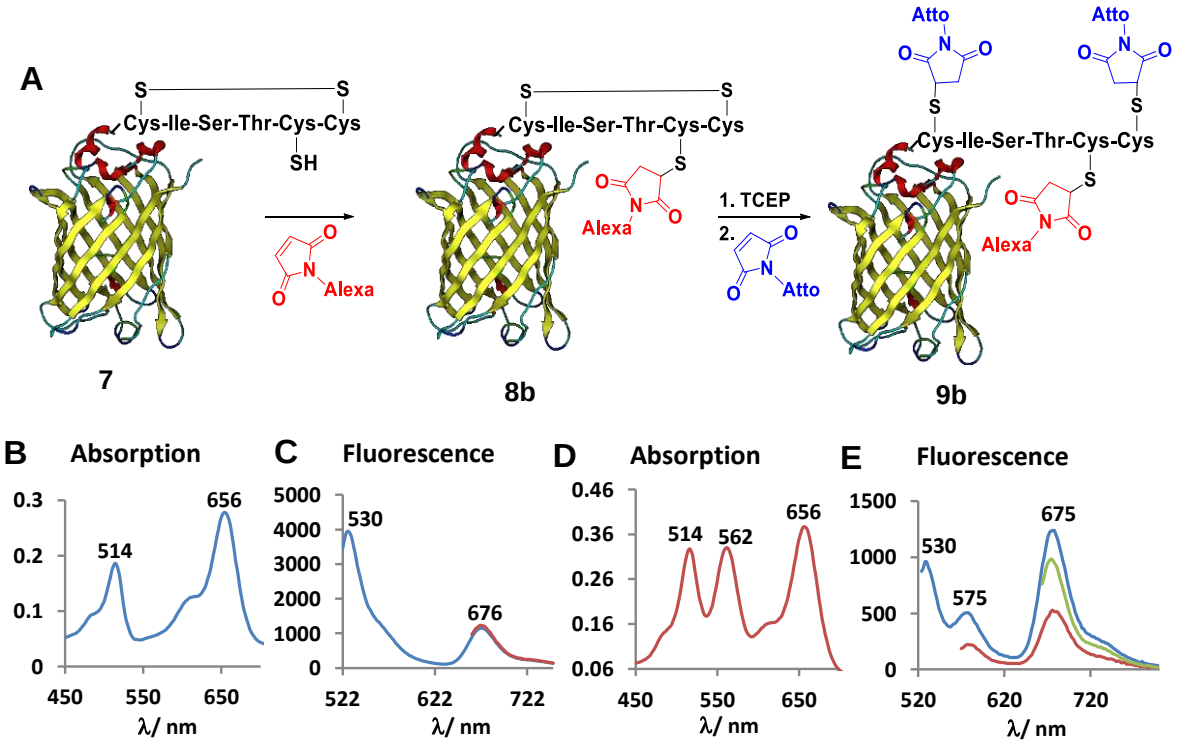


Figure S11 (A) Synthesis of Alexa-Atto-EYFP **9b**. (B) The absorption spectra of Alexa-EYFP **8b** (blue). (C) The emission spectra of the Alexa-EYFP **8b** excited at 516 nm (blue), 562 nm (red); (D) The absorption spectra of the Alexa-Atto-EYFP **9b** (red), (E) The emission spectrum of the Alexa-Atto-EYFP **9b** excited at 516 nm (blue), 562 nm (red) and 652 nm (green).

In order to measure the single-molecule FRET, the Dis-tag-EYFP was labeled with Alexa647-MI and Atto550-MI sequentially (Figure S11). The Alexa647 was labeled first instead of Atto550, since the Dis-tag-EYFP (0.2 mg, 6.9 nmol) was reacted with 3 eq. of Alexa647-MI (0.027 mg, 20.7 nmol) in 50 mM phosphate buffer (pH 7.4) for 3 h under argon atmosphere. The reaction mixture was purified by size exclusion chromatography using Sephadex G-25 matrix. Thereafter, Alexa-EYFP **8b** (0.1 mg, 3.4 nmol) was sequentially reacted with TCEP (9.9 µg, 34.5 nmol) for 1 h and Atto550-MI (8.4 µg, 10.3 nmol) for 3 h. The resulting Alexa-Atto-EYFP **9b** was purified by size exclusion chromatography using Sephadex G-25 matrix. The labeling efficiency for Alexa647-MI and Atto550-MI were 47% and 94 % respectively based on the absorbance spectra (Figure S11). The detailed calculation was shown in the following part.

4.3.2 Calculation of the labeling efficiency.

Alexa647-MI labeling efficiency.

A_{514}(8a) = 0.1861, A_{652}(8a) = 0.2758, A_{514}(Alexa) = 0.8%A_{max}(Alexa), ε_{EYFP} = 83400 $M^{-1}cm^{-1}$, ε_{Alexa} = 265000 $M^{-1}cm^{-1}$.

According to Beer–Lambert law, the concentration of Alexa dye on the EYFP is $A_{652}(8a)/\varepsilon_{Alexa}$ = 0.2758/265000 = $1.04x10^{-6}$, while the concentration of EYFP is $[A_{514}(8a)- A_{514}(Alexa)]/\varepsilon_{EYFP}$ = (0.1861-0.008*0.2758)/83400 = $2.205x10^{-6}$. Therefore, the labeling efficiency of Alexa647-MI is $1.04x10^{-6}/2.205x10^{-6}$= 47%.

Specification: $A_{514}(8a)$ and $A_{652}(8a)$ are the absorbance of Alexa-EYFP 7 at 514 nm and 652 nm respectively.

$A_{514}(Alexa)$ = 0.8% $A_{max}(Alexa)$, means the absorbance of Alexa647-MI at 514 nm is 0.8% of its maximum absorbance according to its absorbance spectrum.

ε_{EYFP} and ε_{Alexa} are the extinction coefficient of EYFP and Alexa647-MI respectively.

Atto550-MI labeling efficiency.

$A_{516}(9b)$ = 0.3224, $A_{562}(9b)$ = 0.33, $A_{656}(9b)$ = 0.363, $A_{562}(Alexa)$ = 6.57%$A_{max}(Alexa)$, $A_{516}(Alexa)$ = 0.8%$A_{max}(Alexa)$, $A_{516}(Atto)$ = 28.5%$A_{max}(Atto)$, ε_{EYFP} = 83400 $M^{-1}cm^{-1}$, ε_{Atto} = 120000 $M^{-1}cm^{-1}$.

According to Beer–Lambert law, the concentration of Atto dye on the EYFP is $[A_{562}(9b) – A_{562}(Alexa)]/\varepsilon_{Atto}$ = (0.33-0.0657*0.363) / 120000 = $2.556x10^{-6}$, while the concentration of EYFP is $[A_{516}(8)-A_{516}(Atto)-A_{516}(Alexa)]/\varepsilon_{EYFP}$ = (0.3224-0.285*0.33-0.008*0.363)/83400 =$2.703x10^{-6}$. Therefore, the labeling efficiency of Atto550-MI is $2.556x10^{-6}/2.703x10^{-6}$= 94%.

Specification: $A_{516}(8b)$, $A_{562}(8b)$ and $A_{652}(9b)$ are the absorbance of Alexa-Atto-EYF P 8 at 516 nm, 562 nm and 656 nm respectively.

$A_{562}(Alexa)$ = 6.57%$A_{max}(Alexa)$, means the absorbance of Alexa647-MI at 562 nm is 6.57% of its maximum absorbance according to its absorbance spectrum.

$A_{516}(Alexa)$ = 0.8% $A_{max}(Alexa)$, means the absorbance of Alexa647-MI at 516 nm is 0.8% of its maximum absorbance according to its absorbance spectrum.

$A_{516}(Atto)$ = 28.5%$A_{max}(Atto)$, means the absorbance of Atto550-MI at 516 nm is 28.5% of its maximum absorbance according to its absorbance spectrum.

ε_{EYFP} and ε_{Atto} are the extinction coefficient of EYFP and Atto550-MI respectively.

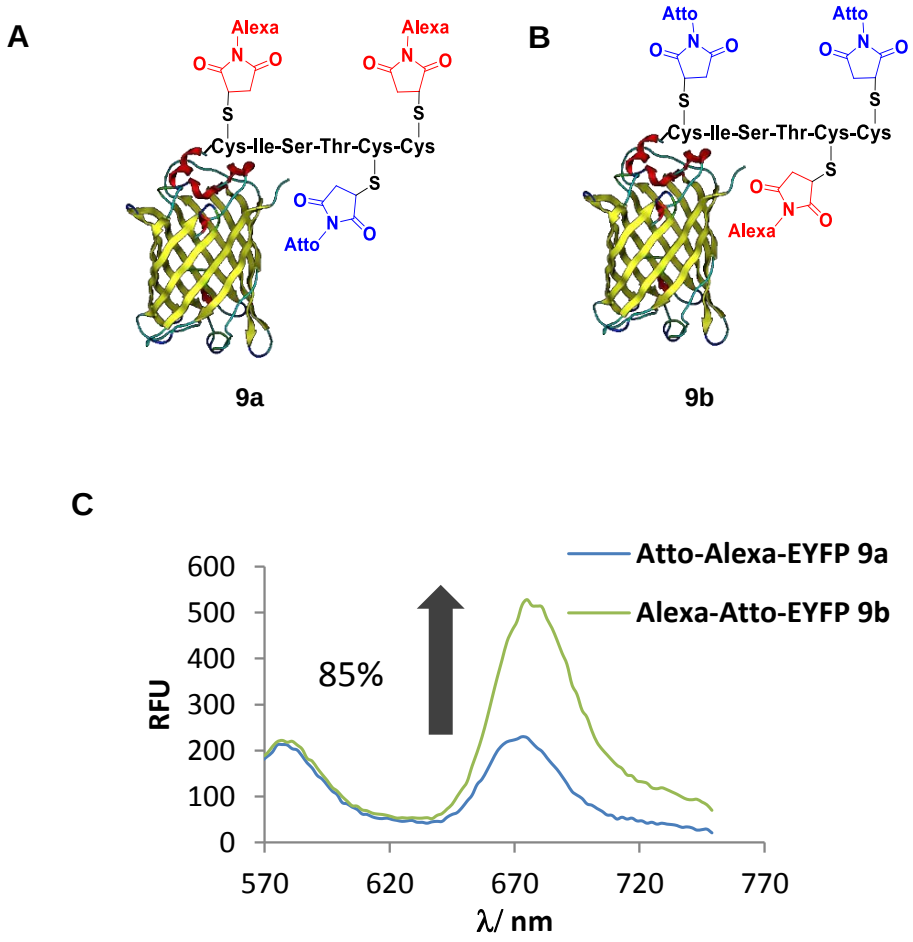


Figure S12 (A) Cartoon representation of Atto-Alexa-EYFP **9a** and (B) Alexa-Atto-EYFP **9b**; (C) Comparison of fluorescent emission from **9a** and **9b** when excited with 562 nm.

4.4 Single-molecule FRET measurement Atto-Alexa-EYFP 9b

Experimental Setup and Data Analysis

Single-molecule (sm)FRET experiments were performed on a homebuilt prism-based total internal reflection fluorescence microscope (TIRFM) as described previously.[2] Briefly, fluorescence was collected through a water-immersion objective and split into a donor and acceptor emission channel using a dichroic mirror. The fluorescence signal of the donor and acceptor emission was each directed to an EMCCD camera (iXon DU-897E-CS0-BV, Andor), images of both cameras where overlaid using fluorescent beads which could be detected simultaneously in both channels for calibration. The donor and acceptor molecules were excited using the alternating laser excitation scheme (ALEX)[3] with an excitation wavelength of 532nm. For direct excitation of the acceptor molecules a diode laser (643 nm, Toptica) was used.

Microfluidic chambers were constructed similar to a procedure published previously.[4] Fluorescent samples were deposited onto a PEG/Biotin-coated chamber surface. All measurements were recorded with an integration time of 100 ms per frame and a total duration of 60 s.

The acquired data was analyzed using custom-written MATLAB (The MathWorks) software as described elsewhere.[5] For the calculation of the FRET efficiency of each individual FRET pair, the following formula was used:

$$E = \frac{I_A - \beta\, I_D}{I_A + \gamma\, I_D}, \text{ where } \quad \gamma = \frac{I_A - I_A'}{I_D' - I_D} \quad \text{and} \quad \beta = \frac{I_A'}{I_D'}.$$

I_A and I_D are the background corrected intensities from acceptor and donor channels and I and I' are the intensities before and after acceptor photobleaching, respectively. The correction factor β accounts for the leakage of the donor emission into the acceptor channel whereas the γ factor corrects for different quantum yields of donor and acceptor dyes and different detection efficiencies of the two channels. Including a corection factor for the direct excitation of the acceptor by the 532nm did not change the result and was therefore omitted for simplicity. All parameters for the calculation of the FRET value were estimated as described previously.[6] Since the observed smFRET efficiencies were larger than 90%, the histograms were fitted with a Beta distribution.[7] The mean FRET efficiency as well as its standard error were determined from the fit. Molecules with γ–values below 0.25 and above 1.5 were excluded from the analysis.

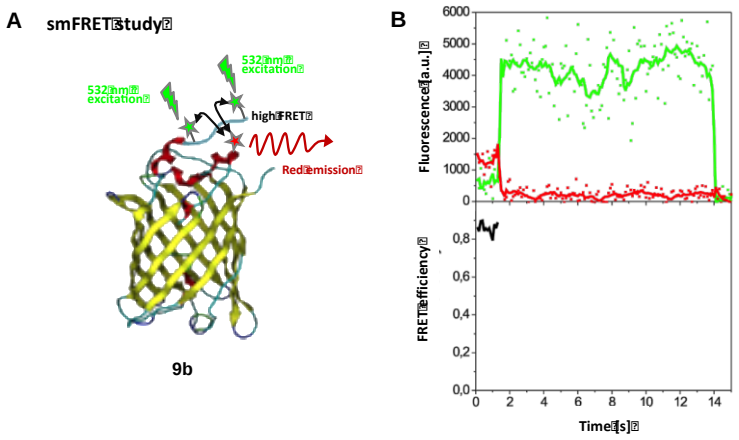
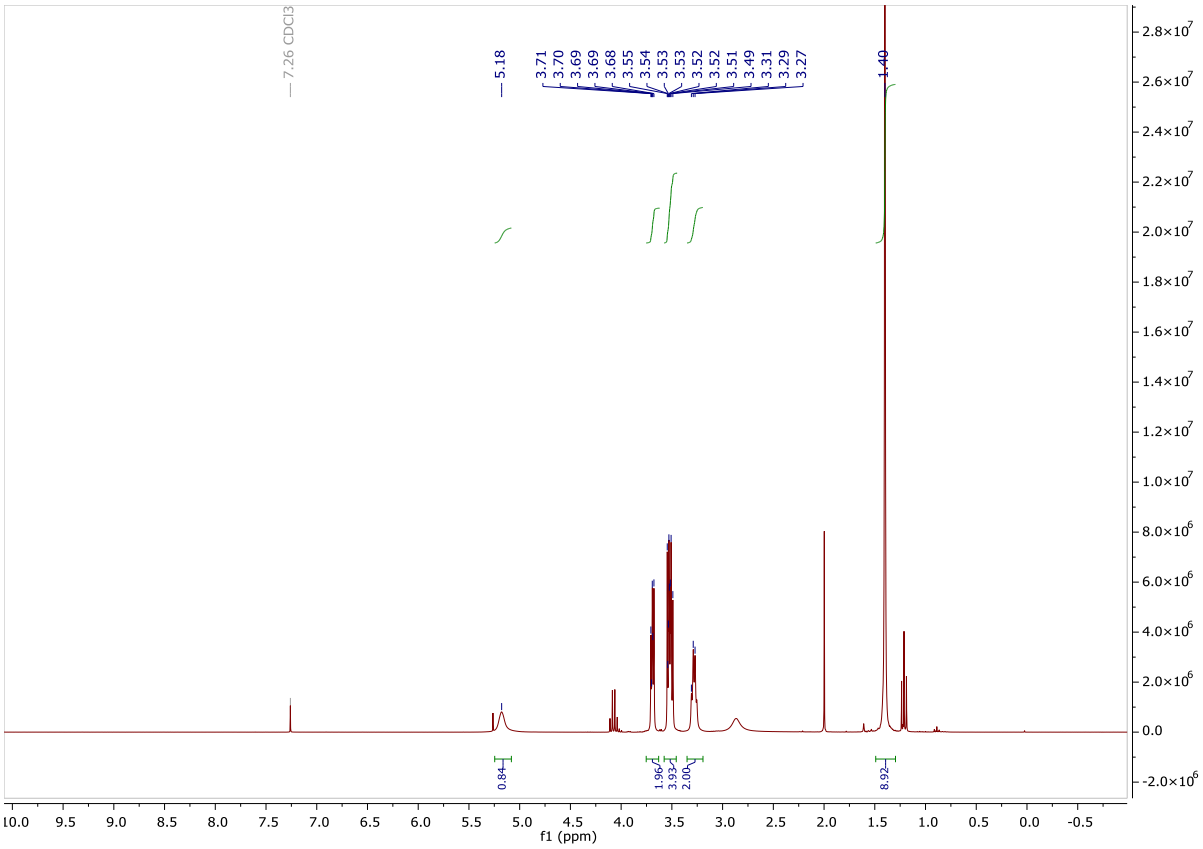


Figure S13 Single molecule FRET experiments (A) Cartoon representation of a sm-FRET event on the chromophore-labelled EYFP-Dis-tag protein **9b**. The positions of ATTO550- (green) and Alexa647-dyes (red) are indicated

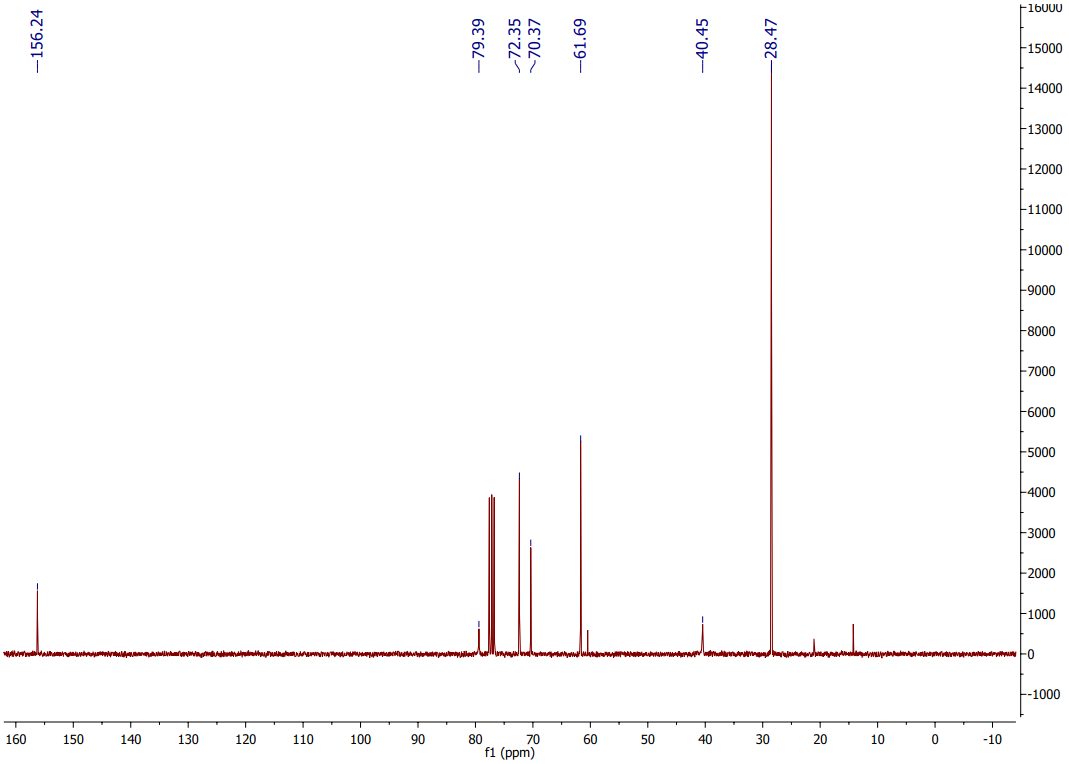
by asterisks. (B) Single-molecule FRET data. Upper panel: Exemplary intensity time trace of a sm-FRET event. Fluorescence intensities of the donor molecule, Atto550 (green), and the acceptor molecule, Alexa647 (red), shown are the raw data (dots) as well as a 10 point sliding average (solid lines). Lower panel: Computed FRET-efficiency.

NMR spectra.

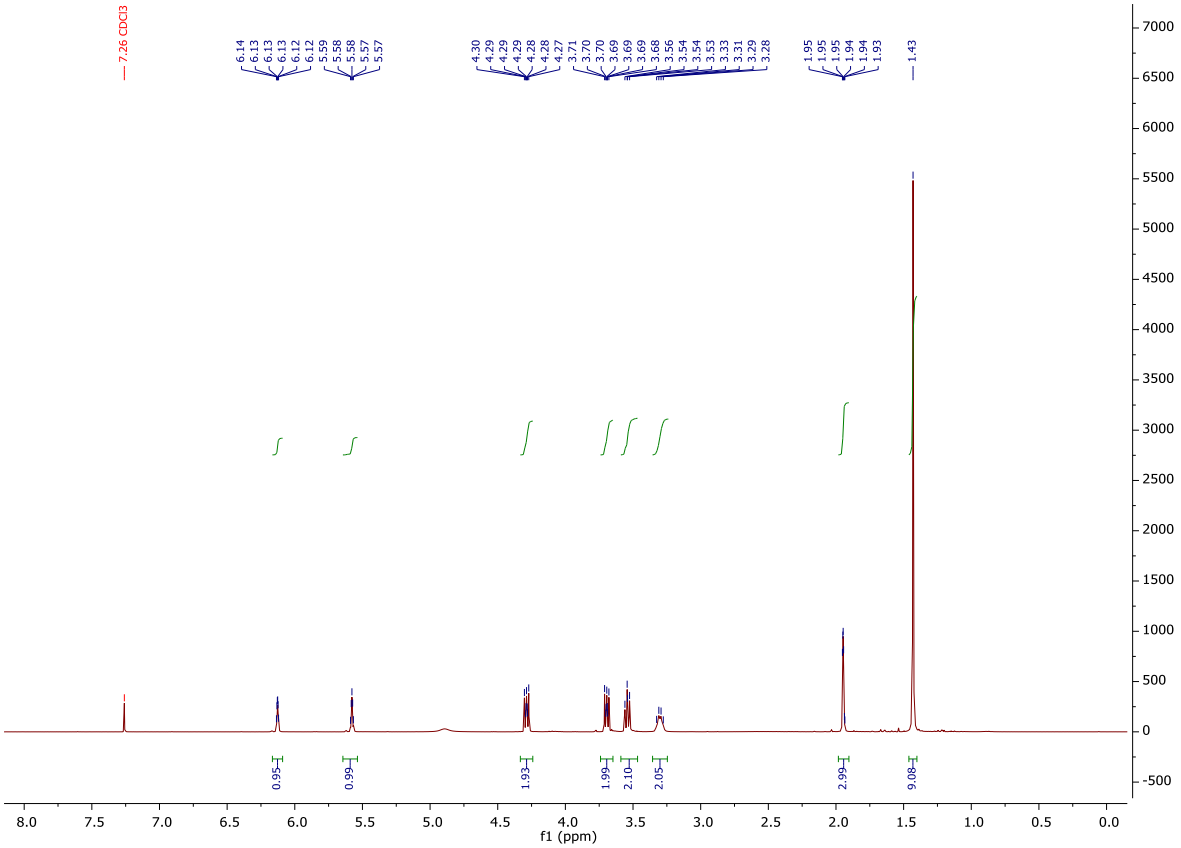
^{1}H NMR of tert-butyl (2-(2-hydroxyethoxy)ethyl)carbamate **10**.



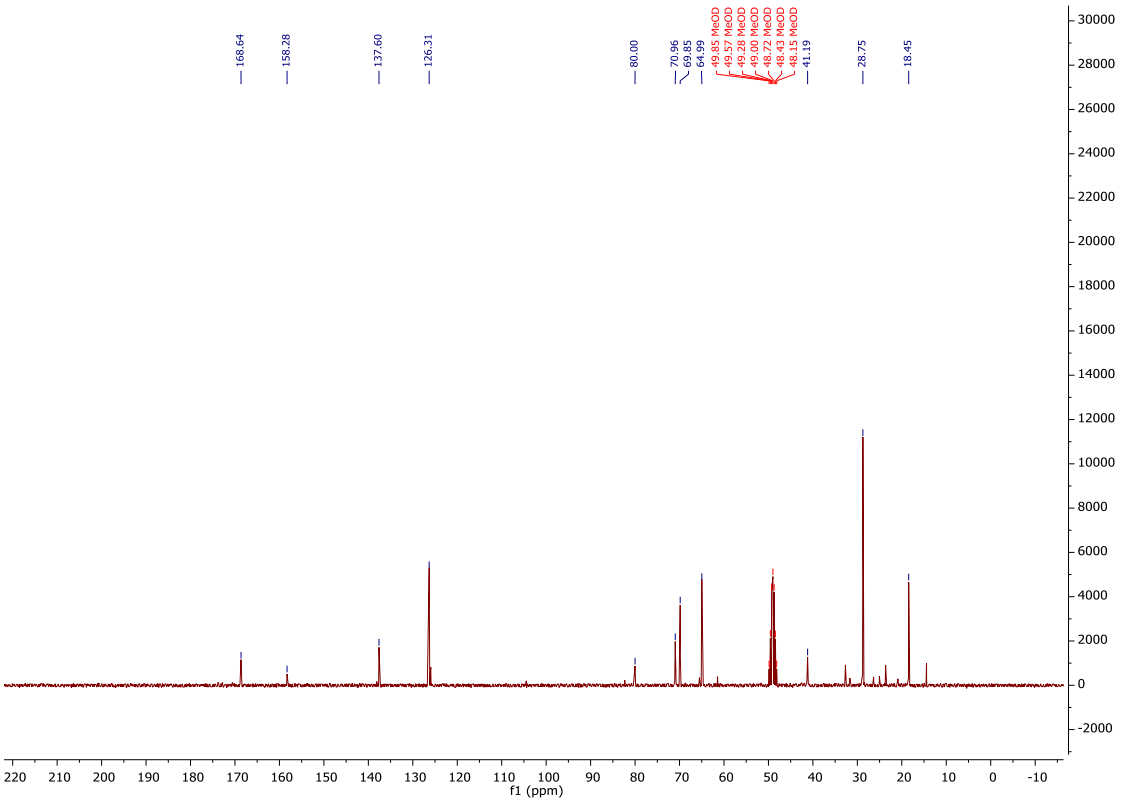
^{13}C NMR of tert-butyl (2-(2-hydroxyethoxy)ethyl)carbamate **10**.



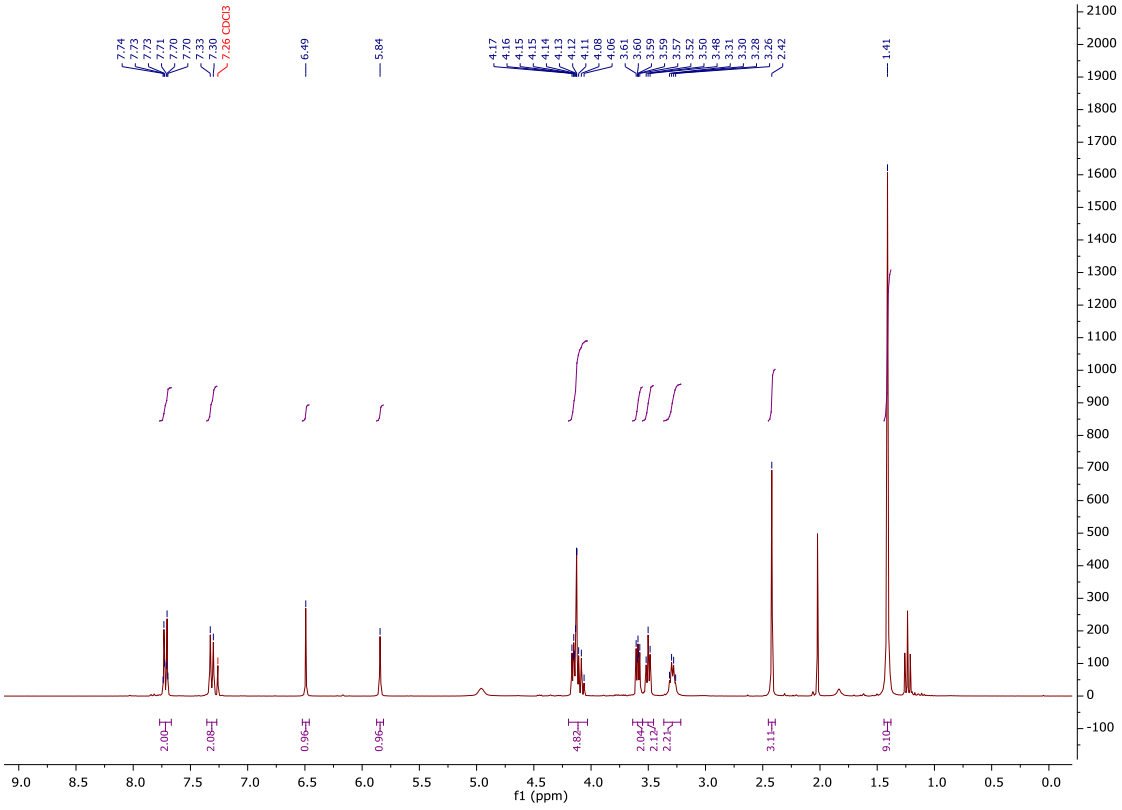
^{1}H NMR 2-(2-((tert-butoxycarbonyl)amino)ethoxy)ethyl methacrylate **11**.



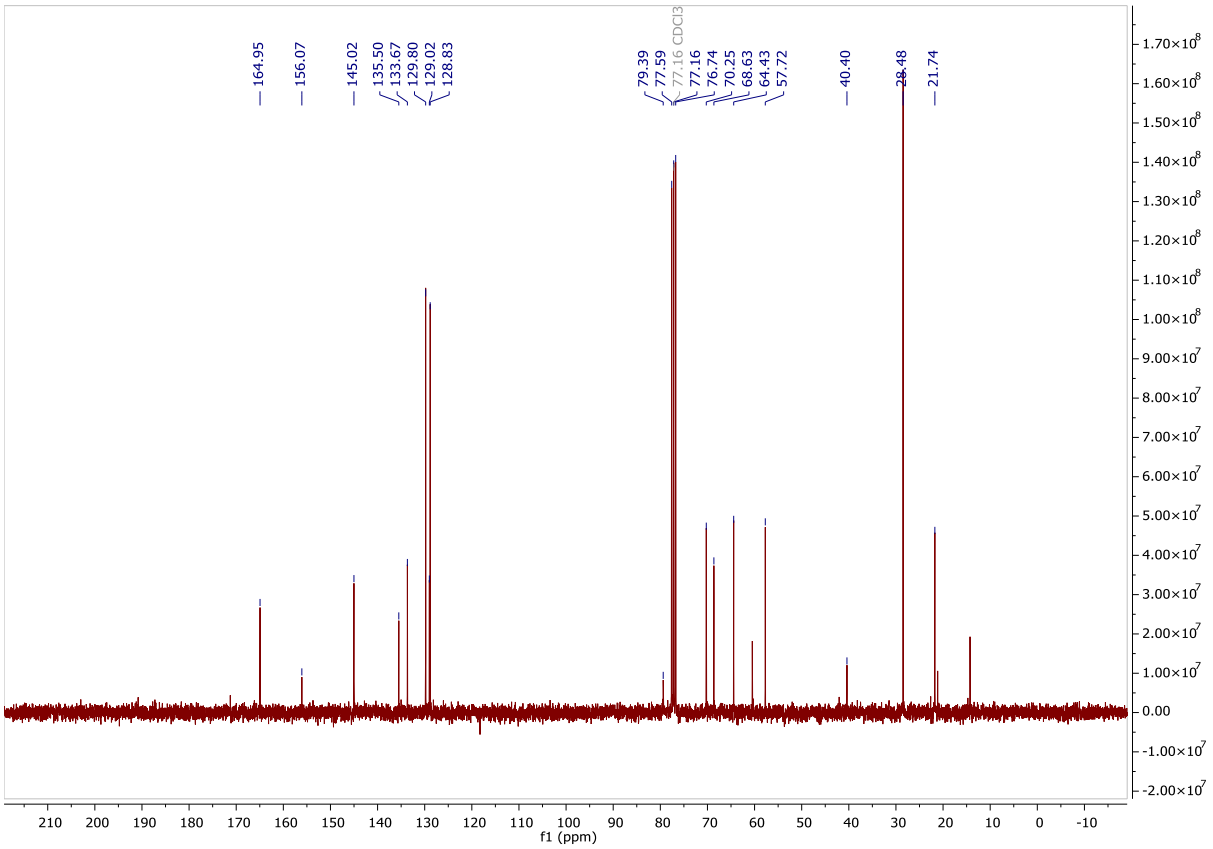
^{13}C NMR 2-(2-((tert-butoxycarbonyl)amino)ethoxy)ethyl methacrylate **11**.



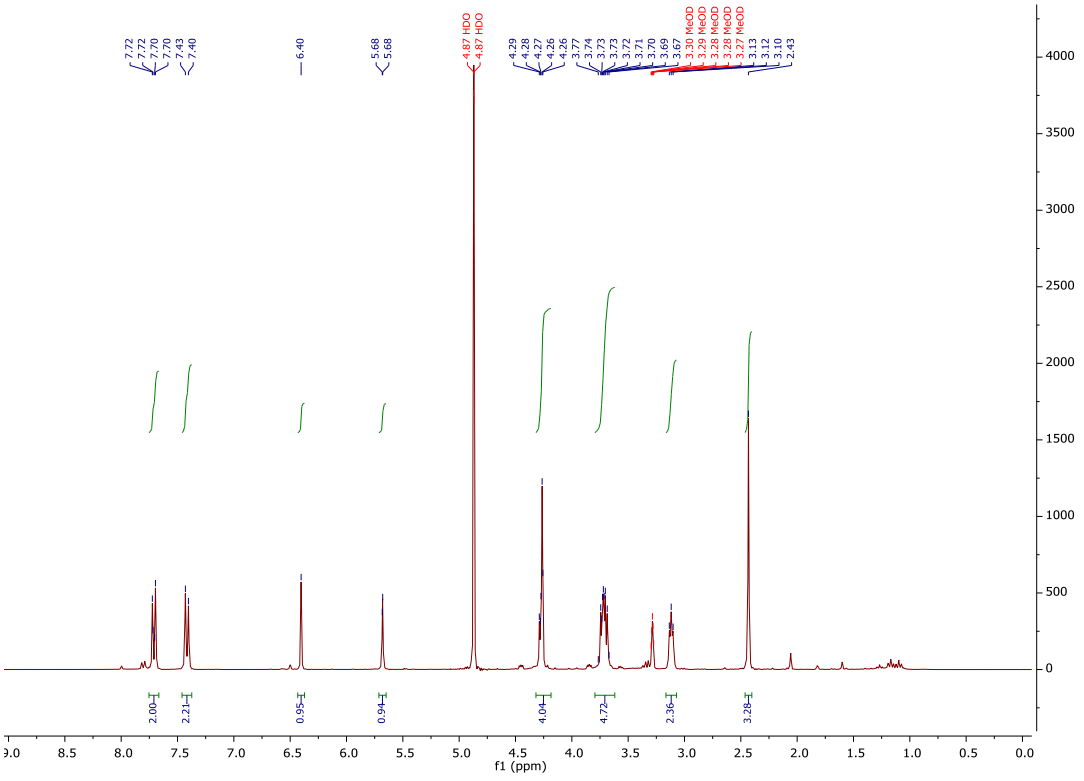
^{1}H NMR 2-(2-((tert-butoxycarbonyl)amino)ethoxy)ethyl 2-(tosylmethyl)acrylate **12**.



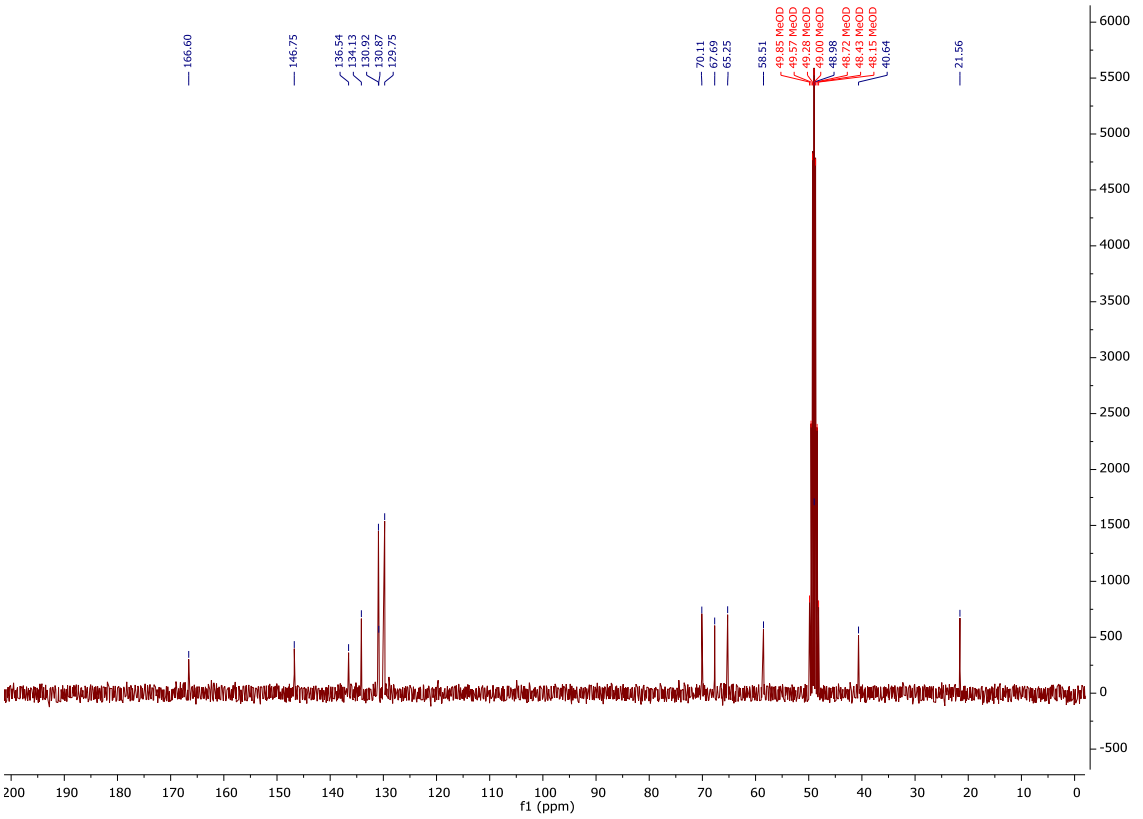
^{13}C NMR 2-(2-((tert-butoxycarbonyl)amino)ethoxy)ethyl 2-(tosylmethyl)acrylate **12**.



^{1}H NMR of 2-(2-aminoethoxy)ethyl 2-(tosylmethyl)acrylate **13**

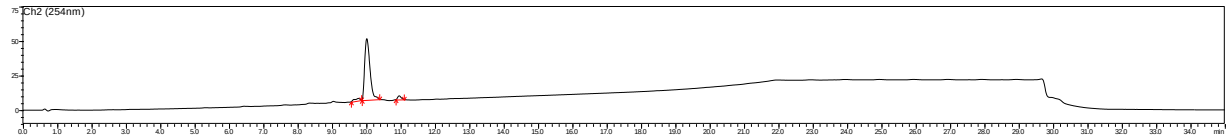


^{13}C NMR of 2-(2-aminoethoxy)ethyl 2-(tosylmethyl)acrylate **13**

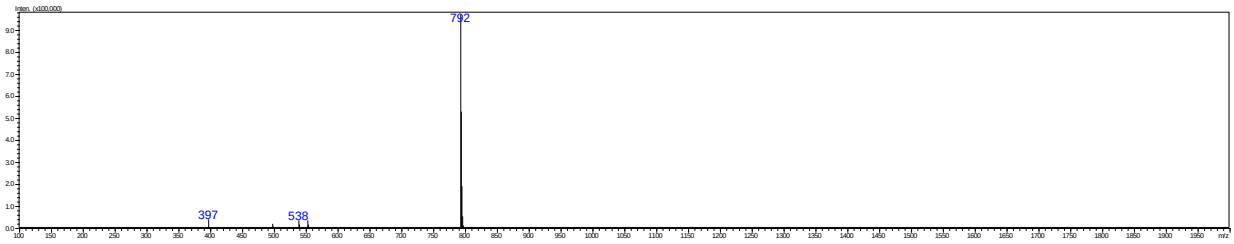


LC-MS **14**

A

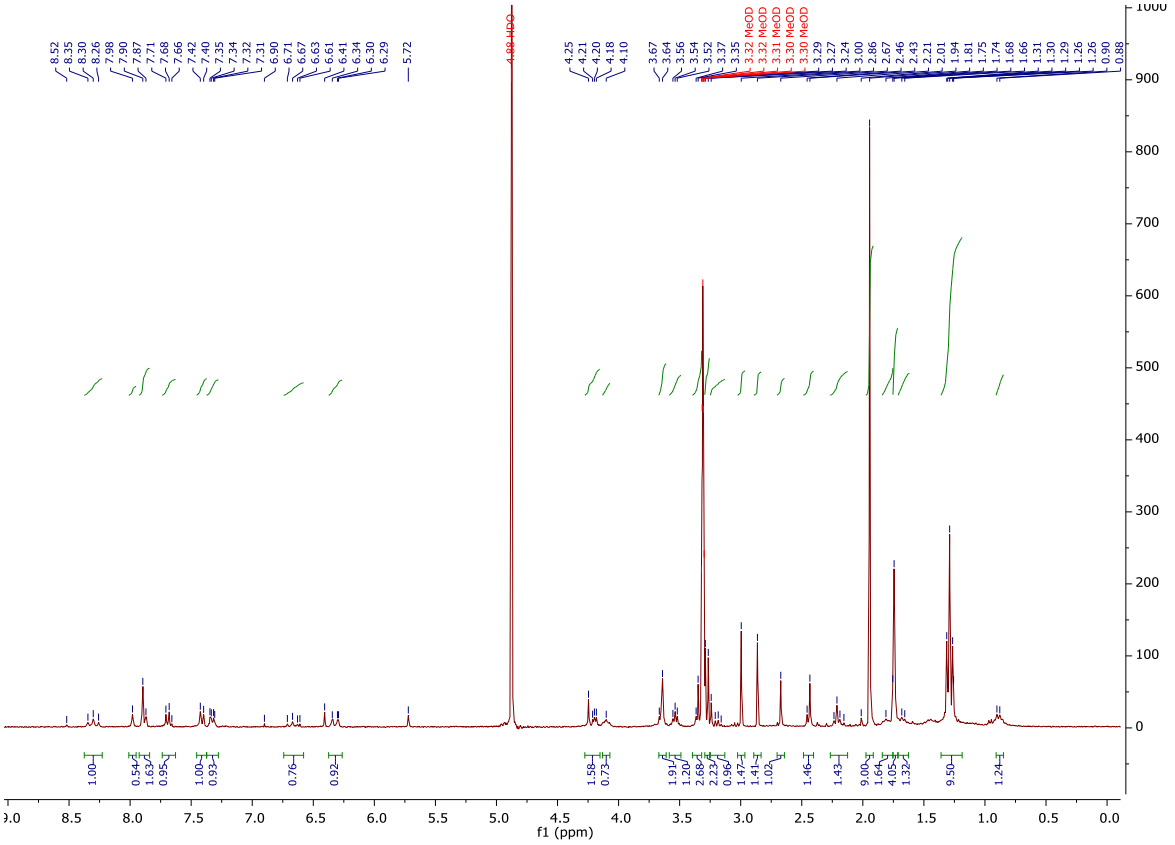


B



(A) Liquid chromatogram of **14** at λ_{max} = 254 nm showing integration traces and (**B**) Mass spectrum of **14** in positive ion mode with peak at m/z = 792 correspnding to [M^+].

^{1}H NMR **14**



REFERENCES

(1) Wang, T., Riegger, A., Lamla, M., Wiese, S., Oeckl, P., Otto, M., Wu, Y. Z., Fischer, S., Barth, H., Kuan, S. L., and Weil, T. (2016) Water-soluble allyl sulfones for dual site-specific labelling of proteins and cyclic peptides. *Chem. Sci. 7*, 3234–3239.

(2) Nagy, J., Grohmann, D., Cheung, A. C. M., Schulz, S., Smollett, K., Werner, F., and Michaelis, J. (2015) Complete architecture of the archaeal RNA polymerase open complex from single-molecule FRET and NPS. *Nat. Commun. 6*, 6161.

(3) Kapanidis, A. N., Laurence, T. A., Nam, K. L., Margeat, E., Kong, X., and Weiss, S. (2005) Alternating-laser excitation of single molecules. *Acc. Chem. Res. 38*, 523–533.

(4) Treutlein, B., Muschielok, A., Andrecka, J., Jawhari, A., Buchen, C., Kostrewa, D., Hög, F., Cramer, P., and Michaelis, J. (2012) Dynamic Architecture of a Minimal RNA Polymerase II Open Promoter Complex. *Mol. Cell 46*, 136–146.

(5) Andrecka, J., Lewis, R., Bruckner, F., Lehmann, E., Cramer, P., and Michaelis, J. (2008) Single-molecule tracking of mRNA exiting from RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A. 105*, 135–140.

(6) Andrecka, J., Treutlein, B., Arcusa, M. A. I., Muschielok, A., Lewis, R., Cheung, A. C. M., Cramer, P., and Michaelis, J. (2009) Nano positioning system reveals the course of upstream and nontemplate DNA within the RNA polymerase ii elongation complex. *Nucleic Acids Res. 37*, 5803–5809.

(7) Dahan, M., Deniz, A. A., Ha, T., Chemla, D. S., Schultz, P. G., and Weiss, S. (1999) Ratiometric measurement and identification of single diffusing molecules. *Chem. Phys. 247*, 85–106.