Supporting Information

Critical Evaluation of Photocrosslinking Parameters for the Implementation of Efficient DNA-Encoded Chemical Library Selections

Alessandro Sannino,[a] Adrián Gironda-Martínez,[a] Émile M.D. Gorre,[a] Luca Prati,[a] Jacopo Piazzi,[a] Jörg Scheuermann,[b] Dario Neri*,[b] Etienne J. Donckele*,[a] and Florent Samain*[a]

[a] Philochem AG, 8112 Otelfingen (Switzerland)
[b] Institute of Pharmaceutical Sciences, ETH Zürich, 8093 Zürich (Switzerland)

Corresponding Authors:

*F.S.: florent.samain@philochem.ch
*E.J.D.: etienne.donckele@philochem.ch
*D.N.: dario.neri@pharma.ethz.ch

Table of Contents

1. Abbreviations

- AAZ: acetazolamide
- ACN: acetonitrile
- BB: building block
- CAIX: carbonic anhydrase IX
- Cl-SABA: 4-chloro-3-sulfamoyl benzoic acid
- Ct: cycle threshold
- DIPEA: N, N'- diisopropylethylamine
- DMSO: dimethylsulfoxide
- DMT-MM: 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride
- dNTP: deoxynucleotide
- EDC: 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride
- 6-FAM: 6-carboxyfluorescein
- FITC: fluorescein isothiocyanate
- Fmoc: 9-fluorenylmethyloxycarbonyl
- His: polyhistidine-tag
- HFIP: 1,1,1,3,3,3-Hexafluoro-2-propanol
- HOAt: 1-hydroxy-7-azabenzotriazole
- HRP: horseradish peroxidase
- Igg: immunoglobulin G
- *m*-SABA: 3-sulfamoylbenzoic acid
- MeOH: methanol
- PBS: phosphase buffered saline
- PBST: phosphate buffered saline containing Tween 20%
- Q-PCR: Quantitative Polymerase Chain Reaction
- S-NHS: Sulfo-N-hydroxysuccinimide
- SABA: p-sulfamoylbenzoic acid
- SDS-PAGE: Sodium Dodecyl Sulphate – Polyacrilamide Gel Electrophoresis

- SM: small molecule
- TEA: triethylamine
- TEAA: triethylammonium chloride
- TFA: Trifluoroacetic acid
- UV: Ultraviolet

2. Materials and Instrumentation

Starting materials, solvents, and reagents: Unless otherwise noted, all solvents were used as supplied by Fisher Chemicals, Merck, VWR or Sigma-Aldrich in HPLC or analytical grade. All other reagents (Carboxylic acids, Amino acids, Iso(thio)cyanates, Sulfonyl chlorides, Amines) were purchased from several commercial suppliers including Enamine (Princeton, United States), ABCR (Karlsruhe, Germany) Sigma Aldrich (Saint-Louis, United States) and BroadPharm (San Diego, United States). The oligonucleotides were provided by LGC Biosearch Technologies (Denmark), and Eurofins Genomics (Ebersberg, Germany). Klenow polymerase kit was acquired from New England Biolabs (Ipswitch, MA). Dynabeads His-Tag Pull-Down and Herring Sperm DNA were purchased from Thermo Fisher Scientific. PCR purification and gel extraction kits were provided by QIAGEN and used according to the protocol provided by the supplier.

Chromatography, spectrometry and instruments: Mass spectrometry (LC-ESI-MS) spectra were recorded on an Agilent 6100 Series Single Quadrupole MS system combined with an Agilent 1200 Series LC. For small molecules a InfinityLab Proshell 120 EC-C18 column (4.6 x 56 mm, Agilent) was used and compounds were eluted by applying a gradient of 0–100% ACN in H_2O containing 0.1% formic acid (FA). For oligonucleotides an ACQUITY UPLC Oligonucleotide BEH C18 column (130 Å, 1.7 µm, 2.1 mm x 50 mm) was used and compounds were eluted by applying gradients of MeOH and 400 mM HFIP / 15 mM TEA in H_2O. Calculated and measured m/z values are reported as dimensionless quantities. Preparative reversed-phase high-pressure liquid chromatography (RP-HPLC) for the oligonucleotide conjugates was performed on an Agilent 1200 Series with an XTerra® C18 semipreparative column using a gradient of eluent A (TEAA 100 mM) and eluent B (TEAA 100 mM in 80% ACN). Ionic exchange high-pressure liquid chromatography (IE – HPLC) for the purification of the single-stranded single pharmacophore library was performed on a DNAPac PA200 (8 µm, 250 mm x 4 mm) column using a gradient of eluent C (25 mM Tris-HCl 6 M Urea in H_2O, pH = 8) and eluent D (25 mM Tris-HCl, 6 M Urea, 0.4 M $NaClO_4$ in H_2O, pH =8). Preparative reversed-phase high-pressure liquid chromatography (RP-HPLC) for the small molecules was performed on an Phenomenex Gemini® 5 µm NX-C18 110 Å 150 x 10 mm semipreparative column using a gradient of eluent H_20 + 0.1% TFA and ACN + 0.1 % TFA. Affinity

selections were performed robotically on a KingFisher Magnetic Particle Processor (Thermofisher). Photocrosslinking reactions were performed using UVP CL-1000 Ultraviolet Crosslinking at a wavelength of 365 nm with an intensity of approximately 1000 μJ /cm^2. The selection fingerprints were analysed using Matlab (Mathworks). Quantitative PCR analysis was performed using Applied Biosystems StepOne Real-Time PCR System (Applied Biosystems, Massachusetts, USA). Fluorescence polarization measurements were performed using Tecan Spark (Tecan). The concentration was checked by UV absorbance at 260 nm measurements using Nanodrop 2000 instrument (Thermofisher). Preparative medium pressure liquid chromatography (MPLC) was performed on CombiFlash NextGen 300+. (Teledyne ISCO).

3. Oligonucleotide and primer sequences

3.1 Sequences of the oligonucleotides

Elib2_1

5’ – GGAGCTTCTGAATTCTGTGTGCTGAAACGTCGAGTCCCATGGCGCAGC – 3’

Modification: 5’-C6-Aminolink

MW: 14994.6 Da

Elib2_2

5’ – GGAGCTTCTGAATTCTGTGTGCTGAAACTGCGAGTCCCATGGCGCAGC –3’

Modification: 5’-C6-Aminolink

MW: 14994.6 Da

Elib2_3

5’ – GGAGCTTCTGAATTCTGTGTGCTGAAAGCTCGAGTCCCATGGCGCAGC – 3’

Modification: 5’-C6-Aminolink

MW: 15043.7 Da

Elib2_4

5’ – GGAGCTTCTGAATTCTGTGTGCTGAAAGAGCGAGTCCCATGGCGCAGC –3’

Modification: 5’-C6-Aminolink

MW: 14994.6 Da

Elib2_5

5’ – GGAGCTTCTGAATTCTGTGTGCTGACATCGCGAGTCCCATGGCGCAGC –3’

Modification: 5’-C6-Aminolink

MW: 14970.7 Da

d-spacer

5’ – CATGGGACTCGddddddCAGCACACAGAATTCAGAAGCTCC – 3’

Modification: 5’ - $(PO_4)^{3-}$; 3’-C6-Aminolink

MW: 12058.74 Da

E4 DNA – RNA chimeric adapter

5’– CGA **GUC** CCA TGG C**GC AG**C TGC – 3’, **bold:** RNA portions

MW: 6520.17 Da

FAM-Elib4

5’ – CCTGCATCGAATGGATCCGTGGTCGAATTGCAGCTGCGC – 3’

Modification: 5’- FAM

MW: 12497.87 Da (no counter ions); 12537.26 Da (fully protonated)

Elib4_1

5’ – TCCTGCATCGAATGGATCGATACTCGCTGGCAGCTGCGC – 3’

MW: 11959.78 Da

Elib4_2

5’ – ATAACTTCACTGCCGTGTCCATGAACAGAGCAGCTGCGC – 3’

MW: 11936.79 Da

Elib4_3

5’ – AAGTCTCGTCAATTCACACTGTTGGCGATGCAGCTGCGC – 3’

MW: 11958.80 Da

Elib4_4

5’ – ACGTACGTCTCATGTGATGCACATGTATCGCAGCTGCGC – 3’

MW: 11958.80 Da

Elib4_5

5’ – GCTTGGGTGTATTGCACTAGCGTCAAGGCGCAGCTGCGC – 3’

MW: 12055.34 Da

Elib6_code

5’–GCTCTGCACGGTCGCCTGAGATGTAGGATCACGCTGCCTGACGCddddddddCGTCGATCCGGCGC – 3’

MW: 19024.57 Da (no counter ions); 19089.21 Da (fully protonated)

DNA adapter_Elib6

5’–CGAGTCCCATGGCGCCGGATCGACG–3’

MW: 7669 Da

Library Construction Code 1

5’– GGAGCTTCTGAATTCTGTGTGCTG**XXXXXX**CGAGTCCCATGGCGC– 3’

Library Construction Code 2

5’ – CGGATCGACG**YYYYYYY**GCGTCAGGCAGC – 3’

Modification: 5’ - $(PO_4)^{3-}$

Library Construction Code 3

5’ –GCTCTGCACGGTCGCCTGAGATGCTGCCTGACGC - 3’

MW. 10411. 8 Da

DNA Adapter Code1_Code2

5’ – CGATCCGGCGCCAT – 3’

MW = 4224.78 Da

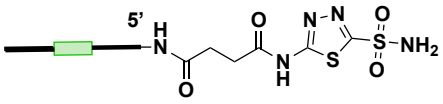
3.2 Sequences of PCR Primers used for Quantitative PCR

LB_FP
5'– GGAGCTTCTGAATTCTGTGTGCTG – 3'
Primer_ESAC qPCR
5' – CCTGCATCGAATGGATCCGTG – 3'
Primer_AAZ_1
5' – TCTGAATTCTGTGTGCTGAAACGT – 3'
Primer_AAZ_2
5' – TCCTGCATCGAATGGATCGATACT – 3'
Primer_SABA_1
5' – TCTGAATTCTGTGTGCTGAAACTG – 3'
Primer_SABA_2
5' – ATAACTTCACTGCCGTGTCCATGA – 3'
PrimerClSABA_1
5' – TCTGAATTCTGTGTGCTGAAAGCT – 3'
PrimerClSABA_2
5' – AAGTCTCGTCAATTCACACTGTTG – 3'
PrimermSABA_1
5' – TCTGAATTCTGTGTGCTGAAAGAG – 3'
PrimermSABA_2
5' – ACGTACGTCTCATGTGATGCACAT – 3'
Primer_NC_1
5' – GGAGCTTCTGAATTCTGTGTGCTGACATCG – 3'
Primer_NC_2
5' – GCTTGGGTGTATTGCACTAGCGTC – 3'

4. Procedures

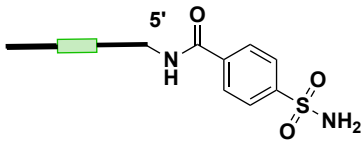
4.1 Synthesis of DNA-tagged ligands for model selection experiments

a) Synthesis of AAZ_Elib2_1



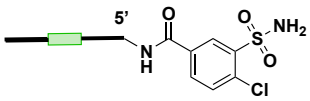
AAZ carboxylic acid derivative (4-Oxo-4-((5-sulfamoyl-1,3,4-thiadiazol-2-yl)amino)butanoic acid) in DMSO (12.5 µl, 200 mM), s-NHS (20 µl, 333 mM) in DMSO/H_2O (2:1), followed by EDC in DMSO (24 µl, 100 mM) were added in DMSO (200 µl) and stirred for 30 min at 30 °C. The oligonucleotide (Elib2_1, 25 µl, 0.66 mM in H_2O) diluted in triethylamine hydrochloride buffer (TEA•HCl, 50 µl, 500 mM, pH = 10.0) was then added. The reaction mixture was stirred for 16 h at 37 °C. The DNA conjugate was purified by EtOH precipitation, followed by RP-HPLC. ESI- LC-MS analysis: theoretical m/z: 15256.87 m/z; found m/z: 15255.98.

b) Synthesis of SABA_Elib2_2



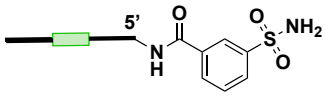
SABA (4-sulfamoyl benzoic acid) in DMSO (12.5 µl, 200 mM), 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methyl-morpholinium chloride (DMT-MM, 32.5 µl, 208 mM in H_2O) were dissolved in 12 µl of DMSO and stirred for 15 min at 25 °C. The oligonucleotide (Elib2_2, 25 µl, 0.66 mM in H_2O) diluted in MOPS buffer (3-(*N*-morpholino)propanesulfonic acid, 35 µl, 100 mM, 1M NaCl, pH = 8.0) and H_2O (12 µl) was added to the mixture and the reaction was stirred for 16 h at 25 °C. The DNA was purified by EtOH precipitation, followed by RP-HPLC. ESI-LC-MS analysis: theoretical m/z: 15176.89; found m/z: 15176.99.

c) Synthesis of Cl-SABA_Elib2_3



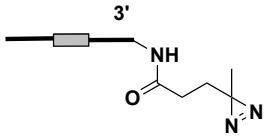
Cl-SABA (4-chloro-3-sulfamoyl benzoic acid) was conjugated to the oligonucleotide Elib2_3, as previously described in **4.1 b)**. ESI-LC-MS analysis: theoretical m/z: 15212.24; found m/z: 15211.63.

d) Synthesis of *m*-SABA_Elib2_4



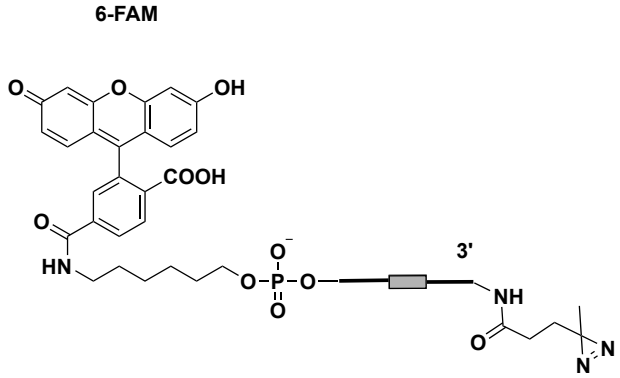
m-SABA (3-sulfamoyl benzoic acid) was conjugated to the oligonucleotide Elib2_4, as previously described in **4.1 b)**. ESI-LC-MS analysis: theoretical m/z: 15226.9; found m/z: 15226.25.

e) Synthesis of DA-d-spacer (DA-d-spacer)



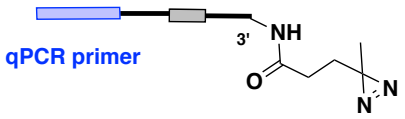
Sulfo-SDA (Sulfo NHS Diazirine, 5 µl, 200 mM in H_2O), d-spacer (65 µl, 304 µM) were dissolved in triethylamine hydrochloride buffer (TEA•HCl, 45 µl, 500 mM, pH = 10.0). The reaction was stirred overnight at 37 °C. The DNA-conjugate was purified by EtOH precipitation and RP-HPLC. ESI-LC-MS analysis: theoretical m/z: 12168; found m/z: 12168.20.

f) Ligation of FAM-Elib4 to DA-d-spacer: synthesis of DA-FAM-Elib4



DA-d-spacer (25 µl, 80 µM), FAM-Elib4 (26 µl, 100 µM), E4 chimeric DNA/RNA adapter (13.3 µl, 300 µM) were dissolved in H_2O (111 µl) and 10 x T4 DNA Ligase Reaction buffer (B0202S, *New England Biolabs*) was added (20 µl). A pre-hybridization step was performed for 2 min at 90 °C, before the mixture was allowed to cool down at 25 °C. T4 DNA Ligase (5 µl, 400 U/ µl, *New England Biolabs*) was added and the reaction was left 14 h at 16 °C, before inactivating the ligase for 10 min at 65 °C. The crude ligation product was checked by ESI-LC-MS (theoretical m/z: 24687 found m/z: 24686.58) and directly used for the next step **h)**.

g) Ligation of unique qPCR primer sequences to d-spacer (DA_Elib4_X)



Procedure **4.1 f)** was used to perform the ligations of different PCR-primers. The encoding strategy is displayed in the Scheme S2.

h) RNAse adapter degradation

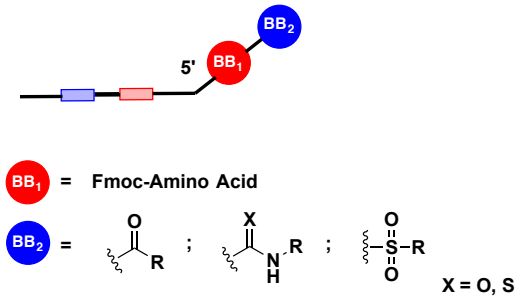
The crude ligation solution (100 µl) was treated with RNAse HII (2 µl) in 10x ThermoPol Reaction Buffer (10 µl, *New England Biolabs*). The degradation of the adapter was confirmed by gel electrophoresis analysis on Novex™ TBE-Urea Gels 15% (*Invitrogen*). The product was purified by nucleotide removal columns (Smartpure DNA Purification Kit, *Eurogenentec*), redissolved in H_2O and the concentration evaluated by Nanodrop measurement.

i) Klenow hybridization

FAM-DNA-CAIX Ligand **DNA-CAIX Ligand**

Elib2 conjugates and **DA-Elib4 derivatives (with/without FAM)** conjugates (reaction scale: 50 pmol) were dissolved in H_2O (final volume: 100 µl) and NEBuffer™ 2 (10 µl, *New England Biolabs*). The code **Elib2_5** was used to construct the respective **FAM-DNA-NC** and **DNA-NC** derivatives. A pre-hybridization step was performed for 2 min at 90°C, before the mixture was allowed to cool down at 25°C. The DNA Polymerase I, Large (Klenow) Fragment (1 µl, *New England Biolabs)* and deoxynucleotide solution mix (4 µl, 5 mM) were added and incubated for 1 h at 25°C. The reaction was checked by gel electrophoresis analysis on Novex™ TBE-Urea Gels 15% (*Invitrogen*). The product was purified by QIAquick PCR Purification Kit (*Qiagen*) and the concentration evaluated by Nanodrop measurement. The corresponding derivatives (FAM-DNA CAIX ligands and DNA-CAIX ligands) were used as input in model selection experiments. The DNA encoding strategy is displayed in Scheme S1 and S2.

4.2 Synthesis of the DNA single-pharmacophore DNA-Encoded chemical library

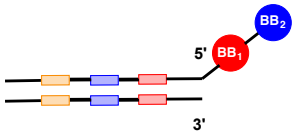


The encoding strategy is displayed in Scheme S3. The coupling procedures of the first set of building block 1 (BB1) and building block 2 (BB2) were performed as described in Sannino *et al.*[1]

a) Ligation Procedure for single-stranded DNA library synthesis

The pool of the code 1-BB1 (1.4 µl, 1 mM H_2O) conjugates, the corresponding **Code 2** (2.1 µl, 1 mM H_2O) and **DNA adapter Code1_Code2** (1.4 µl, 2 mM H_2O), were mixed and incubated at room temperature for 1 hour. In a separate vial, a master mix of the corresponding amount of 10x T4 DNA-ligase buffer (1.7 µl, New England Biolabs), water (0.15 µl) and T4 DNA-ligase (0.2 µl, 400 U/µl, *New England Biolabs*) was prepared and added to the oligo mixture. The ligation process was left for 16 h at 16 °C, before inactivating the ligase for 10 min at 65 °C. The formation of the desired ligated product was confirmed by gel electrophoresis analysis. The mixture was dried over speedvac and the DNA pellet was directly used for the coupling of the second building block (BB2). After reaction, the pool of the DNA conjugates was purified by ionic exchange chromatography (IE-HPLC).

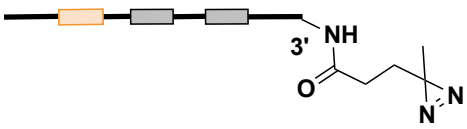
b) Klenow Polymerization



The single-stranded single pharmacophore library (31.7 µl, 3.15 µM) and **Code 3** (2 µl, 50 µM) were dissolved in H_2O (50.3 µl) and NEB Buffer 10x (10 µl). A pre-hybridization step was performed at 95 °C for 2 min, before the mixture was allowed to cool down at 25 °C. The DNA Polymerase I, Large (Klenow) Fragment (2 µl, *New England Biolabs)* and deoxynucleotide solution mix (4 µl, 5 mM) were added and incubated for 1 h at 25 °C. The reaction was checked by gel electrophoresis analysis on Novex™ TBE-Urea Gels 15% (*Invitrogen*). The library was purified by QIAquick PCR Purification Kit (*Qiagen*), redissolved in 100 µl and the concentration evaluated by Nanodrop measurement.

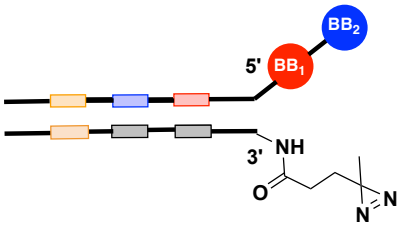
4.3 Synthesis of the Photoreactive Library

a) Ligation procedure: Synthesis of the photoreactive DNA-strand (DA-Elib6)



DA-d-spacer (25 µl, 80 µM), **Elib6_code** (5.20 µl, 500 µM), **DNA adapter_Elib6** (8 µl, 500 µM) were dissolved in H_2O (49 µl) and 10 x T4 DNA Ligase Reaction buffer (B0202S, *New England Biolabs*) was added (10 µl). A pre-hybridization step was performed at 90°C for 2 min, before the mixture was allowed to cool down at 25°C. T4 DNA Ligase (2.5 µl, 400 U/ µl, *New England Biolabs*) was added and the reaction was left for 14 h at 16 °C, before inactivating the ligase for 10 min at 65 °C. The crude ligation product was checked by LC-MS (theoretical m/z: 31239 found m/z: 31238.42). The product was purified by QIAgen PCR Purification Kit (Qiagen) to remove the excess of code and adapter. The concentration was evaluated by Nanodrop measurement and the product was directly used for the next step.

b) Klenow polymerization



The single-stranded single pharmacophore library, prepared as described in Paragraph **4.2.a),** (30 µl, 1.31 µM) and the **DA-Elib6** (19 µl, 2.11 µM) were dissolved in H_2O (35.5 µl) and NEB Buffer 10x (10 µl). A pre-hybridization step was performed for 2 min at 95 °C, before the mixture was allowed to cool down at 25 °C. The DNA Polymerase I, Large (Klenow) Fragment (1.5 µl, *New England Biolabs)* and deoxynucleotide solution mix (4 µl, 5 mM) were added and incubated for 1 h at 25 °C. The reaction was checked by gel electrophoresis analysis on Novex™ TBE-Urea Gels 15% (*Invitrogen*). The library was purified by QIAquick PCR Purification Kit (*Qiagen*) and the concentration evaluated by Nanodrop measurement.

4.4 Screening Methodologies

a) Photocrosslinking Methodology

The photocrosslinking methodology is described in the main text.

b) Affinity selection against polyhistidine-tagged CAIX

Dynabeads His-Tag Pull-Down (12 µl / selection, *Invitrogen*) were washed with 3 x binding buffer (500 ml; 50 mM sodium phosphate, 300 mM NaCl, pH 8.0, 0.01 %/ Tween-20) and resuspended in binding buffer (100 µl). By using the KingFisher magnetic particle processor, the magnetic beads were transferred to a solution of binding buffer and His-tagged target (CAIX) at a certain concentration (2 µM, 100 µl) and incubated for 30 min with continuous gentle mixing. The beads were washed (3 x 3 min, 200 µl) with binding buffer and subsequently transferred to a solution of library (10^6 copies / lib. member) in pull-down buffer (100 µl; 3.25 mM sodium phosphate, 70 mM NaCl, pH 7.4, 0.01 % Tween-20), containing herring sperm (0.3 mg/ml, *Thermofisher*). The beads were incubated for 1 h with continuous gentle mixing. The beads were removed and washed with pull-down buffer (5 x 30 s, 200 µl) and incubated with elution buffer (100 µl, 10 mM NaH_2PO_4, 300 mM NaCl and 200 mM Imidazole) for 30 min. After incubation, the beads were separated. The eluate was amplified by PCR and submitted to the Functional Genomics Center in Zürich for high-throughput DNA sequencing on an Illumina HiSeq 2500 instrument.

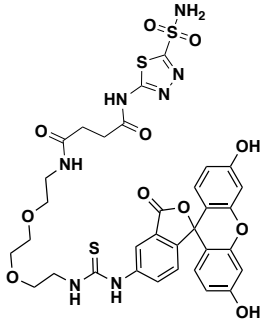
4.5 Quantitative PCR: General Procedure

In an Optical 8-Cap Strip (*Applied Biosystems*), in a volume of 25 µl, the following were combined: DNA template (2.5 µl; selection eluate or water for the no-templace control), forward primer (0.5 µl, 10 µM in H_2O), reverse primer (0.5 µl, 10 µM in H_2O), SensiFAST™ SYBR Hi-ROX PCR MasterMix (12.5 µl, *Bioline)* and H_2O (9 µl). qPCR program: 2 min at 95°C (hold), 5 s at 95 °C, (denaturation) 10 s at 60°C (annealing), 10 s at 72 °C (extension). The number of denaturation/annealing/extension cycles was set at 40. The corresponding Ct values were correlated with the absolute quantity of DNA through standard curve calibration experiments. The

experimental measurements are plotted as absolute quantities of DNA molecules (DNA copies) in 100 µl. The specificity of the reactions was checked by melt curve analysis or gel electrophoresis. Melt curve program: 15 s 95°C, 60°C to 95°C, 0.3 °C/min.

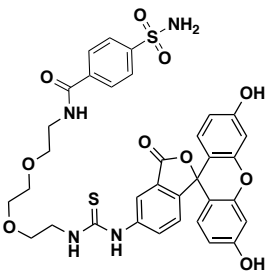
4.6. Hit Validation: Synthesis of small-molecule FITC derivatives

a) Synthesis of AAZ - FITC



To a solution of 4-Oxo-4-((5-sulfamoyl-1,3,4-thiadiazol-2-yl)amino)butanoic acid (0.1 mmol, 1 eq, 28 mg) in DCM (1 ml) trimethylamine (0.3 mmol, 3 eq, 80 µl) was added. Once dissolved, FITC-$(PEG)_2$ amine (0.1 mmol, 1 eq, 57 mg) and HATU (0.11 mmol, 1.1 eq, 42 mg) were added. The reaction was stirred overnight at 25 °C. The crude was evaporated *in vacuo* and purified by inverse phase HPLC (r.t.: 4.65 min, H_2O/MeCN, from 30% to 60%, 5.0 ml/min, 5 min run), obtaining a yellow powder (23 mg, 28%). ^{1}H NMR (400MHz, DMSO-d_6) δ = 10.03 (br. s., 2 H), 8.39 - 8.24 (m, 3 H), 8.09 (br. s., 1 H), 8.00 (dd, J_1 = 4.6 Hz, J_2 = 4.6 Hz, 1 H), 7.74 (d, J = 8.1 Hz, 1 H), 7.23 - 7.15 (m, 1 H), 6.71 - 6.66 (m, 2 H), 6.64 - 6.53 (m, 4 H), 3.64 - 3.54 (m, 8 H), 3.42 (t, J = 6.0 Hz, 2 H), 3.20 (m, 2 H), 2.292 (m, 2H), 2.78 - 2.70 (m, 2 H). ^{13}C NMR (100MHz, DMSO-d_6) δ = 180.78, 172.02, 171.00, 168.73, 167.61, 164.43, 161.32, 159.72, 152.11, 143.75, 141.55, 129.26, 126.74, 124.30, 112.81, 109.94, 102.45, 87.13, 80.14, 69.76, 69.31, 68.66, 45.98, 43.89, 30.48, 29.59. LCMS (ES^+) m/z 800.0 $(M+H)^+$. TOF-MS (ES^+): $C_{33}H_{34}N_7O_{11}S_3$ $[M+H]^+$ calculated: 800.1473, found: 800.1452.

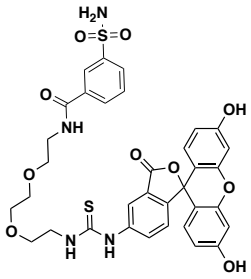
b) Synthesis of SABA - FITC



To a solution of 4-sulfamoylbenzoic acid (0.1 mmol, 1 eq, 20 mg) in DCM (1 ml), trimethylamine

(0.3 mmol, 3 eq, 80 μl) was added. Once dissolved, FITC-$(PEG)_2$ amine (0.1 mmol, 1 eq, 57 mg) and HATU (0.11 mmol, 1.1 eq, 42 mg) were added. The reaction was stirred overnight at 25 °C. The crude was evaporated *in vacuo* and purified by inverse phase HPLC (r.t.: 4.67 min, H_2O/MeCN, from 30% to 62%, 5.0 ml/min, 6 min run), obtaining a yellow powder (33 mg, 45%). ^{1}H NMR (400MHz, DMSO-d6) δ = 10.12 (br. s., 3 H), 8.73 (t, *J* = 5.6 Hz, 1 H), 8.28 (s, 1 H), 8.17 - 8.05 (m, 1 H), 8.03 - 7.96 (m, 2 H), 7.93 - 7.84 (m, 2 H), 7.73 (m, 1 H), 7.48 (s, 2 H), 7.19 (d, *J* = 8.4 Hz, 1 H), 6.68 (d, *J* = 2.5 Hz, 2 H), 6.64 - 6.51 (m, 4 H), 3.69 (br. s., 2 H), 3.64 - 3.53 (m, 8 H), 3.45 (m, 2 H). ^{13}C NMR (100MHz, DMSO-d_6) δ = 180.77, 168.73, 165.54, 159.67, 152.08, 147.93, 147.42, 146.43, 141.53, 137.49, 132.31, 129.25, 128.06, 126.74, 125.83, 124.26, 116.59, 112.77, 109.91, 102.45, 83.20, 72.00, 69.79, 69.01, 68.65, 43.89. LCMS (ES$^+$) m/z 721.1 $(M+H)^+$. TOF-MS (ES$^+$): $C_{34}H_{33}N_4O_{10}S_2$ $[M+H]^+$ calculated: 721.1633, found: 721.1624.

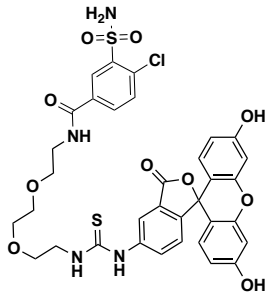
c) Synthesis *m*-SABA - FITC



To a solution of 3-Sulfamoylbenzoic acid (0.1 mmol, 1 eq, 20 mg) in DCM (1 ml), trimethylamine (0.3 mmol, 3 eq, 80 μl) was added. Once dissolved, FITC-$(PEG)_2$ amine (0.1 mmol, 1 eq, 57 mg) and HATU (0.11 mmol, 1.1 eq, 42 mg) were added. The reaction was stirred overnight at 25 °C. The crude was evaporated *in vacuo* and purified by inverse phase HPLC (r.t.: 4.67 min, H_2O/MeCN, from 30% to 62%, 5.0 ml/min, 6 min run), obtaining a yellow powder (30 mg, 41%). ^{1}H NMR (400MHz, DMSO-d_6) δ = 10.32 - 9.95 (m, 3 H), 8.79 (t, *J* = 5.6 Hz, 1 H), 8.32 (m, 1 H), 8.28 (s, 1 H), 8.13 (br. s., 1 H), 8.09 - 8.00 (m, 1 H), 8.00 - 7.92 (m, 1 H), 7.82 - 7.71 (m, 1 H), 7.67 (dd, J_1 = J_2 =7.7 Hz, 1 H), 7.44 (br. s., 2 H), 7.21 - 7.14 (m, 1 H), 6.73 - 6.65 (m, 2 H), 6.64 - 6.50 (m, 4 H), 3.68 (br. s., 2 H), 3.64 - 3.54 (m, 8 H), 3.45 (m, 2 H). ^{13}C NMR (100MHz, DMSO-d_6) δ = 180.56, 168.52, 165.15, 159.54, 151.90, 147.58, 147.02, 144.38, 141.33, 135.05, 132.04, 130.13, 129.13, 128.94, 128.11, 124.73,

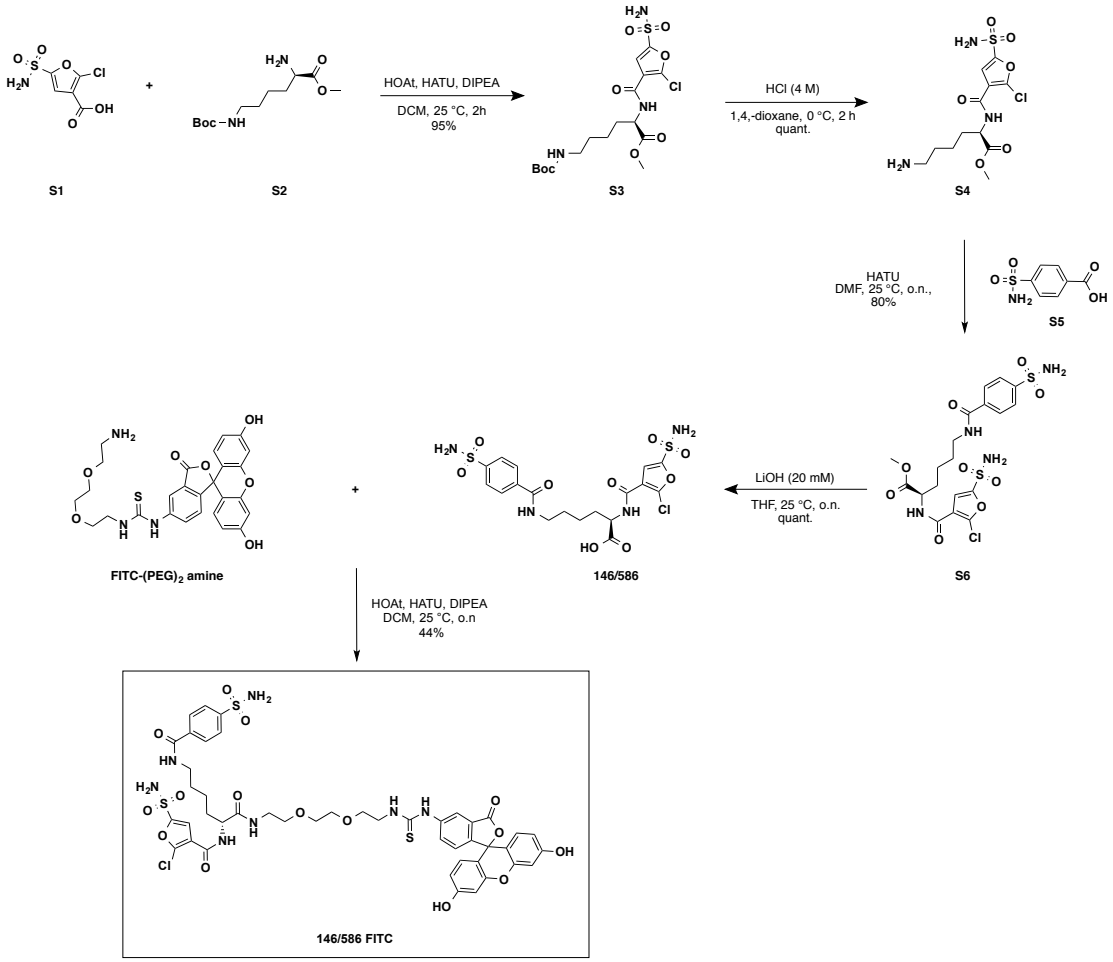
116.38, 112.63, 109.74, 102.23, 82.26, 69.63, 68.80, 68.43, 43.66. LCMS (ES^+) m/z 721.1 $(M+H)^+$. TOF-MS (ES^+): $C_{34}H_{33}N_4O_{10}S_2$ $[M+H]^+$ calculated: 721.1633, found: 721.1630.

d) Synthesis Cl-SABA - FITC



To a solution of 4-chloro-3-sulfamoylbenzoic acid (0.1 mmol, 1 eq, 23 mg) in DCM (1 ml), trimethylamine (0.3 mmol, 3 eq, 80 μl) was added. Once dissolved, FITC-**(**$PEG)_2$ amine (0.1 mmol, 1 eq, 57 mg) and HATU (0.11 mmol, 1.1 eq, 42 mg) were added. The reaction was stirred overnight at 25 °C. The crude was evaporated *in vacuo* and purified by inverse phase HPLC (r.t.: 4.75 min H_2O/MeCN, from 30% to 62%, 5.0 ml/min, 6 min run), obtaining a yellow powder (35 mg, 46%). ^1H NMR (400MHz, DMSO-d_6) δ = 10.02 (br. s., 2 H), 8.85 (t, *J* = 5.4 Hz, 1 H), 8.46 (m, 1 H), 8.28 (s, 1 H), 8.15 - 7.99 (m, 2 H), 7.80 - 7.63 (m, 4 H), 7.19 (d, *J* = 8.4 Hz, 1 H), 6.72 - 6.53 (m, 6 H), 3.72 - 3.65 (m, 2 H), 3.64 - 3.53 (m, 8 H), 3.44 (m, 2 H). ^{13}C NMR (100MHz, DMSO-d_6) δ = 180.56, 176.08, 172.72, 168.51, 164.36, 159.51, 151.89, 141.33, 141.10, 133.32, 133.13, 131.62, 131.44, 129.42, 129.04, 128.13, 126.53, 125.48, 124.08, 116.40, 112.60, 109.72, 102.23, 69.62, 69.57, 68.75, 43.67. LCMS (ES^+) m/z 755.0 $(M+H)^+$. TOF-MS (ES^+): $C_{34}H_{32}N_8ClN_4O_{10}S_2[M+H]^+$ calculated: 755.1243, found: 755.1237.

e) Synthesis of **146/586 - FITC**



Scheme S1. Synthetic route of compound **146/586 FITC**.

- Synthesis of compound **S3**

To a solution of 2-chloro-5-sulfamoylfuran-3-carboxylic acid (**S1**; 0.44 mmol, 1 eq, 100 mg) in DCM (1 ml), DIPEA (1.3 mmol, 3 eq, 250 µl) was added. (*R*)-2-Amino-6-tert-butoxycarbonylamino-hexanoic acid methyl ester (**S2**; 0.44 mmol, 1 eq, 114 mg), HOAt (0.22 mmol, 0.5 eq, 30 mg) and HATU (0.53 mmol, 1.2 eq, 200 mg) were then subsequently added. The reaction was stirred for 2 h at 25 °C . The crude was evaporated *in vacuo*, diluted in DCM and washed with NH_4Cl twice. The organic phases were dried and purified by silica gel chromatography (DCM:MeOH 100:0 to DCM:MeOH 93:7) yielding **S3** as a white powder (194 mg, 95%). ^{1}H NMR (400MHz, CD_3CN) δ = 7.36 (s, 1 H), 7.07 (br. d, *J* = 6.8 Hz, 1 H), 6.10 (br. s., 1 H), 5.29 (br. s., 1 H), 4.52 (td, *J* = 5.1, 7.9 Hz, 1 H), 3.72 (s, 3 H), 3.03 (m, 2 H), 1.93 - 1.71 (m, 2 H), 1.55 - 1.33 (m, 13 H). ^{13}C NMR (100MHz,

DMSO-d_6) δ = 171.20, 168.49, 164.65, 164.07, 159.05, 145.67, 144.86, 143.35, 77.38, 52.90, 45.28, 33.75, 28.31, 25.40, 22.89. LCMS (ES^+) m/z 490.1 $(M+Na)^+$. TOF-MS (ES^+): $C_{17}H_{26}ClN_3NaO_8S$ $[M+Na]^+$ calculated: 490.1221, found: 490.1008

- Synthesis of compound **S4**

HCl 4 M in 1,4-dioxane (1.0 eq.) previously cooled to 0 °C was added to a solution of **S3** in DCM (1 M). The reaction was stirred and slowly warmed to room temperature for 2 h. The mixture was evaporated *in vacuo*, Et_2O was added and the mixture was evaporated several times until a yellowish powder was obtained (**S4**; 290 mg, quantitative). The crude was used for the following step without further purification.

- Synthesis of compound **S6**

To a solution of 4-sulfamoylbenzoic acid (**S5**; 0.2 mmol, 1 eq, 45 mg) in DCM (1 ml), DIPEA (0.6 mmol, 3 eq, 104 μl) was added. **S4** in 1 mL of DMF (0.2 mmol, 1 eq, 90 mg), and HATU (0.24 mmol, 1.2 eq, 91 mg) were then added. The reaction was stirred overnight at 25 °C. The crude was evaporated *in vacuo*, diluted in DCM and washed with NH_4Cl twice. The organic phases were dried and purified by silica gel chromatography (DCM:MeOH 100:0 to DCM:MeOH 95:5) obtaining **S6** as a reddish powder (88 mg, 80%). ^1H NMR (400MHz, CD_3CN) δ = 8.03 - 7.86 (m, 4 H), 7.41 (s, 1 H), 7.26 (br. t, *J* = 6.1 Hz, 1 H), 7.20 (br. d, *J* = 7.1 Hz, 1 H), 5.96 (br. s., 4 H), 4.52 (td, *J* = 4.8, 10.0 Hz, 1 H), 3.71 (s, 3 H), 3.39 (s, 2 H), 1.86 (s, 2 H), 1.64 (s, 2 H), 1.54 - 1.42 (m, 2 H). ^{13}C NMR (100MHz, DMSO-d_6) δ = 171.62, 169.91, 165.07, 164.49, 159.47, 152.68, 151.86, 146.09, 145.28, 139.24, 129.04, 63.67, 53.32, 45.70, 34.17, 28.73, 23.21. LCMS (ES^+) m/z 551.0 $(M+H)^+$, TOF-MS (ES^+): $C_{19}H_{24}ClN_4O_9S_2$ $[M+H]^+$ calculated: 551.0673, found: 551.0664.

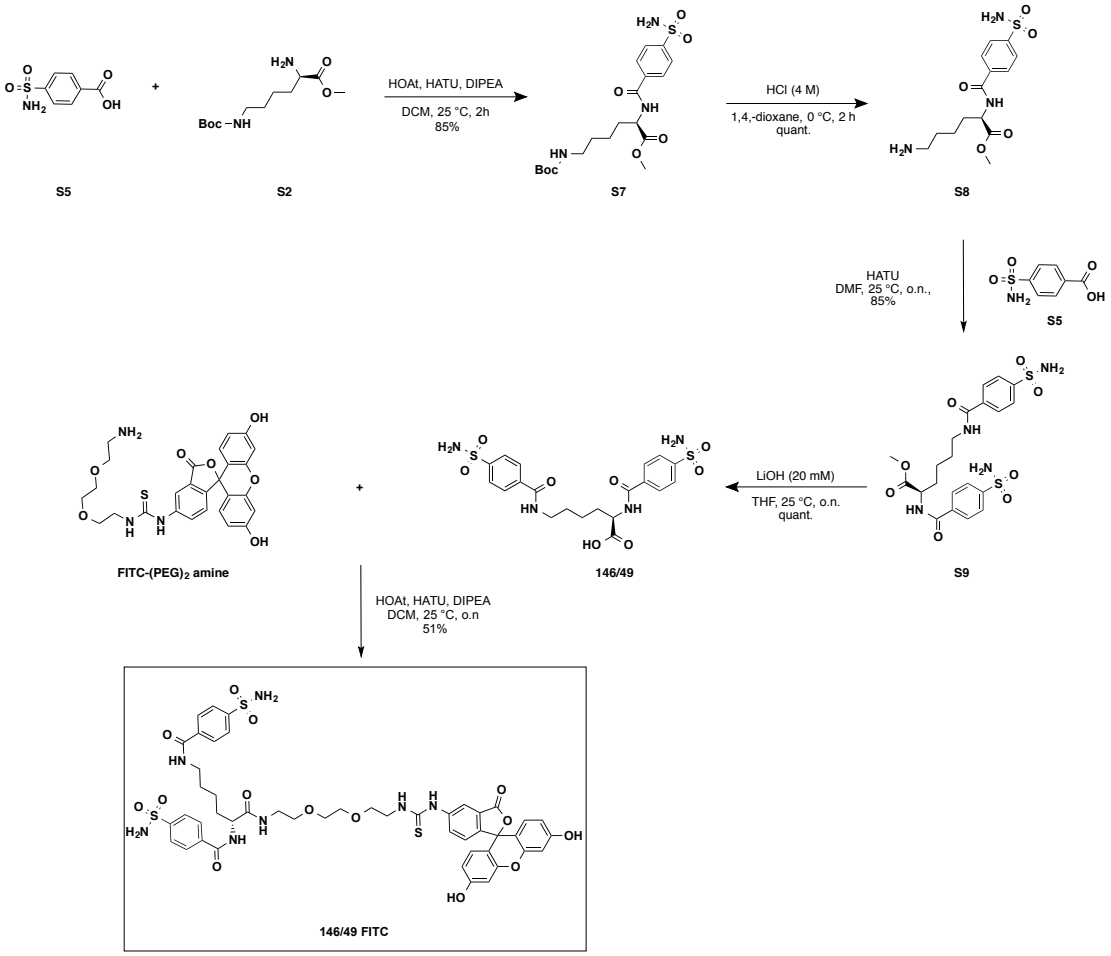
- Synthesis of **146/586**

To a solution of **S6** (0.054 mmol, 1 eq, 30 mg) in THF (20 mM), a solution of 20 mM LiOH(aq) (0.054 mmol, 1 eq, 2.75 ml) was added. The reaction was stirred overnight at 25 °C. The crude was then neutralized with 1 M HCl(aq) (c.a. 50 μl), evaporated *in vacuo*, ACN was added and the mixture was evaporated several times until a white powder is obtained. The product **146/586** was used for the following step without further purification. LCMS (ES^+) m/z 537.0 $(M+H)^+$

- Synthesis of **146/586 - FITC**

To a solution of **146/586** (0.019 mmol, 1 eq, 10 mg) in DMF (1 mL), DIPEA (0.058 mmol, 3 eq, 12 μl), FITC-$(PEG)_2$ amine (0.02 mmol, 1.1 eq, 10 mg), HOAt (0.01 mmol, 0.5 eq, 1.5 mg) and HATU (0.02 mmol, 1.1 eq, 7.6 mg) were added. The reaction was stirred overnight at 25 °C. The crude was evaporated *in vacuo* and purified by inverse phase HPLC (r.t.: 4.74 min, H_2O/MeCN, from 30% to 68%, 5.0 ml/min, 7 min run), obtaining **146/586-FITC** as a yellow powder (9.0 mg, 44%). ^{1}H NMR (400MHz ,DMSO-d_6) δ = 10.15 (br. s., 2 H), 8.64 (s, 1 H), 8.40 - 8.27 (m, 2 H), 8.22 - 8.11 (m, 1 H), 8.05 (s, 1 H), 8.01 - 7.94 (m, 1 H), 7.94 - 7.85 (m, 2 H), 7.75 (m, 2 H), 7.59 (s, 1 H), 7.46 (s, 1 H), 7.18 (d, *J* = 8.4 Hz, 1 H), 6.72 - 6.49 (m, 6 H), 4.36 - 4.24 (m, 1 H), 3.75 - 3.64 (m, 2 H), 3.63 - 3.51 (m, 8 H), 3.45 - 3.37 (m, 2 H), 3.29 - 3.17 (m, 3 H), 3.10 (dd, *J* = 4.6, 7.1 Hz, 2 H), 1.86 - 1.26 (m, 6 H). ^{13}C NMR (100MHz, DMSO-d_6) δ = 179.16, 172.10, 169.39, 165.55, 165.22, 164.97, 159.95, 153.16, 152.34, 146.57, 145.76, 144.25, 139.72, 138.02, 135.70, 134.89, 132.43, 129.51, 128.25, 126.05, 125.09, 120.50, 112.14, 110.17, 107.26, 103.03, 102.70, 81.20, 70.09, 69.42, 53.79, 46.17, 44.59, 34.65, 29.21, 23.68. LCMS (ES^+) m/z 1056.0 $(M+H)^+$. TOF-MS (ES^+): $C_{45}H_{47}ClN_7O_{15}S_3$ $[M+H]^+$ calculated: 1056.1981, found: 1056.1987.

f) Synthesis of **146-49 - FITC**



Scheme S2. Synthetic route of compound **146/49 FITC**.

- Synthesis of compound **S7**

To a solution of 4-sulfamoylbenzoic acid (**S5**; 1 mmol, 1 eq, 201 mg) in DCM (1 ml), DIPEA was added (3 mmol, 3 eq, 500 µl). (*R*)-2-Amino-6-tert-butoxycarbonylamino-hexanoic acid methyl ester (**S2**; 1 mmol, 1 eq, 260 mg), HOAt (0.5 mmol, 0.5 eq, 68 mg) and HATU (1.2 mmol, 1.2 eq, 456 mg) were added. The reaction was stirred for 2 h at 25 °C. The crude was evaporated *in vacuo*, diluted in DCM and washed with NH_4Cl twice. The organic phases were dried and purified by silica gel chromatography (DCM:MeOH 100:0 to DCM:MeOH 93:7) obtaining **S7** as a white powder (376 mg, 85%). ^{1}H NMR (400MHz, DMSO-d_6) δ = 8.91 (br. d, *J* = 7.3 Hz, 1 H), 8.07 - 7.99 (m, 2 H), 7.95 - 7.88 (m, 2 H), 7.63 - 7.35 (m, 2 H), 6.79 (br. t, *J* = 5.6 Hz, 1 H), 4.51 - 4.32 (m, 1 H), 3.66 (s, 3 H), 3.00 - 2.83 (m, 2 H), 1.92 - 1.68 (m, 2 H), 1.36 (s, 13 H). ^{13}C NMR (100MHz, DMSO-d_6) δ = 172.22, 165.66,

165.48, 147.58, 146.50, 137.97, 137.69, 77.97, 53.98, 45.13, 31.86, 29.19, 25.50, 23.70. LCMS (ES^+) m/z 465.1 $(M+Na)^+$. TOF-MS (ES^+): $C_{19}H_{29}N_3O_7SNa$ $[M+Na]^+$ calculated: 466.1618, found: 466.1611.

- Synthesis of compound **S8**

4 M HCl in 1,4-dioxane (1.0 eq.) previously cooled to 0 °C, was added to a solution of **S7** in DCM (1 M). The reaction was stirred and slowly warmed to room temperature for 2 h. The mixture was evaporated *in vacuo*, Et_2O was added and the mixture was evaporated several times until a yellowish powder was obtained (**S8**; 290 mg, quantitative). The crude was used without further purifications for the next step.

- Synthesis of compound **S9**

To a solution of 4-sulfamoylbenzoic acid (**S5**; 0.94 mmol, 1.1 eq, 189 mg) in DCM (1 ml), DIPEA (4.3 mmol, 5 eq, 500 μl) was added. Once dissolved, **2** in 1 mL of DMF (0.85 mmol, 1 eq, 296 mg), and HATU (1 mmol, 1.2 eq, 390 mg) were added. The reaction was left stirring for 2 hours. The crude was evaporated *in vacuo*, diluted in DCM and washed with NH_4Cl_2 two times. The organic phases were dried and purified by silica gel chromatography (DCM:MeOH 100:0 to DCM:MeOH 85:15) obtaining **S9** as a white powder (376 mg, 85%). ^{1}H NMR (400MHz, DMSO-d_6) δ = 8.93 (d, *J* = 7.4 Hz, 1 H), 8.65 (t, *J* = 5.6 Hz, 1 H), 8.06 - 8.00 (m, 2 H), 8.00 - 7.95 (m, 2 H), 7.95 - 7.86 (m, 4 H), 7.47 (br. s., 4 H), 4.54 - 4.39 (m, 1 H), 3.65 (s, 3 H), 3.30 - 3.18 (m, 2 H), 1.94 - 1.77 (m, 2 H), 1.65 - 1.51 (m, 2 H), 1.51 - 1.35 (m, 2 H). ^{13}C NMR (100MHz, DMSO-d_6) δ = 171.80, 165.24,165.06, 159.46, 158.86, 163.16, 152.60, 147.08, 144.35, 140.55, 135.27, 133.01, 129.03, 64.16, 53.56, 45.71, 31.44, 28.77, 23.29. LCMS (ES^+) m/z 527.8 $(M+H)^+$. TOF-MS (ES^+): $C_{21}H_{27}N_4O_8S_2$ $[M+H]^+$ calculated: 527.1265, found: 527.1268**.**

- Synthesis of **146/49**

To a solution of **S9** (0.075 mmol, 1 eq, 40 mg) in THF (0.1 M), a solution of 2 M LiOH aq. (0.23 mmol, 3 eq, 115 μl) was added. The reaction was stirred overnight at 25 °C. The crude was then neutralized with 1 M HCl aq. (about 250 μl) and evaporated *in vacuo*. ACN was added and the

mixture was evaporated several times until a white powder is obtained. The product **146/49** was used for the following step without further purification. LCMS (ES^+) m/z 513.0 $(M+H)^+$

- Synthesis of **146/49 FITC**

To a solution of **146/49** (0.019 mmol, 1 eq, 10 mg) in DMF (1 ml) was added DIPEA (0.058 mmol, 3 eq, 12 µl), FITC-**(**$PEG)_2$ amine (0.02 mmol, 1.1 eq, 10 mg), HOAt (0.01 mmol, 0.5 eq, 1.5 mg) and HATU (0.02 mmol, 1.1 eq, 7.6 mg) were added. The reaction was stirred overnight at 25 °C. The crude was evaporated *in vacuo* and purified by inverse phase HPLC (r.t.: 4.25 min, H_2O/MeCN, from 30% to 68%, 5.0 ml/min, 7 min run), obtaining **146/49 FITC** as yellow powder (10 mg, 51%). ^1H NMR (400MHz, DMSO-d_6) δ = 10.29 - 10.01 (m, 3 H), 8.78 - 8.58 (m, 2 H), 8.30 (br. s., 1 H), 8.15 (br. s., 1 H), 8.11 - 7.93 (m, 4 H), 7.93 - 7.85 (m, 3 H), 7.80 - 7.71 (m, 1 H), 7.54 - 7.42 (m, 3 H), 7.19 (d, *J* = 8.4 Hz, 1 H), 6.68 (d, *J* = 2.3 Hz, 2 H), 6.64 - 6.53 (m, 4 H), 4.44 (br. s., 1 H), 3.69 (br. s., 2 H), 3.59 (br. s., 8 H), 3.44 (br. s., 2 H, overestimated due to H_2O), 3.26 (br. s., 3 H), 3.11 (s, 2 H), 2.83 (s, 2 H), 1.76 (br. s., 2 H), 1.64 - 1.50 (m, 2 H), 1.50 - 1.28 (m, 3 H). ^{13}C NMR (100MHz, DMSO-d_6) δ = 181.05, 172.28, 169.00, 165.85, 165.72, 165.54, 159.94, 152.34, 147.64, 146.78, 146.56, 141.83, 138.03, 137.75, 137.49, 129.51, 128.69, 128.32, 128.25, 126.04, 126.94, 124.51, 116.80, 113.04, 110.18, 102.70, 101.13, 100.99, 100.67, 83.45, 70.09, 70.04, 69.45, 69.26, 54.04, 46.19, 44.13, 31.923, 29.25, 23.76. LCMS (ES^+) m/z 1033.0 $(M+H)^+$. TOF-MS (ES^+): $C_{47}H_{50}N_7O_{14}S_3$ $[M+H]^+$ calculated: 1032.2578, found: 1032.2563.

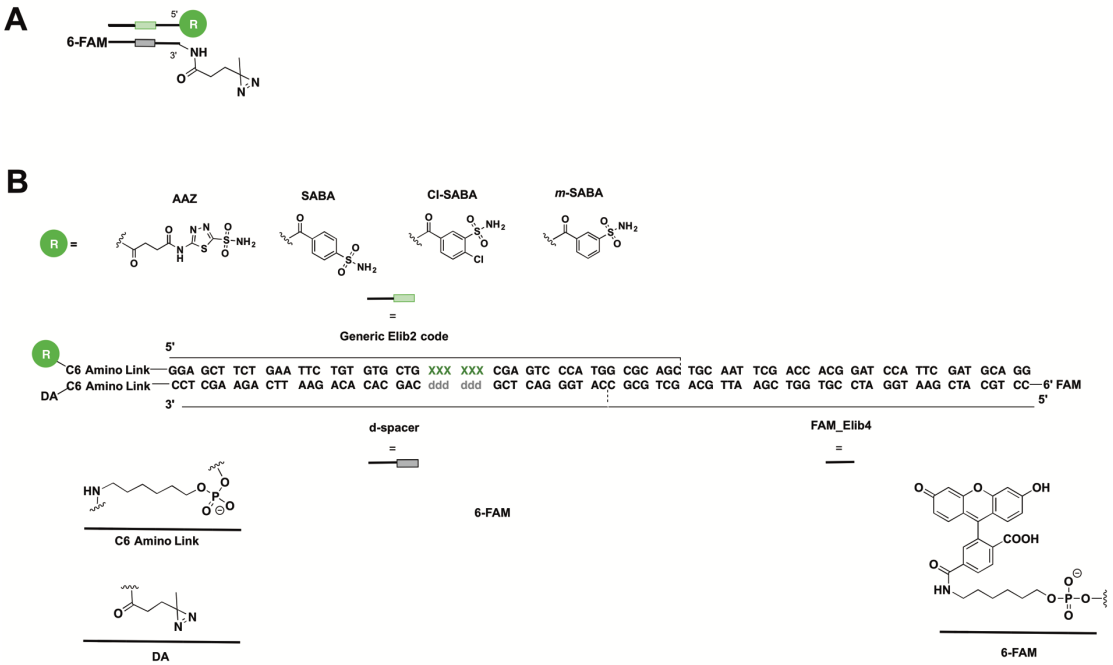
4.7. Fluorescence Polarization procedure

FITC small molecule derivatives were dissolved in DMSO obtaining 10 μM stock solutions. Stocks were serially diluted with PBS Gibco 1X pH = 7.4 at a concentration of 20 nM. A serial dilution of the protein in PBS Gibco 1X (starting from 10^{-4} to 10^{-11} M) was prepared in a 384-wells plate. The FITC-labeled small molecule solution was added in each well at a final concentration of 10 nM. After incubation for 30 min, the fluorescence was measured with 384-well microplate reader. The raw data were exported as anisotropy values and plotted against protein concentration. The data were fitted on Kaleida Graph using the following equation (**1**):

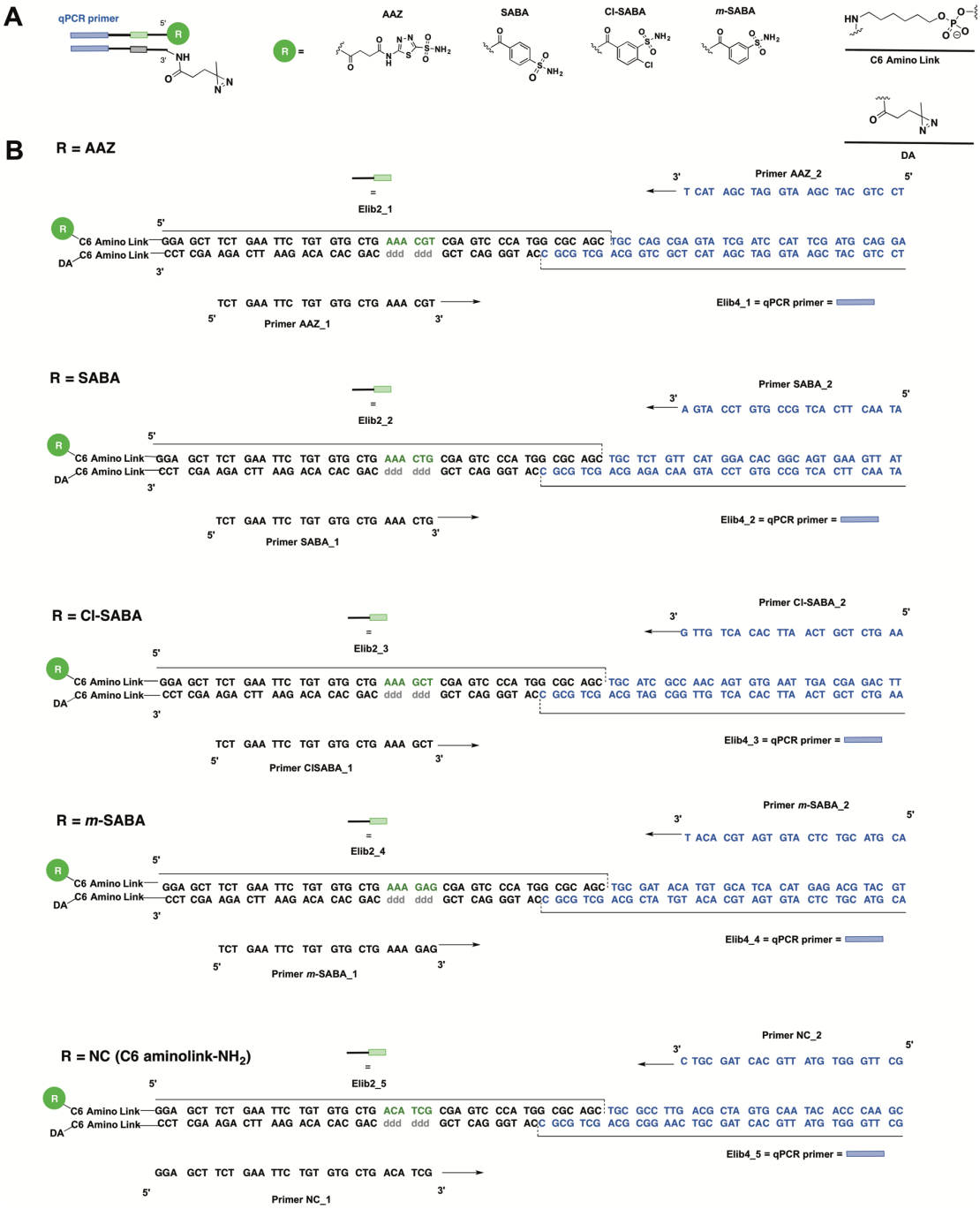
(1) $$A = \alpha \cdot L_0 + (\beta - \alpha) \cdot \frac{1}{2} \cdot \left[(P_0 + L_0 + K_d) - \sqrt{(P_0 + L_0 + K_d)^2 + 4 \cdot P_0 \cdot L_0}\right]$$

A = measured anisotropy, [P] = protein concentration, L_0 = initial ligand concentration, α,β = coefficients.

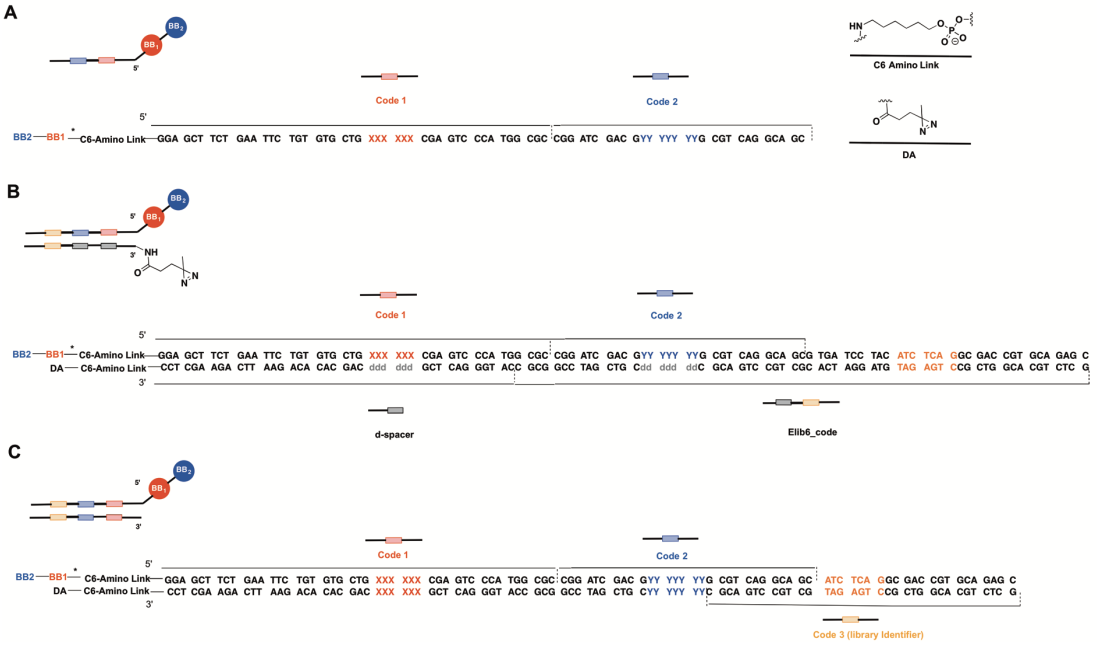
5. Supplementary Tables, Figures and Schemes



Scheme S3 Graphical representation (**A**) and oligonucleotide sequence (**B**) of the **FAM-DNA-CAIX** ligands used for the model selections against CAIX using the photocrosslinking procedure.



Scheme S4 Graphical representation (A) and oligonucleotide sequences (B) of the **DNA-CAIX** ligands used as input in selection experiments. The sequences of the primers used for the qPCR quantification are displayed in (B).



Scheme S5 A) DNA encoding scheme of the single-stranded single pharmacophore library synthetized using a splint-ligation procedure. DNA-encoded compounds are displayed at the 5' end of a 45-mer oligonucleotide carrying a sequence codon (Code 1, red), which is used to encode the first set of building blocks (BB1, red). After split and pool of the first set of conjugates, a 29-mer oligonucleotide carrying the second sequence codon (Code 2, blue) was ligated through splint-ligation, using a 14-mer DNA adapter. B) DNA-encoding scheme of the photoreactive library. C) Double-stranded DNA library analogue of the DECL above mentioned (A).

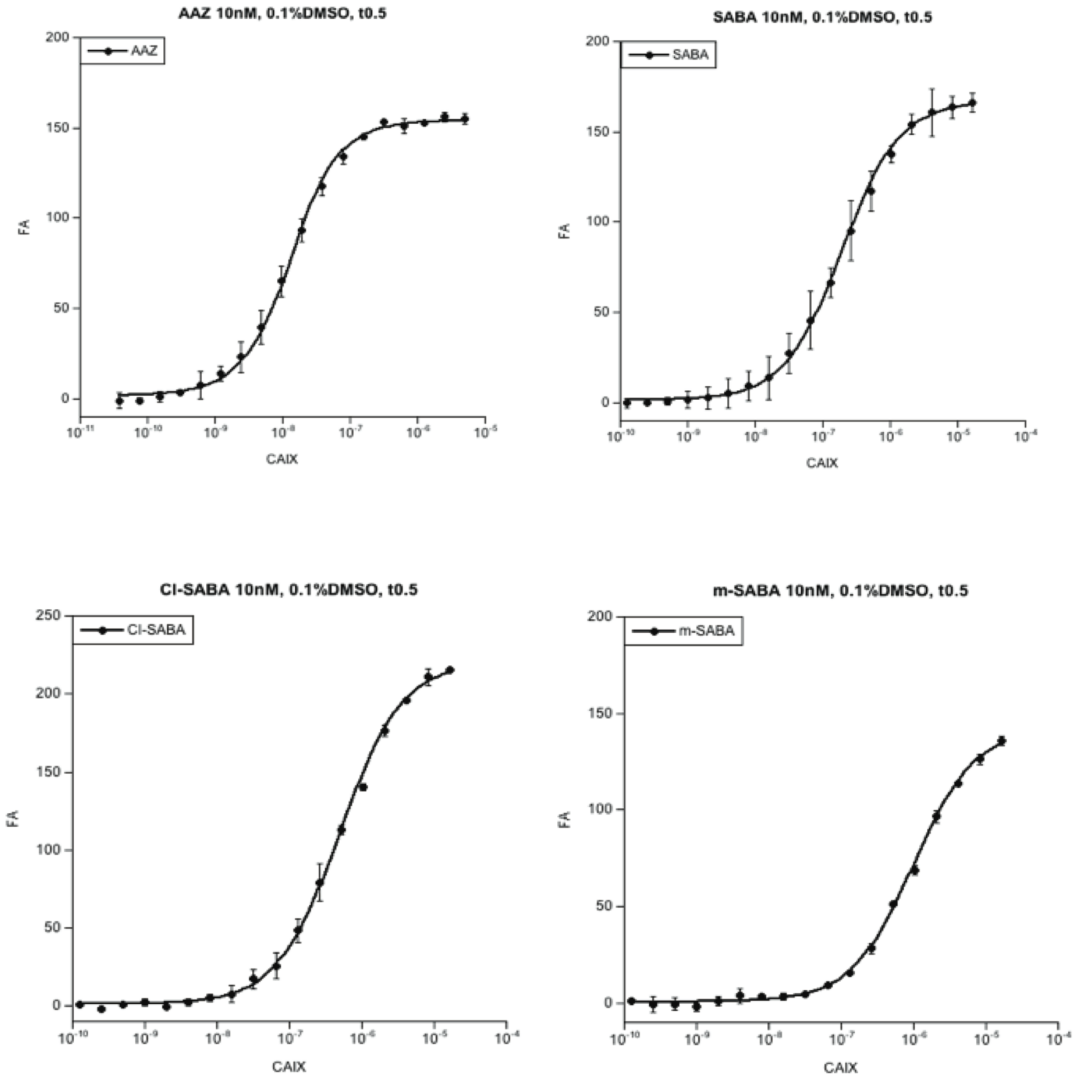


Figure S1. Fluorescence polarization measurements of the CAIX-ligands-FITC (AAZ - FITC, SABA - FITC, Cl-SABA - FITC, and m-SABA - FITC) used as positive controls in the photocrosslinking model selection setting. K_d meaurements: **AAZ - FITC** (K_d = 8.7 ± 0.5 nM; *R*= 0.99936), **SABA - FITC** (K_d = 192 ± 7 nM; *R* = 0.99956), **Cl–SABA - FITC** (K_d = 503 ± 21 nM; *R* = 0.99943), and ***m*-SABA - FITC** (K_d = 1.029 ± 0.045 μM; *R* = 0.99942).

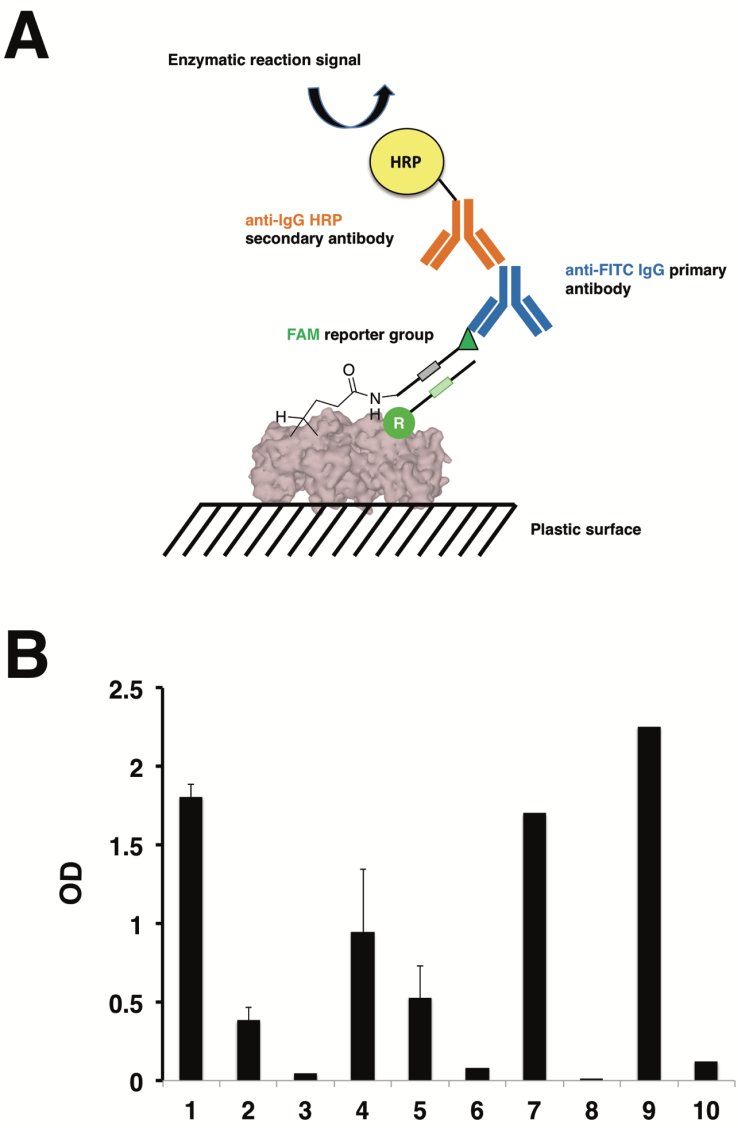
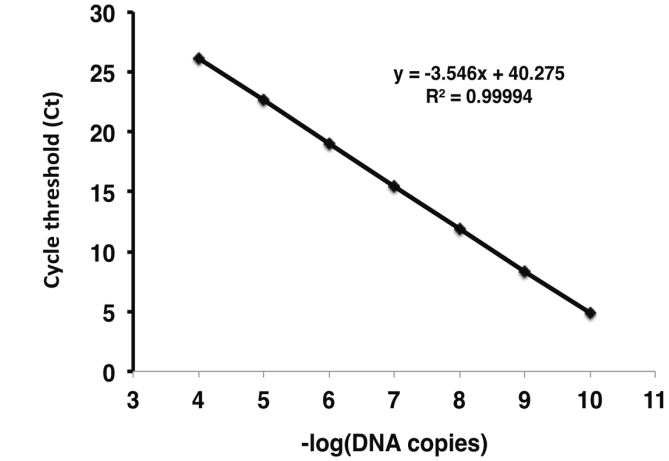


Figure S2. The formation and stability of a covalent adduct between his-CAIX and the FAM-tagged oligonucleotide carrying the photoreactive moiety using an ELISA procedure. A monoclonal antibody specific to fluorescein was used as primary reagent, eventually leading to the concentration of horseradish peroxidase onto the target protein immobilized on a solid support and to the conversion of a chromogenic substrate. Lane **1**: CAIX-AAZ-FAM-complex after UV Irradiation at 365 nm (IR) (Input: FAM-DNA-AAZ: 5 pmol) Protein concentration: 2 µM, Lane 2: no UV irradiation control (NIR). Lane 3: no protein control (NP). Lane 4: CAIX-NC-FAM-complex after UV Irradiation at 365 nm (IR) (Input: FAM-DNA-NC: 5 pmol) Protein concentration: 2 µM. Lane 5: NIR control. Lane 6: NP control. Lane 7 – 10: controls. (lane 7: his-CAIX treated with anti-6X His tag (HRP) (coating control); lane 8: his-CAIX treated with anti-FITC IgG/anti-IgG-HRP antibody. Lane 9: F8-FITC treated with anti-FITC HRP/anti-IgG-HRP. Lane 10: blank of the method.



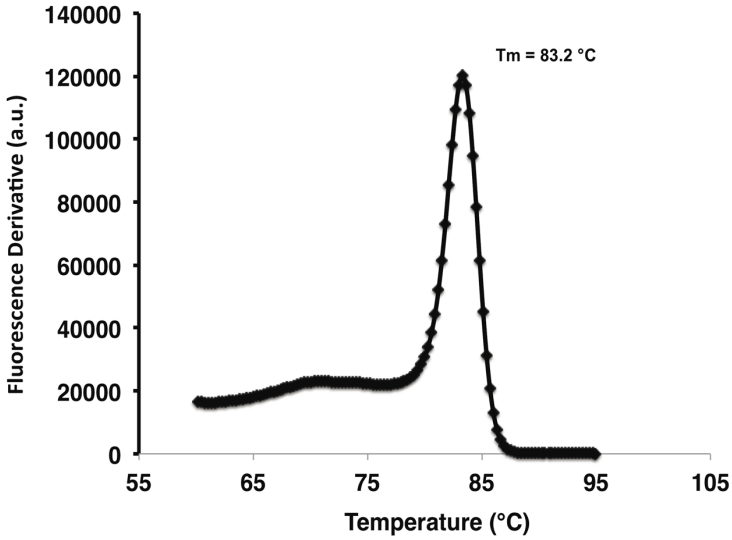


Figure S3. Representative calibration curves and melt curve profile used for the qPCR quantification of FAM-DNA CAIX ligand derivatives in model selection experiments. Calibration curve is plotted in semi-logarithmic scale.

Primer sets used for the quantification:

- FAM_DNA_CAIX ligand **Primers:** LB_FP; ESAC_q_PCR RP

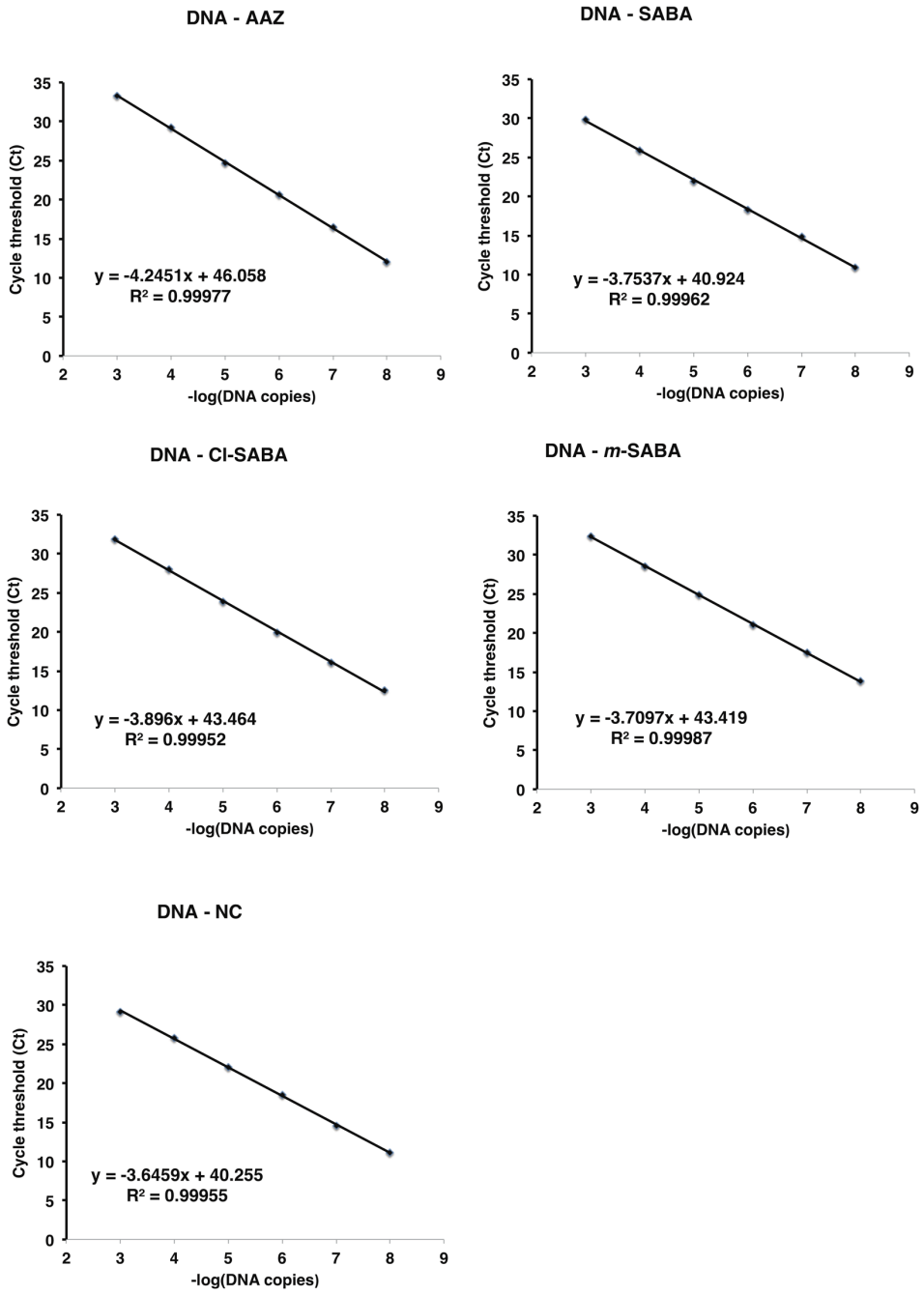


Figure S4. Representative calibration curves used for the qPCR quantification of DNA-CAIX ligand derivatives after model selection experiments. Calibration curves are plotted in semi-logarithmic scale.

Primers sets used for the quantification:

- DNA_AAZ **Primers:** PrimerAAZ_1, PrimerAAZ_2
- DNA_SABA **Primers:** PrimerSABA_1/PrimerSABA_2
- DNA_Cl_SABA **Primers:** PrimerClSABA_1/PrimerClSABA_2
- DNA_*m*-SABA **Primers:** PrimermSABA_1 / PrimermSABA_2
- DNA_NC **Primers:** PrimerNC_1 / PrimerNC_2

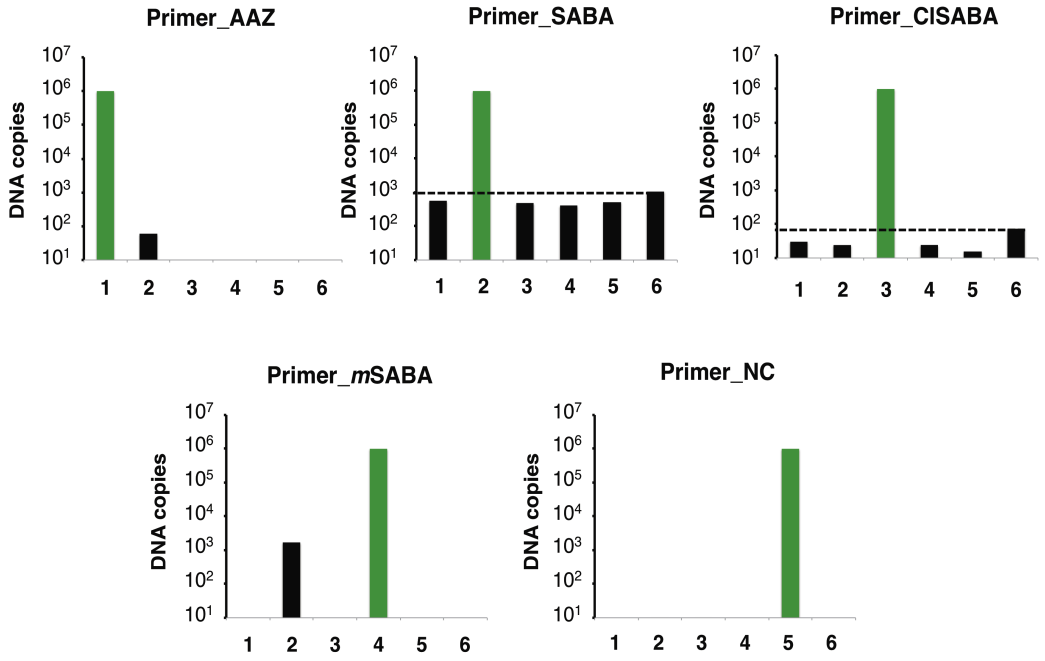
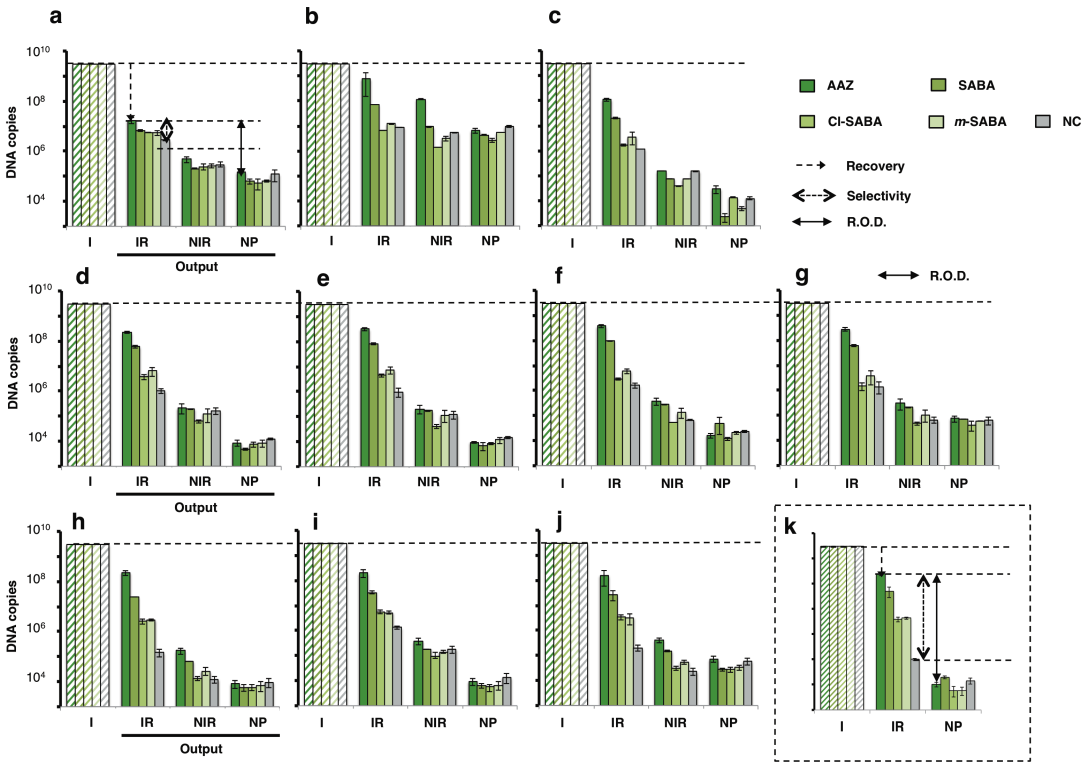


Figure S5. Orthogonality validation of qPCR primer sets used for the quantification of DNA-CAIX ligands. Each DNA-CAIX ligand (quantity: 10^6 copies of DNA molecules) was quantified with various set of primers (Primers_AAZ, Primers_SABA, Primers_Cl-SABA, Primers_mSABA, Primers_NC). 1: DNA-AAZ; 2: DNA-SABA, 3: DNA-Cl-SABA, 4: DNA-*m*SABA, 5: DNA-NC. The lowest limit of detection is also indicated (6, no template control). Results show that each ligand is efficiently amplified only by the corresponding primer set, respectively (green bar). Cross-amplification is not efficient when primer sets are not compatible with a template having non-complementary primer binding sites (black bars).



Condition	Quenching	Incubation Buffer	Ionic strength	Blocking additives
a	YES	H_2O		
b	NO	PBS Gibco 1x, pH 7.4		
c	YES	PBS Gibco 1x, pH 7.4		
d	YES	PBS Gibco 1x, pH 7.4	0.05 M NaCl	
e	YES	PBS Gibco 1x, pH 7.4	0.1 M NaCl	
f	YES	PBS Gibco 1x, pH 7.4	0.5 M NaCl	
g	YES	PBS Gibco 1x, pH 7.4	1 M NaCl	
h	YES	PBS Gibco 1x, pH 7.4		0.3 mg/ml HS
i	YES	PBS Gibco 1x, pH 7.4		0.1% Tween-20
j	YES	PBS Gibco 1x, pH 7.4		4% non-fat dry milk
k	YES	PBS Gibco 1x, pH 7.4	0.1 M NaCl	0.3 mg/ml HS

Figure S6. A) qPCR quantification of model selection experiments performed using a pool of DNA-CAIX ligands as input. I = Input, IR = Protein Selection with UV-irradiation, NIR = Protein Selection, no UV-irradiation NP = Selection in the absence of protein. Different experimental conditions were evaluated to optimize the recovery of DNA-CAIX ligands, selectivity of the screening procedure and R.O.D. (range of detection). Results are plot in logarithmic scale. Conditions b-c. Effect of denaturing quenching conditions adding 5% SDS in the incubation buffer after UV-irradiation. Conditions d-g. Effect of ionic strength, adding various quantities of NaCl(aq) in the incubation buffer. Conditions h-j. Effect of blocking additives in the incubation buffer (HS: Herring Sperm) k. Optimized conditions for the screening methodology.

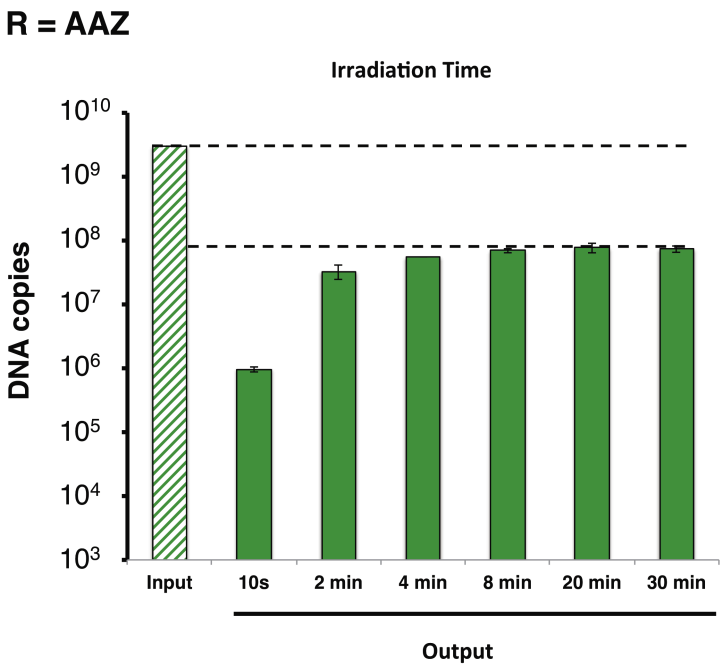


Figure S7. Evaluation of the optimal UV irradiation time for the photocrosslinking reaction. Model selections were performed against his-CAIX (2 μM). Photoreactive DNA-AAZ ligand (3 x 10^9 copies of DNA molecules) was used as input. The incubation mixtures of DNA-AAZ and CAIX were irradiated for different time frames. After denaturation of the protein, capture on magnetic beads/washes/elution, the amount of DNA molecules in the output solution was analyzed by qPCR. The quantity of the DNA after selection is plotted in logarithmic scale (green bars). After 20 minutes, no improvement of the recovery yield was observed.

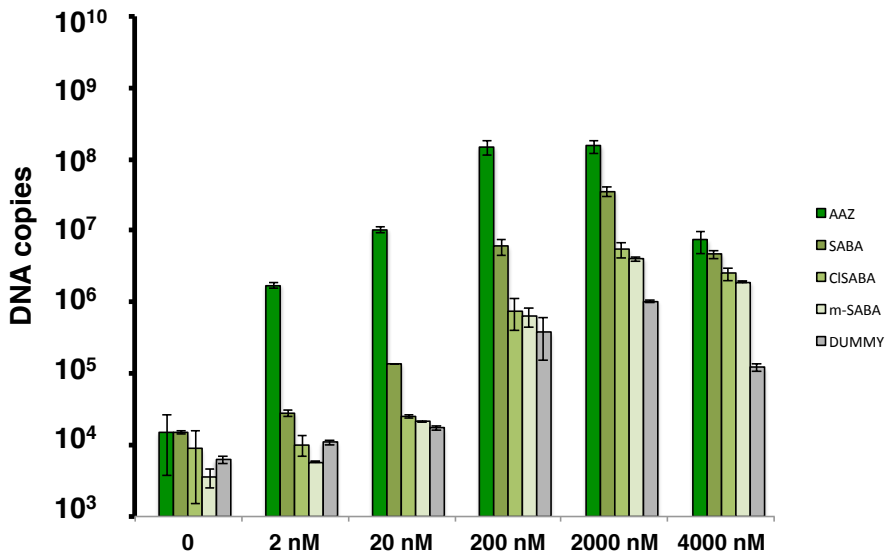


Figure S8. Evaluation of the recovery yield and selectivity after model selections against his-CAIX, used at different protein concentrations. (2 nM – 20 nM – 200 nM – 2000 nM – 4000 nM). A pool of the photoreactive DNA-CAIX ligands (3 x 10^9 copies for each conjugate) was used as input. The pool of conjugates was incubated with different concentration of protein before UV irradiation. After denaturation of the protein, capture on magnetic beads/washes/elution, the amount of DNA molecules in the output solution was analyzed by qPCR. The quantity of the DNA after selection is plotted in logarithmic scale (green bars). Selection against empty beads (no protein, 0) was included as a negative control. We observed a correlation between recovery and binding constants of the ligands. The best recovery yields were obtained with a protein concentration of 2 μM. At 4 μM protein concentration, we observed a decrease of the recovery yield, probably determined by oversaturation of the beads (hook effect).

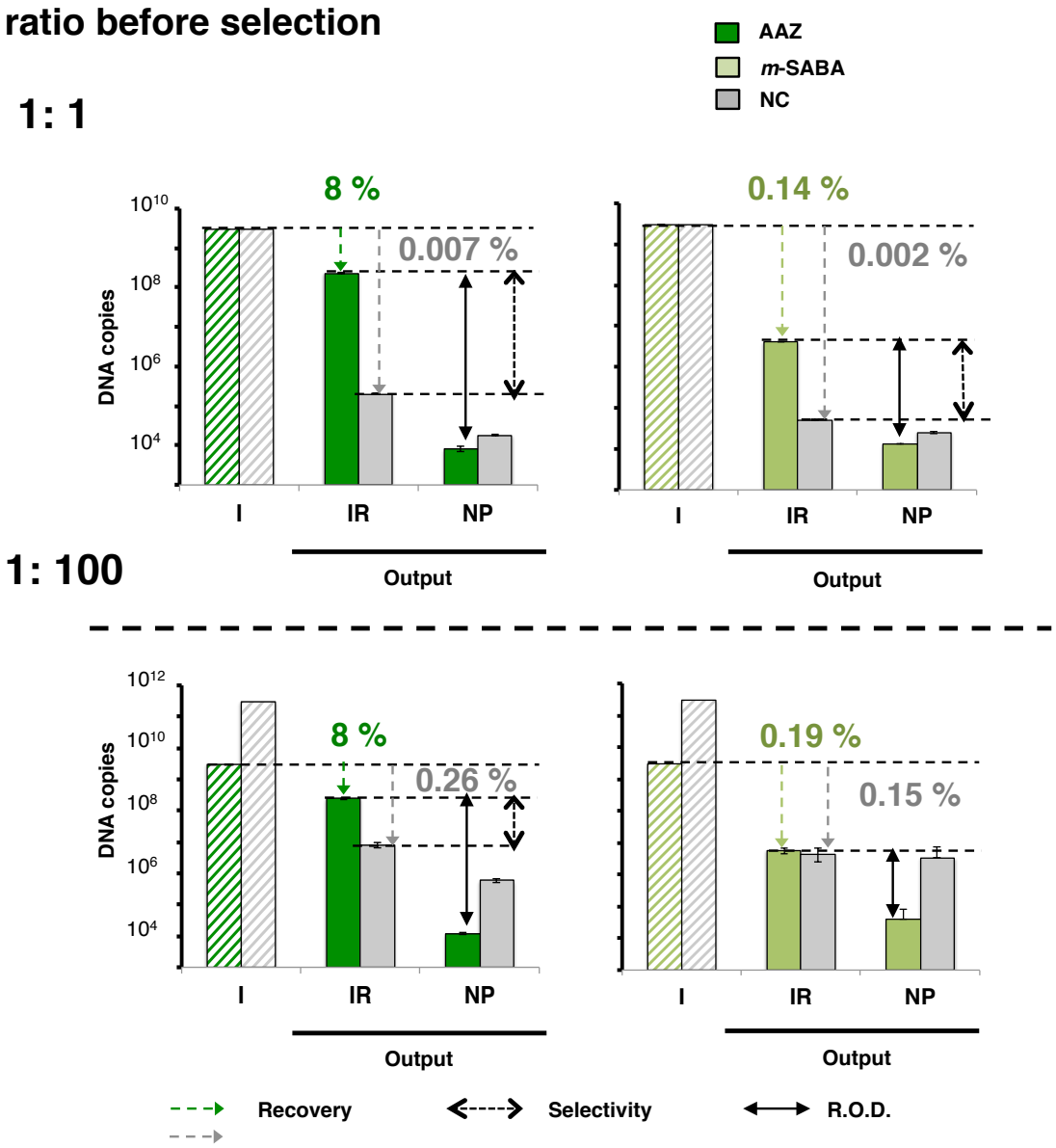


Figure S9. Experimental data of the set of experiments described in Figure 6.

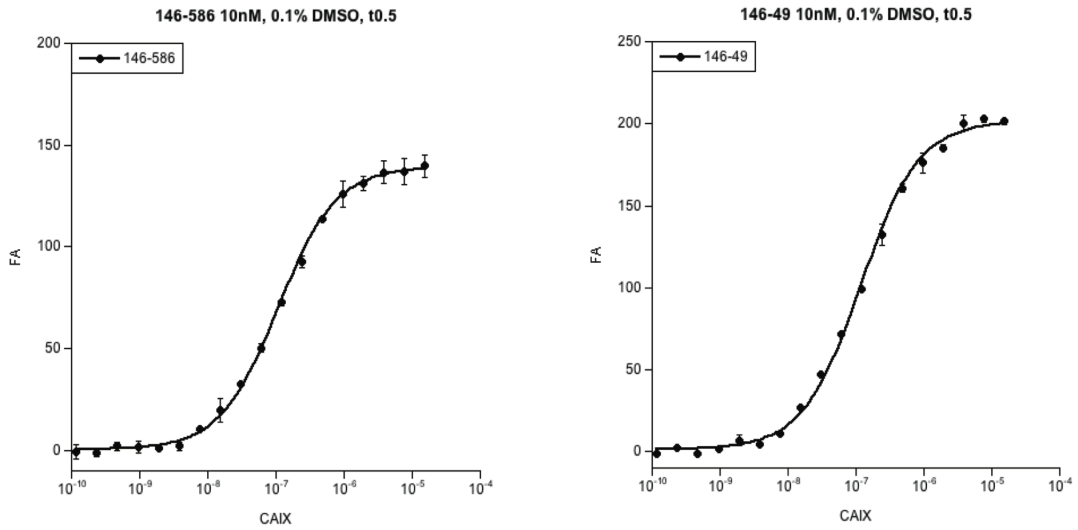
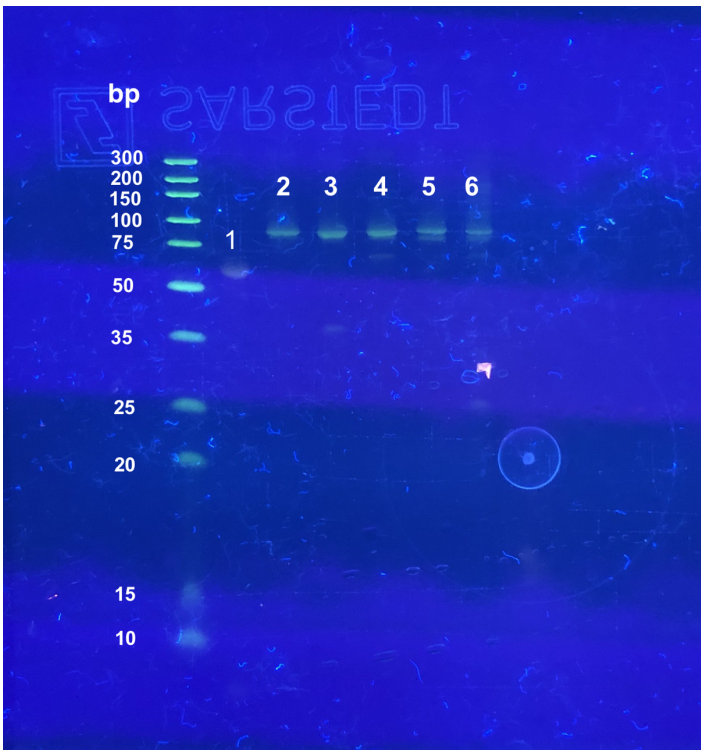


Figure S10. Fluorescence polarization measurements of **146_586 FITC** and **146_49 FITC** derivatives. K_d measurements: **146/586_FITC**: Kd = 105 ± 4 nM**,** *R*= 0.99951**, 146/49_FITC:** Kd= 116 ± 6 nM *R*= 0.99912.

A



B

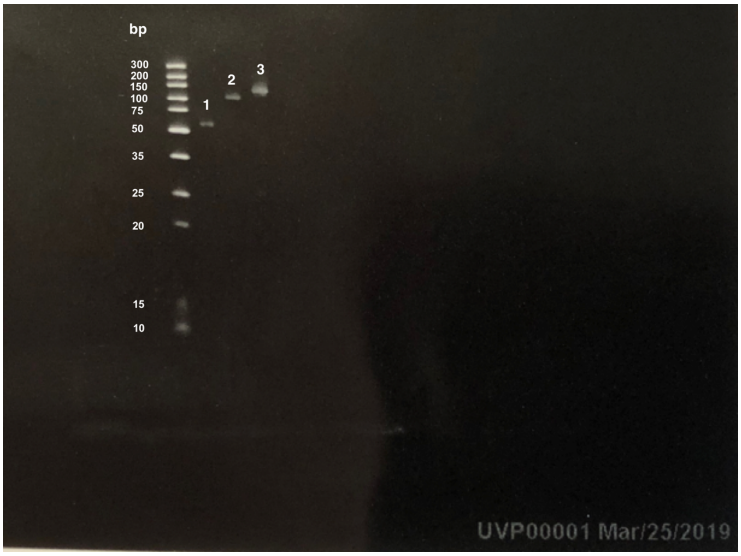
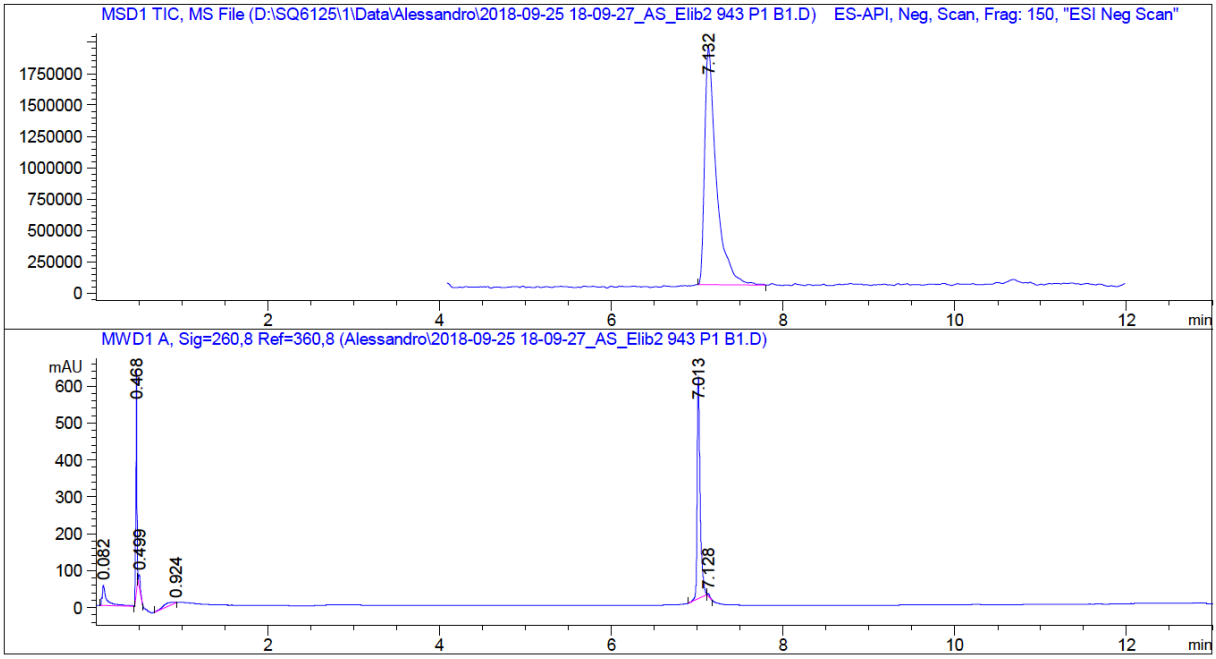
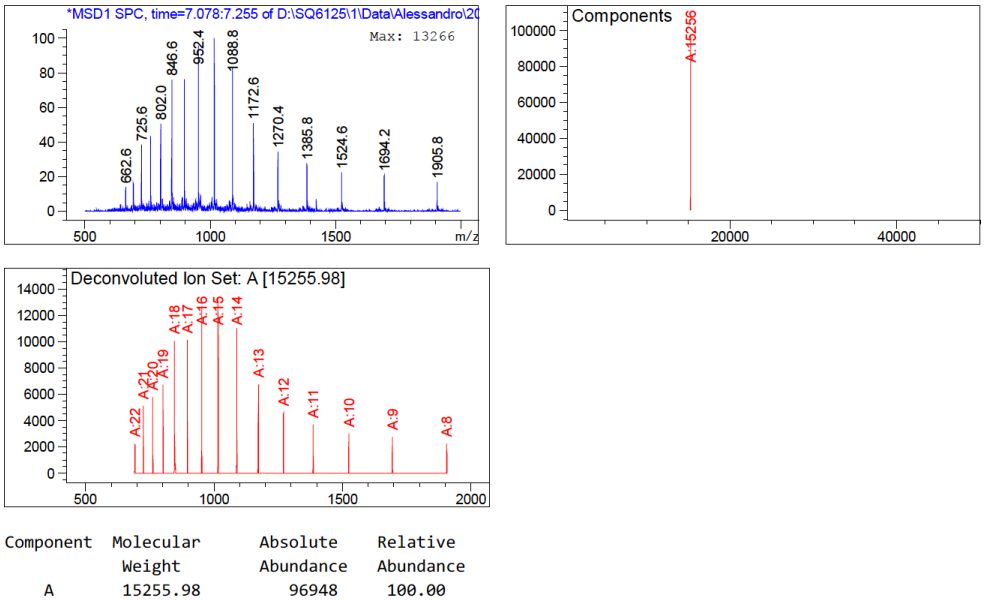


Figure S11. (A) Representative 15% TBE-Urea gel electrophoresis of the model DNA-CAIX ligand photoreactive derivatives. Lane 1: DA_Elib4_X before Klenow fill-in. Lanes 2-6: DNA-CAIX ligand derivatives after Klenow polymerization (B) Representative 15% TBE-Urea gel Electrophoresis analysis of the photocrosslinking library. Lane 1: Single-stranded single pharmacophore library Lane 2: Photoreactive strand (DA_Elib6) Lane 3: Photoreactive library after Klenow fill-in polymerization.

6. Characterization

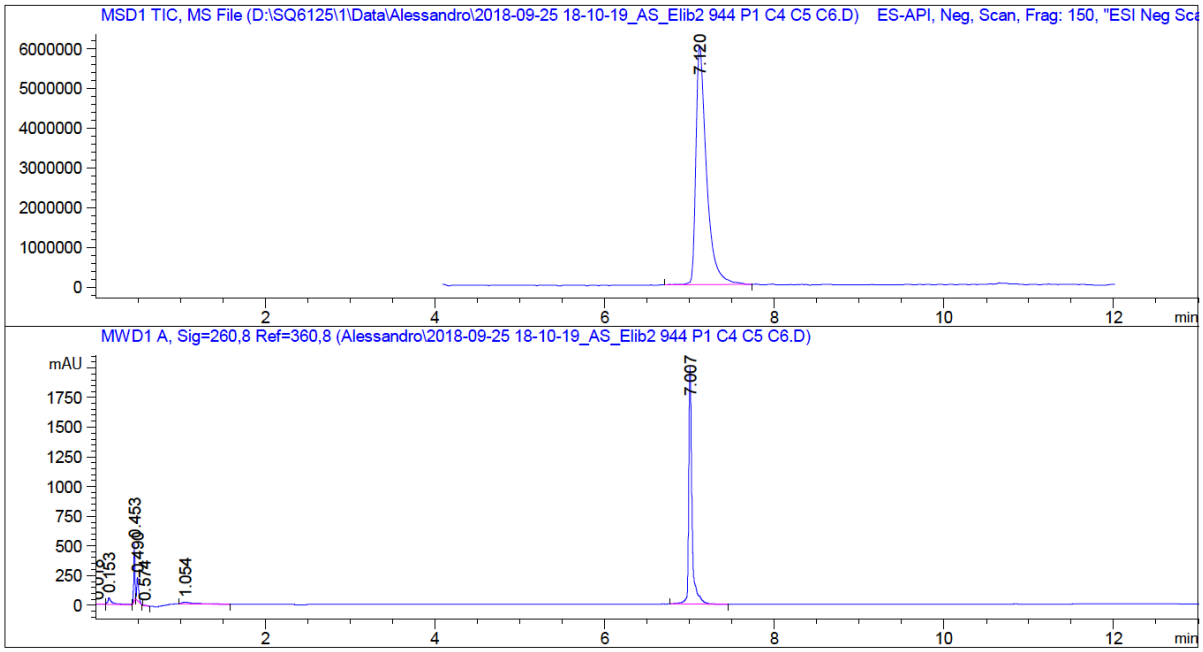
6.1. LC-MS analysis

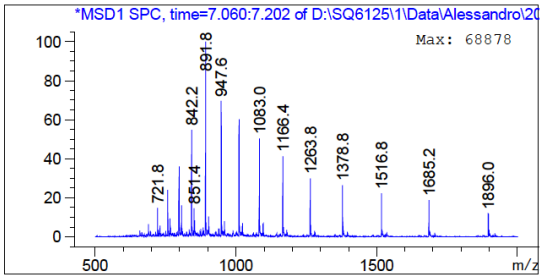


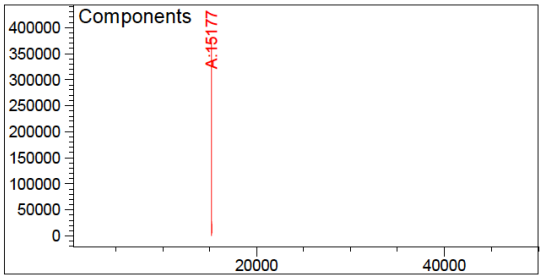


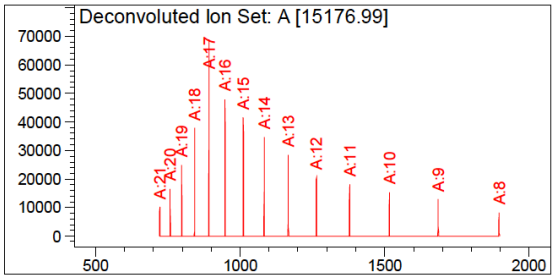
Component	Molecular Weight	Absolute Abundance	Relative Abundance
A	15255.98	96948	100.00

Figure S12. LC-ESI-MS chromatogram of compound **AAZ_Elib2_1**, t_R: 7.132 min (DAD), 7.013 min (TIC). Expected mass 15256.87, observed 15255.98.



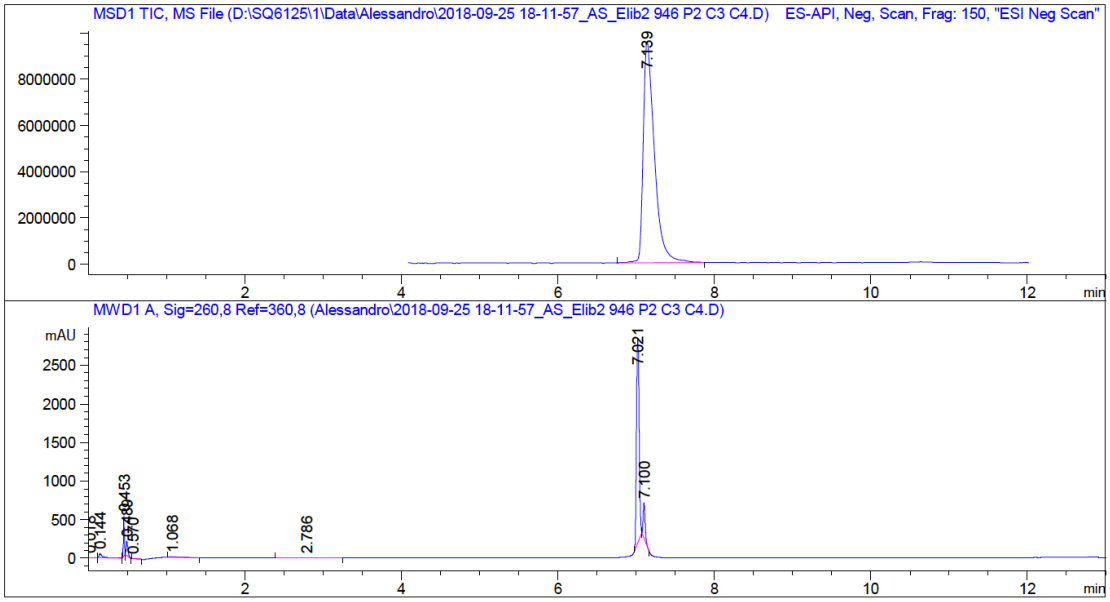


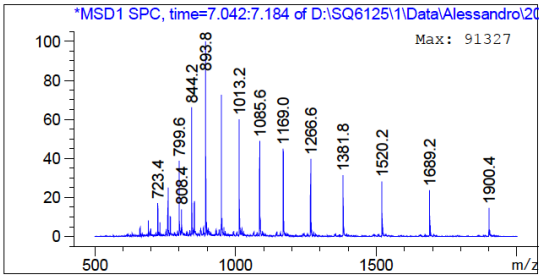


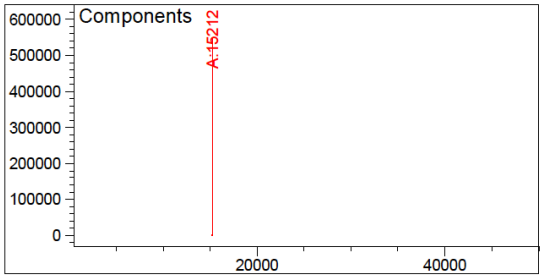


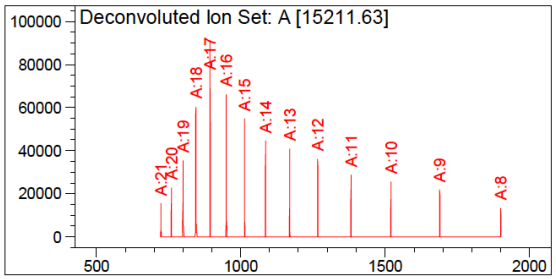
Component	Molecular Weight	Absolute Abundance	Relative Abundance
A	15176.99	378050	100.00

Figure S13. LC-ESI-MS chromatogram of compound **SABA_Elib2_2**, t_R: 7.120 min (DAD), 7.007 min (TIC). Expected mass 15176.89, observed 15176.99.



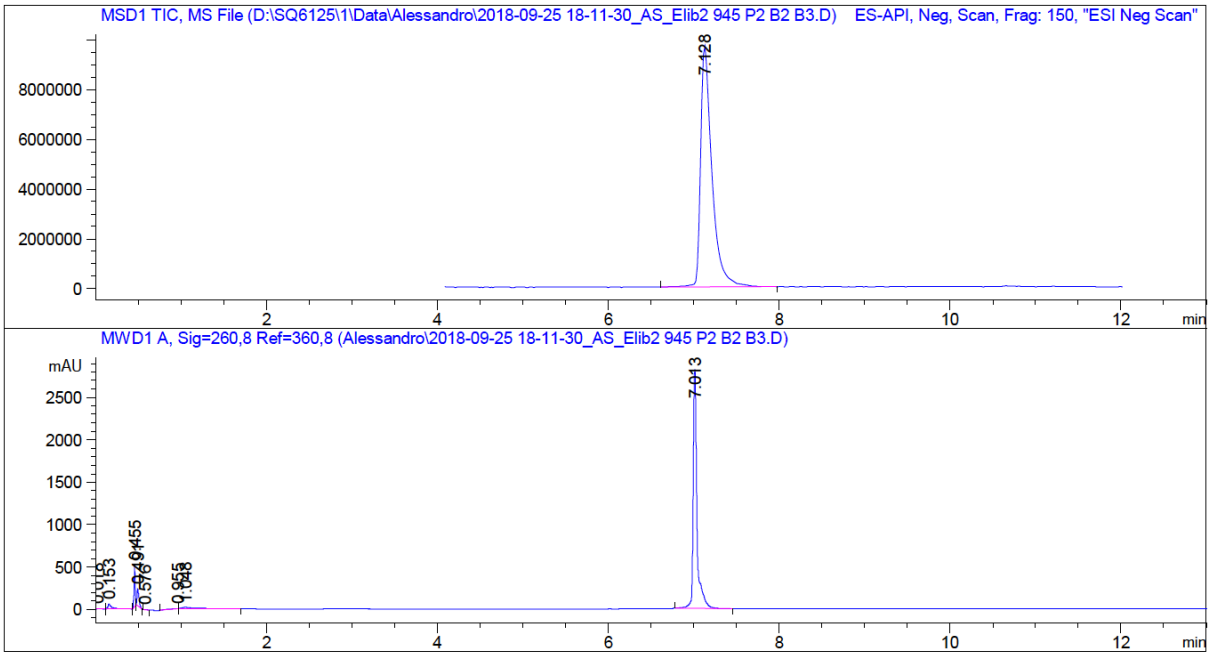


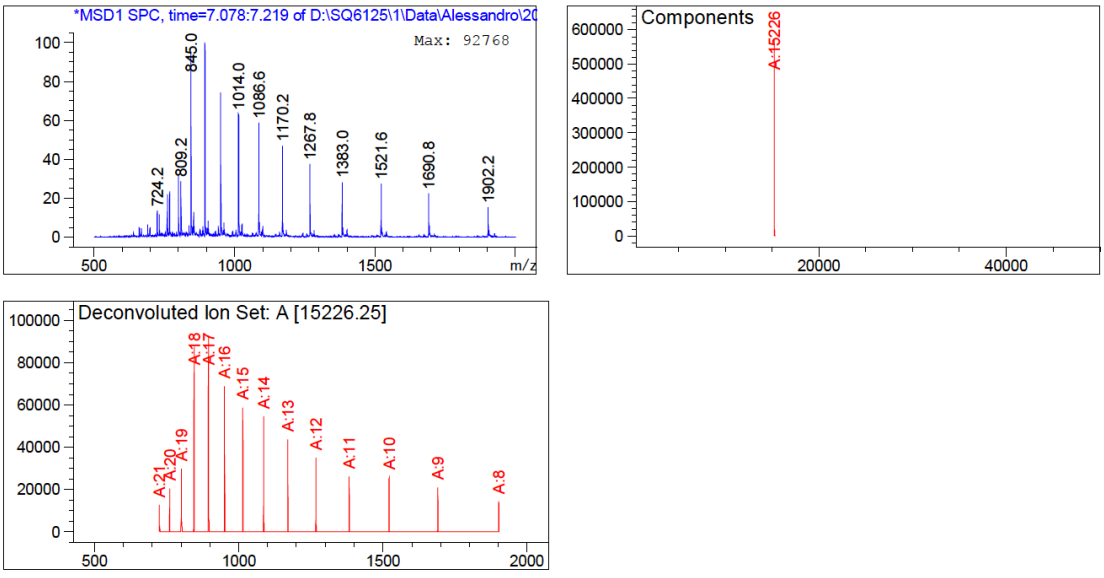




Component	Molecular Weight	Absolute Abundance	Relative Abundance
A	15211.63	546442	100.00

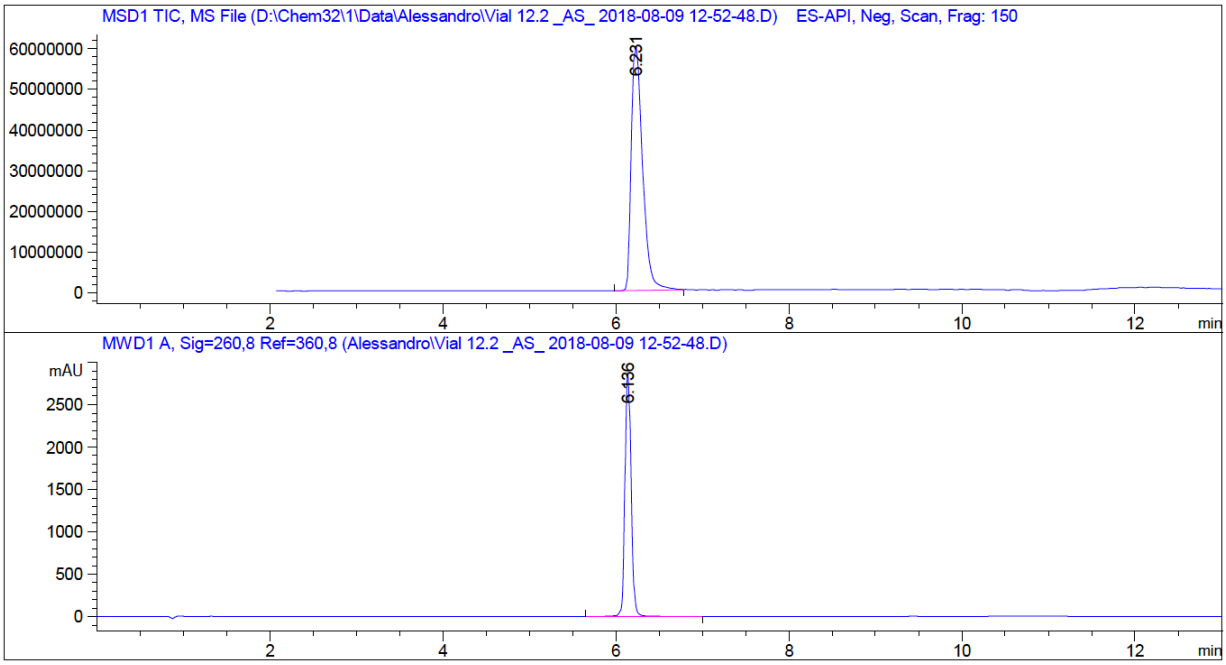
Figure S14. LC-ESI-MS chromatogram of compound **Cl-SABA_Elib2_3**, t_R: 7.139 min (DAD), 7.021 min (TIC). Expected mass 15212.24, observed 15211.63.

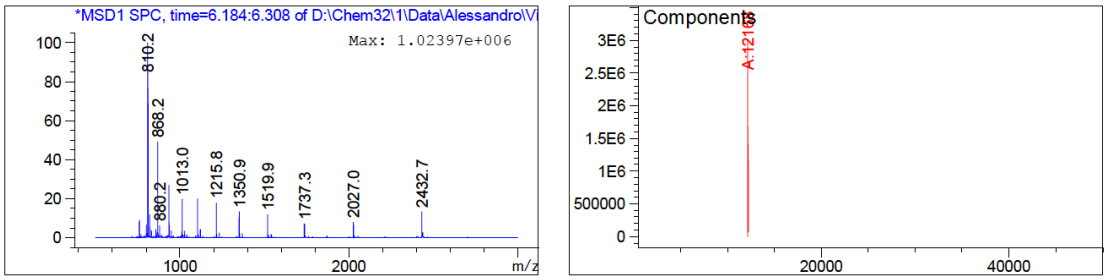


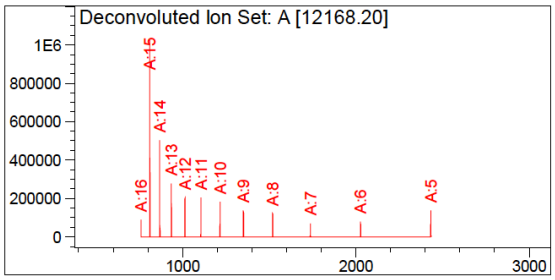


Component	Molecular Weight	Absolute Abundance	Relative Abundance
A	15226.25	570387	100.00

Figure S15 LC-ESI-MS chromatogram of compound ***m*-SABA_Elib2_4**, t_R: 7.128 min (DAD), 7.013 min (TIC). Expected mass 15226.9, observed 15226.25.

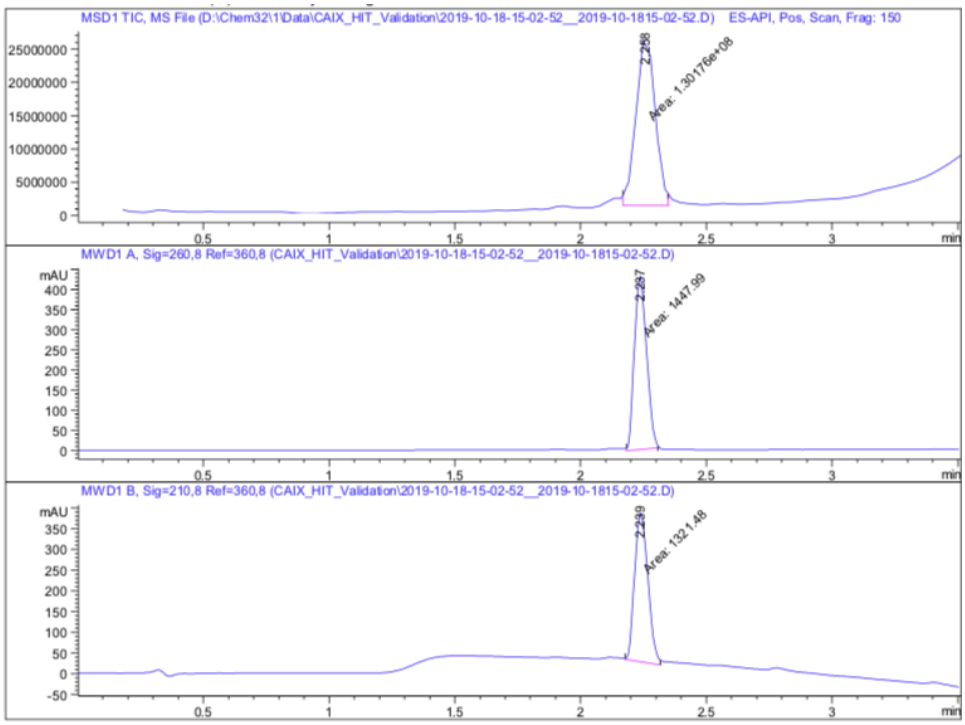






Component	Molecular Weight	Absolute Abundance	Relative Abundance
A	12168.20	3004967	100.00

Figure S16. LC-ESI-MS chromatogram of compound **DA_d-spacer**, t_R: 6.231 min (DAD), 6.136 min (TIC). Expected mass 12168, observed 12168.20.



Retention Time (MS)	MS Area	Mol. Weight or Ion
2.258	130175896	898.10 I
		802.00 I
		801.00 I
		800.00 I

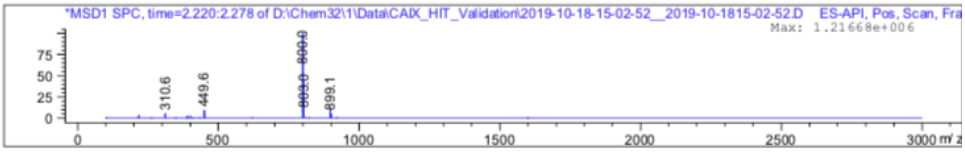
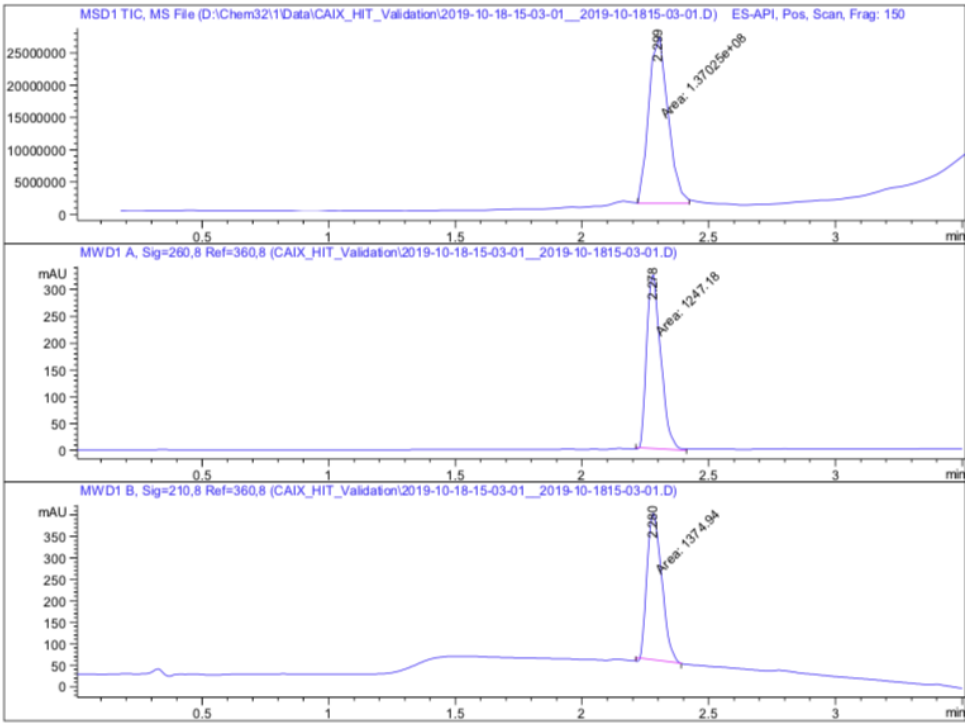


Figure S17. LC-ESI-MS chromatogram of compound **AAZ-FITC**, t_R: 2.237 min (DAD), 2.258 min (TIC).



Retention Time (MS)	MS Area	Mol. Weight or Ion
2.299	137024816	723.00 I
		722.10 I
		721.10 I
		435.00 I
		390.00 I

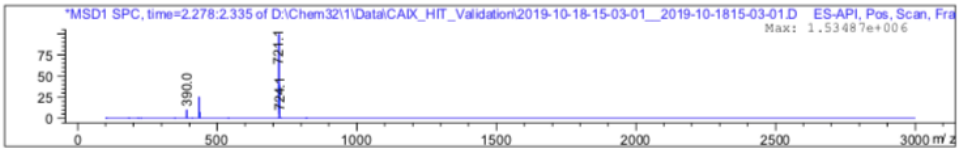
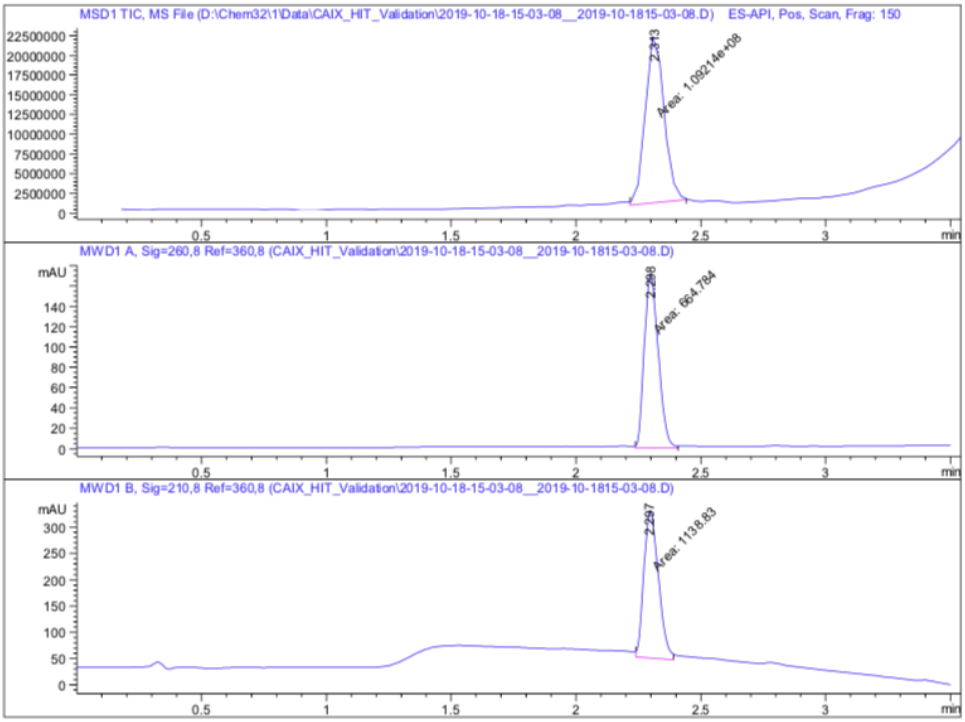


Figure S18. LC-ESI-MS chromatogram of compound **SABA-FITC**, t_R: 2.278 min (DAD), 2.299 min (TIC).



Retention Time (MS)	MS Area	Mol. Weight or Ion
2.313	109214000	723.00 I
		722.10 I
		721.10 I
		436.10 I
		435.00 I
		390.00 I

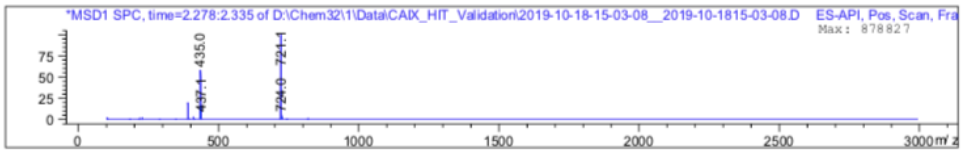
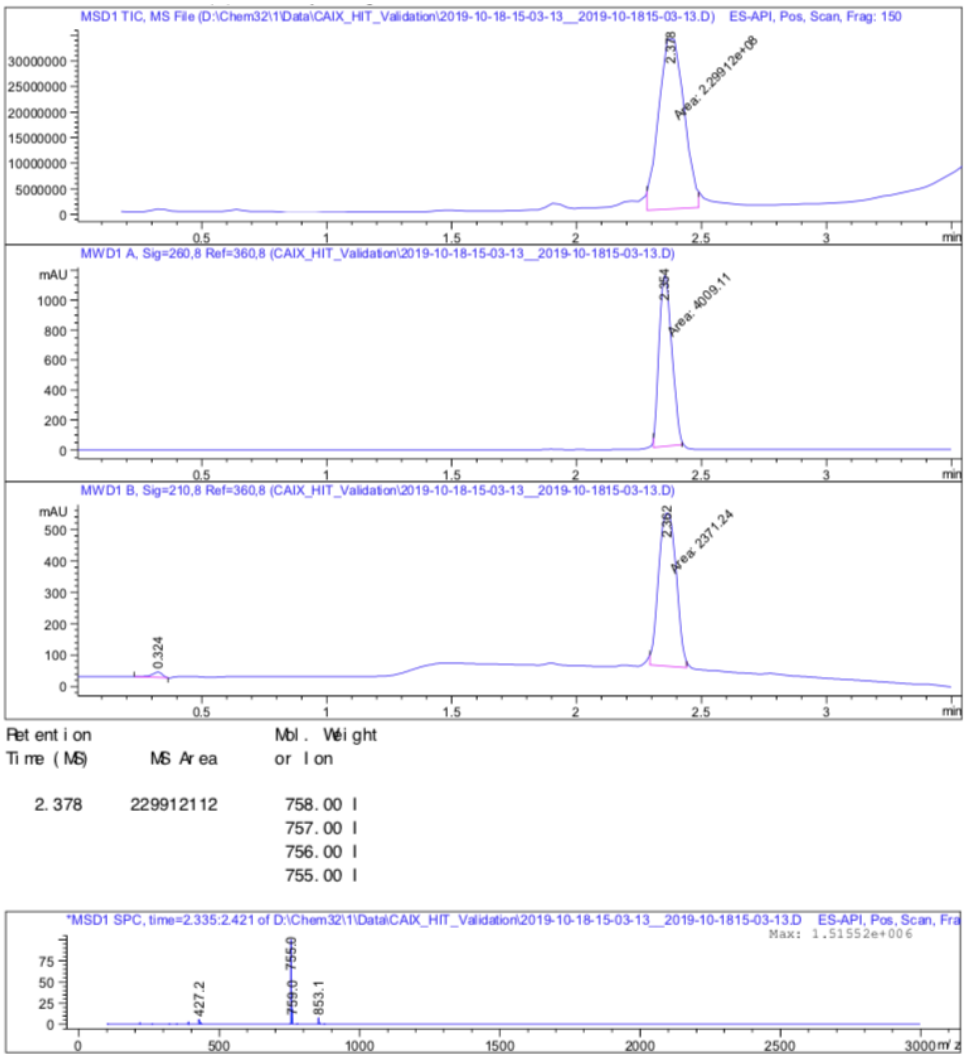
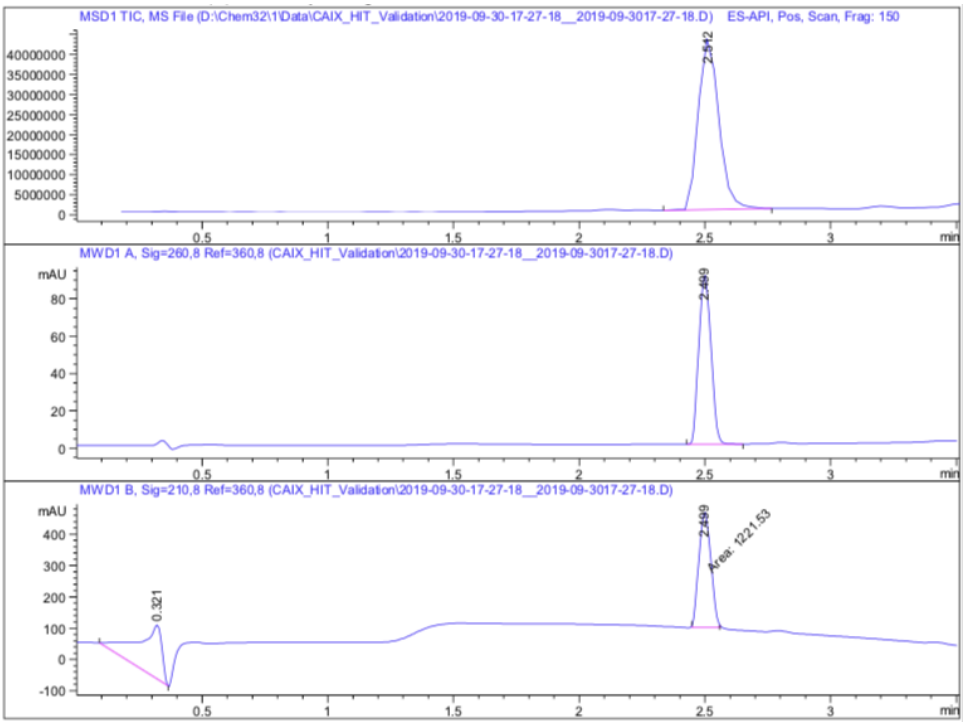


Figure S19. LC-ESI-MS chromatogram of compound ***m*-SABA-FITC**, t_R: 2.298 min (DAD), 2.313 min (TIC).



Retention Time (MS)	MS Area	Mol. Weight or Ion
2.378	229912112	758.00 I
		757.00 I
		756.00 I
		755.00 I

Figure S20. LC-ESI-MS chromatogram of compound **Cl-SABA-FITC**, t_R: 2.354 min (DAD), 2.378 min (TIC).



Retention Time (MS)	MS Area	Mol. Weight or Ion
2.512	241154384	959.00 I
		957.00 I
		837.10 I
		835.10 I
		490.10 I
		370.00 I
		369.10 I
		368.10 I

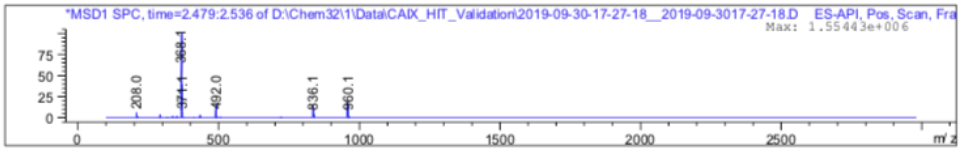
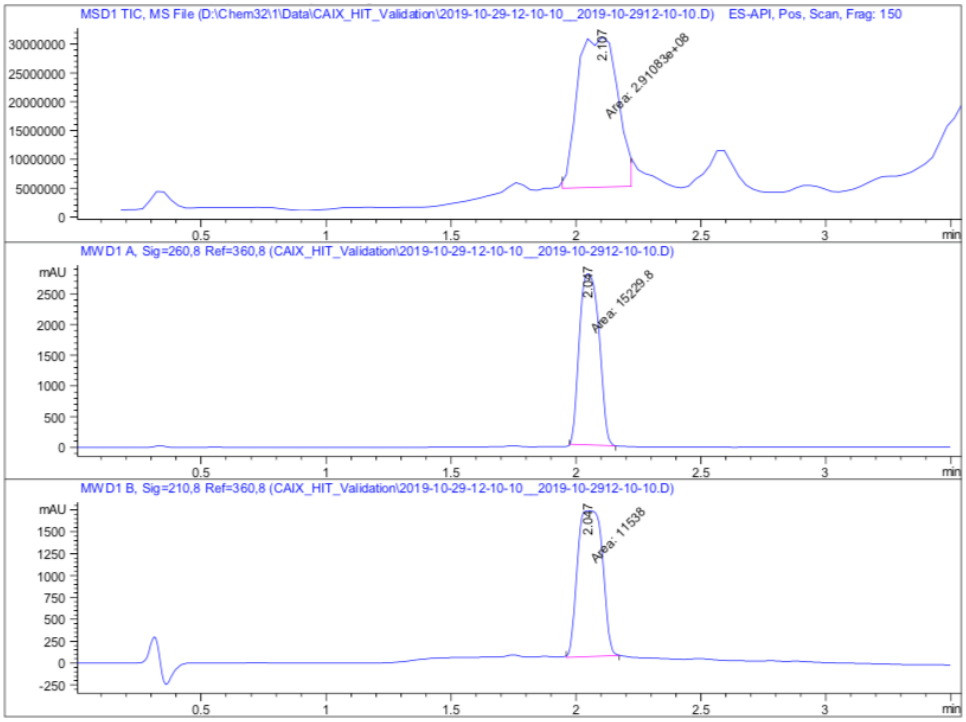


Figure S21. LC-ESI-MS chromatogram of compound **S3**, t_R: 2.499 min (DAD), 2.512 min (TIC).



Retention Time (MS)	MS Area	Mol. Weight or Ion
2.107	291082944	1104.90 I
		1103.90 I
		1102.90 I
		1101.90 I
		1100.90 I
		554.00 I
		553.00 I
		552.00 I
		551.00 I

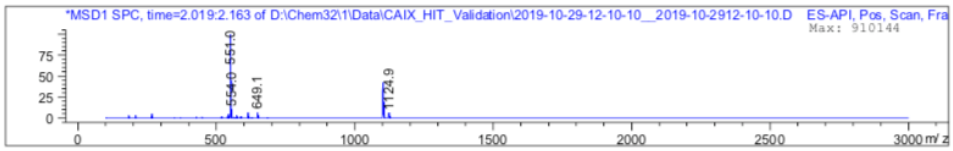
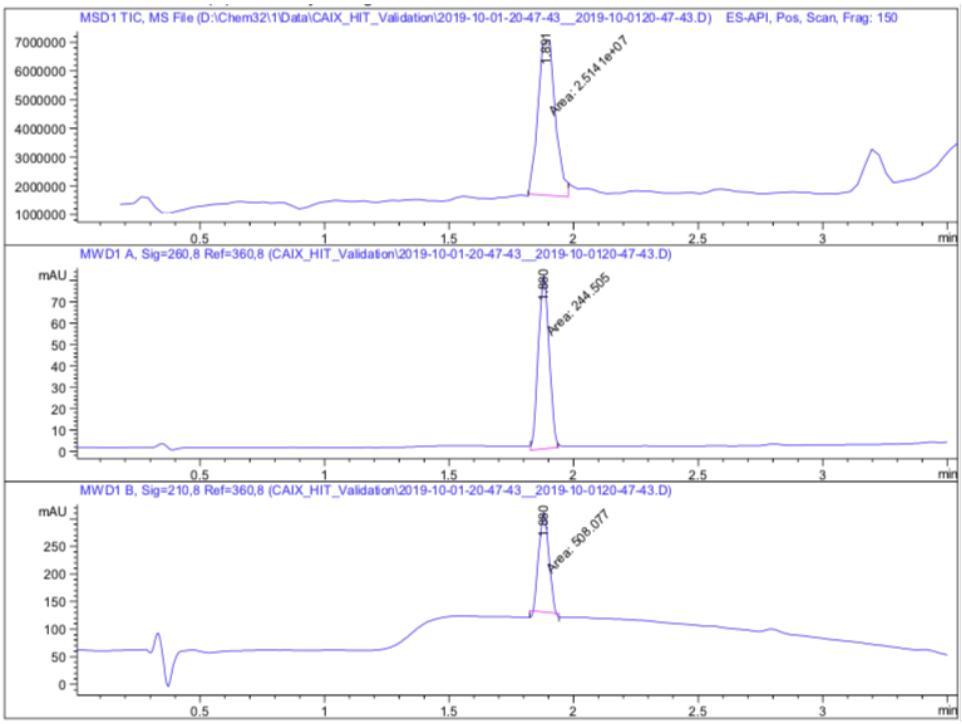


Figure S22. LC-ESI-MS chromatogram of compound **S6**, t_R: 2.047 min (DAD), 2.107 min (TIC).



Retention Time (MS)	MS Area	Mol. Weight or Ion	
1.891	25141026	1074.90	I
		1072.90	I
		539.00	I
		538.00	I
		537.00	I
		130.20	I

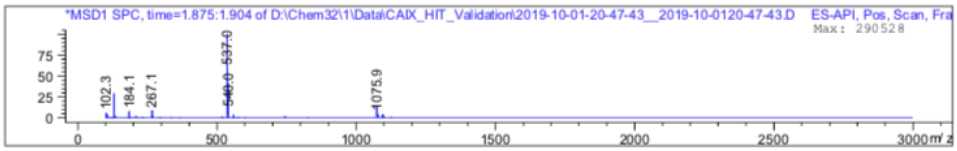
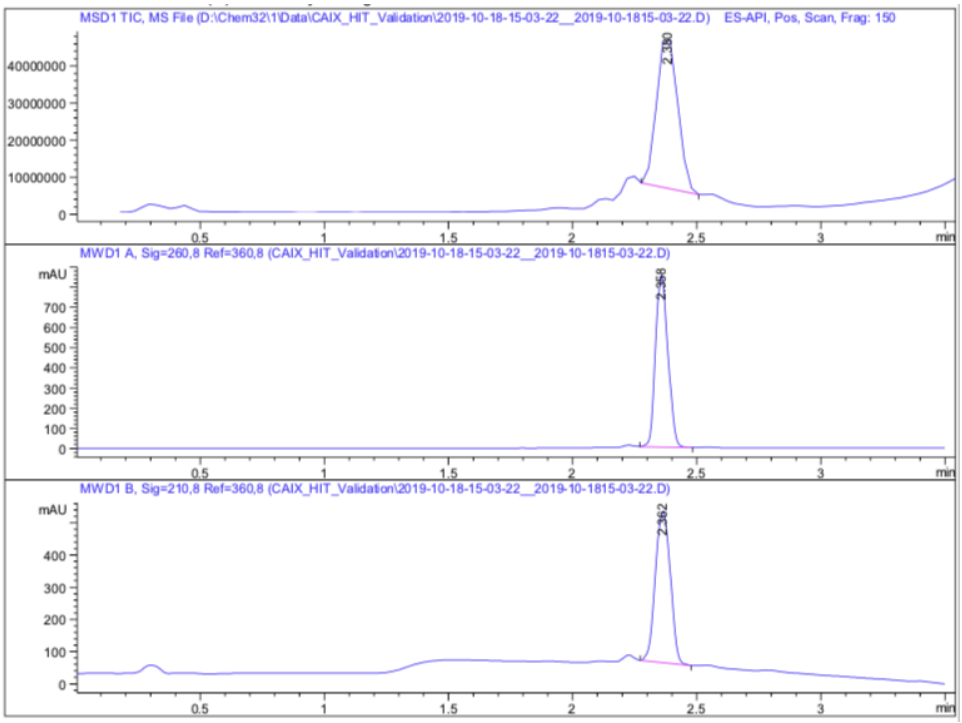


Figure S23. LC-ESI-MS chromatogram of compound **146/586**, t_R: 1.880 min (DAD), 1.891 min (TIC).



Retention Time (MS)	MS Area	Mol. Weight or Ion
2.380	239737968	1155.10 I
		1154.10 I
		1153.10 I
		1152.10 I
		1058.10 I
		1057.00 I
		1056.00 I
		1055.10 I
		1054.10 I
		627.30 I
		626.70 I
		578.20 I
		577.70 I
		435.10 I
		390.00 I

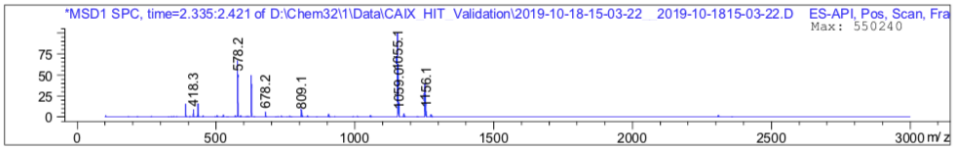
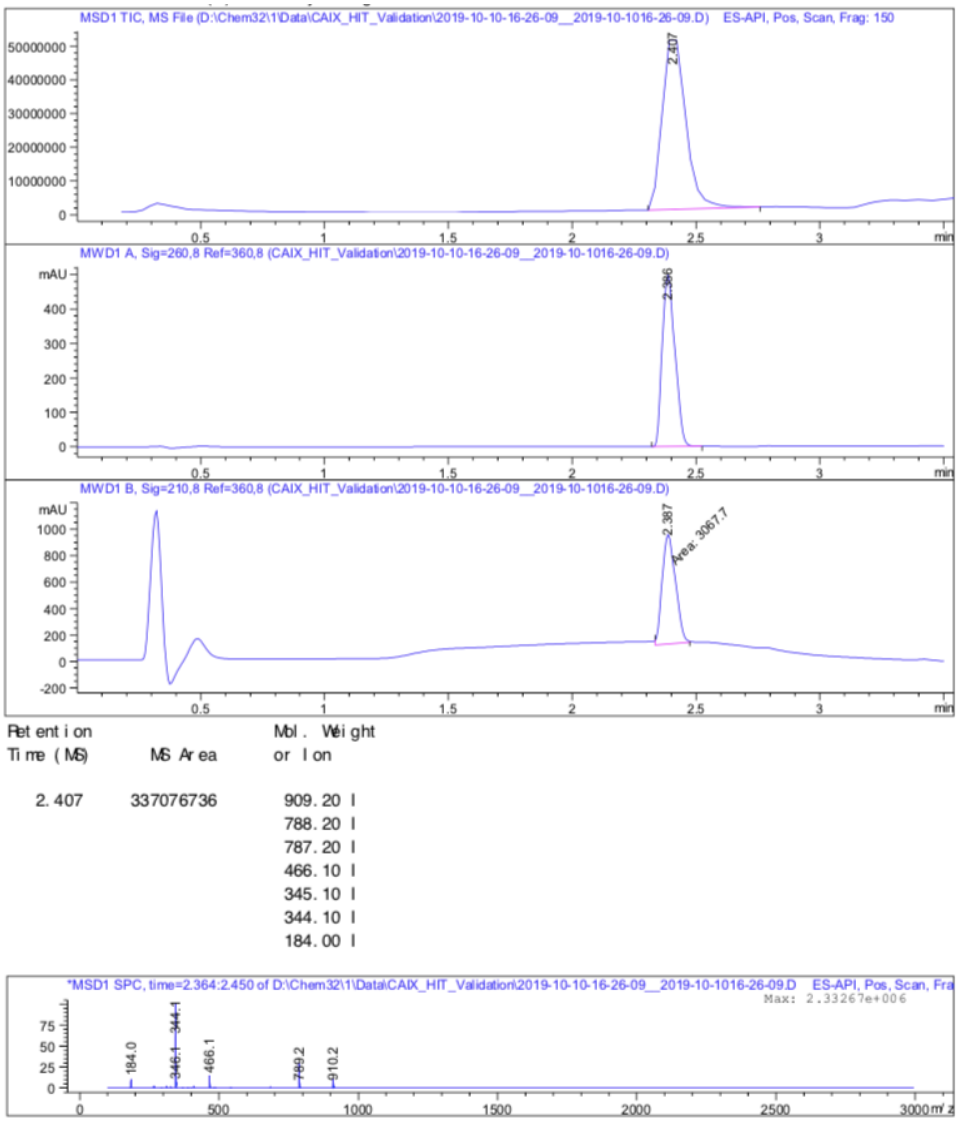
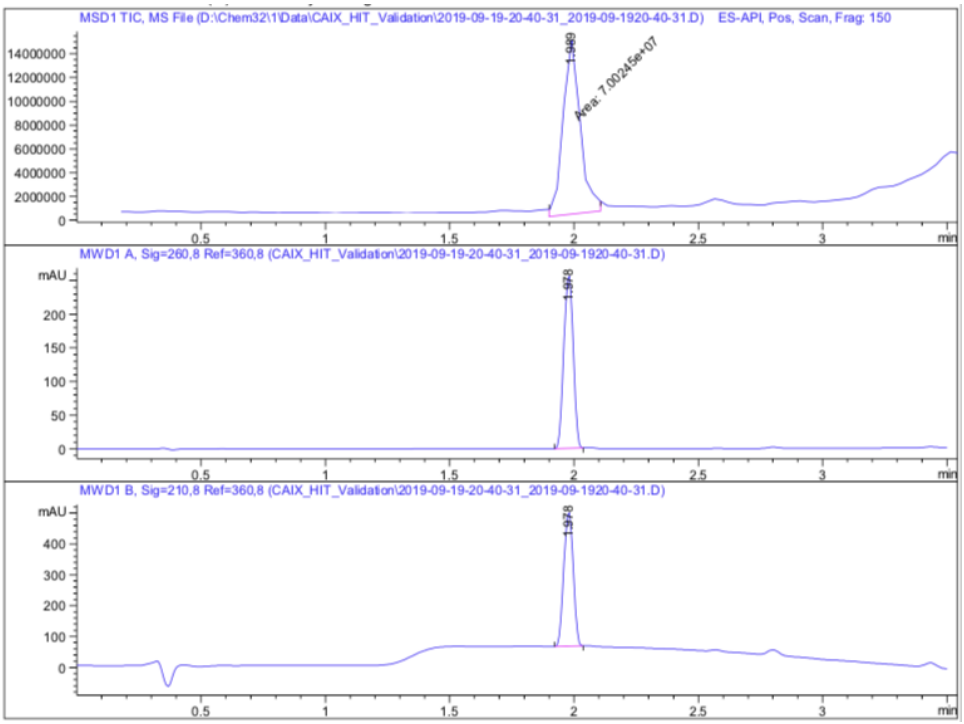


Figure S24. LC-ESI-MS chromatogram of compound **146/586-FITC**, t_R: 2.358 min (DAD), 2.380 min (TIC).



Retention Time (MS)	MS Area	Mol. Weight ion
2.407	337076736	909.20 I
		788.20 I
		787.20 I
		466.10 I
		345.10 I
		344.10 I
		184.00 I

Figure S25. LC-ESI-MS chromatogram of compound **S7**, t_R: 2.358 min (DAD), 2.380 min (TIC).



Retention Time (MS)	MS Area	Mol. Weight ion
1.989	70024544	1053.40 I
		1052.40 I
		528.80 I
		527.80 I
		526.80 I

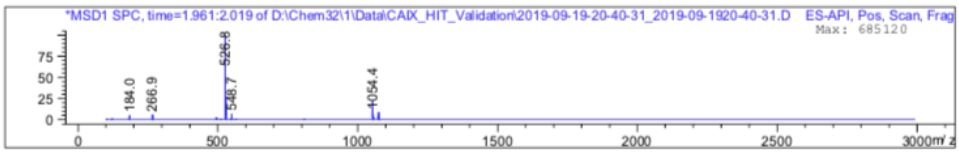
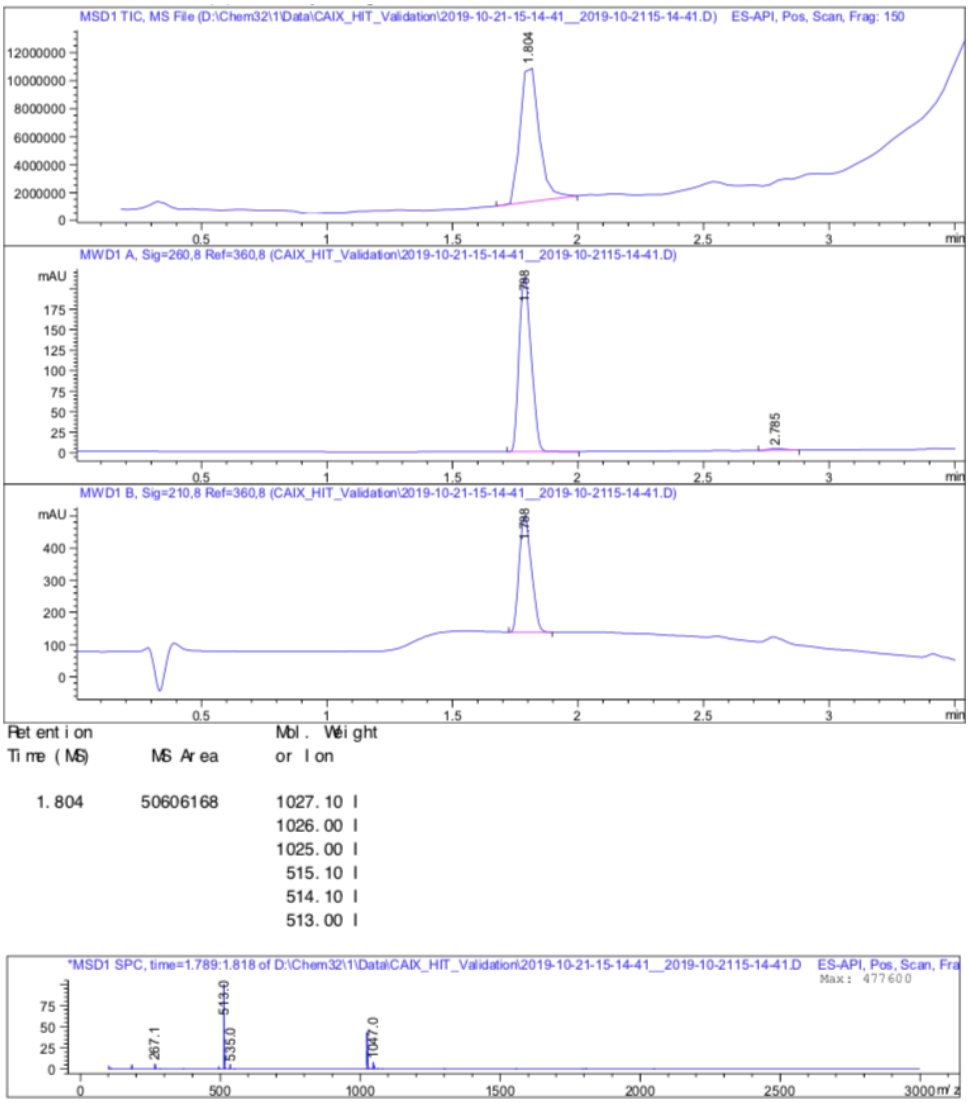
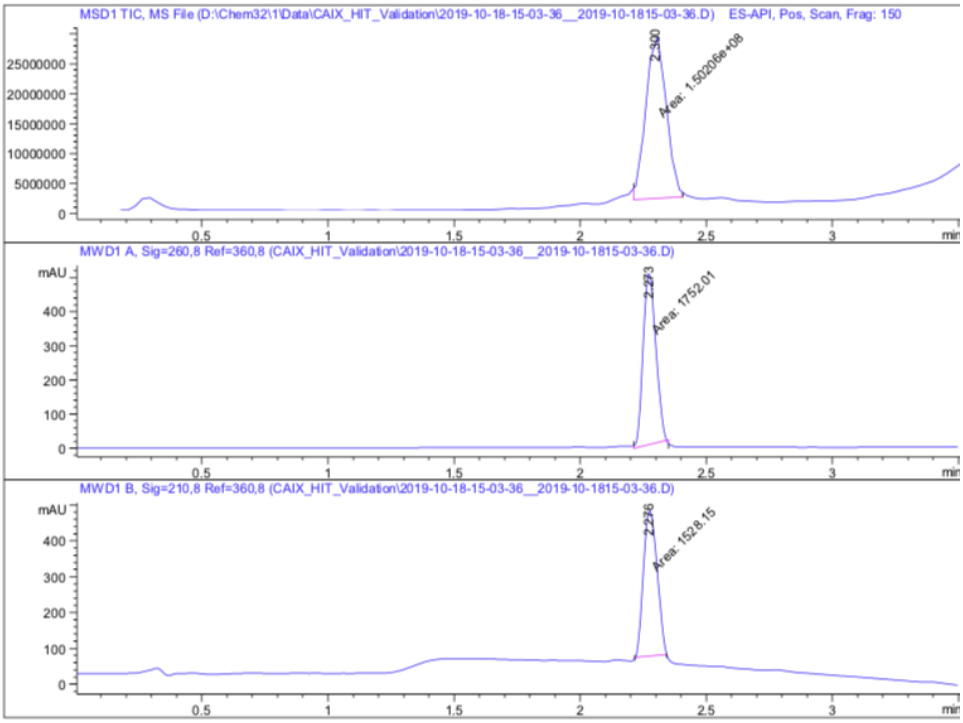


Figure S26. LC-ESI-MS chromatogram of compound **S9**, t_R: 1.978 min (DAD), 1.989 min (TIC).



Retention Time (MS)	MS Area	Mol. Weight or Ion
1.804	50606168	1027.10 I
		1026.00 I
		1025.00 I
		515.10 I
		514.10 I
		513.00 I

Figure S27. LC-ESI-MS chromatogram of compound **146/49**, t_R: 1.798 min (DAD), 1.804 min (TIC).



Retention Time (MS)	MS Area	Mol. Weight ion
2.300	150206240	1130.10 I
		1034.10 I
		1033.00 I
		1032.00 I
		819.10 I
		723.10 I
		722.10 I
		721.10 I
		565.70 I
		410.20 I
		390.00 I

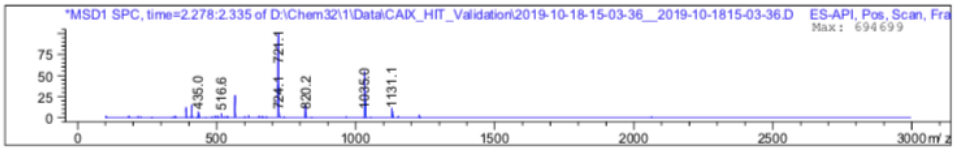


Figure S28. LC-ESI-MS chromatogram of compound **146/49-FITC**, t_R: 2.273 min (DAD), 2.300 min (TIC).

6.2 NMR spectra

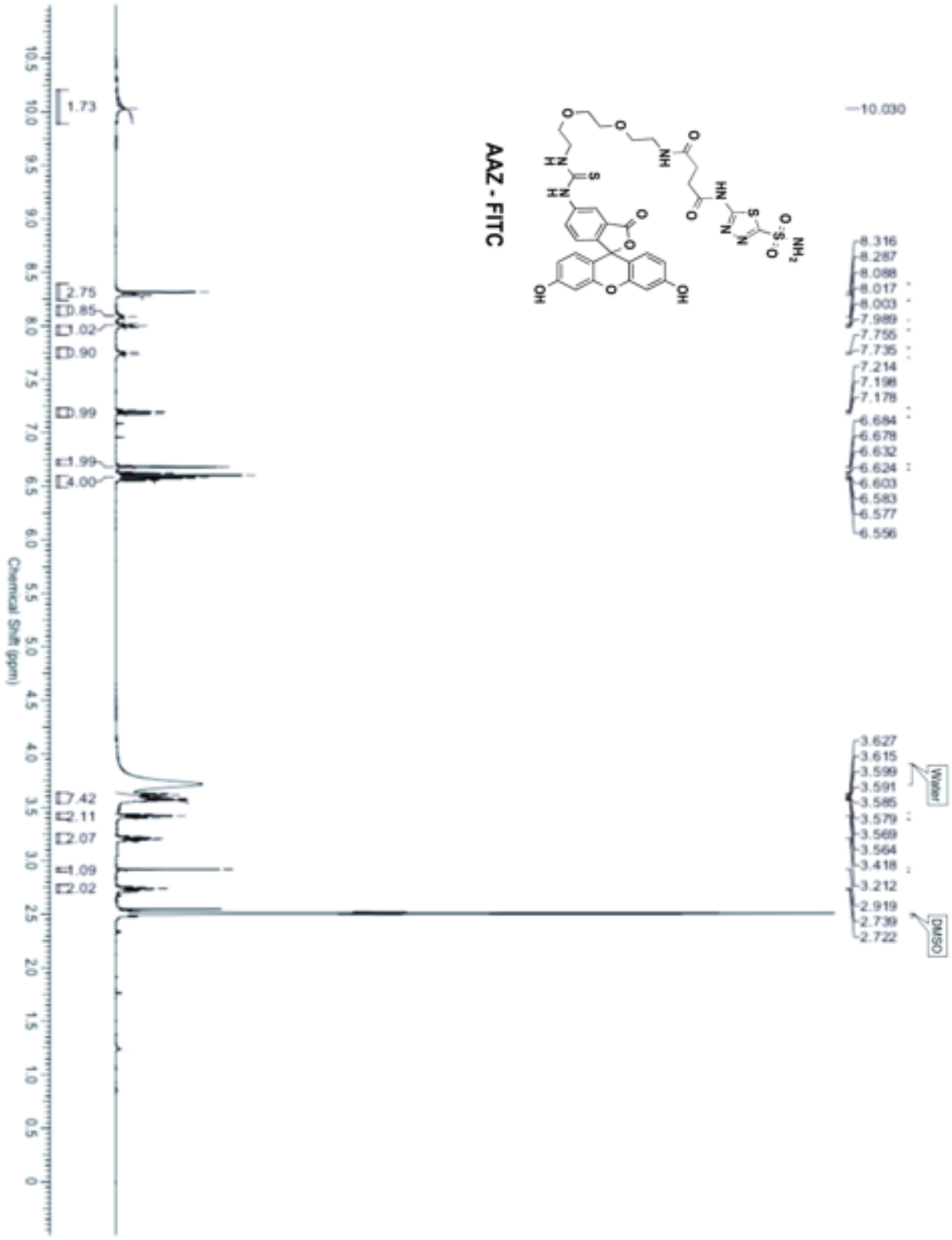


Figure S29.[1]H-NMR spectra of compound **AAZ - FITC** in DMSO d_6, 400 MHz

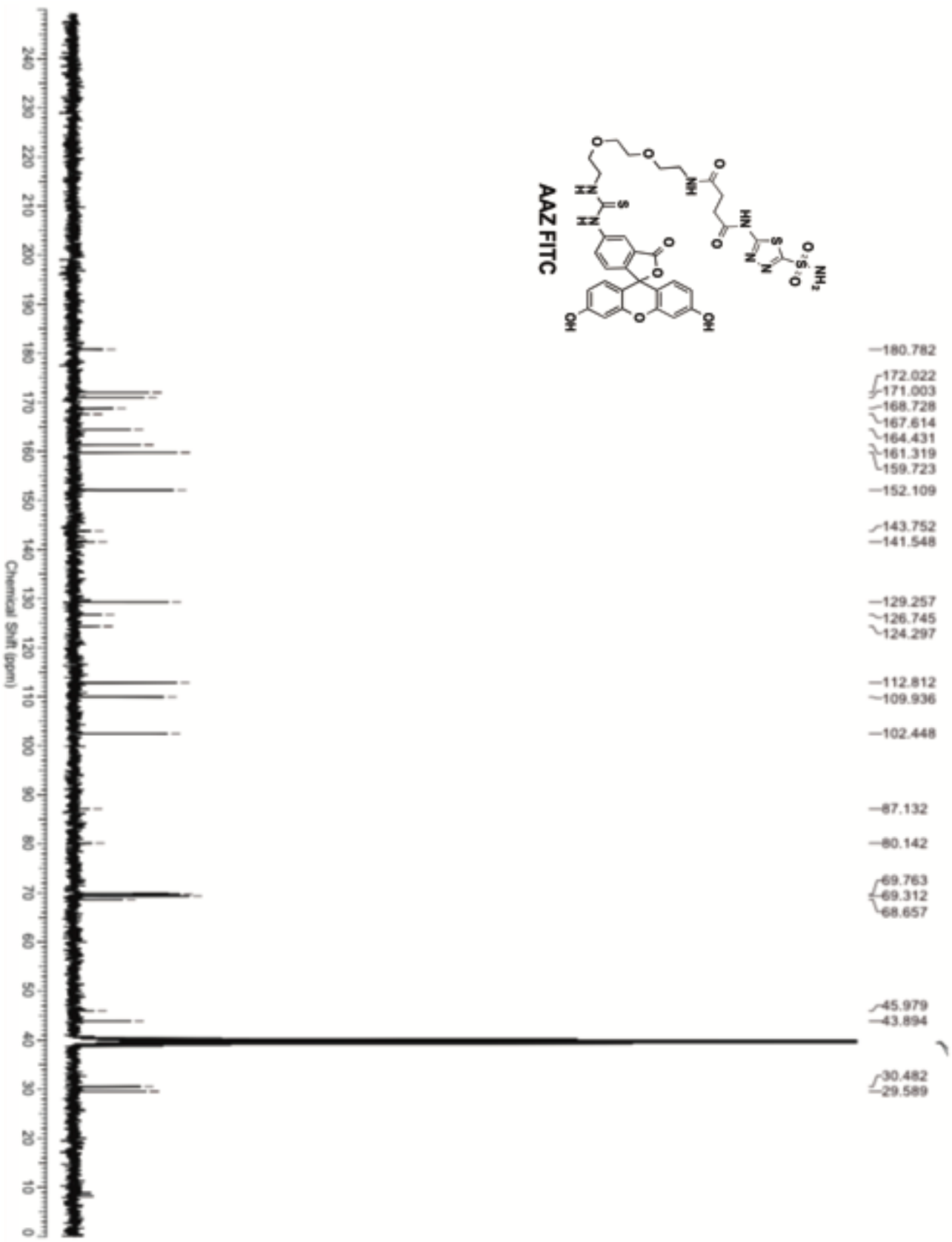


Figure S30. ^{13}C-NMR spectra of compound **AAZ - FITC** in DMSO d_6, 100 MHz

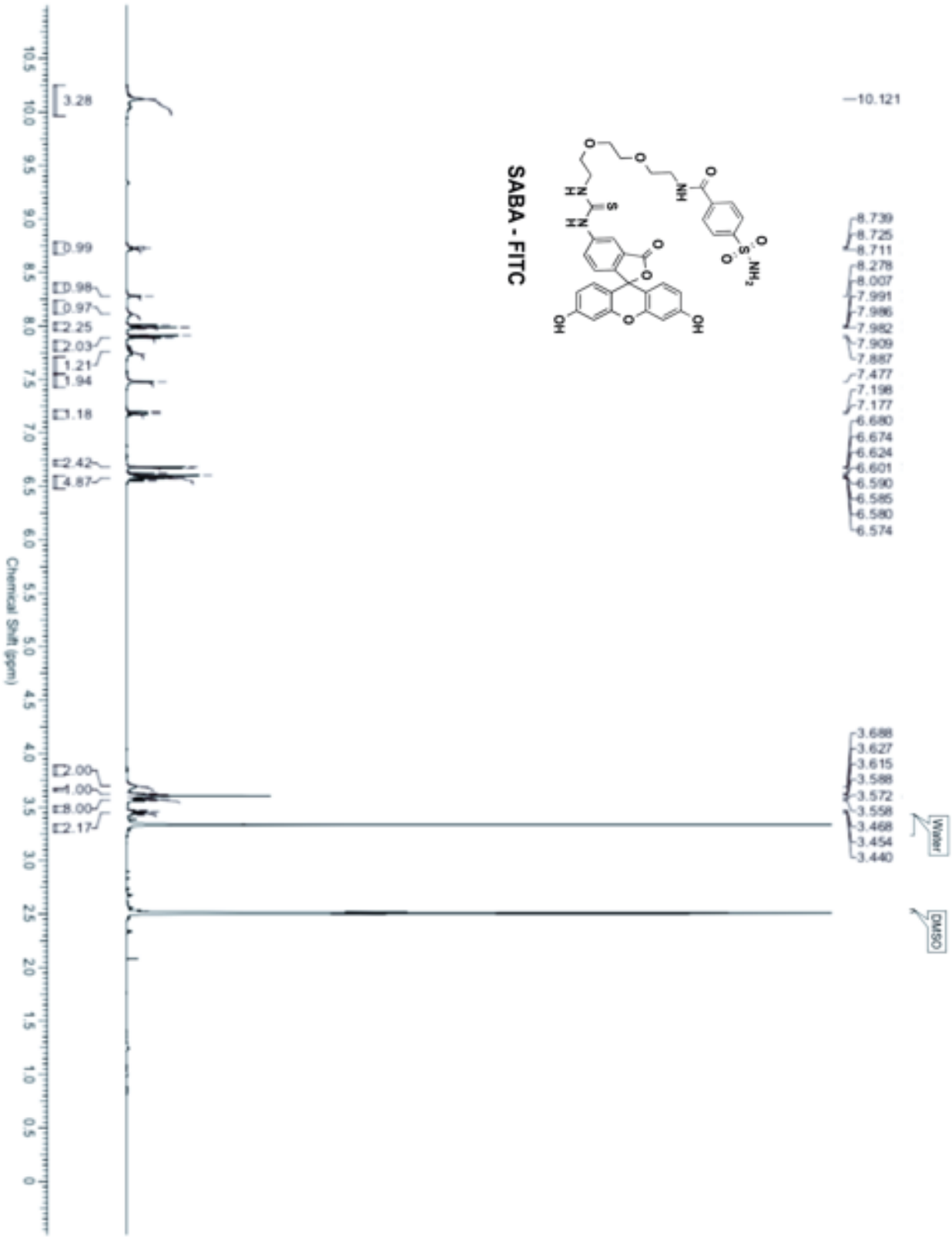


Figure S31. ^{1}H-NMR spectra of compound **SABA - FITC** in DMSO d_6, 400 MHz.

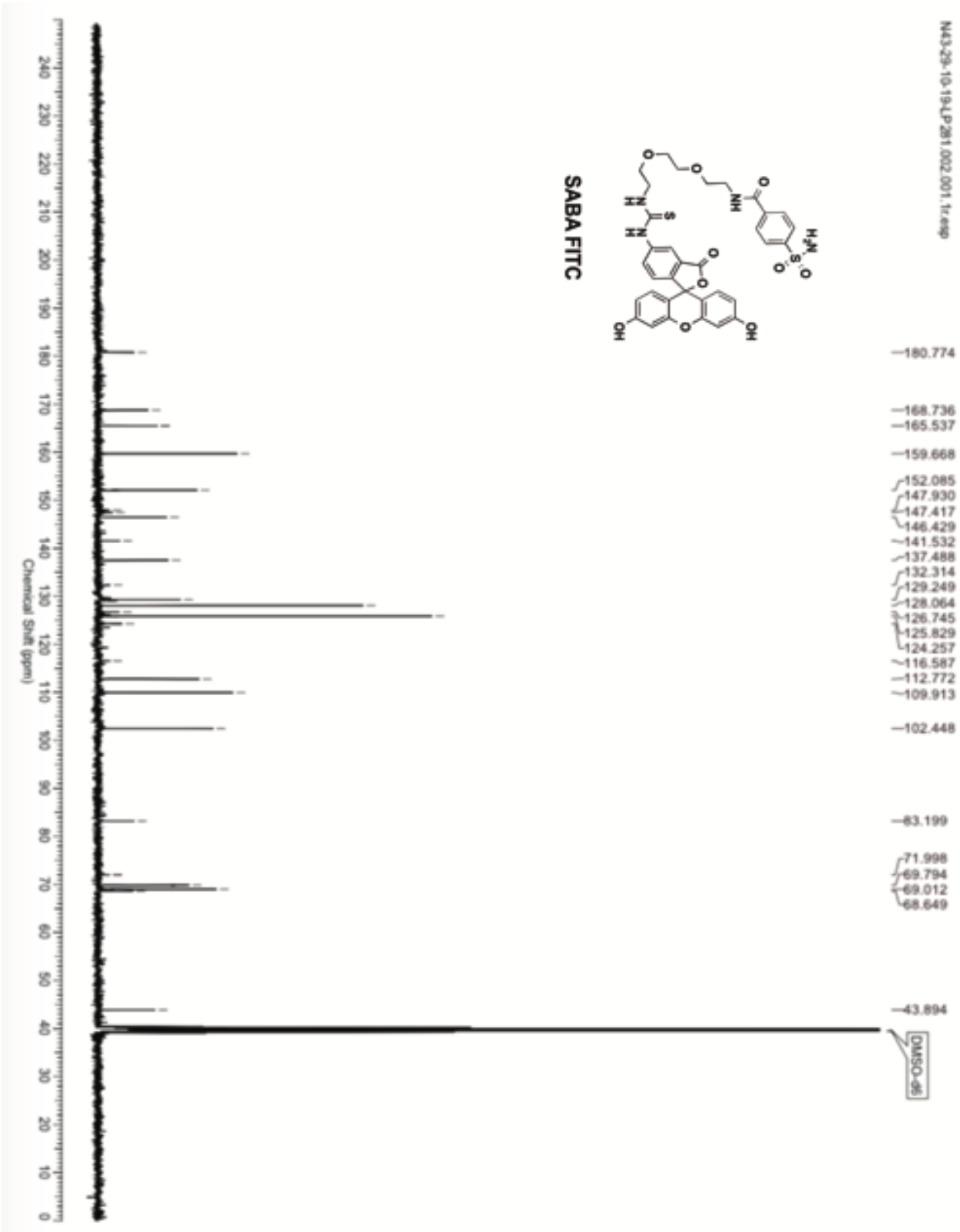


Figure S32. ^{13}C-NMR spectra of compound **SABA - FITC** in DMSO d_6, 100 MHz.

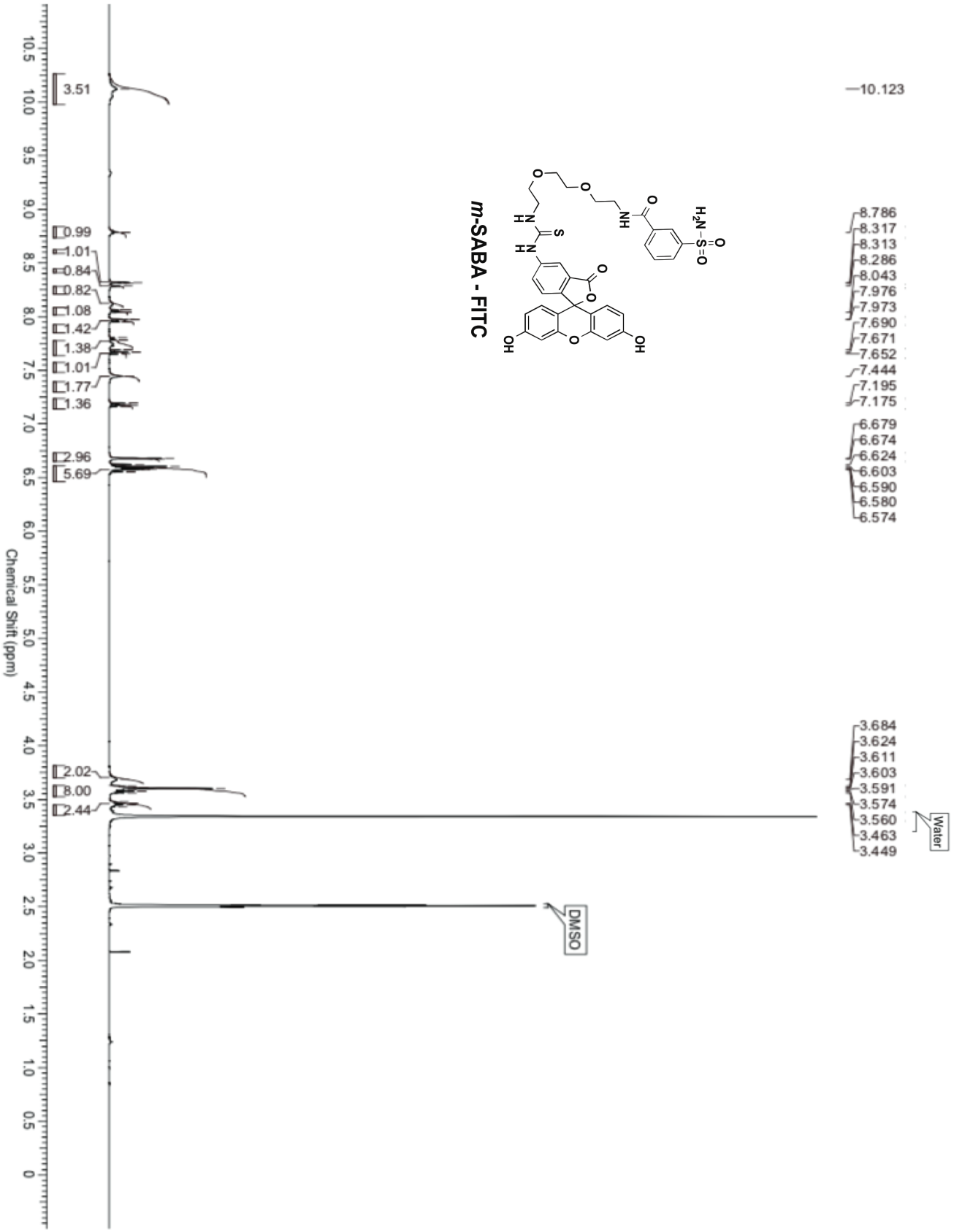


Figure S33. ^{1}H-NMR spectra of compound ***m*-SABA - FITC** in DMSO d_6, 400 MHz.

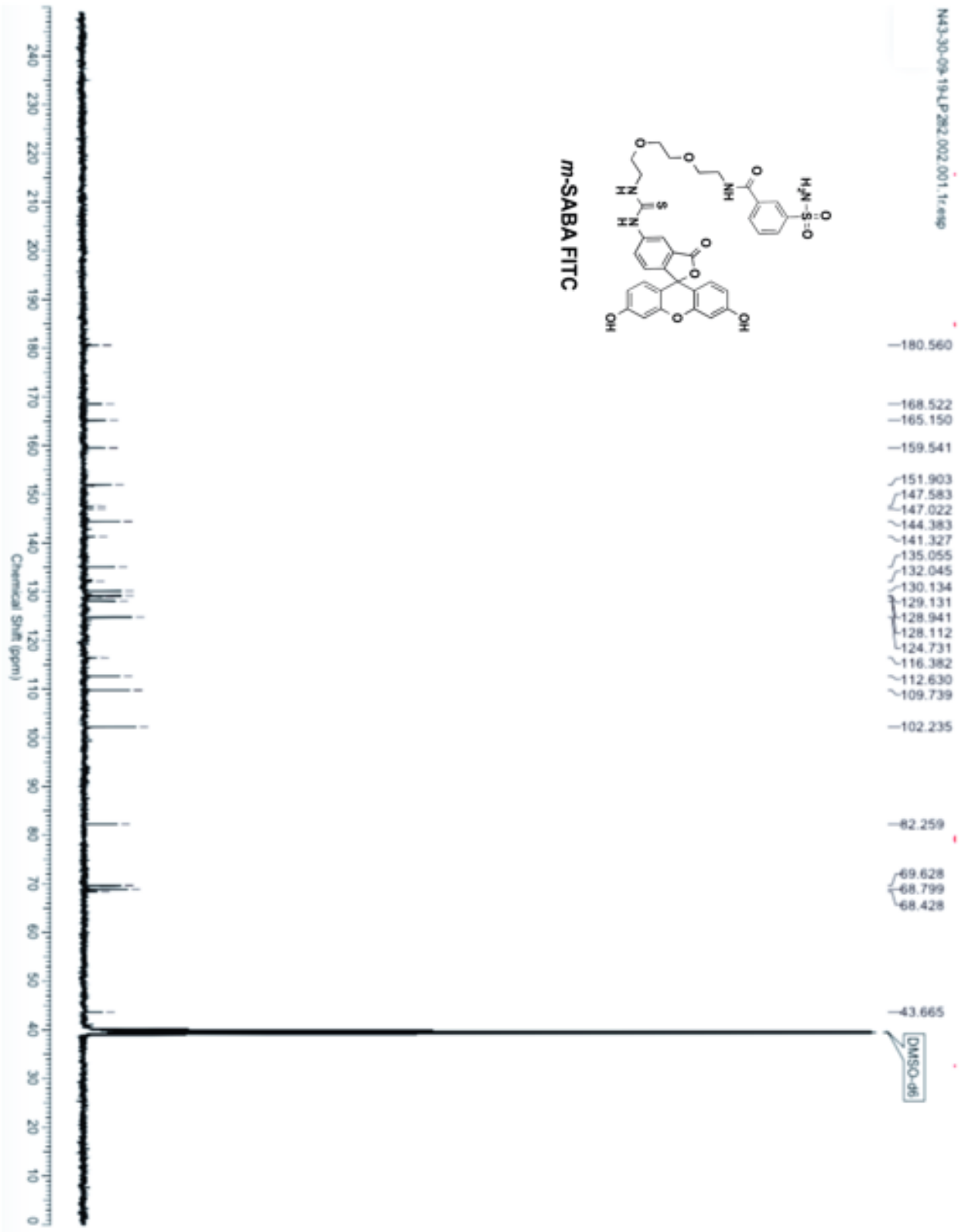


Figure S34. ^{13}C-NMR spectra of compound ***m*-SABA - FITC** in DMSO d_6, 100 MHz.

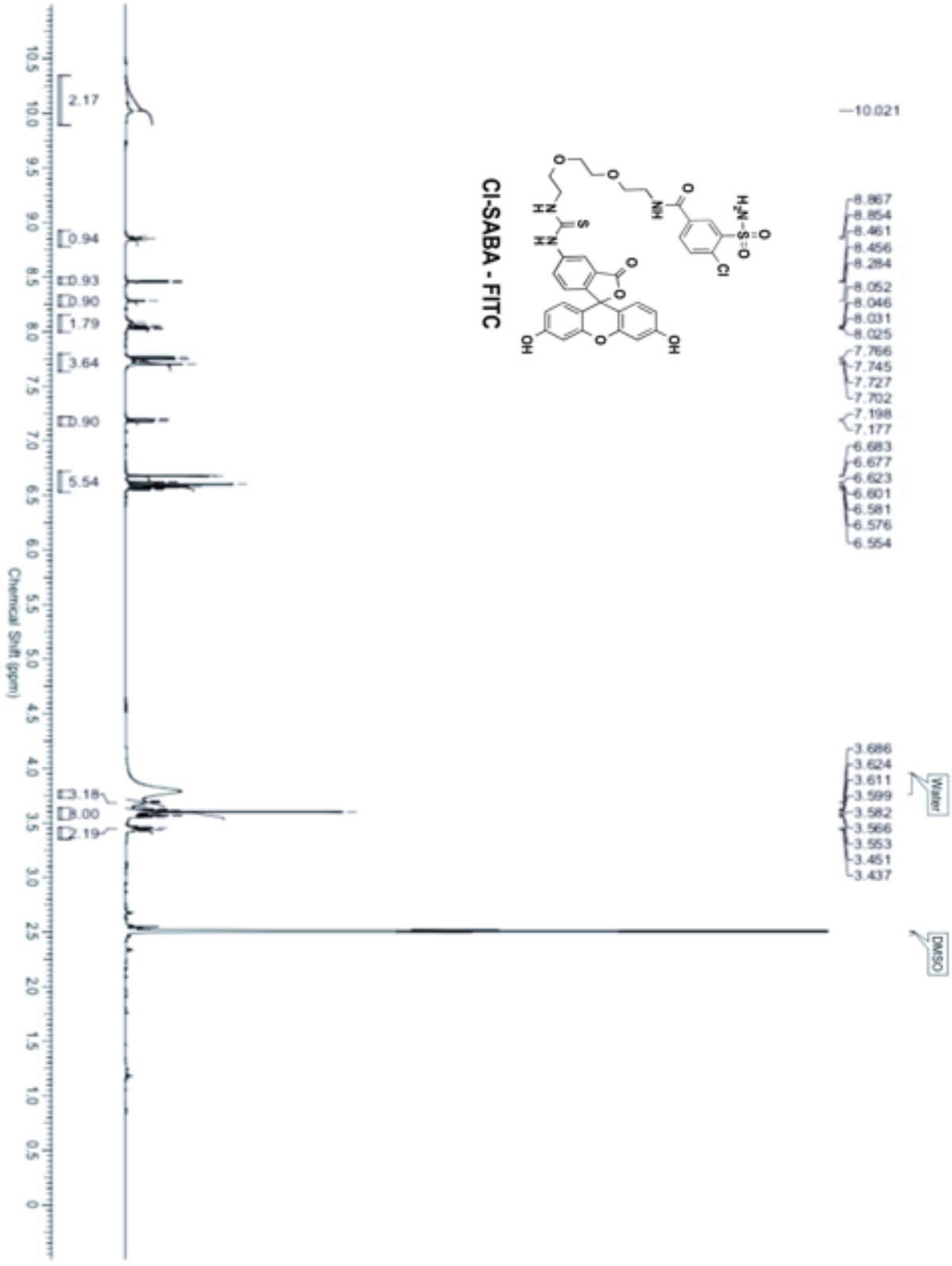


Figure S35. ^{1}H-NMR spectra of compound **Cl-SABA - FITC** in DMSO d_6, 400 MHz.

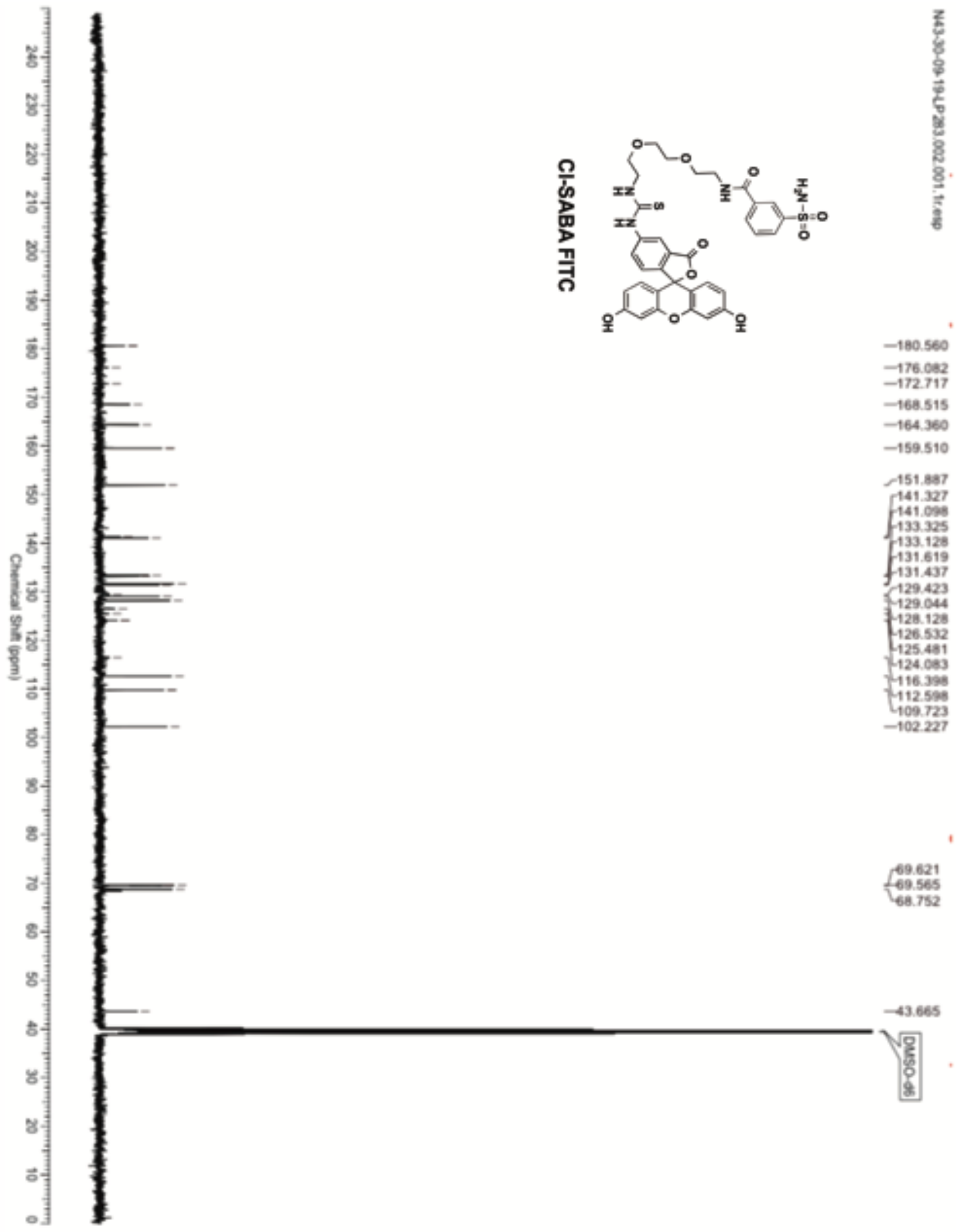


Figure S36. ^{13}C-NMR spectra of compound **Cl-SABA - FITC** in DMSO d_6, 100 MHz.

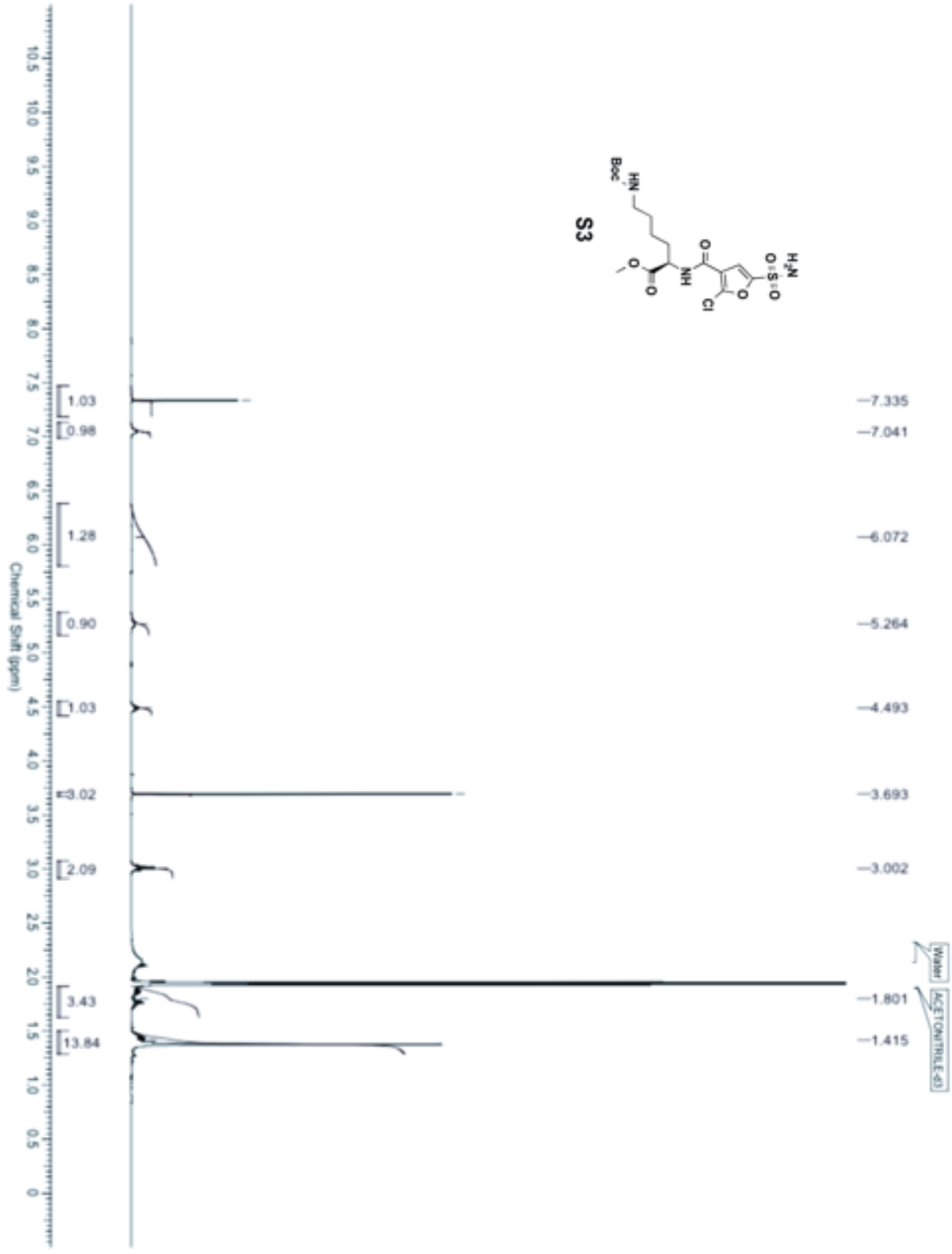


Figure S37. ^{1}H-NMR spectra of compound **S3** in CD_3CN, 400 MHz.

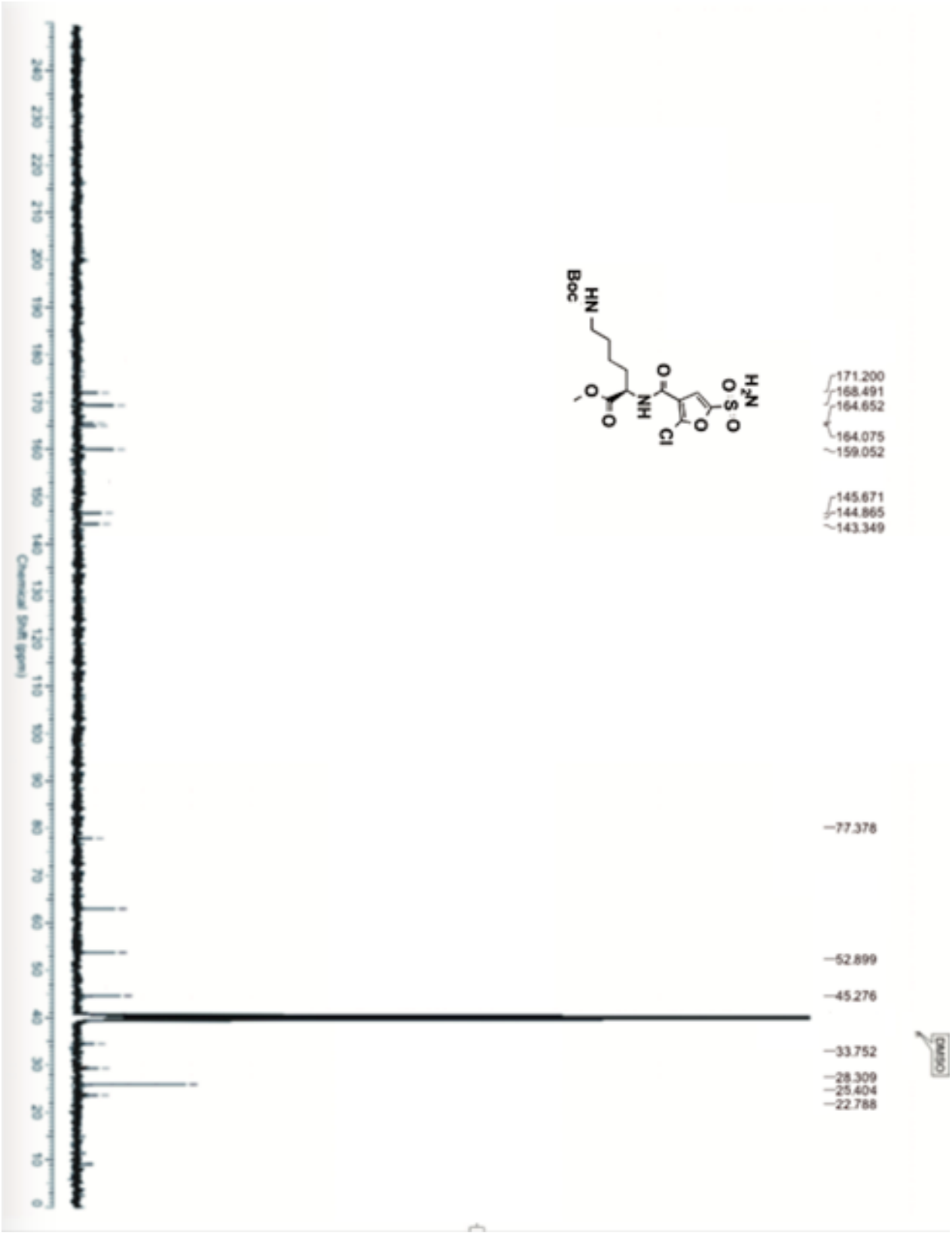


Figure S38. ^{13}C-NMR spectra of compound **S3** in DMSO-d6, 100 MHz.

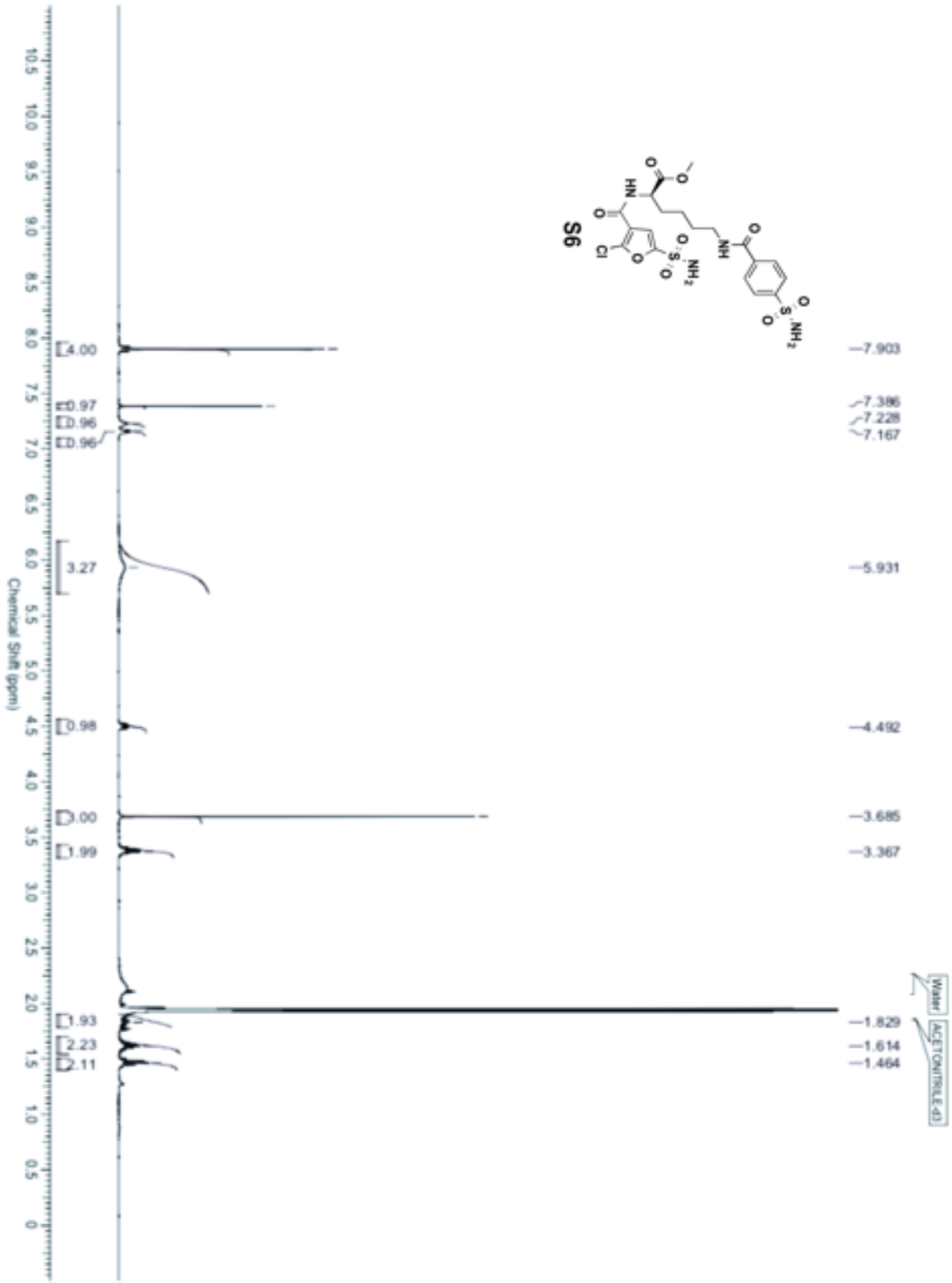


Figure S39. 1H-NMR spectra of compound **S6** in CD_3CN, 400 MHz.

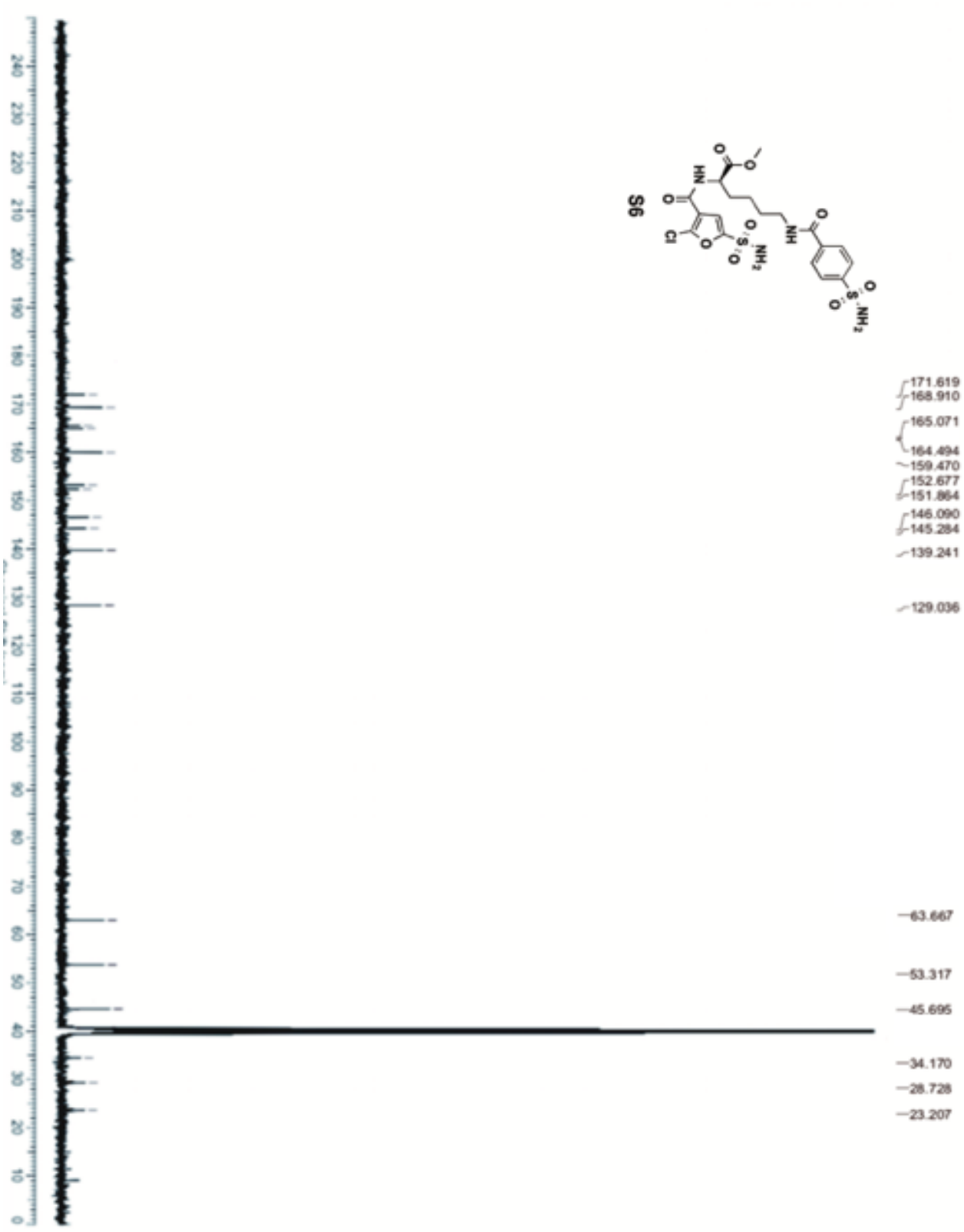


Figure S40. ^{13}C-NMR spectra of compound **S6** in DMSO-d6, 100 MHz.

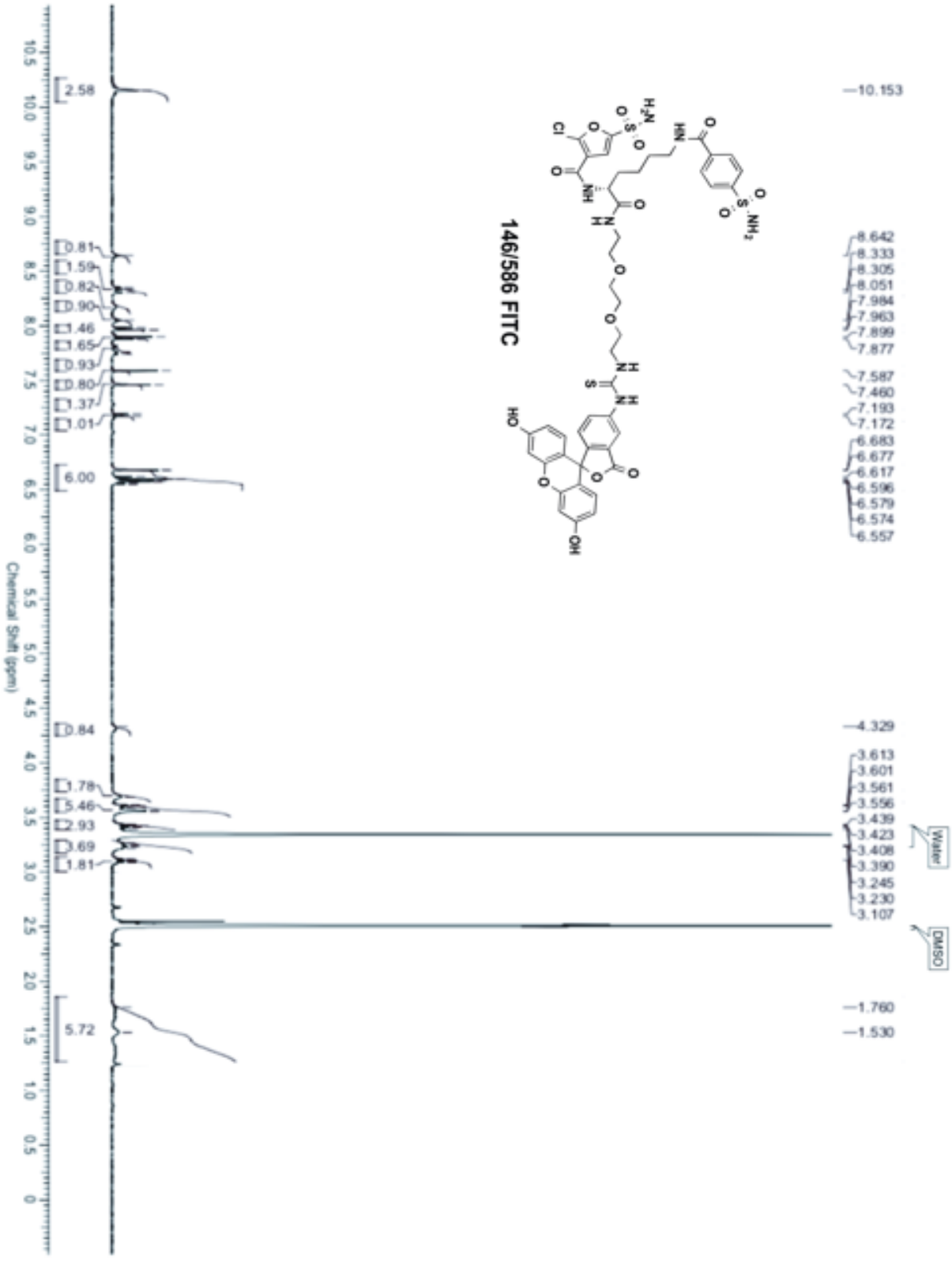


Figure S41. ^{1}H-NMR spectra of compound **146/586 FITC** in DMSO d_6, 400 MHz.

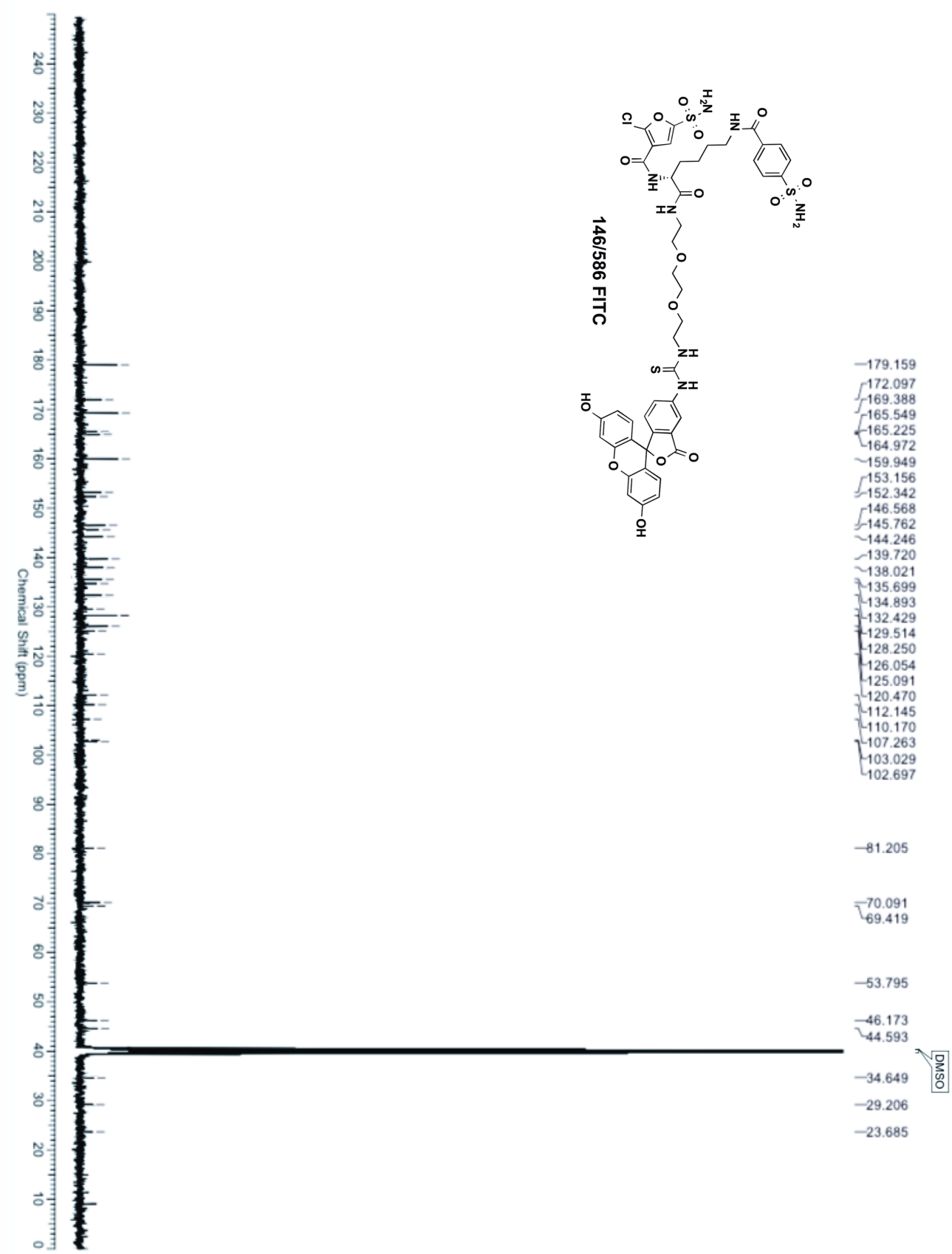


Figure S42. ^{13}C-NMR spectra of compound **146/586 FITC** in DMSO d_6, 400 MHz.

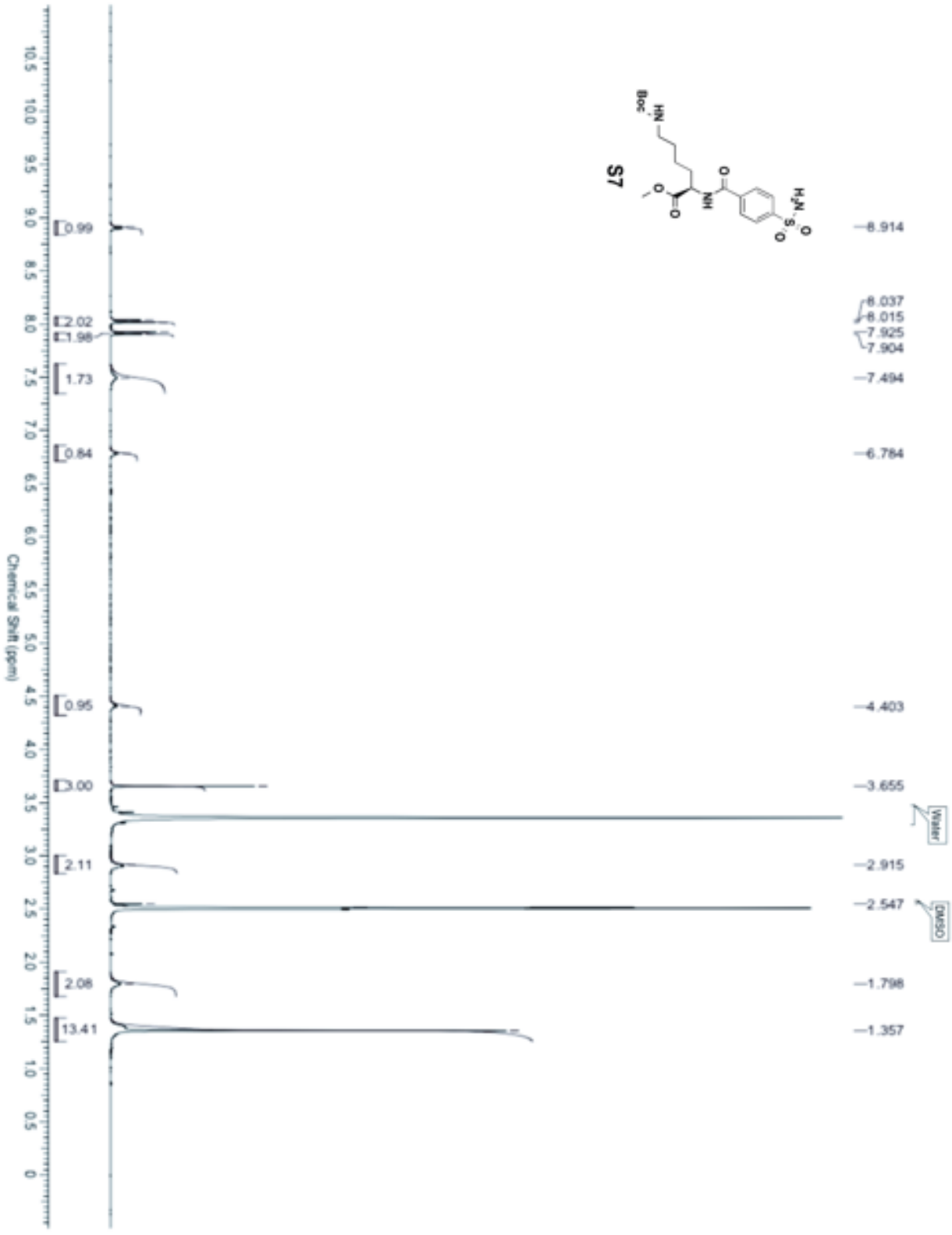


Figure S43. 1H-NMR spectra of compound S7 in DMSO d6, 400 MHz.

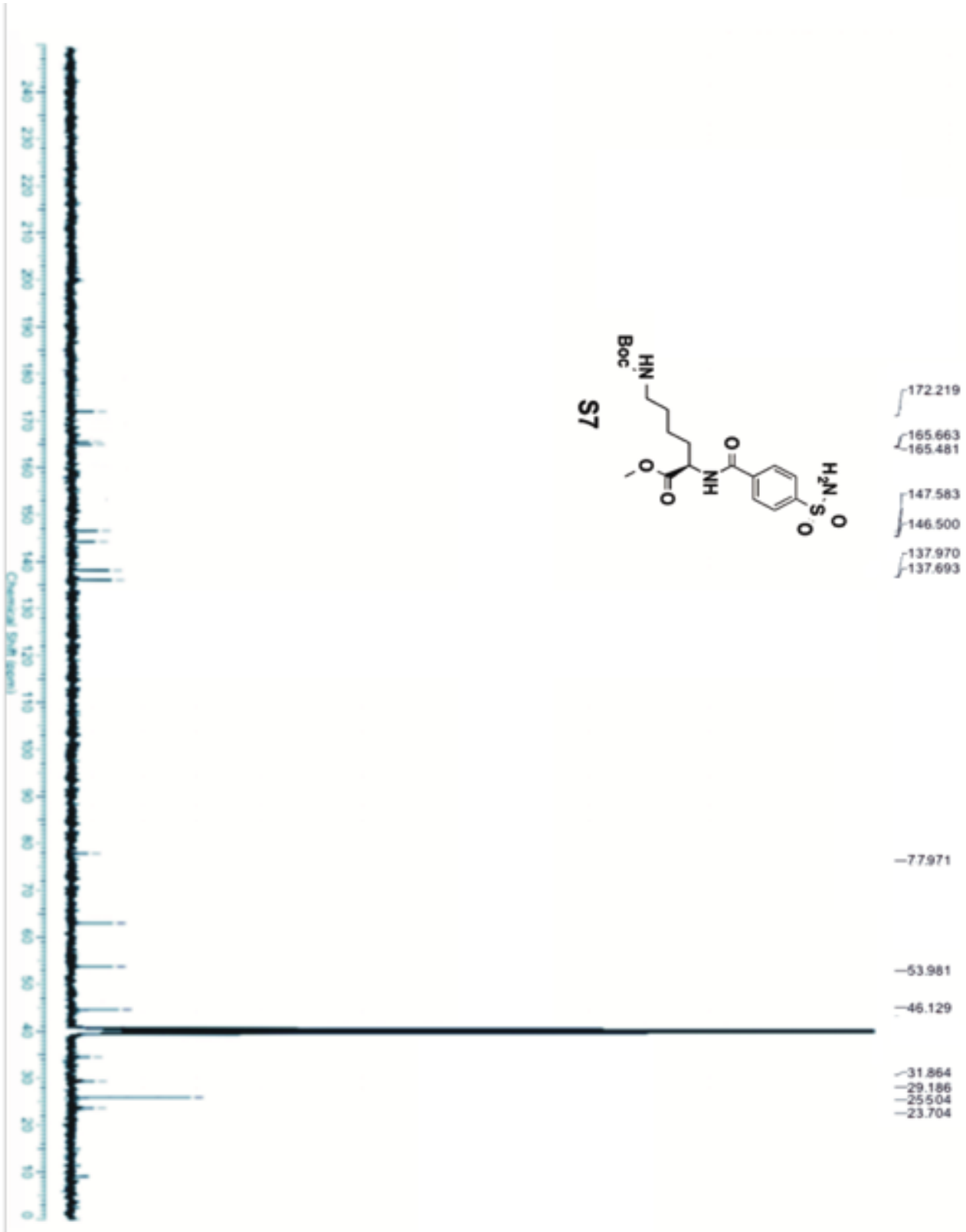


Figure S44. ^{13}C-NMR spectra of compound **S7** in DMSO d_6, 100 MHz.

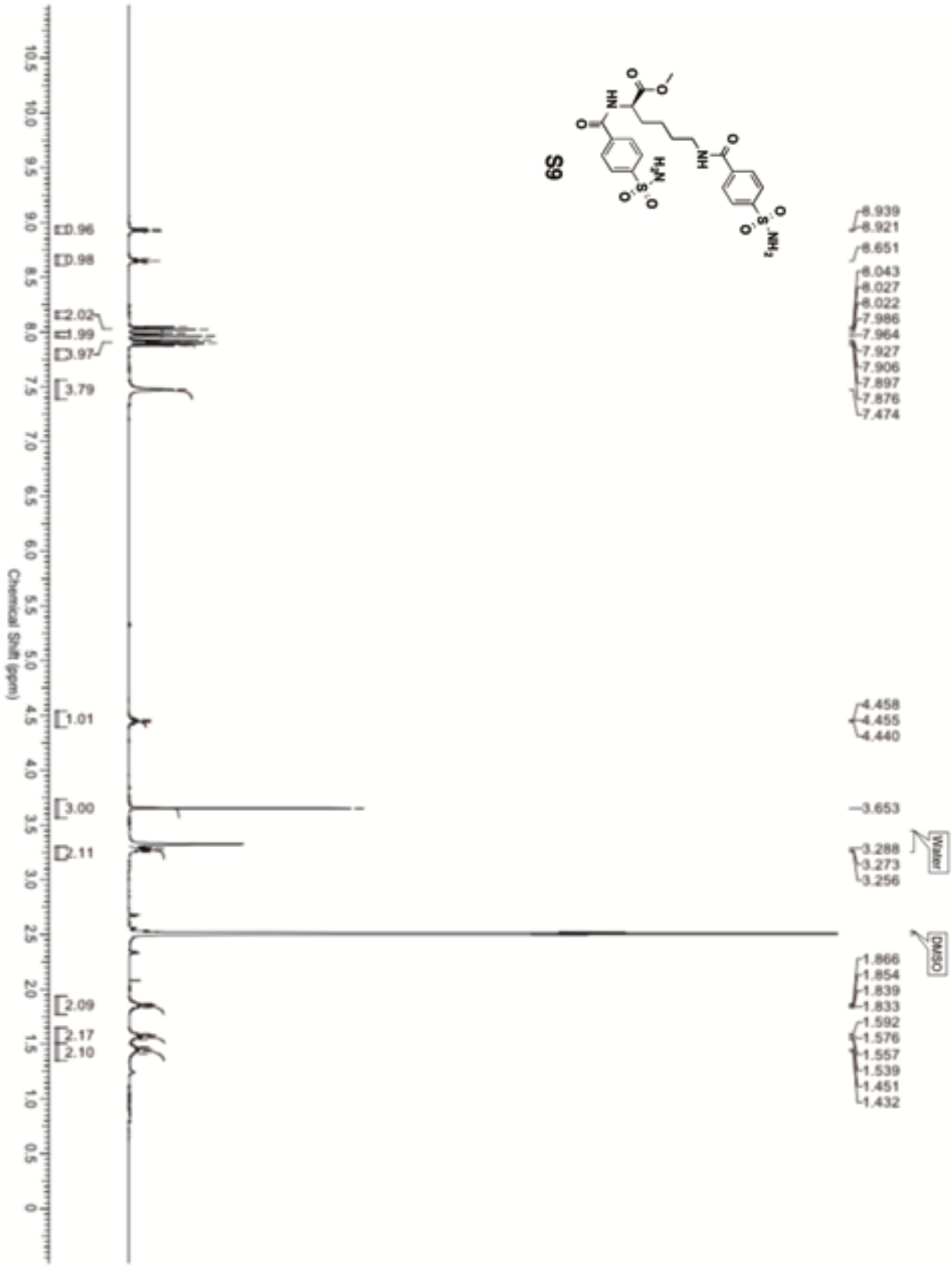


Figure S45. ^{1}H-NMR spectra of compound **S9** in DMSO d_6, 400 MHz**.**

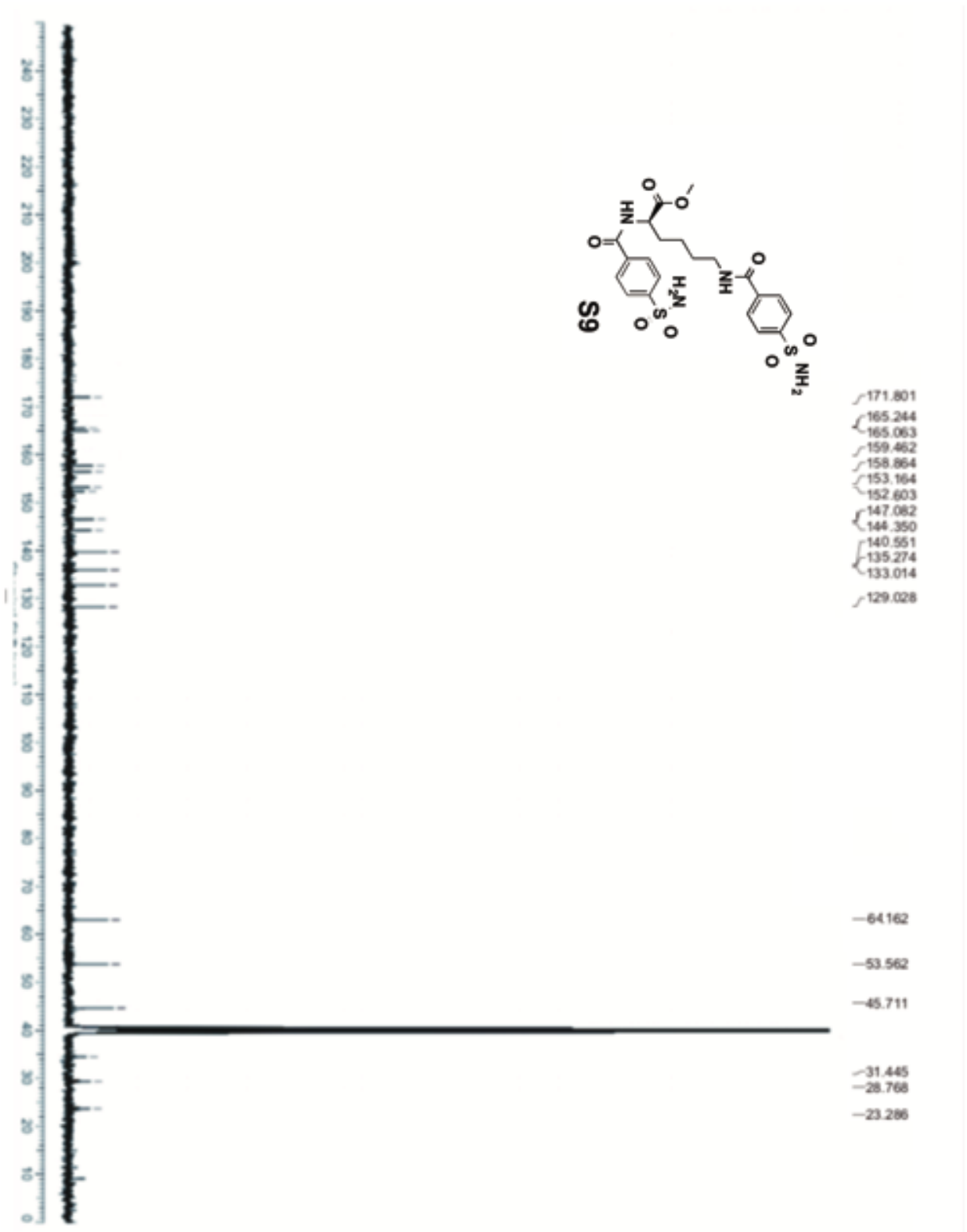


Figure S46. ^{13}C-NMR spectra of compound **S9** in DMSO d_6, 100 MHz**.**

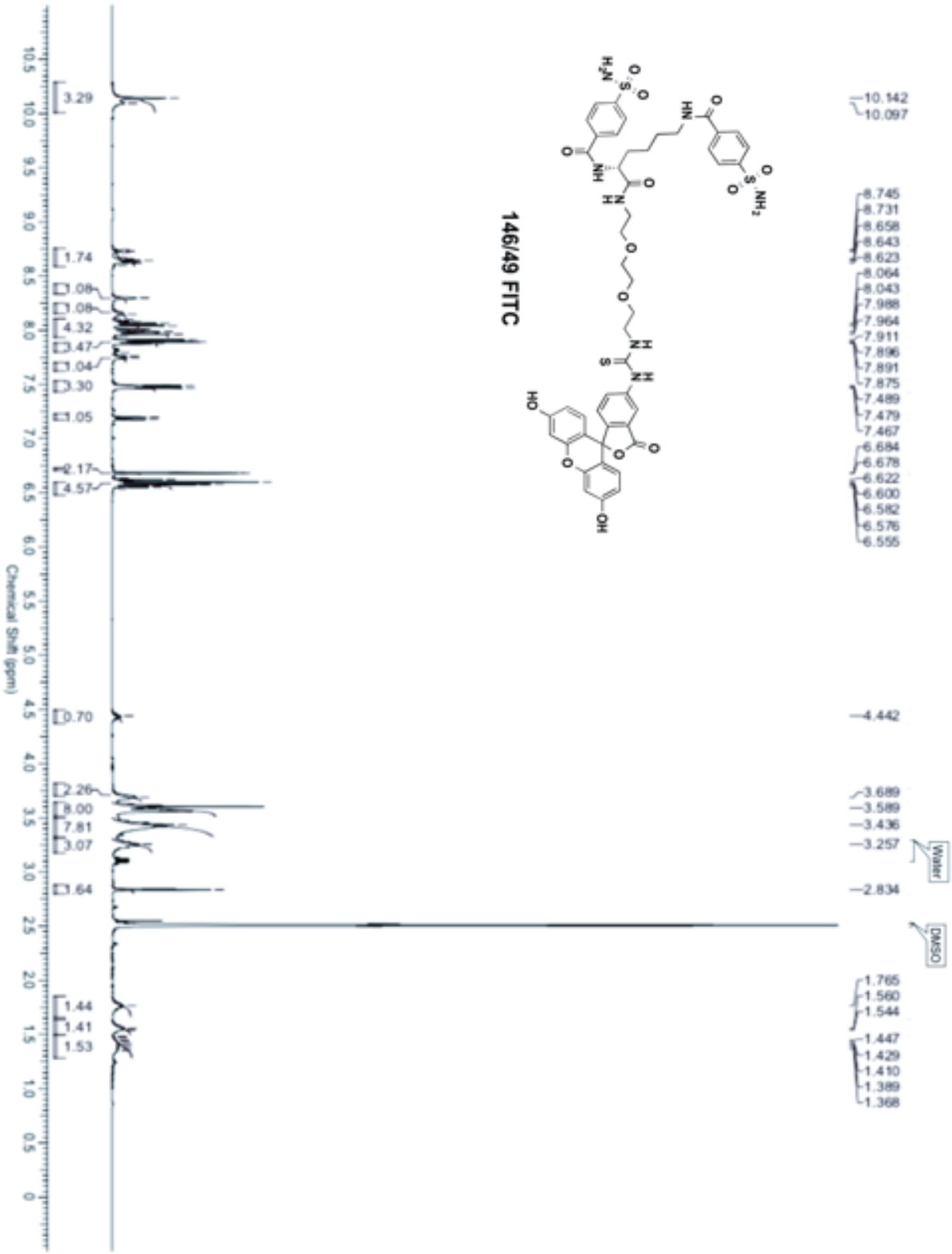


Figure S47. ^{1}H-NMR spectra of compound **146/49 - FITC** in DMSO d_6, 400 MHz

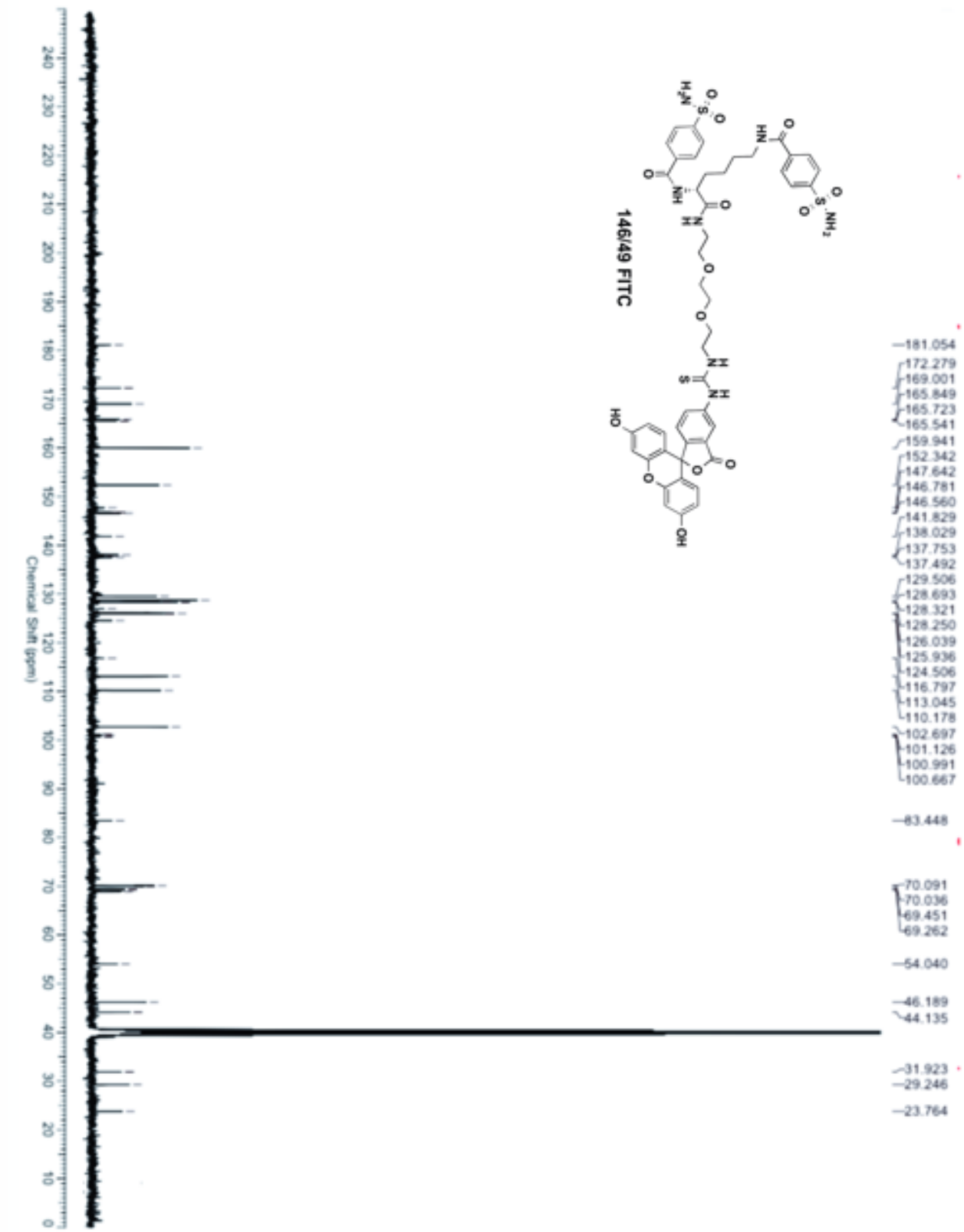


Figure S48. ^{13}C-NMR spectra of compound **146/49 - FITC** in DMSO d_6, 100 MHz

7. References

(1) Sannino, A.; Gabriele, E.; Bigatti, M.; Mulatto, S.; Piazzi, J.; Scheuermann, J.; Neri, D.; Donckele, E. J.; Samain, F. Quantitative Assessment of Affinity Selection Performance by Using DNA-Encoded Chemical Libraries. *ChemBioChem* **2019**, *20*, 955–962.