Supplementary Information

Discovery of a potent and selective covalent inhibitor and activity-based probe for the deubiquitylating enzyme UCHL1, with anti-fibrotic activity

Nattawadee Panyain,[†] Aurélien Godinat,[†] Thomas Lanyon-Hogg,[†] Sofía Lachiondo-Ortega,[†,‖] Edward Will,[†] Christelle Soudy,[‡] Milon Mondal,[†] Katie Mason,[§] Sarah Elkhalifa,[§] Lisa M. Smith,[§] Jeanine A. Harrigan,[§] and Edward W. Tate[*,†,‡]

[†] Department of Chemistry, Molecular Sciences Research Hub, Imperial College London, London, W12 0BZ, UK
[‡] The Francis Crick Institute, London, NW1 1AT, UK
[§] Mission Therapeutics Ltd, Moneta, Babraham Research Campus, Cambridge, CB22 3AT, UK

* Corresponding author: e.tate@imperial.ac.uk

Table of Contents

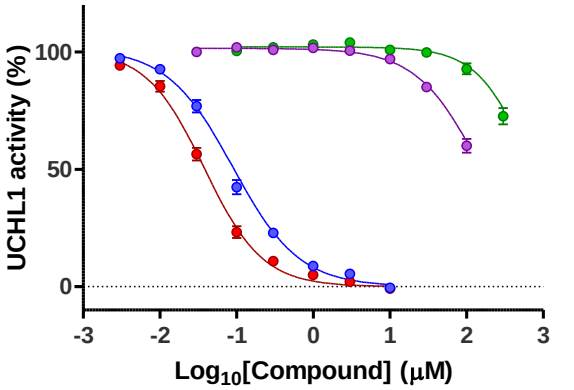
Supplementary Results

Supplementary Data Files (Separate Excel files)

Supplementary Data S1	Peptide lists of compound-modified peptide identification
Supplementary Data S2	Protein lists of ABP 2 target engagement
Supplementary Data S3	Protein lists of whole proteome analysis
Supplementary Data S4	Protein lists of competitive ABPP

Supplementary Figures

a)



Compound	FP IC_{50} (95% CI)
1	90 (79-100) nM
2	38 (32-45) nM
3	>100 μM*
LDN-57444	>300 μM*

*Maximum concentration used in the assay

b)

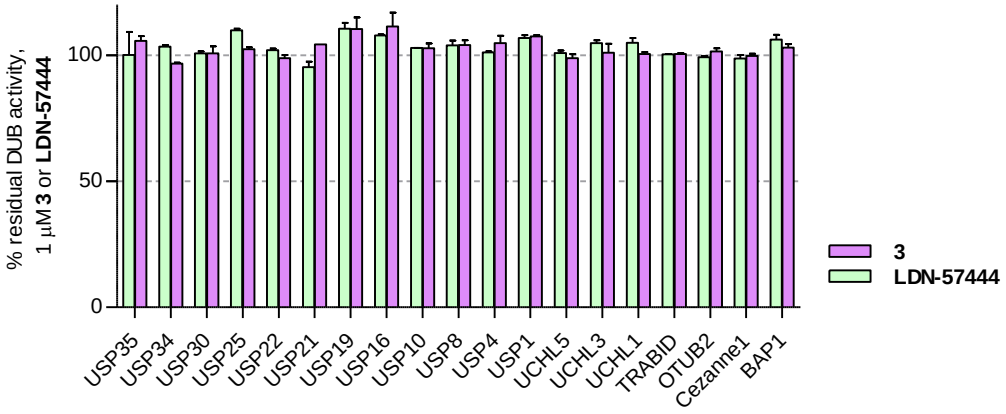


Figure S1. Biochemical enzyme assay.
(a) UCHL1 activity was measured *via* FP with Ub-Lys-TAMRA. UCHL1 (5 nM) was incubated with 8-point dilutions of test compounds for 30 min, and measurement initiated by adding Ub-Lys-TAMRA (50 nM). Data represents mean ± SEM (n=3). IC_{50} and 95% confidence intervals (95% CI) were calculated from a four-parameter dose-response curve. (b) Biochemical selectivity profiling of compound **3** and LDN-57444 at 1 μM using FP and FI assays. For each DUB enzyme, the substrate and enzyme concentration were optimized, as described in Materials and Methods. Data represent mean ± SEM of assay in duplicate (n=1-3).

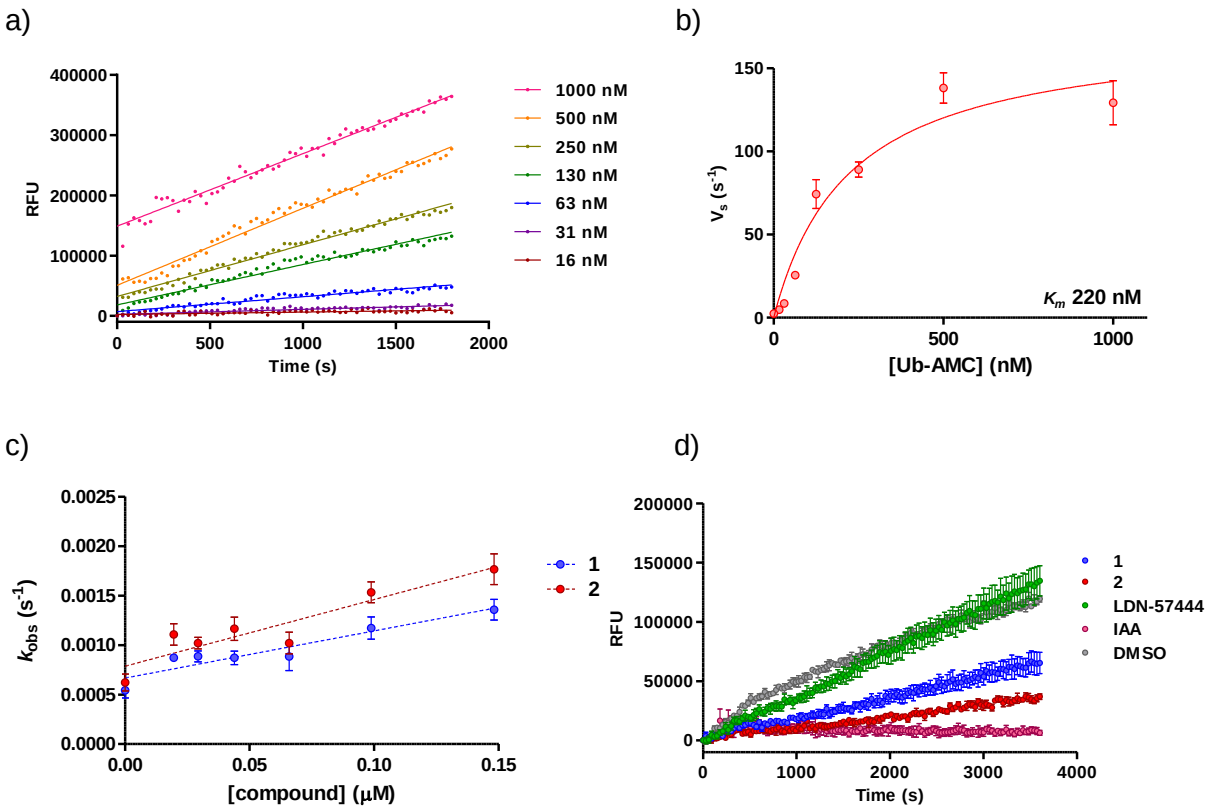
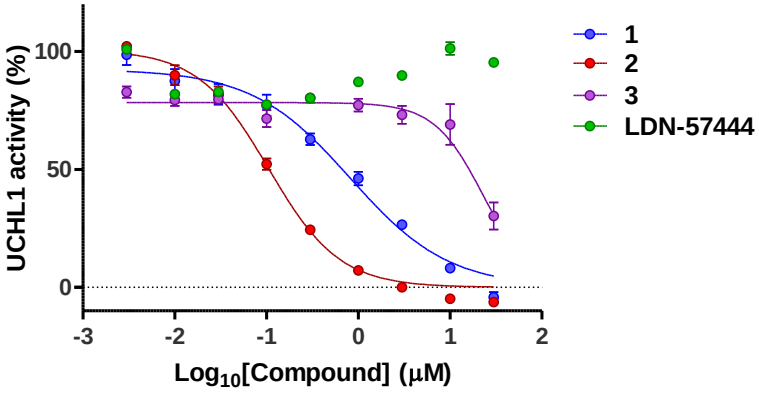


Figure S2. UCHL1 kinetics
(a) Kinetic analysis of UCHL1 (3 nM)-catalyzed cleavage of varied concentrations of Ub-AMC. (b) Non-linear fit of Michaelis–Menten equation to initial reaction velocities versus Ub-AMC concentrations, giving K_m of 220 nM (95% CI 140-310 nM). (c) Apparent rate of inhibition (k_a) of compound **1** and **2** at varying concentrations. Ub-AMC (130 nM) fluorescence intensity was measured in the presence of UCHL1 (3 nM). The apparent rate constant for inhibition, k_a/I was determined by linear regression, giving the absolute value of the rate constant k_{obs}/I values for compounds **1** and **2** of 7400 (95% CI 5200-9700) $M^{-1}s^{-1}$ and 11000 (95% CI 7700-13000) $M^{-1}s^{-1}$, respectively. (d) Reaction progression curves for Ub-AMC cleavage after 100x dilution of UCHL1 (300 nM) treated with DMSO (positive control), compound **1**, **2**, LDN-57444 and iodoacetamide (IAA, negative control).



Compound	HTRF IC_{50} (95% CI)
1	820 (500-1100) nM
2	110 (86-130) nM
3	23 (18-29) μM*
LDN-57444	>30 μM*

*Maximum concentration used in the assay

Figure S3. Compound characterization in Cal51 cells.
In-cell UCHL1 activity profiling of **1**, **2**, **3** and LDN-57444 by HTRF. Cal51 cells overexpressing UCHL1 were treated with compounds **1**, **2**, **3** or LDN-57444 for 1 h at 37ºC before lysis. Lysates were combined with HA-Ub-VME (100 nM) and incubated for 30 min at room temperature. Detection antibodies were added to the reaction plate and incubated overnight at 4ºC before measurement. HTRF ratios were calculated as described in Materials and Methods. Data represents mean ± SEM (n=4). IC_{50} and 95% CI values were calculated from a four-parameter dose-response curve.

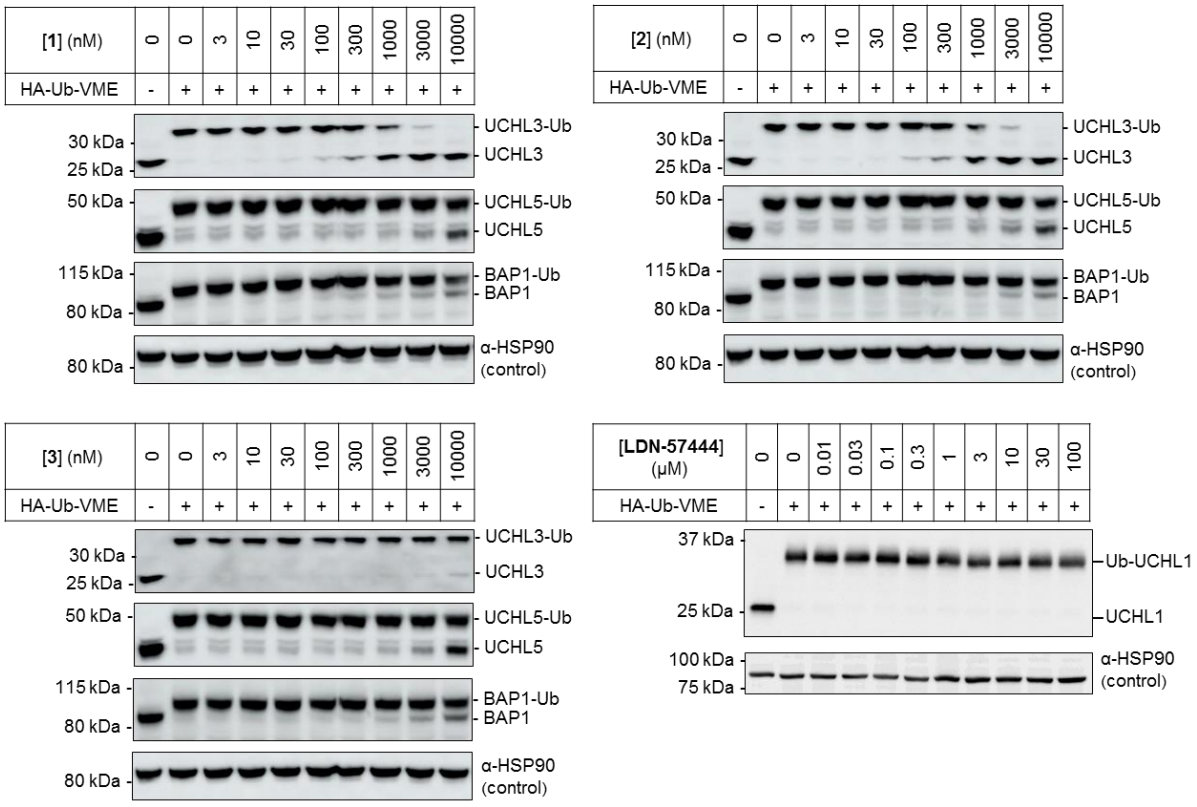


Figure S4. In-cell target engagement of 1, 2, 3 and LDN-57444 using HA-Ub-VME ABP. Immunoblot analysis of HEK293T cells incubated with different concentrations of **1**, **2**, **3** and LDN-57444 for 1 h. The compound-treated lysates were labeled with HA-Ub-VME ABP for 15 min, separated by SDS-PAGE and probe-labeled proteins visualized by immunoblotting using the indicated antibodies.

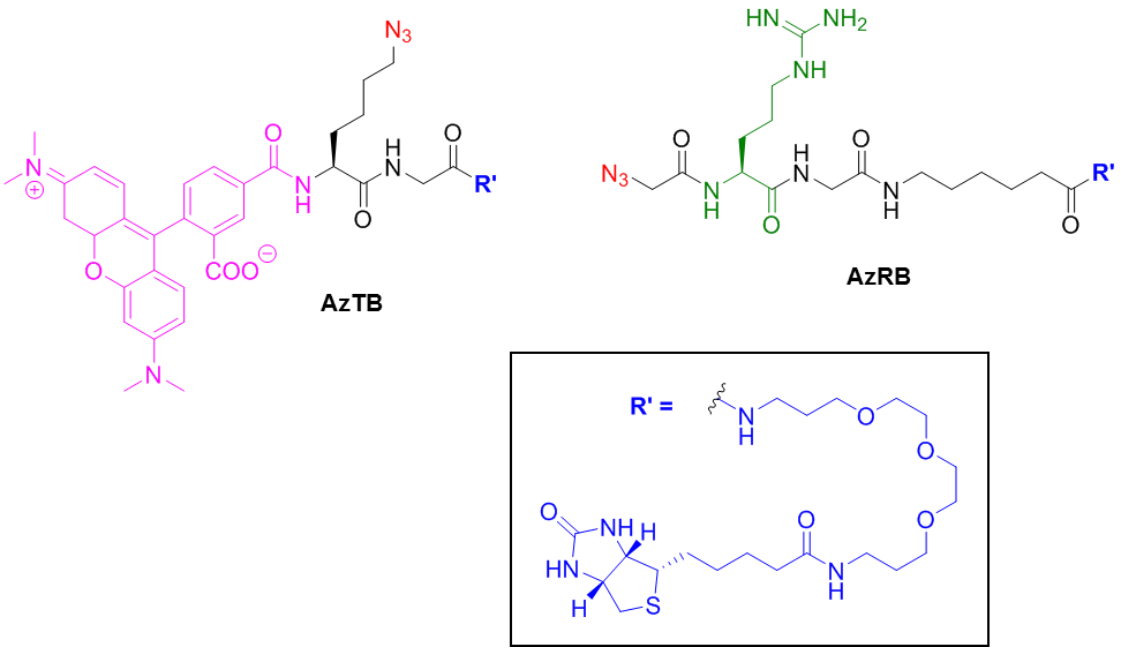


Figure S5. Structures of 'capture' reagents.
Structure of azido-TAMRA-biotin (AzTB) and azido-arginine-biotin (AzRB). Capture reagents contain key functional moieties; an azide (red) for copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) reaction, a TAMRA fluorophore (pink) for in-gel fluorescence visualization, a biotin affinity handle (blue) for enrichment and purification, and an arginine residue (green) for on-bead trypsin digestion.

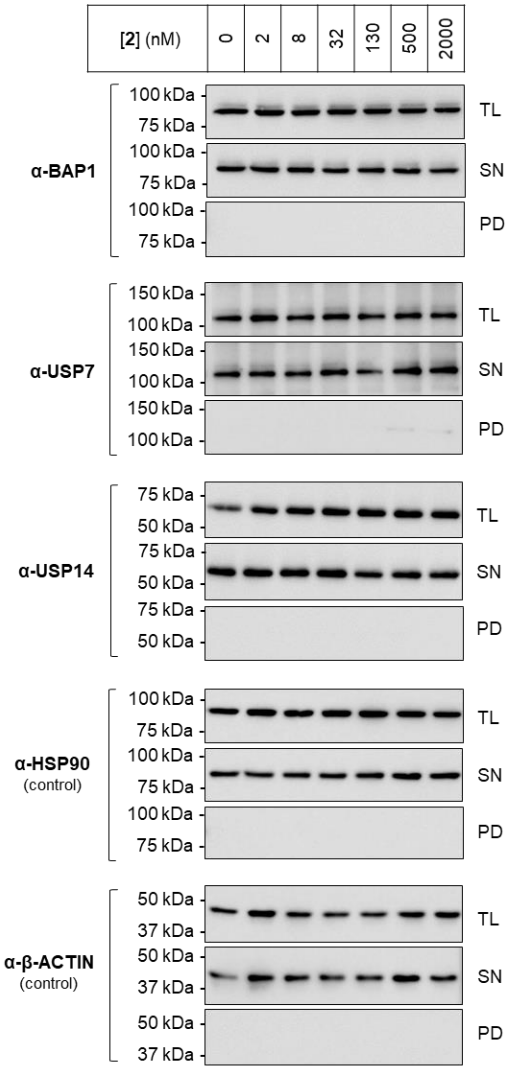


Figure S6. Immunoblot analysis of ABP 2 labeling in HEK293 cells.
HEK293 cells were incubated with 2, 8, 32, 130, 500 or 2000 nM of **2** for 1 h then lysed. The lysate was ligated to AzTB capture reagent and labeled proteins were enriched on Streptavidin magnetic beads. The total lysate (TL), pulldown (PD), and supernatant (SN) samples were separated by SDS-PAGE and analyzed by immunoblotting using the indicated antibodies.

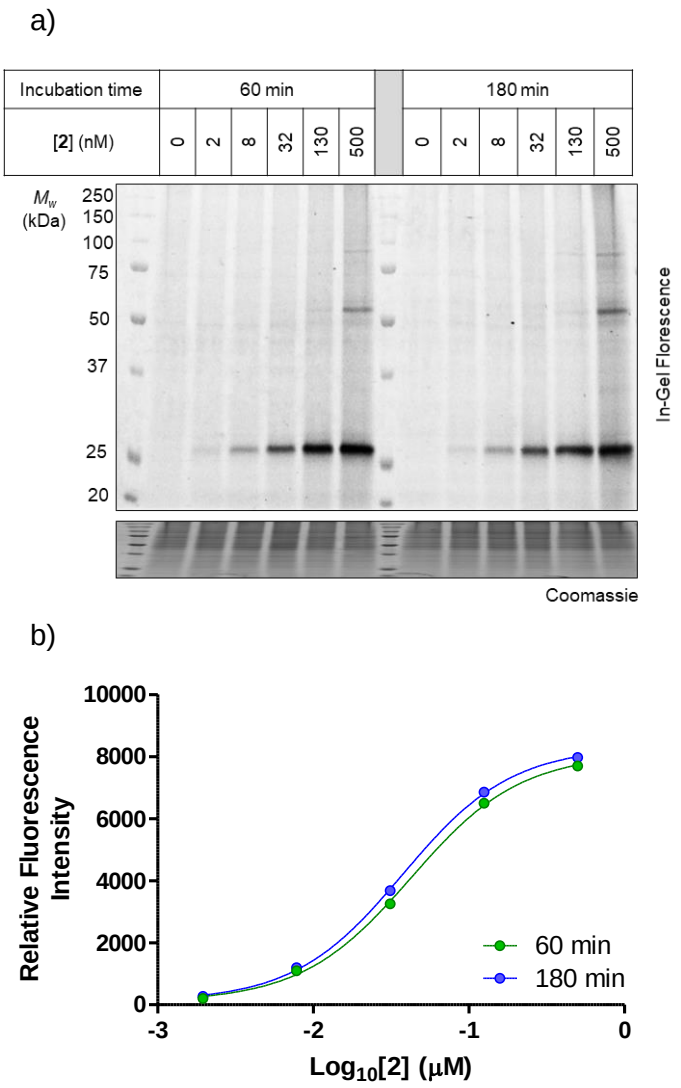


Figure S7. In-gel fluorescence analysis of ABP 2 labeling.
(a) HEK293 cells were incubated with 2, 8, 32, 130 or 500 nM of ABP **2** for 1 or 3 h. The lysate was ligated to AzTB capture reagent, separated by SDS-PAGE and visualized by in-gel fluorescence. (b) The labeling intensity on the gel was quantified using ImageJ and plotted against **2** concentration, demonstrating little further labeling after 1 h. Values were background corrected to DMSO vehicle control.

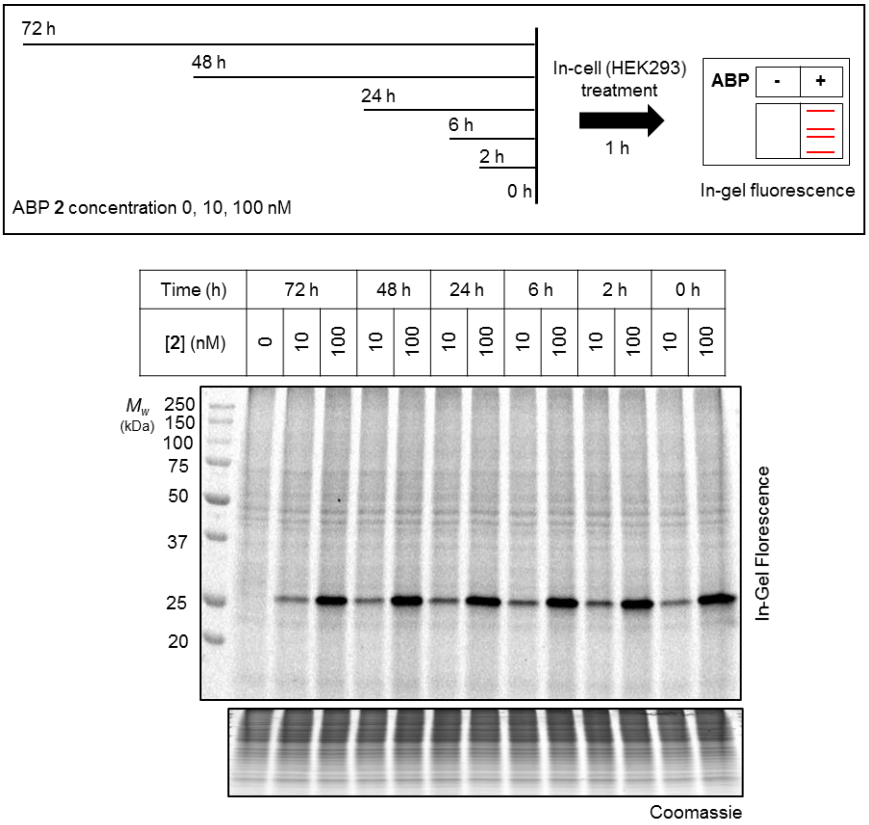
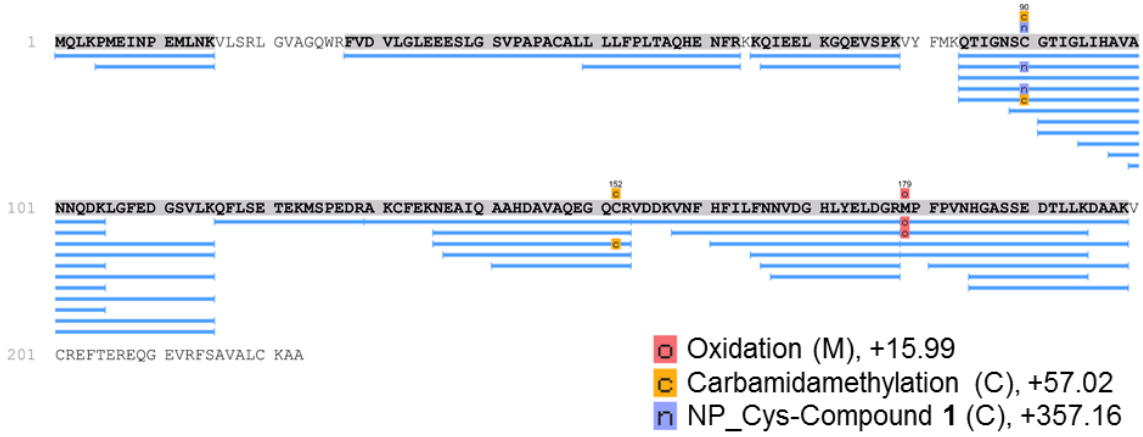
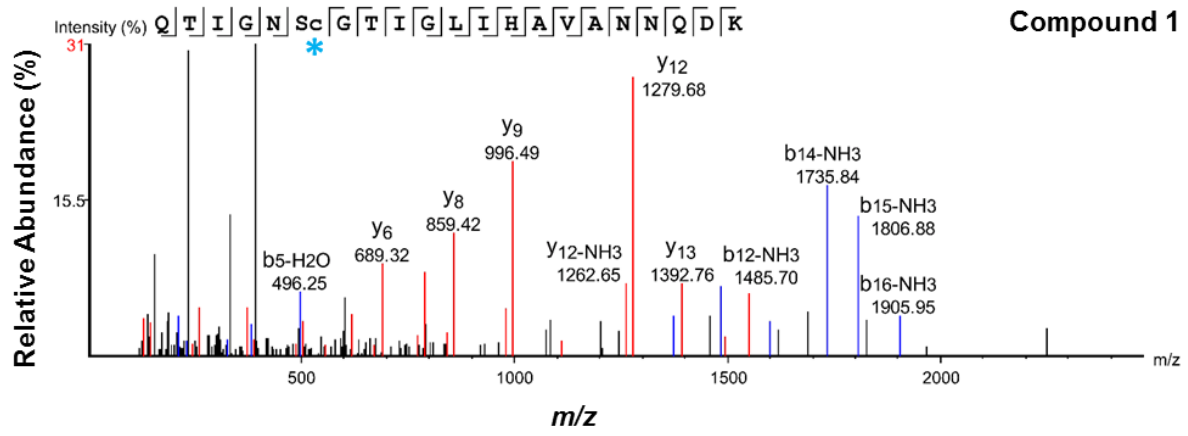


Figure S8. In-gel fluorescence analysis of ABP 2 stability.
Probe stability was assessed by measuring incorporation *via* in-gel fluorescence after incubation in cell culture medium. ABP **2** was incubated in medium for 2, 6, 24, 48 or 72 h. Probe-containing medium was then transferred to HEK293 cells and further incubated for 1 h, flowed by lysis, AzTB functionalization and in-gel fluorescence analysis. Protein loading was assessed by Coomassie Blue staining.

a) **1** modification: 82% peptide sequence coverage





b) **2** modification: 96% peptide sequence coverage

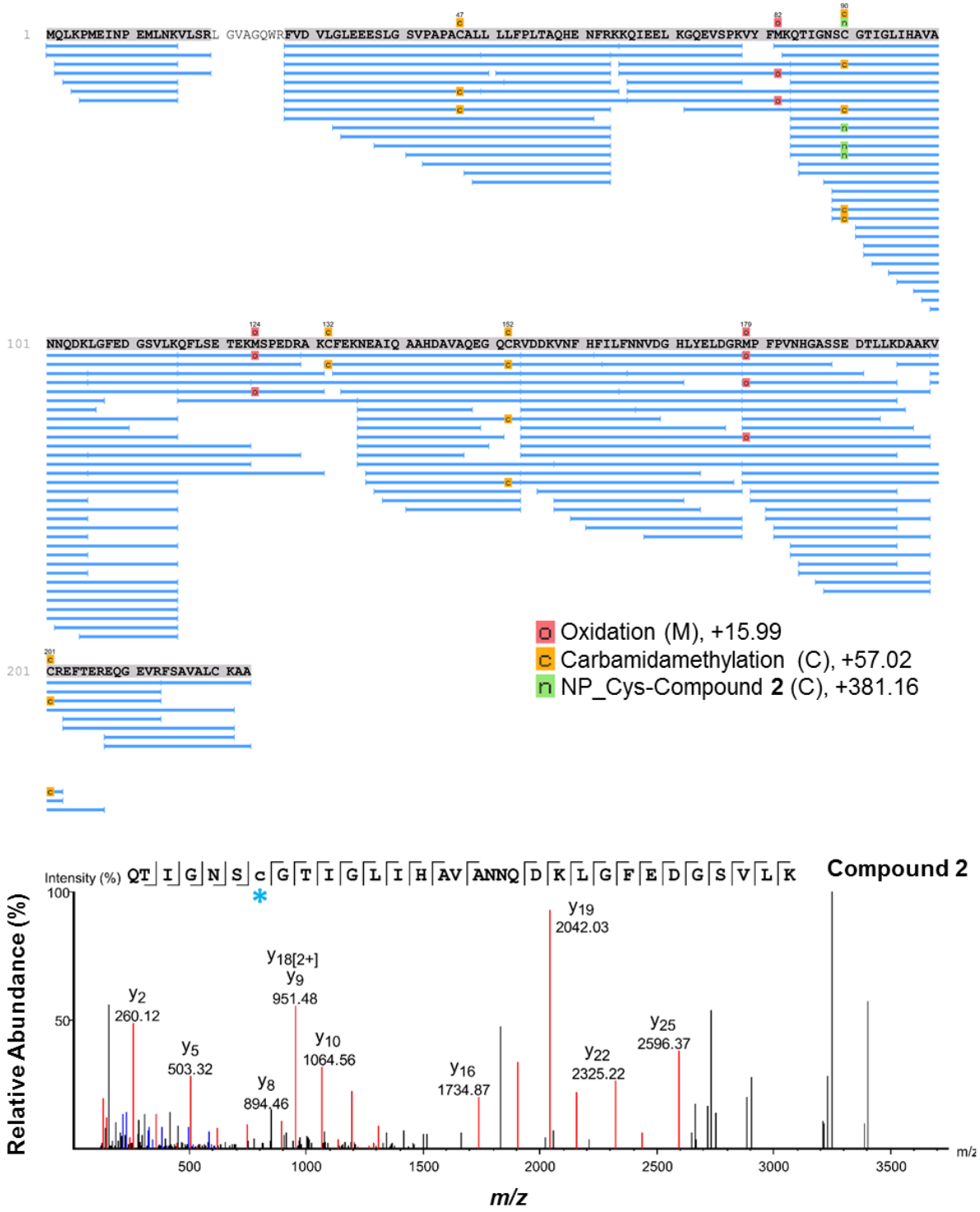


Figure S9. Identification of modification sites by nanoLC-MS/MS.
Recombinant UCHL1 (5 µM) was incubated with compound **1** or **2** (13 µM) for 1 h at room temperature. The labeled protein was digested with trypsin and analyzed by nanoLC-MS/MS. The data was processed in PEAKS 8.0. 82% and 96% of UCHL1 peptide sequence coverage for (a) **1**, or (b) **2** modification, respectively. The high confidence modification of compound **1** or **2** was observed on catalytic cysteine 90(*). Statistical data and peptide lists from PEAKS are detailed in Supplementary Data S1.

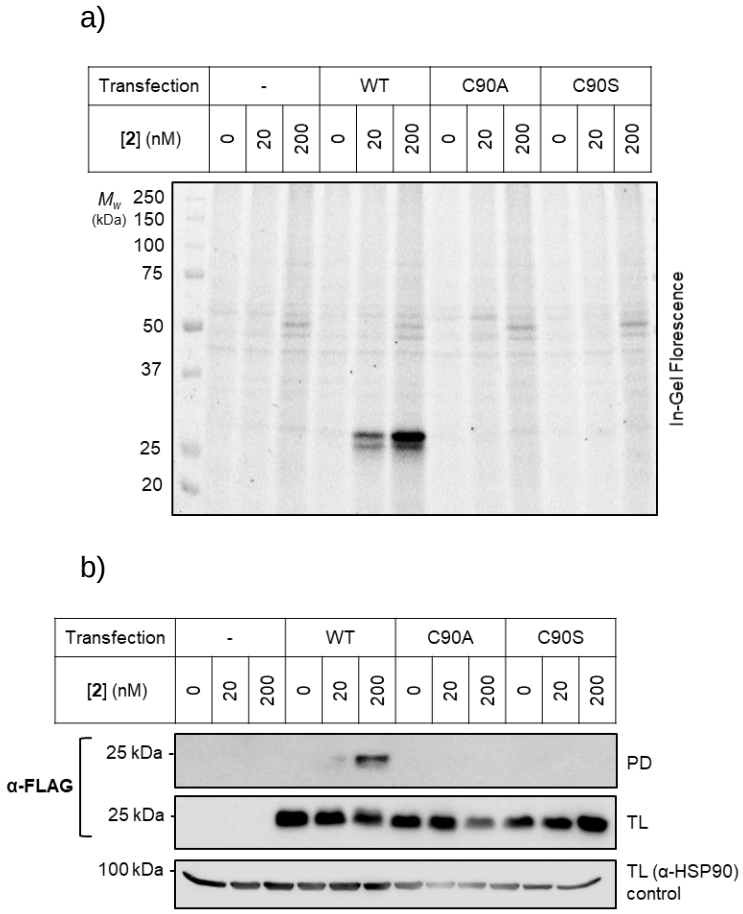


Figure S10. In cell UCHL1 modification site identification.
HeLa cells overexpressing FLAG-UCHL1 (WT, C90A or C90S) were incubated with 20 or 200 nM of **2** for 1 h. (a) Lysates were functionalized with AzTB, labeled proteins separated by SDS-PAGE and visualized by in-gel fluorescence. (b) The labeled proteins were enriched using Streptavidin magnetic beads. The total lysate (TL) and pulldown (PD) were separated by SDS-PAGE and analyzed by immumoblotting using α-FLAG antibody.

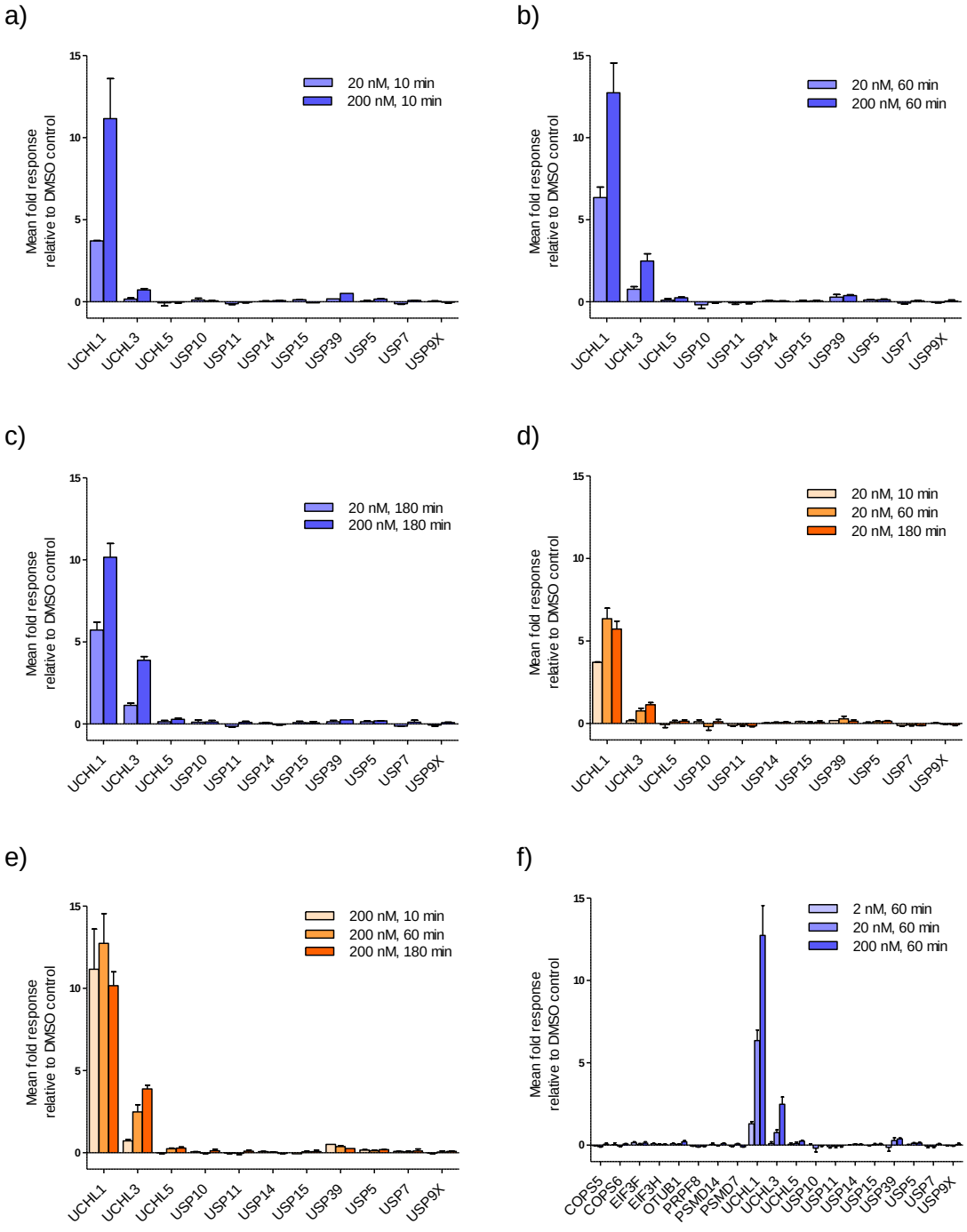


Figure S11. ABP 2 shows concentration- and time-dependent DUB incorporation. HEK293 cells were treated with ABP **2** (20, 200 nM) or vehicle (DMSO) for 10, 60 or 180 min. Following CuAAC ligation to the AzRB cleavable capture reagent, probe-labeled proteins were enriched on dimethylated NeutrAvidin agarose beads. Enriched proteins were digested on-bead with LysC followed by in solution tryptic digestion. Peptides were labeled with 10-plex tandem mass tag (TMT) reagents, desalted and analyzed by nanoLC-MS/MS. Data was processed in MaxQuant and analyzed in Perseus as described in Materials and Methods. Data represent mean ± SEM (n=3). Selected DUBs were normalised to DMSO controls and compared side-by-side. (a-c) Concentration-dependent incorporation of **2**; (d-e) Time-dependent incorporation of **2**; (f) DUBs identified in this experiment in HEK293 cells. The protein lists from MaxQuant and Perseus analysis are detailed in Supplementary Data S2.

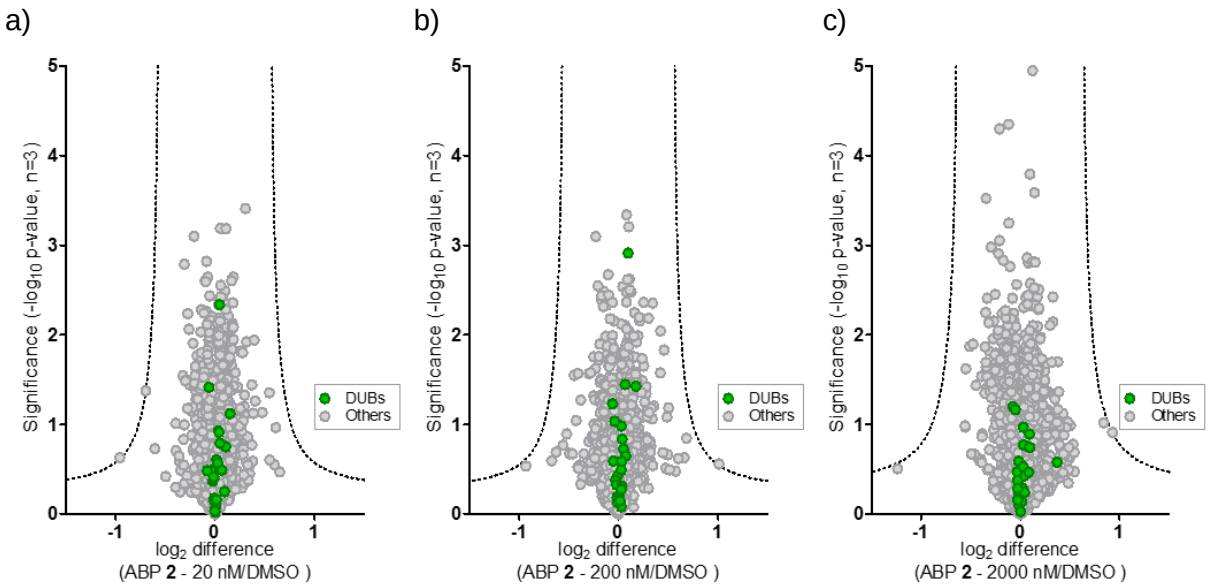
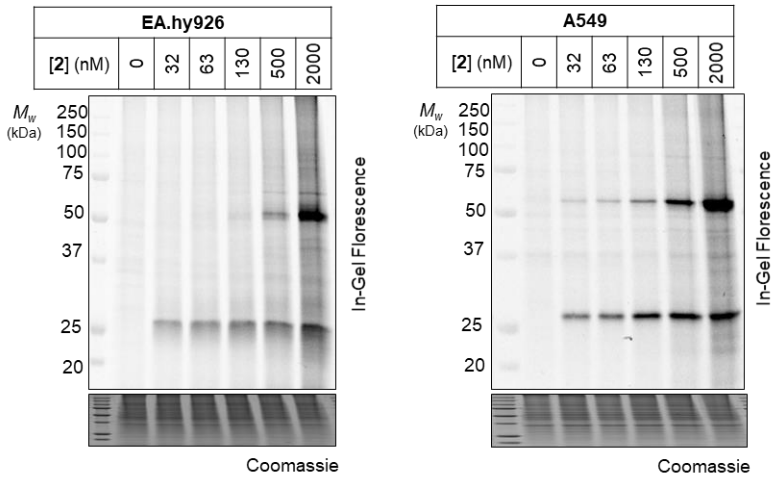


Figure S12. ABP 2 does not affect global protein levels.
HEK293 cells were treated with ABP **2** ((a) 20 nM, (b) 200 nM or (c) 2000 nM) or vehicle (DMSO) for 60 min. Following cell lysis, reduction and alkylation, the samples were tryptic digested and labeled with 10-plex tandem mass tag (TMT) reagents, desalted and analyzed by nanoLC-MS/MS. Data was processed in MaxQuant and analyzed in Perseus as described in Materials and Methods. Volcano plots showing the $\log_2$ difference (fold change) and significance ($-\log_{10}$ p-value) between protein enrichment at different concentration of ABP **2** (20, 200 or 2000 nM) and DMSO control (two sample t-test, n=3, Permutation-based FDR=0.01, S0=1). ABP **2** treatment does not affect protein abundance across the whole proteome. The protein lists from MaxQuant and Perseus analysis are detailed in Supplementary Data S3.

a)



b)

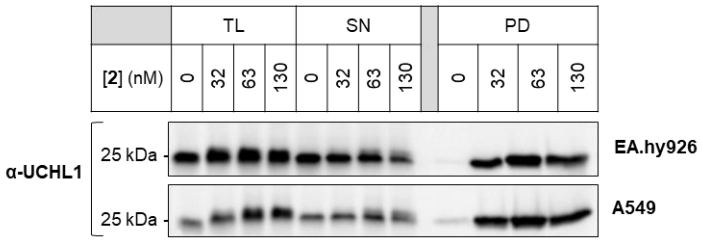


Figure S13. ABP 2 target engagement in A549 and EA.hy926 cells.
EA.hy926 and A549 cells were incubated with 32, 63, 130, 500 or 2000 nM ABP **2** for 3 h, and labeled proteins ligated to AzTB capture reagent. (a) Labeled proteins were separated by SDS-PAGE and visualised by in-gel fluorescence. (b) Labeled proteins were enriched by NeutrAvidin magnetic beads. Total lysate (TL), supernatant (SN) and pulldown (PD) samples were separated by SDS-PAGE and analyzed by immunoblotting using anti-UCHL1 antibody.

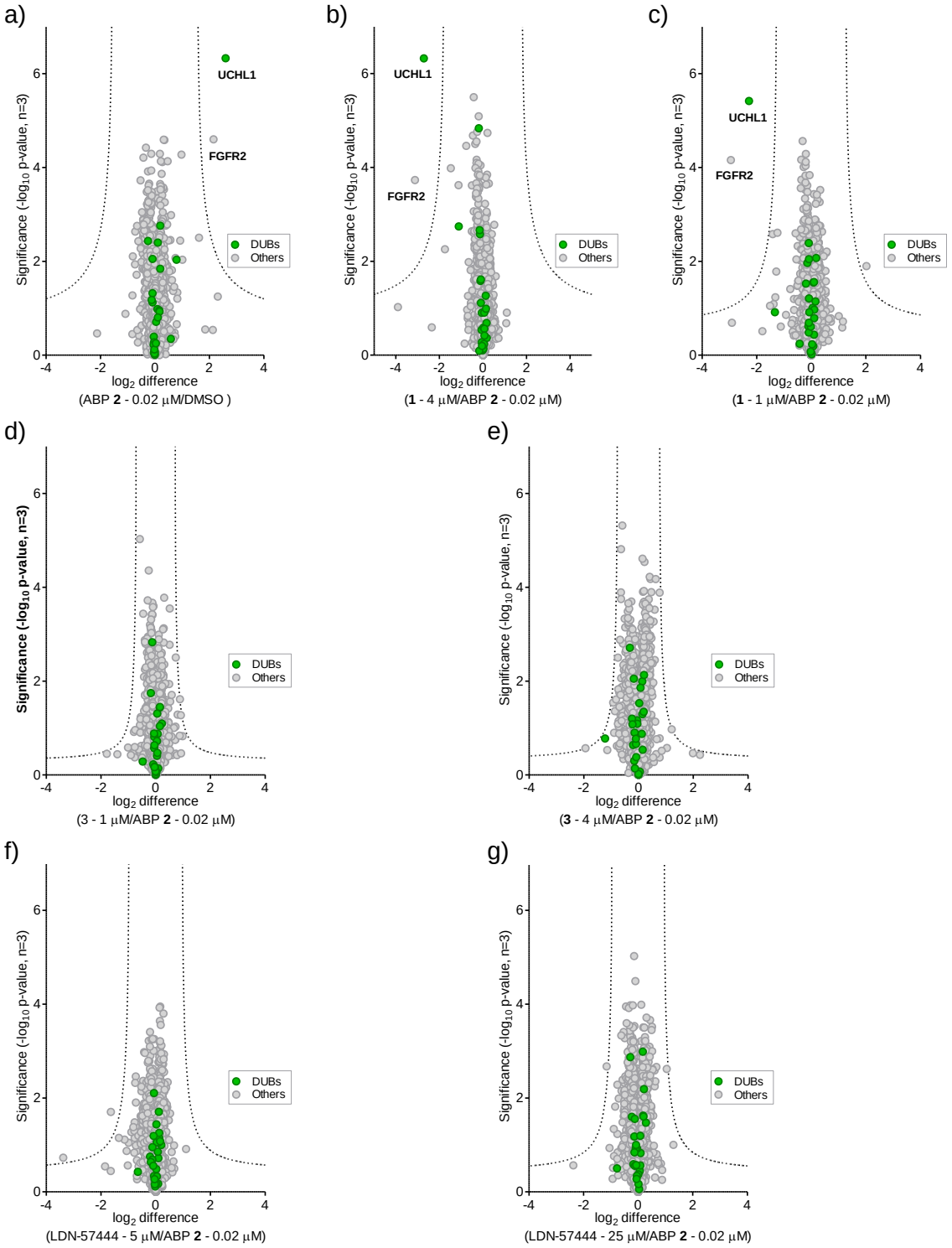


Figure S14. Competitive activity-based proteome profiling.
Competitive activity-based proteome profiling of ABP **2** with UCHL1 inhibitor **1**, control compound **3** or LDN-57444 competition using quantitative proteomics. HEK293 were incubated with (a) vehicle (DMSO), compound (b, c) **1** (1, 4 µM), (d, e) **3** (1, 4 µM) or (f, g) LDN-57444 (5, 25 µM) for 60 min and further incubated with ABP **2** (20 nM) for 10 min. Following CuAAC ligation to AzRB cleavable capture reagent, probe-labeled proteins were enriched on dimentylated NeutrAvidin agarose beads and on-bead digested with LysC followed by in solution tryptic digestion. Peptides were labeled with 10-plex TMT reagents and analyzed by nanoLC-MS/MS. Data was processed in MaxQuant and analyzed in Perseus as

described in Materials and Methods n. Volcano plots show $\log_2$ difference (fold change) and significance ($-\log_{10}$ p-value) of inhibitor-treated samples versus non-treated samples (two sample t-test, n=3, permutation-based FDR=0.01, S0=1). The protein lists from MaxQuant and Perseus analysis are detailed in Supplementary Data S4.

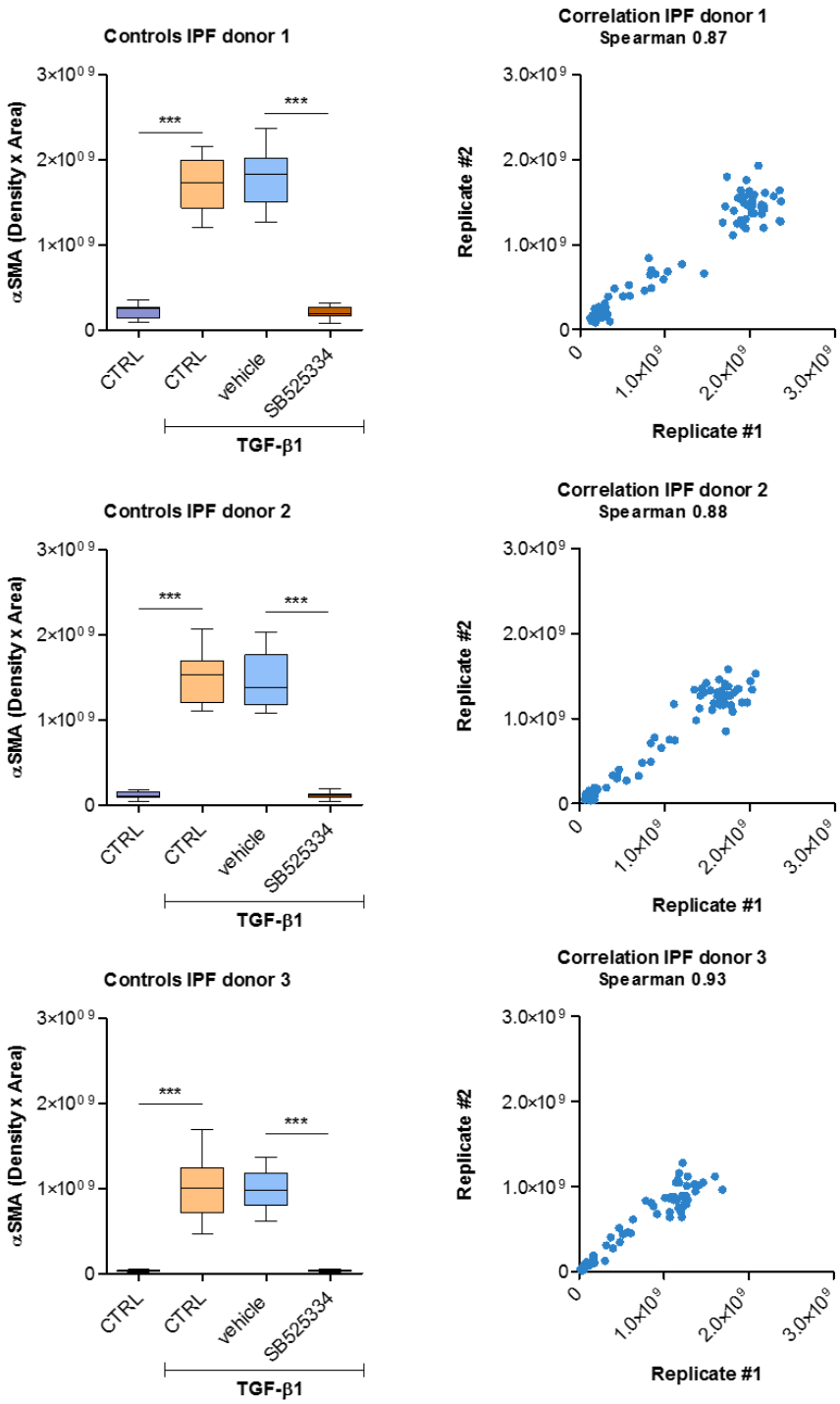
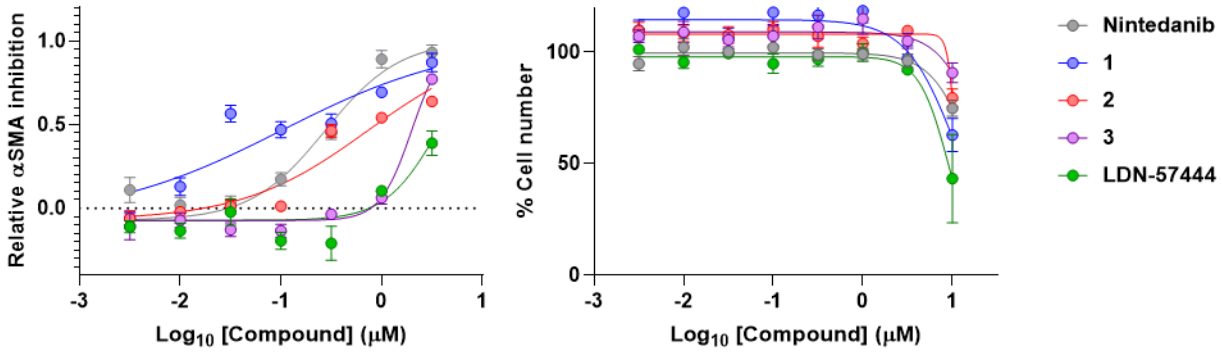


Figure S15. Quality control of fibroblast-to-myofibroblast transition (FMT) assay. Quality control of FMT assay in human lung cells derived from idiopathic pulmonary fibrosis (IPF) patients. Patient-derived primary human bronchial fibroblasts from three donors were FMT triggered by 1.3 ng/mL TGF-β1 in the presence or absence of positive control inhibitor (1 µM SB525334). After 3 days incubation, cells were fixed and stained using αSMA for imaging using high content analysis (HCA). αSMA immunoreactivity was quantified by an HCA algorithm with density x area output. A Kruskal–Wallis test with Dunn's post-test was performed (***$P \leq 0.01$). Plots represent median values (center lines) and 25th/75th percentiles (box limits) with Tukey whiskers.



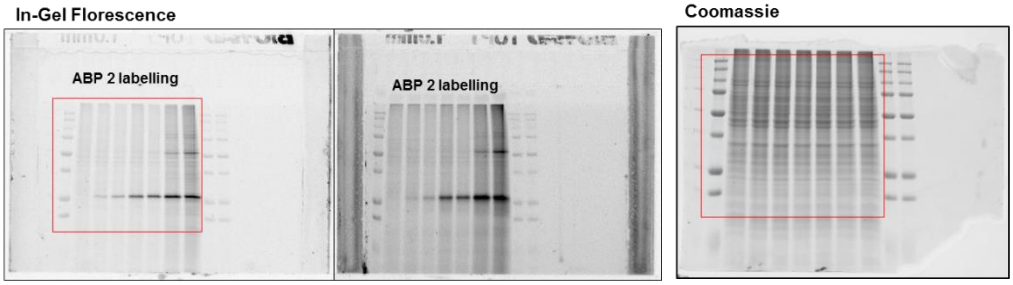
Compound	αSMA inhibition IC_{50} (95% CI)	Cell number EC_{50}
Nintedanib	0.28 (0.20-0.40) μM	~18 μM
1	0.10 (0.048-0.22) μM	~11 μM
2	0.74 (0.42-1.3) μM	~11 μM
3	2.0 (1.6-2.5) μM	~23 μM
LDN-57444	3.7 (2.4-5.5) μM	~9.2 μM

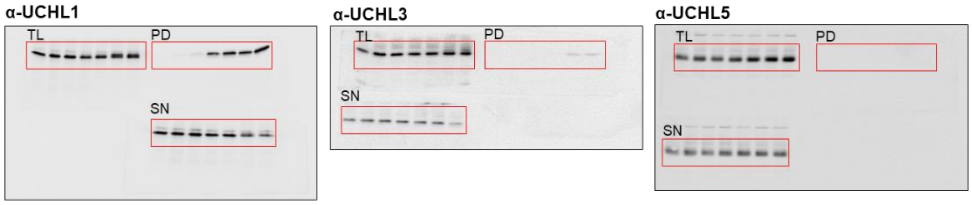
Figure S16. Effect of UCHL1 inhibitors on fibrotic activity and cell number.
Patient-derived primary human bronchial fibroblasts from three donors were FMT trigged by 1.3 ng/mL TGF-β1 in combination with 8-point serial dilution of reference compound nintedanib, compound **1**, **2**, **3** or LDN-57444. After 3 days incubation, the cells were fixed and stained using αSMA and DAPI staining and imaged using HCA. Data represent mean ± SEM (n=6) of relative αSMA inhibition and percentage of cell number. αSMA IC_{50} and EC_{50} values were calculated from a four-parameter dose-response curve.

Figure S17. Uncropped gels/blots

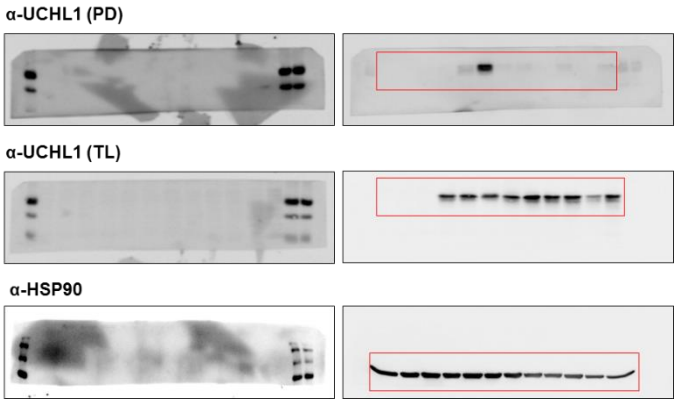
- Uncropped gels/blots for Figure 1

- Uncropped gels/blots for Figure 2

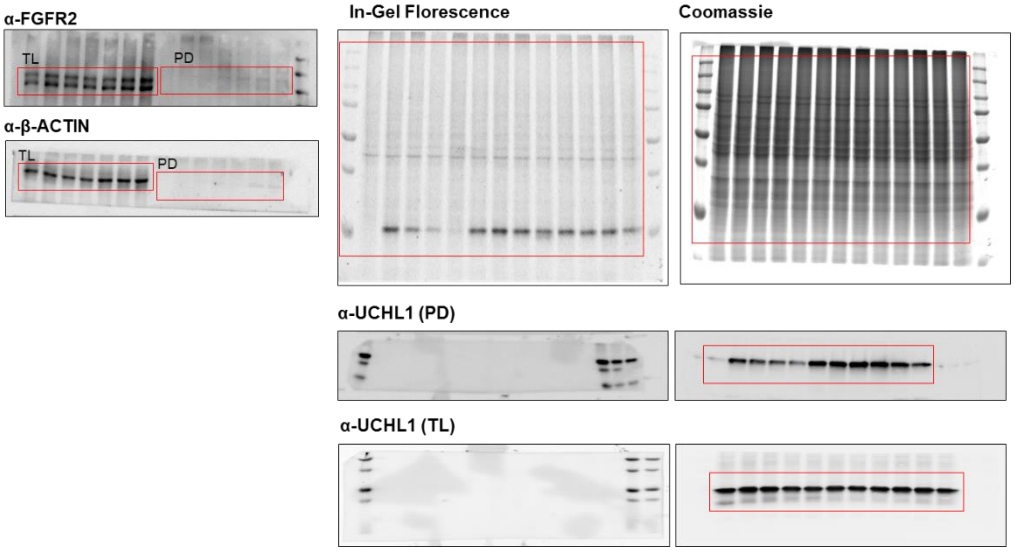




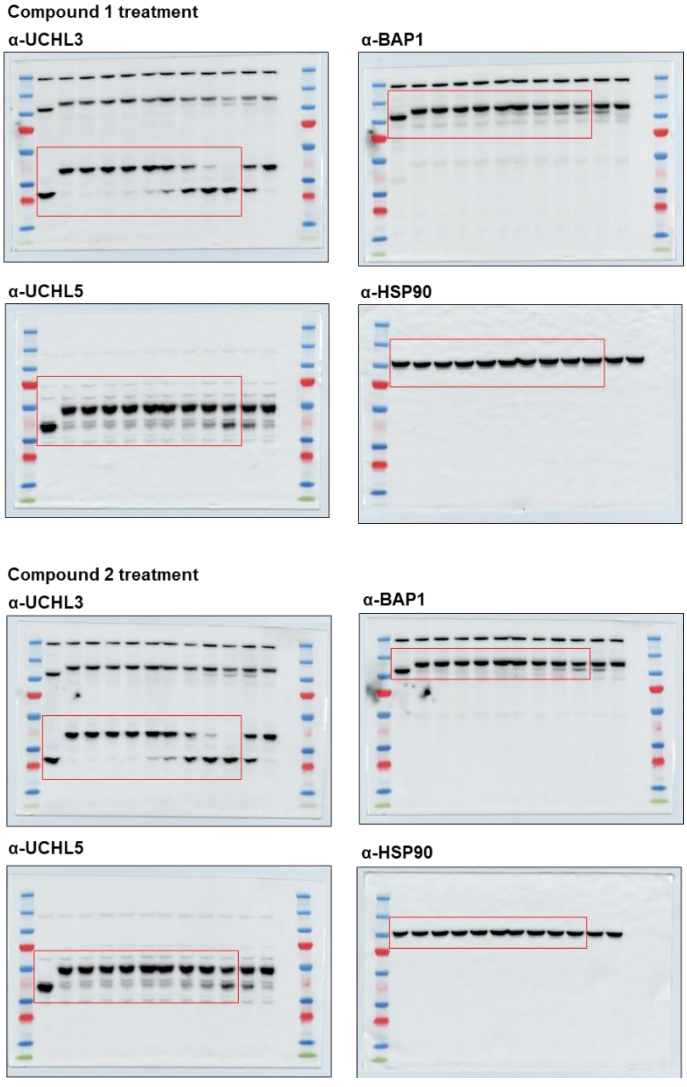
- Uncropped gels/blots for Figure 3



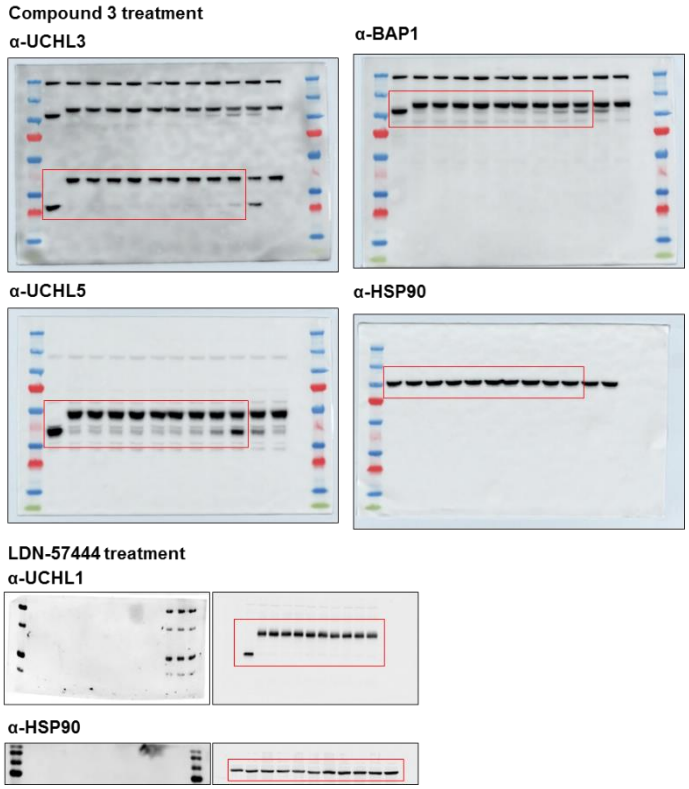
- Uncropped gels/blots for Figure 4



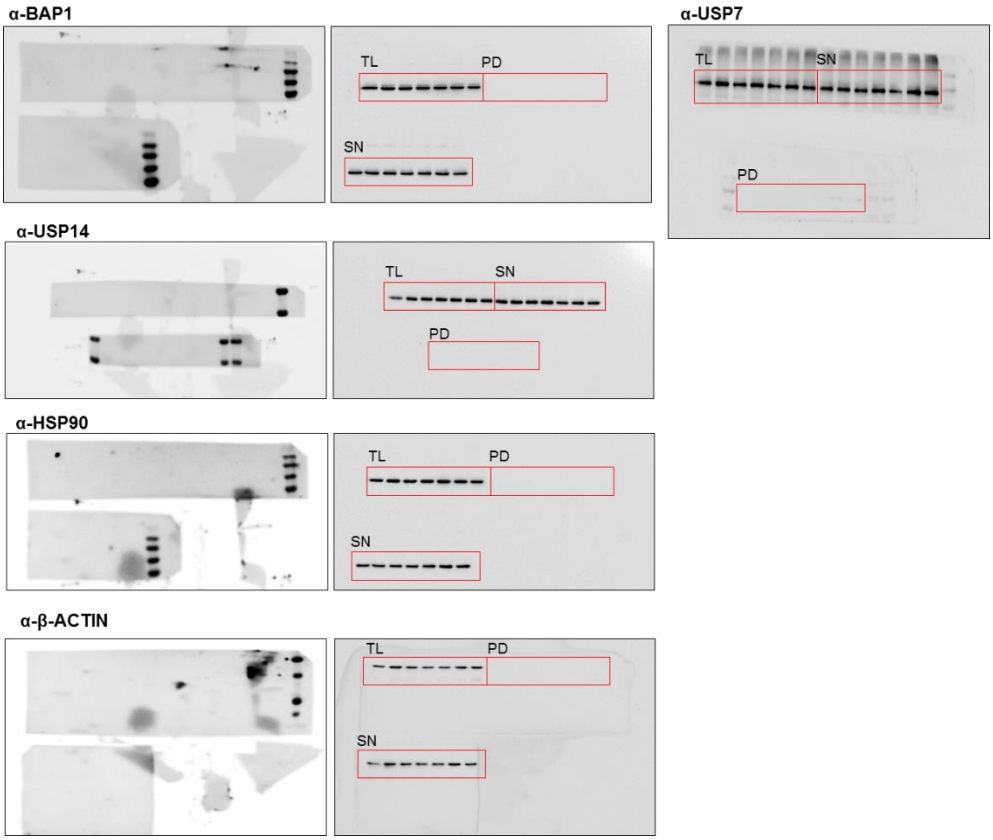
- Uncropped gels/blots for Figure S4



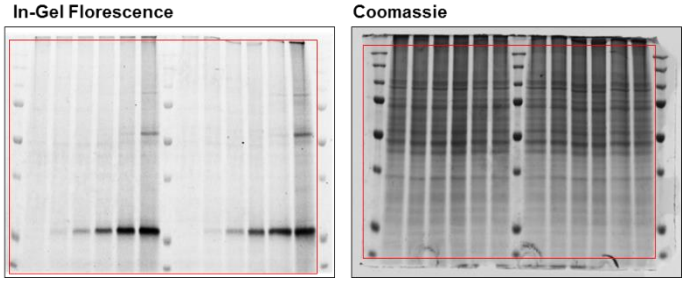
- Uncropped gels/blots for Figure S4 (cont.)



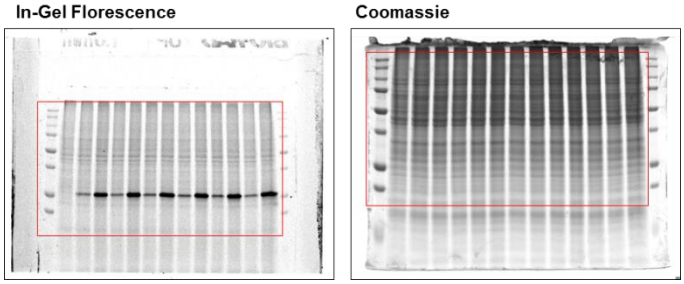
- Uncropped gels/blots for Figure S6



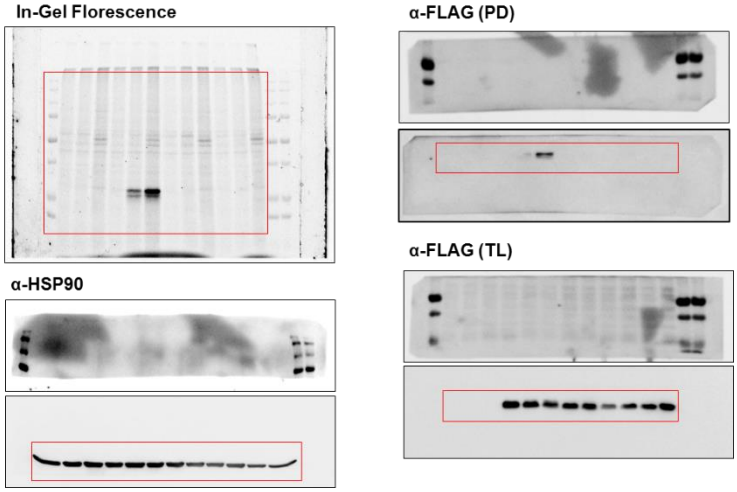
- Uncropped gels/blots for Figure S7



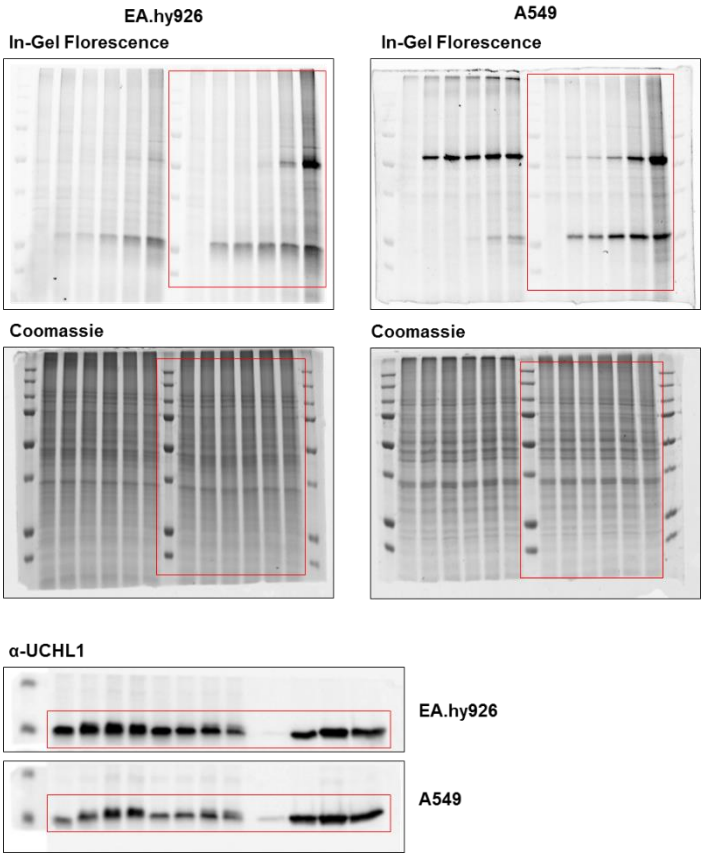
- Uncropped gels/blots for Figure S8



- Uncropped gels/blots for Figure S10



- Uncropped gels/blots for Figure S13



Supplementary Tables

Ranges: **A** < 0.1 µM
0.1 < **B** < 1 µM
1 < **C** < 10 µM
D > 10 µM

DUBs	Biochemical IC_{50} range				
	1	**2**	**3**	**LDN-57444**	
USP1	D	D	D	D	USP family
USP4	D	D	D	D	
USP8	D	D	D	D	
USP10	D	D	D	D	
USP16	D	D	NA	D	
USP19	D	D	D	D	
USP21	D	D	D	D	
USP22	D	D	D	D	
USP25	D	D	D	D	
USP28	D	D	NA	D	
USP30	C	C	D	D	
USP34	C	C	D	D	
USP35	D	D	D	D	
BAP1	C	C	D	D	UCH family
UCHL1	**A**	**A**	**D**	**D**	
UCHL3	C	C	D	D	
UCHL5	D	D	D	D	
Cezanne1	D	D	D	D	OTU family
OTUB2	D	D	D	D	
TRABID	D	D	D	D	

Table S1. Activity profiling of 1, 2, 3 and LDN-57444 against 20 DUBs.
Biochemical IC_{50} values were measured by fluorescence polarization (FP) and fluorescence intensity (FI) assays using Ub-Lys-TAMRA, Ub-FANCD2-TAMRA, and Ub-Rho110 substrates (NA=Not assessed).

Target	Species	Dilution	Catalogue no.	Supplier
BAP1	Mouse	1:500	sc-28383	Santa cruz
FLAG, HRP	Mouse	1:1000	A8592	Sigma-Aldrich
HA-tag (C29F4)	Rabbit	1:1000	3724	Cell Signaling Technology
HSP90	Mouse	1:1000	sc-69703	Santa Cruz Biotechnology
UCHL1/PGP9.5	Rabbit	1:2000	14730-1-AP	ProteinTech
UCHL3	Mouse	1:1000	sc-100340	SantaCruz
UCHL5/UCH37	Rabbit	1:1000	ab124931	Abcam
USP14	Rabbit	1:2000	11931S	Cell Signalling
USP7/HAUSP	Rabbit	1:1000	ab4080	Abcam
β-ACTIN	Rabbit	1:5000	Ab8227	Abcam
Mouse IgG, HRP	Mouse	1:10000	R-05071-500	Advansta
Rabbit IgG, HRP	Rabbit	1:10000	R-05072-500	Advansta

Table S2. Primary and secondary antibodies used.
Primary and secondary antibodies were used at indicated concentrations for immunoblotting as described in Materials and Methods.

DUB	Assay Buffer	Substrate	Substrate Concentration (nM)
BAP1	A	Ub-Lys-TAMRA	50
Cezanne1	B	Ub-Lys-TAMRA	50
OTUB2	B	Ub-Lys-TAMRA	50
TRABID	B	Ub-Rho110	400
UCHL1	A	Ub-Lys-TAMRA	50
UCHL3	A	Ub-Lys-TAMRA	50
UCHL5	A	Ub-Lys-TAMRA	50
USP1	A	Ub-FANCD2-TAMRA	50
USP4	A	Ub-Lys-TAMRA	50
USP8	B	Ub-Lys-TAMRA	50
USP10	B	Ub-Rho110	400
USP16	B	Ub-Lys-TAMRA	50
USP19	A	Ub-Lys-TAMRA	50
USP21	A	Ub-Lys-TAMRA	50
USP22	B	Ub-Rho110	400
USP25	B	Ub-Lys-TAMRA	50
USP28	B	Ub-Lys-TAMRA	50
USP30	B	Ub-Lys-TAMRA	50
USP34	B	Ub-Rho110	400
USP35	B	Ub-Lys-TAMRA	50

Table S3. Conditions of DUB selectivity profiling assays.
The selectivity profile of compounds **1**, **2**, **3** and LDN-57444 for inhibition of the DUB panel was assessed by FP (TAMRA) or FI (Rho110) using the indicated substrates and concentrations in buffer A (20 mM Tris, pH 7.5, 100 mM NaCl, 0.05% Tween-20, 0.5 mg/mL BSA, 5 mM BME) or buffer B (40 mM Tris, pH 7.5, 0.005% Tween-20, 0.5 mg/mL BSA, 5 mM BME).

Materials and Methods

Abbreviations

AzRB (azido-arginine-biotin); AzTB (azido-TAMRA-biotin); BME (β-mercaptoethanol); BSA (bovine serum albumin); CuAAC (copper(I)-catalyzed azide-alkyne cycloaddition); DCM (dichloromethane); DMF (dimethylformamide); DMSO (dimethyl sulfoxide); DIPEA (*N*,*N*-diisopropylethylamine); EDTA (ethylenediaminetetraacetic acid); FI (fluorescence intensity); FP (fluorescence polarization) GST (glutathione-*S*-transferase); HATU (1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate); HA-Ahx-Ahx-Ub-VME (HA-tagged ubiquitin vinyl methylester); HCD (higher-energy collisional dissociation); HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid); IPTG (isopropyl-β-D-thiogalactoside); LB (lysogeny broth); PBS (phosphate buffered saline); SDS (sodium dodecyl sulfate); SDS-PAGE (SDS-Polyacrylamide gel electrophoresis); TBTA (tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine); TCEP (tris(2-carboxyethyl)phosphine hydrochloride); TFA (trifluoroacetic acid); THF (tetrahydrofuran); TLC (thin-layer chromatography); TMT (tandem mass tag); TRIS (tris(hydroxymethyl)aminomethane); Ub-AMC (ubiquitin-7-amino-4-methylcoumarin) Ub-FANCD2-TAMRA (ubiquitin-Fanconi anemia group D2-TAMRA); Ub-Lys-TAMRA (ubiquitin-ε-Lys-TAMRA); Ub-Rho110 (ubiquitin-Rhodamine 110); XPhos Pd G2 (2nd Generation XPhos Precatalyst, chloro(2-dicyclohexylphosphino-2′,4′,6′-triisopropyl-1,1′-biphenyl)[2-(2′-amino-1,1′-biphenyl)]palladium(II)).

Reagents

Reagents were purchased from Sigma-Aldrich Chemical Co. Ltd., Acros Organics Ltd., VWR international Ltd. or Fluorochem Ltd. and were used without further purification. Common organic solvents used in chemical reactions (THF, DMF, DCM, 1,4-dioxane, and MeOH) were purchased anhydrous from Acros Organics Ltd. over molecular sieves (where available) in AcroSeal® bottles and handled under nitrogen. Water used for chemical reactions was obtained from a Sartorius Arium® Mini Plus lab water system (ASTM Type 1 ultrapure). For purifications, HPLC-grade solvents (≥ 99% purity) were used as purchased from Sigma-Aldrich Chemical Co. Ltd. or Fisher Scientific UK. For reaction work-up, technical grade solvents (≥ 95% purity) were used as purchased from VWR or Fisher. LDN-57444 was obtained from Tocris Bioscience (Cat. No. 3998). Recombinant DUBs were purchased from Boston Biochem or Ubiquigent or produced in-house as described below. The synthesis of capture reagents AzTB, and AzRB have been described previously.[1] All compounds were dissolved in DMSO and stored at -20°C.

Biological and Biochemical Methods and Proteomics

Recombinant UCHL1 production

UCHL1 expression plasmid (GST-tagged UCHL1) was a generous gift from Dr C. Das (Purdue University, USA). Full-length human UCHL1 DNA (residues 1-223, Uniprot P09936) was produced as previously described,[2] with the following modifications. The GST-UCHL1 DNA construct in pGEX-6P-1 vector was transformed into *E. coli* Rosetta™ 2(DE3) Singles™ competent cells (Novagen). A single transformed colony was inoculated into LB (50 mL) supplemented with chloramphenicol (34 µg/mL) and ampicillin (100 µg/mL) and cells grown overnight at 37 °C with shaking (220 rpm). The starter culture (40 mL) was expanded in LB (2 L) supplemented with chloramphenicol (34 µg/mL) and ampicillin (100 µg/mL) and incubated at 37 °C with shaking (220 rpm). After reaching an OD_{600} of 0.6, protein expression was induced by addition of IPTG (0.5 mM) at 18 °C for 18 h with shaking (220 rpm). Cells were harvested by centrifugation (4,000 rpm, 4 °C, 30 min). The resulting cell pellet was re-suspended in lysis buffer (50 mL PBS, pH 7.4 (Sigma-Aldrich) containing cOmplete™ EDTA-free protease inhibitor cocktail (Roche) and DNase I (Invitrogen)) and lysed *via* one passage at 25 kPSI through a Cell Disruptor (Constant Systems). The lysate was centrifuged (15,000 rpm, 4 °C, 45 min) and the supernatant was affinity purified using Glutathione

Sepharose® 4B resin (GE Healthcare). The UCHL1 protein was cleaved from GST resin using PreScission™ Protease (GE Healthcare) in cleavage buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.5) and concentrated using a PES Protein Concentrator (Pierce, 10 kDa molecular weight cut-off (MWCO)). The protein was separated from the resin and cleared by centrifugation (13,000 rpm, 4 °C, 5 min) and purified using a Superdex® S75 10/300 GL column (GE Healthcare), eluting with cleavage buffer. Fractions containing protein were pooled and concentrated using Amicon® Ultra-0.5 Centrifugal Filter Units (10 kDa MWCO (Merck Millipore), flash frozen, and stored at -80 °C. Protein concentration was estimated by measuring the absorbance at 280 nm on a NanoDrop™ One UV-Vis spectrophotometer (Thermo Fisher Scientific), using a UCHL1 molar extinction coefficient of 8,480 $M^{-1}cm^{-1}$.

Biochemical UCHL1 fluorescence polarisation (FP) assay using Ub-Lys-TAMRA
Reactions were performed in duplicate in black 384 well plates (Greiner #784076) in a final reaction volume of 21 µL. Compound dilutions were prepared at 21× final concentration (2100 µM for a final concentration of 100 µM) in 50% DMSO/water in a 96-well polypropylene V-bottom plate (Greiner #651201). 50% DMSO (1 µL) or diluted compound (1 µL) was added to the plate. UCHL1 was diluted in reaction buffer A (20 mM Tris, pH 7.5, 100 mM NaCl, 0.05% Tween-20, 0.5 mg/mL BSA, 5 mM BME) and 10 µL of diluted UCHL1 was added to the compound (final concentration 5 nM). Enzyme and compound were incubated for 30 min at room temperature. Reactions were initiated by the addition of 50 nM of TAMRA labeled peptide linked to ubiquitin *via* an iso-peptide bond (Ub-Lys-TAMRA, UbiQ Bio). Reactions were read immediately after substrate addition and following a 2 h incubation at room temperature using a Pherastar Plus (BMG Labtech; λ_{Ex} 540 nm, λ_{Em} 590 nm). IC_{50} values were determined by plotting percentage UCHL1 activity against compound concentration and fitting a four-parameter dose-response curve (Graphpad Prism 5.03).

Biochemical DUB selectivity profiling assays
DUB selectivity assays were performed as outlined in
Table S3. For USP10, USP22, USP34 and TRABID fluorescence intensity assays, enzymes were diluted in reaction buffer B (40 mM Tris, pH 7.5, 0.005% Tween-20, 0.5 mg/mL BSA, 5 mM BME), reactions were initiated by the addition of 400 nM Ub-Rho110 (Boston Biochem, U-555), and read using a Pherastar Plus (λ_{Ex} 487 nm, λ_{Em} 535 nm). Enzyme concentrations were optimized for each DUB to ensure activity was in the linear range.

UCHL1 kinetic studies
UCHL1 was added to Ub-AMC substrate (UbiQ Bio) in black flat bottom 384-well microplates (Corning #3575) in reaction buffer C (100 mM sodium phosphate, pH 8.5, 0.01% Tween-20) in 40 µL total reaction volume. The florescence intensity was read at room temperature using a 2104 EnVision™ Multilabel reader (Perkin Elmer; λ_{Ex} 380 nm, λ_{Em} 460 nm).

Michaelis-Menten kinetic analysis: Michaelis constant (K_m) determination was performed by preparing Ub-AMC substrate (20 µL) at 2× final concentration in reaction buffer C. Reactions were initiated by adding UCHL1 (20 µL, 6 nM) and florescence intensity recorded immediately over 1 h. K_m was calculated by plotting initial velocities against Ub-AMC concentration with a non-linear fit of the Michaelis–Menten equation (Graphpad Prism 5.03).

Inhibitor kinetic analysis: Inhibition kinetics for compound **1** and **2** were determined by preparing compound dilution plates at 40× final concentration in DMSO. DMSO (1 µL) or diluted compound (1 µL) was added to reaction buffer C (9 µL) in the reaction plate. Substrate Ub-AMC (500 nM, 10 µL) was prepared in buffer C and added to the reaction plate. Reactions were initiated by the addition of UCHL1 (6 nM, 20 µL), and florescence intensity read immediately for 1 h. The apparent *rate of inhibition (k_a) at each compound concentration was determined by curve fitting using non-linear regression analysis in GraphPad Prism 5.03. k_a

values were plotted against compound concentration to afford the apparent second order rate constant for inhibition, k_a/I. The absolute value of the second order rate constant k_{obs}/I was calculated by using $k_{obs}/I = k_a/I \times (1+[S]/K_m)$ where K_m of UCHL1 and Ub-AMC is ~0.22 μM.[3-4]

Reversible inhibition assay: Jump dilution assays were performed by incubating UCHL1 (300 nM) with compound **1, 2** (400 nM), LDN-57444 (1700 nM), or positive (DMSO) and negative controls (iodoacetamide (IAA) 10 mM) for 30 min at room temperature. The reaction mixtures were diluted 100× with buffer C containing Ub-AMC (125 nM), and the florescence intensity read immediately for 1 h.

UCHL1 covalent adduct determination

UCHL1 (5 μM, 50 μL) in buffer D (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.5) was incubated with compound **1**, **2** (2.5 equiv.) or DMSO for 1 h at room temperature. The excess ligand was removed by washing with buffer D in an Amicon® Ultra MWCO 10 kDa centrifugation spin filter three times. The protein-adducts were desalted using 3× SDB-XC polystyrene-divinylbenzene copolymer extraction disks, following published methods.[5] Samples were eluted with 50% (v/v) MeCN 0.5% (v/v) formic acid in in LC-MS grade water and analyzed by LC-ESI (ACQUITY UPLC H-Class PLUS instrument (Waters)) using an ACQUITY UPLC BEH C_4 column (Waters) coupled with Waters LCT Premier MS detector.

UCHL1 covalent adduct site identification

Compound-protein adduct samples were prepared as described above and digested with trypsin (Promega, 1:50 trypsin:UCHL1 ratio) at 37 °C, overnight. The resulting peptides were desalted using SDB-XC disks and eluted with 80% (v/v) MeCN in MilliQ water and dried in a Savant SPD1010 SpeedVac® Concentrator (Thermo Fisher Scientific). Dried peptides were dissolved in 2% (v/v) MeCN, 0.5% (v/v) TFA in LC-MS grade water and analyzed on a Q Exactive Orbitrap nanoLC-MS/MS *via* an EASY-Spray™ source (Thermo Fisher Scientific). The MS data was analyzed by PEAKS Studio 8.5 (Bioinformatics Solutions Inc.).

Cell Culture

Cal51 and HEK293T cells were culture in Dulbecco's Modified Eagle's Medium (DMEM) high glucose (Gibco #41965) supplemented with 10% fetal bovine serum (FBS, Gibco® #10500) and penicillin/streptomycin (P/S, Gibco #15070). HEK293, A549, EA.hy926 and HeLa cells were culture in DMEM low glucose (Sigma-Aldrich #D6046) supplemented with 10% FBS. All cells were maintained in a humidified incubator at 37 °C and 5% CO_2. Cells were seeded on plates at least 24 h before experiment.

Homogeneous time resolved fluorescence (HTRF) assay

Cal51 cells stably expressing wild-type or Cys90Ala mutant FLAG-UCHL1 were seeded in a 96-well plate (Nunc #167008) at 45,000 cells/well and incubated overnight at 37 °C, or seeded at 75000 cells/well and incubated for 4 h at 37 °C. Compounds were diluted in medium and incubated with cells for 1 h at 37 °C. Cells were washed twice with PBS, lysed in 20 μL of cold lysis buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.1% NP-40, 0.5% CHAPS, 5 mM $MgCl_2$, 5 mM BME, cOmplete™ EDTA-free protease inhibitor cocktail (Roche) and phosphatase inhibitor (Roche)) and incubated on ice for 30 min. Lysate (5 μL) was dispensed (in duplicate) in a white 384-well plate (Greiner #784075), combined with Ub-VME probe (100 nM, 5 μL) in HTRF buffer (20 mM Tris pH 7.5, 100 mM NaCl, 0.05% Tween-20, 0.5 mg/mL BSA, 5 mM BME) and incubated for 30 min at room temperature. Detection antibodies (10 μL) were diluted in detection buffer (50 mM HEPES pH 7.5, 400 mM KF, 0.5 mg/mL BSA), 0.033 μL of anti-FLAG cryptate antibody (Cisbio #61FG2KLA) and 0.1 μL of anti HA-XL antibody (Cisbio #610HAXLB), and added to plates. The plates were sealed and incubated overnight at 4 °C then read using a PHERAstar (BMG), and the Delta F calculated:

$$\%\ \text{Delta F} = 100 \times \left(\frac{\text{ratio}^{sample} - \text{ratio}^{negative\ control}}{\text{ratio}^{negative\ control}}\right)$$

$$\text{Ratio} = 10000 \times \left(\frac{\text{fluorescence emission at 665 nm}}{\text{fluorescence emission at 620 nm}}\right)$$

The IC_{50} values were determined by plotting the HTRF signal (% Delta F normalised against the vehicle treated sample) against concentration with a four-parameter dose-response curve (Graphpad Prism 5.03).

In-cell target engagement using HA-Ub-VME

HEK293T cells were seeded in 6-well plates (VWR #734-1596) at 1,000,000 cells/well and incubated overnight at 37 °C. Compounds **1**, **2**, **3** and LDN-57444 were diluted in media and incubated with cells for 1 h at 37 °C (final concentration 0.3% (v/v) DMSO). The media was removed and cells were washed with PBS prior to lysing in TRIS-lysis buffer (50 mM Tris pH 7.5, 50 mM NaCl, 0.1% (v/v) NP-40, 0.5% (w/v) CHAPS, 2 mM $MgCl_2$, 5 mM BME, 10 U Benzonase, cOmplete™ EDTA-free protease inhibitor cocktail (Roche) and phosphatase inhibitors (Roche)). HA-Ahx-Ahx-Ub-VME (0.5 µg) was added to lysate (20 µg) and incubated for 15 min at room temperature. 5× sample loading buffer (160 mM Tris pH 8, 5% (w/v) SDS, 0.025% (w/v) bromophenol blue, 25% (v/v) glycerol) containing 4% (v/v) BME was added and samples heated for 5 min at 95 °C prior to SDS-PAGE and immunoblotting. Blots were developed with Amersham ECL Prime Western Blotting Reagent and imaged using Amersham Imager 680. Inhibition was observed as a loss of HA-Ub-VME-labeled upper band intensity, and percentage inhibition was calculated using Image Quant (GE Healthcare).

Preparation of lysates for ABP 2 profiling and competition

Near-confluent (ca. 90%) plates of HEK293 cells were incubated with media containing **2** (2-2,000 nM, 0.1% (v/v) DMSO) or vehicle for indicated times. For competition experiments, HEK293 cells were pre-incubated with indicated concentrations of compound **1**, **3**, LDN-57444 or DMSO control for 1 h before incubation with 20 nM of **2** for 10 min. In both cases, treated cells were washed twice with PBS (500 µL) and lysed in PBS-lysis buffer (1% (v/v) Triton X-100, 0.1% (w/v) SDS in PBS supplemented with 1× cOmplete™ EDTA-free protease inhibitor cocktail). The lysate was clarified by centrifugation (17,000 *g*, 4°C, 10 min) and the supernatant recovered. The protein concentration was determined using the DC™ protein assay (Bio-Rad) following the manufacturer's protocol. After cell lysis and CuAAC ligation, samples were analyzed by in-gel fluorescence analysis, immunoblotting or proteomics (described below).

Preparation of lysates for ABP 2 stability evaluation

ABP **2** was pre-incubated in DMEM medium for 2, 6, 24, 48 or 72 h in a humidified incubator (37 °C, 5% CO_2) before being transferred to near-confluent HEK293 cells and incubated for 1 h. After cell lysis and CuAAC ligation, samples were analyzed by in-gel fluorescence analysis and Coomassie blue staining (described below).

Preparation of UCHL1 DNA constructs

WT UCHL1 was cloned as a N-terminal FLAG tag fusion protein in the multiple cloning site of pFLAG-CMV-6a plasmid using the HindIII and BglII restriction sites. UCHL1 mutants C90A and C90S were generated by PCR using the QuikChange Lightening Site-Directed Mutagenesis kit (Agilent) and the following primers:

UCHL1 C90A forward: 5'- GCAGACCATTGGGAATTCCGCTGGCACAATCGGACTTATT-3'
UCHL1 C90A reverse: 5'- AATAAGTCCGATTGTGCCAGCGGAATTCCCAATGGTCTGC -3'
UCHL1 C90S forward: 5'- GACCATTGGGAATTCCAGTGGCACAATCGGACT-3'
UCHL1 C90S reverse: 5'- AGTCCGATTGTGCCACTGGAATTCCCAATGGTC -3

Preparation of UCHL1 overexpression cell lysates
HeLa cells were transfected with plasmid encoding for UCHL1 WT, C90A or C90S using Lipofectamine™ 2000 Transfection Reagent (Invitrogen) and Opti-MEM™ reduced serum medium (Gibco™) according to the manufacturer's instructions with the following modifications. DNA (0.5 µg) was mixed with lipofectamine (3 µL) and added to cells in a 6-well plate. After 6 h transfection, medium was removed and replaced by fresh culture medium supplemented with **2** (10 and 100 nM) and incubated for 1 h. After cell lysis and CuAAC ligation, samples were analyzed by in-gel fluorescence and immunoblotting (described below).

Click chemistry ligation
Lysates were normalized to 1-2 mg/mL in PBS-lysis buffer and incubated with pre-mixed click reagents, containing AzTB (100 µM) for pulldown and in-gel fluorescence or AzRB (100 µM) for proteomics, $CuSO_4$ (1 mM), TCEP (1 mM) and TBTA (100 µM), with shaking for 1 h at room temperature. The reaction was quenched with addition of EDTA (5 mM final concentration). Protein was precipitated by adding MeOH/$CHCl_3$/water at a ratio of 2:0.5:1, relative to the sample volume. The precipitate was washed twice with cold MeOH, and air-dried for 5 min. The protein pellet was first resuspended in 2% (w/v) SDS in PBS to 10 mg/mL and then diluted to 1 mg/mL and 0.2% (w/v) SDS by addition of PBS.

SDS-PAGE and in-gel fluorescence imaging
4× sample loading buffer (250 mM Tris-HCl pH 6.8, 30% (v/v) glycerol, 10% (w/v) SDS, 0.05% (w/v) bromophenol blue, supplemented fresh with 1.3% (v/v) BME) was added to samples (10 µg) which were incubated for 5 min without heating. The protein samples were separated on 12% (w/v) acrylamide Tris-HCl gels using Tris/glycine/SDS running buffer (0.25 M Tris, 0.2 M glycine, 0.1% (w/v) SDS) at 90 V for 30 min followed by 120 V for 1 h. Gel were imaged on a Typhoon imager (GE Healthcare; λ_{Ex} 532 nm, λ_{Em} 575 nm for TAMRA fluorophore; λ_{Ex} 635 nm, λ_{Em} 575 nm for protein MW markers (Precision Plus All Blue Standards, Bio-Rad)). For protein loading quantification, gels were stained with Coomassie blue staining solution overnight, rinsed with water and imaged on an ImageQuant™ LAS 4000 imager (GE Healthcare).

Biotin affinity enrichment for immunoblot analysis
Protein samples (150 µg) were ligated to the capture reagent (AzTB), precipitated and re-suspended as described previously. An aliquot (10 µg) was kept for total lysate (TL) input before enrichment. The samples (100 µg) were incubated with pre-washed Streptavidin Magnetic Beads (New England Biolabs, 30 µL beads per 100 µg protein) for 2 h at room temperature. Beads were washed with 0.2% (w/v) SDS in PBS (0.5 mL, three times) and eluted by boiling in 1x sample loading buffer (95 °C, 10 min). Proteins were separated by SDS-PAGE and transferred to nitrocellulose membranes (Amersham™ Protran®, GE Healthcare) by wet-tank transfer (Bio-RAD) in Tris-Glycine transfer buffer supplemented with 20% (v/v) MeOH for 1 h at 100 V. Membranes were blocked in 3% (w/v) BSA in Tris-buffer (50 mM Tris pH 7.4, 150 mM NaCl) containing 0.01% (v/v) Tween-20 (TBST) for 1 h before incubation with primary antibody in 0.3% (w/v) BSA in TBST for 1.5 h at room temperature or overnight at 4°C. The membrane was washed three times with TBST for 5 min and incubated with the corresponding HRP-conjugated secondary antibody (α-mouse-HRP or α-rabbit-HRP) in 0.3% (w/v) BSA in TBST for 1 h at room temperature. After three washes with TBST (5 min each), the membrane was incubated with HRP substrate (Luminata Crescendo, Millipore) and the chemiluminescence signal captured with an ImageQuant™ LAS 4000 imager.

Sample preparation for MS-based proteomic analysis
Proteomics samples were prepared in LoBind microcentrifuge tubes (Eppendorf®). All buffer solutions were prepared fresh and filtered with a 0.22 µm syringe filter prior to use. NeutrAvidin® Agarose Beads (Thermo Fisher Scientific) were dimethylated prior to the enrichment procedure as describe previously[6] and washed twice with 0.2% (w/v) SDS in PBS before use. Protein lysates (400 µg, in biological triplicates) were ligated with AzRB capture

reagent *via* CuAAC reaction, followed by protein precipitation and re-suspension as previously described. The protein solution (in 0.2% (w/v) SDS, PBS) was cleared by centrifugation (17,000 *g*, 10 min, room temperature) and incubated with dimethylated NeutrAvidin Agarose Beads (60 µL beads per starting 1 mg protein) for 2 h at room temperature with gentle shaking. The supernatant was removed by centrifugation (3,000 *g*, 2 min) and the protein-immobilised beads were sequentially washed three times with 1% (w/v) SDS in PBS (0.5 mL) and three times with 50 mM HEPES pH 8.0 (0.5 mL). After the final wash, the beads were resuspended in 50 mM HEPES pH 8.0 (50 µL) before addition of Lysyl endopeptidase®/ LysC (0.4 µg, FUJIFILM Wako) and incubated for 2 h at 37 °C. Beads were pelleted and the supernatant containing free peptides released by LysC cleavage was transferred to a new LoBind tube. TCEP (5 mM) and chloroacetamide (10 mM) were added to each sample and samples incubated for 10 min at room temperature, before addition of sequence grade trypsin (0.2 µg, Promega) and incubation overnight at 37°C. Digested peptides were then chemically labeled with 10-plex TMT reagents (Thermo Fisher Scientific). Each peptide sample (40 µL) was labeled with different TMT reagents (0.08 mg, dissolved in 40 µL MeCN) and incubated for 2 h at room temperature with gentle shaking. The reaction was quenched by addition of hydroxylamine (1 µL, 5% (w/v)) and the labeled peptides combined and dried in SpeedVac Concentrator at 45 °C. The dried TMT-labeled peptides were re-suspended in 150 µL 1% (v/v) TFA in MilliQ water and cleared *via* centrifugation, followed by fractionation using 3× SDB-RPS styrenedivinylbenzene-reverse phase sulfonated disks and separated into three fractions as previously described.[7] The samples were dried in a SpeedVac Concentrator and stored at -80°C. Peptides were re-suspended in 2% (v/v) MeCN and 0.5% (v/v) TFA in LC-MS grade water for analysis by nanoLC-MS/MS on a Q Exactive Orbitrap mass spectrometer.[6, 8]

Preparation of samples for MS-based proteomic analysis of whole proteomes

Protein lysates (10 µg) were reduced and alkylated with TCEP (5 mM) and chloroacetamide (10 mM), respectively, for 45 min at room temperature. Lysates were then precipitated using $CHCl_3$/MeOH/water as described previously. The protein pellets were re-suspended in 50 mM HEPES pH 8.0 to 1 mg/mL and further digested with trypsin (1:100, trypsin:lysate) at 37 °C overnight. After digestion, the peptides were labeled with TMT reagents (0.08 µg TMT reagent per peptide sample), fractionated using SDB-RPS disks and analyzed by nanoLC-MS/MS.

NanoLC-MS/MS analysis of compound-UCHL1 adduct samples

Peptides were separated on an EASY-Spray™ Acclaim PepMap C18 column (50 cm × 75 µm inner diameter, Thermo Fisher Scientific) using a 70-min linear gradient separation of 0-100% solvent B (80% (v/v) MeCN supplemented with 0.1% (v/v) formic acid): solvent A (2% (v/v) MeCN supplemented with 0.1% (v/v) formic acid) at a flow rate of 250 nL/min. The liquid chromatography was coupled to a Q Exactive mass spectrometer *via* an easy-spray source (Thermo Fisher Scientific) which operated in data-dependent mode with survey scans acquired at a resolution of 70,000 at *m/z* 200. Scans were acquired from 350 to 1650 *m/z*. Up to 10 of the most abundant isotope patterns with charge +2 or higher from the survey scan were selected with an isolation window of 2.0 *m/z* and fragmented by HCD with normalized collision energy of 25. The maximum ion injection times for the survey scan and the MS/MS scans (acquired with a resolution of 17,500 at *m/z* 200) were 20 and 120 ms, respectively. The ion target value for MS was set to 10^6 and for MS/MS to 10^5, and the intensity threshold was set to 8.3×10^2.

NanoLC-MS/MS analysis of TMT-labeled samples

Peptides were separated on an EASY-Spray™ Acclaim PepMap C18 column (50 cm × 75 µm inner diameter, Thermo Fisher Scientific) using a 3-hour linear gradient separation of 0-100% solvent B (80% (v/v) MeCN supplemented with 0.1% (v/v) formic acid): solvent A (2% (v/v) MeCN supplemented with 0.1% (v/v) formic acid) at a flow rate of 250 nL/min. The liquid chromatography was coupled to a Q Exactive mass spectrometer *via* an easy-spray source which operated in data-dependent mode with survey scans acquired at a resolution of 70,000 at *m/z* 200. Scans were acquired from 350 to 1800 *m/z*. Up to 10 of the most abundant isotope

patterns with charge +2 or higher from the survey scan were selected with an isolation window of 1.6 *m/z* and fragmented by HCD with normalized collision energy of 25. The maximum ion injection times for the survey scan and the MS/MS scans (acquired with a resolution of 35,000 at *m/z* 200) were 20 and 120 ms, respectively. The ion target value for MS was set to 10^6 and for MS/MS to 10^5, and the intensity threshold was set to 8.3×10^2.

Proteomics data analysis and TMT-labeling quantification

MS raw files were processed in MaxQuant (version 1.6.4.0) using the built-in Andromeda search engine and searched against the human reference proteome including isoforms (Uniprot, Taxonomy 9606, accessed 17 January 2019).[9] For protein quantification, MS2 based quantification was used, selecting pre/defined TMT 10-plex labeling on lysine side chains and peptide N-termini. Cysteine carbamidomethylation was used as a fixed modification while methionine oxidation and N-terminus acetylation were used variable modifications. Trypsin/P was selected as digestion enzyme, with a maximum of 2 missed cleavages and minimum 7 residues of peptide length allowed. Unique and razor peptides were selected for protein identification and quantification. Other default parameters were used. MaxQuant-processed data was further analyzed using Perseus (version 1.6.2.3) and GraphPad Prism (version 5.03).

TMT data analysis in Perseus

Protein groups and the corresponding corrected TMT intensity values were analyzed in Perseus (version 1.6.2.3)[10] and filtered against Only ID by site, Reverse and Potential contaminant. The TMT intensity values were grouped per TMT multiplex set (TMT set - A/B/C) and per condition (TMT channel - 126-131). The values were $\log_2$ transformed and filtered to require at least 2 valid values per condition and 2 unique and razor peptides for each protein. The mean values within each row (TMT set) were subtracted followed by subtraction of the median values across TMT channels (conditions). A two-sample t-test was performed across all quantified proteins to compare the TMT intensity values of condition 1 (eg. ABP **2**-treated samples) versus condition 2 (eg. DMSO control) and to determine the significantly enriched proteins in each condition (n=3, Permutation-based FDR=0.01, S0=1). Fold-changes for each comparison were computed and plotted against calculated p-values in a Volcano plot. The protein groups were matched with the known DUBs database and further processed in GraphPad Prism 5.03.

UCHL1 modified site identification analysis in PEAKS

The MS raw files were processed in PEAKS (version 8.0)[11] as described. Trypsin was selected as the digestion enzyme, allowing for maximum 2 missed cleavages. Cysteine carbamidomethylation, methionine oxidation and cysteine-compound **1** or **2** adducts (compound **1** $C_{21}H_{19}N_5O$ exact mass 357.16, compound **2** $C_{23}H_{19}N_5O$ exact mass: 381.16) were used as variable modifications. Other default software parameters were used as described in Supplementary Data 3. The peptide $-\log_{10}P$ was set as ≥60 for the modified site identification analysis.

Mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) *via* the PRIDE partner repository[12] with the dataset identifier PXD015825.

Fibroblast-to-myofibroblast transition (FMT) assay

FMT assays in human lung cells derived from Idiopathic Pulmonary Fibrosis (IPF) patients were performed by Charles River (Assay OTS101-FMT-LUNG-IPF). Briefly, patient-derived primary human bronchial fibroblasts were seeded at 750 cells/well in 384-well plates and subsequently pre-treated with reference compounds SB525334 (ALK-5 inhibitor, αSMA inhibition positive control) and nintedanib (clinically approved IPF drug), test compounds (**1**, **2**, **3** and LDN-57444) or vehicle control (0.1% (v/v) DMSO) for 1 h, followed by triggering of FMT with 1.3 ng/mL TGF-β1. After 3 days, the cells were fixed and stained for αSMA and DAPI and imaged using high content analysis (HCA).

The αSMA immunoreactivity was quantified by an HCA algorithm with density x area (DxA) output. A percentage of αSMA inhibition was calculated:

$$\%\ \alpha\text{SMA inhibition } = 100 - \left(\frac{\mu_p - X_i}{\mu_p - \mu_n}\right) \times 100$$

Where μ_p is an average αSMA value of the positive control (TGF-β1 + 1 μM SB525334)
μ_n is an average αSMA value of the vehicle control (TGF-β1 + 0.1% (v/v) DMSO)
X_i are the αSMA values of the compound treatment

The nuclei were stained with DAPI and quantified by an HCA-based quantification of the number of imaged cells. A percentage of cell number was calculated:

$$\%\ \text{cell number} = \left(\frac{X_i}{\mu_n}\right) \times 100$$

Where μ_n is an average nuclei number of vehicle control (TGF-β1 + 0.1% (v/v) DMSO)
X_i are the nuclei numbers of the compound treatment

The relative αSMA inhibition was calculated:

$$\text{Relative } \alpha\text{SMA inhibition} = \frac{\%\ \alpha\text{SMA inhibition}}{\%\ \text{cell number}}$$

The IC_{50} and EC_{50} values were determined by plotting the relative αSMA inhibition ≤1 μM or the percentage of cell number≤10 μM, respectively, against compound concentration with a four-parameter variable slope curve fit with shared upper and lower plateaus between all datasets (Graphpad Prism 5.03).

Chemical Methods

General synthetic methods

Air and moisture sensitive reactions were performed under nitrogen atmosphere using standard Schlenk manifold techniques with anhydrous solvent unless otherwise stated. All glassware was oven dried at 200°C and allowed to cool to room temperature under positive pressure of nitrogen. Solids were dried under vacuum and solvents were sparged with nitrogen for 30 min. The reaction containers were kept under positive nitrogen pressure for the whole reaction time. Compound **1**, **2** and **3** were synthesized as described in chemical synthesis section. Compound **1** was synthesized based on the original report, following a modified procedure.[13]

Analytical techniques

Reactions were monitored by TLC on Merck silica gel 60 Å aluminum plates and visualized with UV light, unless otherwise stated. Preparative TLC was performed using Merck PLC silica gel 60 Å glass plates (20 × 20 cm, 1 mm thickness). Flash column chromatography was performed using Merck Geduran® 60 Å silica gel (40-63 μm particle size) or on a Biotage Isolera™ One (SNAP KP-Sil columns, 60 Å pore size, 45-50 μm particle size). NMR spectra were recorded in $CDCl_3$, CD_2Cl_2, $(CD_3)_2SO$ or CD_3OD (Acros Organics) on a 400 MHz Bruker AV NMR spectrometer at 298 K (400 MHz for ^1H and 101 MHz for ^{13}C). Chemical shifts (δ) are reported in parts per million (ppm) and multiplicities reported as: s, singlet; d, doublet; dd, doublet of doublets; t, triplet; q, quartet; m, multiplet; br., broad. Coupling constants are reported in hertz (Hz). Spectra were analyzed using MestReNova 12 software. Residual

solvent signals were used as an internal standard and the characteristic solvent peaks were corrected to the data reported in *Organometallics*, **2010**, 29 (9), 2176-2179.[14]
All high-resolution mass spectrometry and LC-MS analyses were performed by the Imperial College Mass Spectrometry Service using a Waters Acquity UPLC i-Class system comprised of a Waters BEH Acquity C18 column (2.1 mm × 50 mm, operating at 0.5 mL/min), a Waters LCT Premier ES-ToF mass spectrometer, and a photodiode array (detecting from 210 to 280 nm). Eluents of MeCN and water (each with 0.1% (v/v) formic acid) were used as linear gradients (% (v/v) MeCN in water) of: 5% to 95% (0 to 3.2 min), 95% to 5% (3.2 to 3.5 min), and 5% (3.5 to 4.0 min).

Preparative LC-MS of all final compounds was performed on a Waters RP-HPLC system comprised of a Waters 2767 auto-sampler, a Waters 515 HPLC pump, a XBridge C18 column (5 µm, 4.6 mm × 100 mm, operating at 1.2 mL/min), a Waters 3100 ESI mass spectrometer, and a Waters 2998 photodiode array (detecting from 190 to 700 nm). MeCN and water (each with 0.1% (v/v) formic acid) were used as eluent, with the method described in the table below.

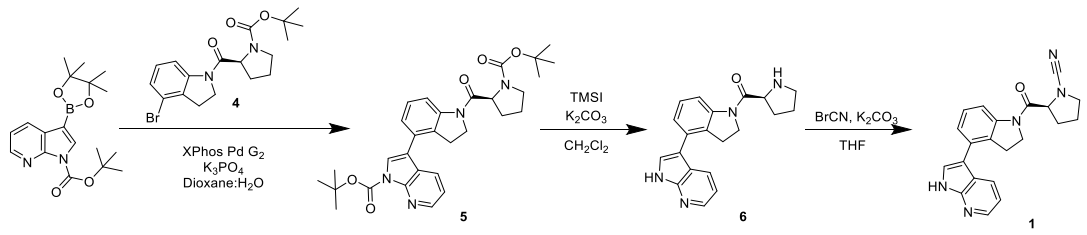
Chiral HPLC to determine % enantiomeric excess (ee) was performed on a PerkinElmer Series 200 HPLC system with a Chiralpak AD column operating at 0.7 mL/min. *n*-Hexane/*i*-PrOH (75:25) were used as eluent, with detection at 220 nm.

Linear gradients used on the Waters 2767 LC-MS system (% (v/v) MeCN in water, with 0.1% (v/v) formic acid)

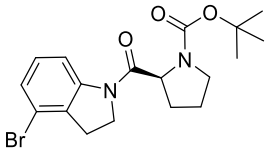
Method, 18 min	0–0.2 min	0.2–10 min	10–13 min	13–14 min	14–18 min
A (40-98%)	40%	40–98%	98%	98–40%	40%
B (20–98%)	20%	20–98%	98%	98–20%	20%
C (2–98%)	2%	2–98%	98%	98–2%	2%

Chemical synthesis

Scheme S1. Synthetic route to Compound 1

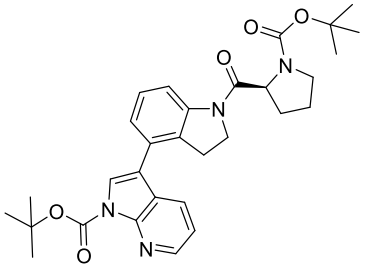


Synthesis of *tert*-butyl (*S*)-2-(4-bromoindoline-1-carbonyl)pyrrolidine-1-carboxylate (4)



To a solution of *N*-(*tert*-butoxycarbonyl)-*L*-proline (260 mg, 1.2 mmol, 1.2 equiv.) in THF (5 mL) was added HATU (570 mg, 1.5 mmol, 1.5 equiv.) and DIPEA (350 µL, 2.0 mmol, 2.0 equiv.). The reaction mixture was stirred at room temperature for 1 h then 4-bromoindoline (130 µL, 1.0 mmol, 1.0 equiv.) was added and the solution stirred at room temperature for 15 h. The resulting reaction mixture was added to saturated aqueous $NaHCO_3$ (15 mL) and extracted with EtOAc (3 × 15 mL). The combined organic phases were washed with brine (15 mL), dried over Na_2SO_4, filtered and concentrated under reduced pressure to afford a crude material that was purified by automated flash chromatography (Biotage SNAP KP-Sil 50 g cartridge, 6-48% EtOAc in *n*-Hexane). Collected fractions were combined and concentrated under reduced pressure to yield an off-white solid. The resulting product was triturated in saturated aqueous $NaHCO_3$, then washed with water (1 mL) and *n*-Hexane (3 mL) to yield compound **4** as a fine white powder (270 mg, 0.68 mmol, 68% yield). $\boldsymbol{R_f}$ = 0.25 (25% EtOAc in *n*-Hexane); **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.18 (dd, *J* = 13.3, 8.0 Hz, 1H), 7.21-6.97 (m, 2H), 4.66-4.52 (m, 0.5H), 4.48-4.34 (m, 1H), 4.31-4.17 (m, 0.5H), 4.14-4.02 (m, 1H), 3.75-3.57 (m, 1H), 3.56-3.40 (m, 1H), 3.28-3.08 (m, 2H), 2.32-1.83 (m, 4H), 1.45 and 1.33 (s, 9H); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 171.70, 171.44, 154.68, 153.77, 144.31, 144.10, 131.81, 131.54, 129.65, 129.35, 126.75, 126.59, 119.34, 119.21, 116.19, 115.94, 80.00, 79.92, 58.79, 58.56, 47.25, 47.02, 46.83, 30.56, 29.84, 29.69, 28.61, 28.42, 24.41, 23.85; **LC-MS** (method A: 7.2 min); **HRMS** (ES-ToF): *m/z* calc. for $C_{20}H_{26}N_3O_3NaBr$ $[M+CH_3CN+Na]^+$: 458.1055, found: 458.1050.

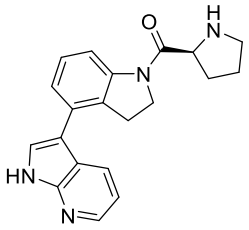
Synthesis of *tert*-butyl (*S*)-2-(4-bromoindoline-1-carbonyl)pyrrolidine-1-carboxylate (5)



Synthesis followed literature protocols with the following adaptions.[15] Compound **4** (180 mg, 0.45 mmol, 1.0 equiv.), tert-butyl 3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-1*H*-pyrrolo[2,3-*b*]pyridine-1-carboxylate (200 mg, 0.58 mg, 1.3 equiv.), K_3PO_4 (190 mg, 0.89 mmol, 2.0 equiv.) and XPhos-Pd-G2 (5.3 mg, 0.006 mmol, 0.015 equiv.) were dissolved in 4.4 mL of 5:1 degassed 1,4-dioxane: water, and the mixture was stirred at 60 °C overnight. After cooling to room temperature, water (5 mL) was added and the solution extracted with EtOAc (3 × 10 mL). The combined organic layers were washed with brine (10 mL), dried over $MgSO_4$, filtered and concentrated under reduced pressure. The recovered material was purified by automated flash chromatography (Biotage SNAP KP-Sil 50 g cartridge, 13-100% EtOAc in *n*-Hexane) to yield compound **5** as an off-white solid (180 mg, 0.33 mmol, 75% yield). **^{1}H NMR** (400 MHz, $CDCl_3$) δ (ppm): 8.62-8.55 (m, 1H), 8.38-8.29 (m, 1H), 7.93-7.84 (m, 1H), 7.67 (d, *J* = 12.8 Hz, 1H), 7.36 (t, *J* = 7.9 Hz, 0.5H), 7.32 (t, *J* = 7.9 Hz, 0.5H), 7.26 (ddd, *J* = 7.6, 4.7, 2.5 Hz, 1H), 7.20 (dd, 1H), 4.68-4.61 (m, 0.5H), 4.55-4.42 (m, 1H), 4.29-4.18 (m, 0.5H), 4.17-4.06 (m, 1H), 3.78-3.63 (m, 1H), 3.62-3.44 (m, 1H), 3.34-3.11 (m, 2H), 2.36-1.85 (m, 4H), 1.71 (s, 5H), 1.49 (s, 4H), 1.39 (s, 4H), 1.26 (s, 5H); **^{13}C NMR** (101 MHz, $CDCl_3$) δ

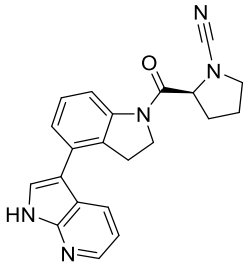
171.34, 171.14, 154.59, 153.71, 148.31, 147.87, 145.61, 145.54, 143.88, 143.70, 129.82, 129.44, 128.50, 128.10, 124.61, 123.86, 122.16, 122.11, 118.72, 117.55, 117.42, 116.76, 116.47, 84.41, 84.33, 79.80, 79.74, 75.01, 58.81, 58.49, 47.75, 46.94, 46.74, 30.50, 29.64, 28.52, 28.34, 28.14, 24.86, 24.33, 23.73; **LC**-**MS** (method B: 8.32 min); **HRMS** (ES-ToF): *m/z* calc. for $C_{30}H_{37}N_4O_5$ $[M+H]^+$: 533.2764, found: 533.2770.

Synthesis of (*S*)-3-(1-prolylindolin-4-yl)-1*H*-pyrrolo[2,3-*b*]pyridine (6)



To a solution of **5** (100 mg, 0.19 mmol, 1.0 equiv.) in DCM (0.5 mL) was added TFA (0.3 mL, 3.9 mmol, 20 equiv.) and the solution stirred at room temperature for 3 h. The reaction mixture was diluted with DCM (2 mL) and concentrated under reduced pressure to yield a crude oil that was triturated with Et_2O. The recovered solid was dried under high vacuum to yield **6** as an off-white solid which was used without further purification (58 mg, 0.17 mmol, 93% yield). **^{1}H NMR** (400 MHz, CD_3OD) δ (ppm): 8.28 (dd, *J* = 4.8, 1.5 Hz, 1H), 8.16 (dd, *J* = 5.8, 3.4 Hz, 1H), 8.09 (dd, *J* = 8.0, 1.5 Hz, 1H), 7.55 (s, 1H), 7.40-7.31 (m, 2H), 7.19 (dd, *J* = 7.9, 4.8 Hz, 1H), 4.71 (dd, *J* = 8.8, 6.5 Hz, 1H), 4.24-4.12 (m, 2H), 3.59-3.53 (m, 1H), 3.47-3.39 (m, 1H), 3.31 (t, *J* = 8.2 Hz, 2H), 2.71-2.58 (m, 1H), 2.24-2.07 (m, 3H); **^{13}C NMR** (101 MHz, CD_3OD) δ (ppm): 167.34, 149.39, 144.02, 143.80, 133.07, 131.35, 129.65, 128.85, 126.67, 125.49, 120.51, 117.14, 116.37, 114.78, 61.15, 48.95, 47.62, 29.79, 29.40, 25.28; **LC**-**MS** (method B: 1.50 min); **HRMS** (ES-ToF): *m/z* calc. for $C_{20}H_{21}N_4O$ $[M+H]^+$: 333.1715, found: 333.1723.

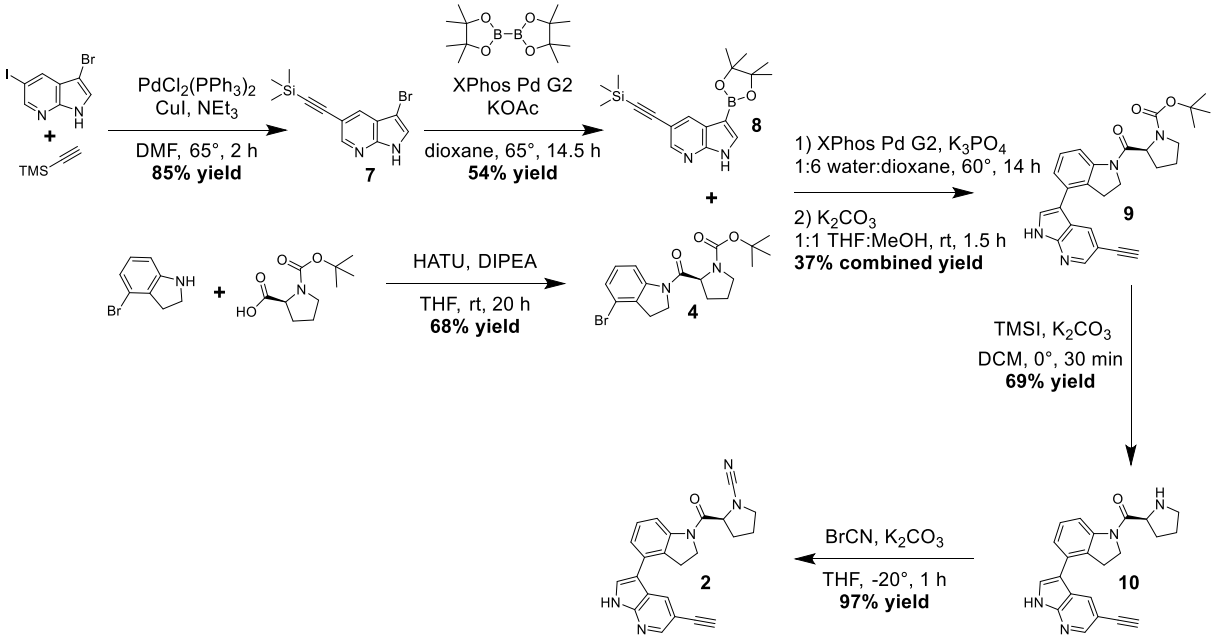
Synthesis of (*S*)-2-(4-(1*H*-pyrrolo[2,3-*b*]pyridin-3-yl)indoline-1-carbonyl)pyrrolidine-1-carbonitrile (1)



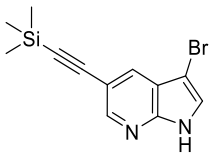
A solution of **6** (60 mg, 0.18 mmol, 1.0 equiv.) and K_2CO_3 (74 mg, 0.54 mmol, 3.0 equiv.) was prepared in dry THF (1.5 mL) and cooled to 0°C. A solution of BrCN in DCM (3 M, 72 µL, 0.22 mmol, 1.2 equiv.) was added dropwise and the reaction mixture allowed to warm to room temperature and stirred for 1 h. The resulting mixture was added to water (5 mL) and extracted with EtOAc (3 × 5 mL). The combined organic phases were washed with brine (5 mL), dried over $MgSO_4$, filtered and concentrated under reduced pressure. The crude material was purified by automated flash chromatography (Isolera, SNAP KP Sil cartridge 10 g, gradient 0 to 6% MeOH in DCM). Fractions were combined, concentrated under vacuum and further purified by preparative LCMS using method A to yield compound **1** as a white solid (24 mg, 0.067 mmol, yield 37%). **^{1}H NMR** (400 MHz, $(CD_3)_2SO$) δ (ppm): 11.95 (s, 1H), 8.28 (d, *J* = 3.7 Hz, 1H), 8.08 (d, *J* = 6.4 Hz, 1H), 8.05 (d, *J* = 7.8 Hz, 1H), 7.69 (d, *J* = 2.3 Hz, 1H), 7.36-7.23 (m, 2H), 7.13 (dd, *J* = 7.9, 4.6 Hz, 1H), 4.75 (dd, *J* = 7.9, 3.6 Hz, 1H), 4.29 (q, *J* = 9.6 Hz, 1H), 4.05 (q, *J* = 9.8 Hz, 1H), 3.53 (m, *J* = 8.3 Hz, 2H), 3.38-3.16 (m, 2H), 2.32 (dq, *J* = 15.1, 7.7 Hz, 1H), 2.04-1.83 (m, 3H); **^{13}C NMR** (101 MHz, $(CD_3)_2SO$) δ (ppm): 168.70,

149.05, 143.87, 143.49, 132.03, 130.00, 128.00, 127.88, 125.18, 124.49, 118.41, 116.76, 116.41, 114.67, 112.90, 62.27, 51.49, 47.72, 29.96, 28.48, 24.31; **HRMS** (ES-ToF): *m/z* Calc. for $C_{21}H_{20}N_5O$ $[M+H]^+$: 358.1662, found 358.1658; **ee** >99%, t_R = 27.93.

Scheme S2. Synthetic route to IMP1710 (2)

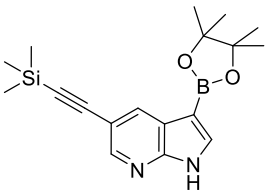


Synthesis of 3-bromo-5-((trimethylsilyl)ethynyl)-1*H*-pyrrolo[2,3-*b*]pyridine (7)



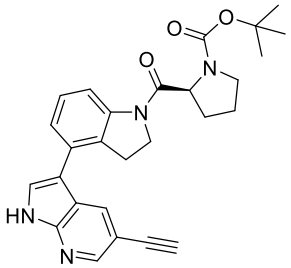
To a solution of 3-bromo-5-iodo-1*H*-pyrrolo[2,3-b]pyridine (500 mg, 1.5 mmol, 1.0 equiv.) in degassed, dry DMF (4 mL) was added dichlorobis(triphenylphosphine)palladium (33 mg, 0.046 mmol, 0.03 equiv.), copper iodide (3.0 mg, 0.015 mmol, 0.02 equiv.), triethylamine (1.1 mL, 7.7 mmol, 5.0 equiv.) and ethynyl(trimethyl)silane (230 µL, 1.6 mmol, 1.1 equiv.). The reaction mixture was heated at 65 °C for 2 h then allowed cool to room temperature and quenched by addition of water (4 mL). The aqueous mixture was extracted with EtOAc (3 × 15 mL) and combined organic extracts were washed with brine, dried over Na_2SO_4 and concentrated under reduced pressure. The crude material was purified by automated flash chromatography (Biotage SNAP KP-Sil 100 g cartridge, 5-40% EtOAc in *n*-Hexane) to yield compound **7** as a pale beige solid (330 mg, 1.1 mmol, 73%). $\boldsymbol{R_f}$ = 0.30 (30% EtOAc in *n*-Hexane); **^{1}H NMR** (400 MHz, $CDCl_3$) δ (ppm): 11.38 (s, 1H), 8.47 (d, *J* = 1.8 Hz, 1H), 8.04 (d, *J* = 1.8 Hz, 1H), 7.40 (d, *J* = 1.8 Hz, 1H), 0.29 (s, 9H); **^{13}C NMR** (101 MHz, $CDCl_3$) δ (ppm): 147.07, 146.57, 131.52, 125.50, 119.54, 113.07, 102.81, 95.68, 89.81, 0.14; **HRMS** (ES-ToF): *m/z* Calc. for $C_{12}H_{13}N_2BrSi$ $[M+H]^+$: 293.0110, found 293.0101.

Synthesis of 3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-5-((trimethylsilyl)ethynyl)-1*H*-pyrrolo[2,3-*b*]pyridine (8)



Synthesis followed literature protocols with the following adaptions.[16] A solution of **2** (90 mg, 0.30 mmol, 1 equiv.), bis(pinacolato)diboron (230 mg, 0.92 mmol, 3.0 equiv.), KOAc (90 mg, 0.92 mmol, 3.0 equiv.) and XPhos-Pd-G2 (9.7 mg, 0.012 mmol, 0.04 equiv.) in 0.7 mL of degassed, dry 1,4-dioxane was heated to 65 °C and stirred overnight. The reaction mixture was allowed to cool to room temperature, filtered through a silica plug and the silica was washed with EtOAc (3 × 5 mL). The recovered solution was concentrated under reduced pressure and purified using automated flash chromatography (Biotage SNAP KP-Sil 25 g cartridge, 7-40% EtOAc in *n*-Hexane) to yield compound **3** as a pale-yellow solid (69 mg, 0.20 mmol, yield 68%). $\boldsymbol{R_f}$ = 0.25 (30% EtOAc in *n*-Hexane); **^{1}H NMR** (400 MHz, $CDCl_3$) δ (ppm): 10.93 (s, 1H), 8.44 (d, *J* = 1.8 Hz, 1H), 8.40 (d, *J* = 1.8 Hz, 1H), 7.81 (d, *J* = 2.4 Hz, 1H), 1.37 (s, 12H), 0.29 (s, 9H); **^{13}C NMR** (101 MHz, $CDCl_3$) δ (ppm): (ppm): 149.03, 146.56, 135.65, 134.31, 123.55, 112.86, 103.90, 94.68, 83.36, 25.05, 0.23; **HRMS** (ES-ToF): *m/z* $[M+H]^+$ Calc. for $C_{18}H_{25}BN_2O_2Si$ $[M+H]^+$: 341.1857, found 341.1864.

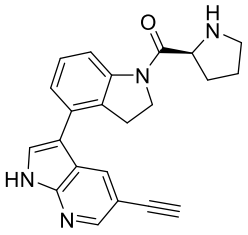
Synthesis of *tert*-butyl (*S*)-2-(4-(5-ethynyl-1*H*-pyrrolo[2,3-*b*]pyridin-3-yl)indoline-1-carbonyl)pyrrolidine-1-carboxylate (9)



Synthesis followed literature protocols with the following adaptions.[15] A solution of **8** (43 mg, 0.13 mmol, 1.3 equiv.), **4** (46 mg, 0.12 mmol, 1.0 equiv.), K_3PO_4 (49 mg, 0.23 mmol, 2.0 equiv.) and XPhos-Pd-G2 (1.4 mg, 0.0017 mmol, 0.015 equiv.) in 6:1 degassed 1,4-dioxane:water (0.7 mL) was stirred at 60 °C overnight. The reaction mixture was allowed to cool to room temperature, filtered through a silica plug, and the silica washed with EtOAc (3 × 5 mL). The filtrate concentrated under reduced pressure to yield a black solid. The crude residue and K_2CO_3 (32 mg, 0.23 mmol, 2.0 equiv.) were dissolved in 1:1 THF:MeOH (1.3 mL), and stirred at room temperature for 1.5 h. The mixture was added to saturated aqueous NH_4Cl (3 mL), then extracted with EtOAc (3 × 5 mL). The combined organic phases were dried with Na_2SO_4, filtered, and concentrated under reduced pressure. The residue was dissolved in EtOAc, filtered through a silica plug and the silica washed with EtOAc (3 × 5 mL). The filtrate was concentrated under reduced pressure to yield a black solid which was purified by preparative TLC (75% EtOAc in *n*-Hexane) to yield compound **9** as an off-white amorphous solid (15 mg, 0.043 mmol, yield 37%). $\boldsymbol{R_f}$ = 0.15 (75% EtOAc in *n*-Hexane); **^{1}H NMR** (400 MHz, Chloroform-d) δ 10.95 (d, *J* = 16.8 Hz, 1H), 8.55 (dd, *J* = 6.4, 1.9 Hz, 1H), 8.30 (dd, *J* = 19.9, 8.0 Hz, 1H), 8.17 (dd, *J* = 14.2, 1.8 Hz, 1H), 7.44 (dd, *J* = 24.5, 2.1 Hz, 1H), 7.41 – 7.16 (m, 2H), 4.72 – 4.64 (m, 0.5H), 4.56 – 4.49 (m, 0.5H), 4.48 – 4.39 (m, 0.5H), 4.30 – 4.19 (m, 0.5H), 4.18 – 4.03 (m, 1H), 3.79 – 3.45 (m, 2H), 3.37 – 3.14 (m, 3H), 2.37 – 1.85 (m, 3H), 1.50 (s, 4.5H), 1.40 (s, 4.5H); **^{13}C NMR** (101 MHz, $CDCl_3$) δ (ppm): 171.42, 171.30, 171.16, 154.79,

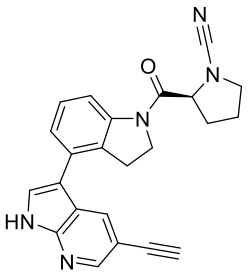
153.91, 147.94, 146.76, 146.66, 143.94, 143.79, 132.30, 130.59, 130.42, 129.41, 129.10, 128.47, 128.17, 124.77, 124.64, 124.35, 118.61, 116.12, 115.91, 114.78, 114.70, 111.53, 111.47, 82.12, 82.09, 79.98, 79.93, 78.09, 78.05, 60.54, 58.95, 58.69, 47.91, 47.09, 46.90, 30.64, 29.83, 29.76, 28.66, 28.48, 24.44, 23.88, 14.34; **LC-MS** (method A: 4.6 min); **HRMS** (ES-ToF): *m/z* calc. for $C_{27}H_{29}N_4O_3$ $[M+H]^+$: 457.2234, found: 457.2227.

Synthesis of (*S*)-5-ethynyl-3-(1-prolylindolin-4-yl)-1*H*-pyrrolo[2,3-*b*]pyridine (10)



Synthesis followed literature protocols with the following adaptions.[17] A suspension of **9** (32 mg, 0.069 mmol, 1 equiv.) and K_2CO_3 (96 mg, 0.69 mmol, 10 equiv.) in degassed, dry DCM (1.0 mL) was cooled to 0°C. TMSI (19.0 µL, 0.136 mmol, 2.0 equiv.) was slowly supplemented to the reaction mixture using a micro syringe. The solution was then stirred at room temperature for 30 min. The mixture was then added to a saturated solution of $NaHCO_3$ (5 mL) and extracted with DCM (3 × 5 mL). The combined organic layers were washed with brine (5 mL), dried over Na_2SO_4, filtered and concentrated under reduced pressure. The crude residue was purified by flash column chromatography (10% 2 M methanolic ammonia in DCM) to yield compound **10** as a pale-beige solid (17 mg, 0.048 mmol, 69% yield). $\boldsymbol{R_f}$ = 0.30 (10% 2 M methanolic ammonia in DCM); **^{1}H NMR** (400 MHz, CD_3OD) δ(ppm): 8.34 (d, *J* = 1.8 Hz, 1H), 8.13 (d, *J* = 7.7 Hz, 1H), 8.10 (d, *J* = 1.9 Hz, 1H), 7.54 (s, 1H), 7.28 (t, *J* = 7.8 Hz, 1H), 7.23 (dd, *J* = 7.7, 1.2 Hz, 1H), 4.19 (q, *J* = 8.5 Hz, 1H), 4.09 (q, *J* = 9.9 Hz, 1H), 4.00-3.93 (m, 1H), 3.53 (s, 1H), 3.26-3.16 (m, 3H), 2.92-2.81 (m, 1H), 2.35-2.23 (m, 1H), 1.94-1.76 (m, 3H); **^{13}C NMR** (101 MHz, CD_3OD) δ(ppm): 173.31, 148.65, 147.13, 144.73, 132.69, 132.20, 131.24, 128.82, 126.71, 125.86, 119.70, 116.48, 115.06, 112.66, 82.18, 60.92, 48.13, 31.12, 29.27, 27.23; **LC-MS** (method C: 9.4 min); **HRMS** (ES-ToF): *m/z* calc. for $C_{22}H_{21}N_4O$ $[M+H]^+$: 357.1710, found: 357.1704.

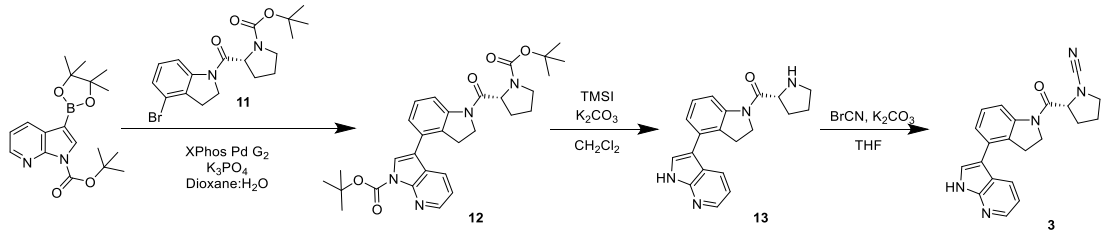
Synthesis of (*S*)-2-(4-(5-ethynyl-1*H*-pyrrolo[2,3-*b*]pyridin-3-yl)indoline-1-carbonyl)pyrrolidine-1-carbonitrile (2)



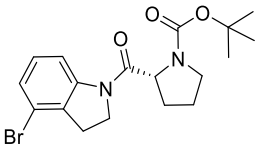
To a stirred solution of **10** (18 mg, 0.049 mmol, 1.0 equiv.) in THF (1 mL) was added K_2CO_3 (27 mg, 0.20 mmol, 4.0 equiv.). The reaction mixture was cooled to -20°C and a solution of BrCN in DCM (3 M, 20 µL, 0.059 mmol, 1.2 equiv.) was added dropwise. The reaction mixture was allowed to warm to room temperature and stirred for 1 h. The resulting mixture was added to water (5 mL) supplemented with K_2CO_3 (25 mg) and extracted with DCM (3 × 5 mL). The combined organic layers were dried over $MgSO_4$, filtered and concentrated under reduced pressure. The crude material was purified by automated flash column chromatography (SNAP KP-Sil 10 g, 1 to 6% MeOH in DCM, 36 mL/min) to yield compound **2** as a white solid (18 mg, 0.048 mmol, 97% yield).$\boldsymbol{R_f}$ = 0.30 (8% MeOH in DCM); **^{1}H NMR** (400 MHz, CD_2Cl_2) δ (ppm):

9.89 (s, 1H), 8.47 (s, 1H), 8.24 (d, *J* = 8.0 Hz, 1H), 8.18 (d, *J* = 1.7 Hz, 1H), 7.47 (d, *J* = 2.1 Hz, 1H), 7.34 (t, *J* = 7.8 Hz, 1H), 7.27 (dd, *J* = 7.7, 1.2 Hz, 1H), 4.48 (dd, *J* = 8.3, 3.6 Hz, 1H), 4.19 (q, *J* = 9.2 Hz, 1H), 4.05 (q, *J* = 9.2 Hz, 1H), 3.70 (q, *J* = 7.6, 7.1 Hz, 1H), 3.61-3.50 (m, 1H), 3.31-3.21 (m, 2H), 3.18 (s, 1H), 2.40-2.27 (m, 1H), 2.16-1.94 (m, 3H); **^{13}C NMR** (101 MHz, CD_2Cl_2) δ (ppm): 168.41, 147.96, 147.13, 143.92, 132.62, 130.87, 129.88, 128.50, 125.37, 125.25, 124.80, 116.53, 116.20, 114.99, 82.05, 78.24, 62.84, 54.31, 52.05, 48.19, 30.55, 28.89, 24.69; **LC-MS** (method B: 9.4 min); **HRMS** (ES-ToF): *m/z* calc. for $C_{23}H_{20}N_5O$ $[M+H]^+$: 382.1662, found: 382.1657; **ee** >94%, t_R = 19.77.

Scheme S3. Synthetic route to IMP1711 (3)

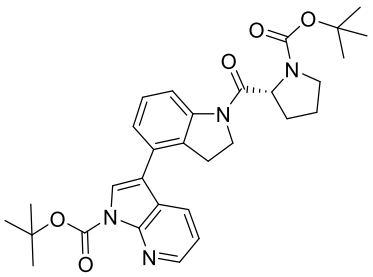


Synthesis of *tert*-butyl (*R*)-2-(4-bromoindoline-1-carbonyl)pyrrolidine-1-carboxylate (11)



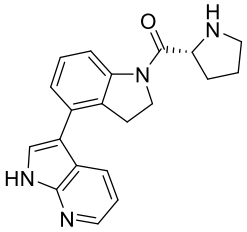
To a solution of *N*-(*tert*-butoxycarbonyl)-*D*-proline (260 mg, 1.2 mmol, 1.2 equiv.) in THF (5 mL) was added HATU (570 mg, 1.5 mmol, 1.5 equiv.) and DIPEA (350 µL, 2.0 mmol, 2.0 equiv.). The reaction mixture was stirred at room temperature for 1 h. 4-Bromoindoline (130 µL, 1.0 mmol, 1.0 equiv.) was dissolved in dry THF (0.5 mL) and added to the reaction mixture. The solution was stirred at room temperature overnight. The resulting reaction mixture was added to saturated aqueous $NaHCO_3$ solution (15 mL), and extracted with EtOAc (3 × 15 mL). The combined organic phases were washed with brine (15 mL), dried over Na_2SO_4, filtered and concentrated under reduced pressure. The crude material was purified by automated flash chromatography (Biotage SNAP KP--Sil 25 g cartridge, 6 to 48% EtOAc in *n*-Hexane). The collected fractions were combined and concentrated under reduced pressure to yield an off-white solid. The resulting product was triturated in saturated aqueous $NaHCO_3$, and washed with water (1 mL) and *n*-Hexane (3 mL) to yield compound **11** as a fine white powder (220 mg, 0.55 mmol, 54% yield) . **^{1}H NMR** (400 MHz, $CDCl_3$) δ (ppm): 8.18 (dd, *J* = 13.1, 8.0 Hz, 1H), 7.22-6.95 (m, 2H), 4.58 (dd, *J* = 7.7, 3.2 Hz, 0.5H), 4.43 (dt, *J* = 11.7, 5.8 Hz, 1H), 4.23 (q, *J* = 9.5 Hz, 0.5H), 4.16-4.01 (m, 1H), 3.75-3.58 (m, 1H), 3.58-3.41 (m, 1H), 3.29-3.09 (m, 2H), 2.35-2.06 (m, 2H), 2.06-1.81 (m, 2H), 1.46 (s, 4.5H), 1.34 (s, 4.5H).; **^{13}C NMR** (101 MHz, $CDCl_3$) δ(ppm): 171.57, 171.31, 154.55, 153.63, 144.20, 143.99, 131.67, 131.40, 129.53, 129.22, 126.61, 126.47, 119.22, 119.08, 116.08, 115.82, 79.86, 79.78, 58.67, 58.44, 47.14, 46.90, 46.72, 30.45, 29.73, 29.58, 28.50, 28.31, 24.30, 23.73. **HRMS** (ES-ToF): *m/z* calc. for $C_{20}H_{26}N_3O_3NaBr$ $[M+CH_3CN+Na]^+$: 460.1035, found: 460.1027.

Synthesis of *tert*-butyl (*R*)-2-(4-bromoindoline-1-carbonyl)pyrrolidine-1-carboxylate (12)



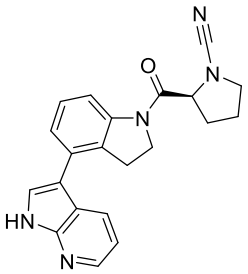
Synthesis followed literature protocols with the following adaptions.[15] Compound **11** (180 mg, 0.45 mmol, 1.0 equiv.), tert-butyl 3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)-1*H*-pyrrolo[2,3-*b*]pyridine-1-carboxylate (200 mg, 0.58 mg, 1.3 equiv.), K_3PO_4 (190 mg, 0.89 mmol, 2.0 equiv.) and XPhos-Pd-G2 (5.3 mg, 0.006 mmol, 0.015 equiv.) were dissolved in 5:1 degassed 1,4-dioxane: water (2.4 mL), and the mixture stirred at 60 °C overnight. After cooling to room temperature, water (5 mL) was added and the solution was extracted with EtOAc (3 × 10 mL). The combined organic layers were washed with brine (10 mL), dried over $MgSO_4$, filtered and concentrated under reduced pressure. The crude material was purified by automated flash chromatography (Biotage SNAP KP-Sil 50 g cartridge, 7 to 100% EtOAc in *n*-Hexane) to yield compound **12** as an off-white solid which was used without further purification (220 mg, 0.42 mmol, 93% yield). $\boldsymbol{R_f}$ = 0.26 (55% EtOAc in *n*-hexane); **LC**-**MS** (method B: 8.70 min); **MS** (Scan ES+): *m/z* calc. for $C_{30}H_{37}N_4O_5$ $[M+H]^+$: 533.2764, found: 533.40.

Synthesis of (*R*)-3-(1-prolylindolin-4-yl)-1*H*-pyrrolo[2,3-*b*]pyridine (13)



To a solution of **12** (200 mg, 0.38 mmol, 1.0 equiv.) in DCM (0.5 mL) was added TFA (0.5 mL, 6.5 mmol, 17 equiv.). The solution was stirred at room temperature for 3 h, then diluted with DCM (2 mL) and concentrated under reduced pressure to yield a yellowish oil. The crude material was triturated with Et_2O and the recovered solid dried under high vacuum to yield compound **13** as an off-white solid which was used without further purification (100 mg, 0.30 mmol, 80% yield).**LC**-**MS** (method B: 2.43 min); **MS** (Scan ES+): *m/z* calc. for $C_{20}H_{21}N_4O$ $[M+H]^+$: 333.1715, found: 333.24.

Synthesis of (*R*)-2-(4-(1*H*-pyrrolo[2,3-*b*]pyridin-3-yl)indoline-1-carbonyl)pyrrolidine-1-carbonitrile (3)



A solution of **13** (100.0 mg, 0.30 mmol, 1.0 equiv.) and K_2CO_3 (120 mg, 0.90 mmol, 3.0 equiv.) was prepared in dry THF (2 mL) and cooled to 0°C. A solution of BrCN in DCM (3 M, 120 µL, 0.36 mmol, 1.2 equiv.) was added dropwise and the reaction mixture allowed to warm to room temperature and stirred for 1 h. The resulting reaction mixture was added to water (5 mL) and extracted with EtOAc (3 × 5 mL). The combined organic phases were washed with brine (5 mL), dried over $MgSO_4$, filtered and concentrated under reduced pressure. The crude material was purified by automated flash chromatography (Biotage SNAP KP Sil cartridge 10 g, gradient 0 to 6% MeOH in DCM). The collected fractions were combined and concentrated under reduced pressure. The recovered material was dissolved in 1:1 MeCN:H_2O (1 mL) supplemented with TFA (0.1% (v/v)). The solution was filtered and lyophilized to yield **3** as white solid (35 mg, 0.098 mmol, yield 33%). **^{1}H NMR** (400 MHz, $(CD_3)_2SO$) δ (ppm): 11.97 (s, 1H), 8.28 (dd, *J* = 4.6, 1.4 Hz, 1H), 8.08 (dd, *J* = 6.6, 2.3 Hz, 1H), 8.05 (dd, *J* = 8.0, 1.3 Hz, 1H), 7.70 (d, *J* = 2.5 Hz, 1H), 7.35-7.25 (m, 2H), 7.13 (dd, *J* = 7.9, 4.6 Hz, 1H), 4.75 (dd, *J* = 8.3, 3.9 Hz, 1H), 4.35-4.23 (m, 1H), 4.10-3.98 (m, 1H), 3.60-3.45 (m, 2H), 3.32-3.16 (m, 2H), 2.38-2.24 (m, 1H), 2.03-1.83 (m, 3H); **^{13}C NMR** (101 MHz, $(CD_3)_2SO$) δ (ppm): 168.70, 149.05, 143.87, 143.49, 132.03, 130.00, 128.00, 127.88, 125.18, 124.49, 118.41, 116.76, 116.41, 114.67, 112.90, 62.27, 51.49, 47.72, 29.96, 28.48, 24.31; **HRMS** (ES-ToF): *m/z* Calc. for $C_{21}H_{19}N_5O$ $[M+H]^+$: 358.1668, found 358.1673. **ee** >99%, t_R = 24.42.

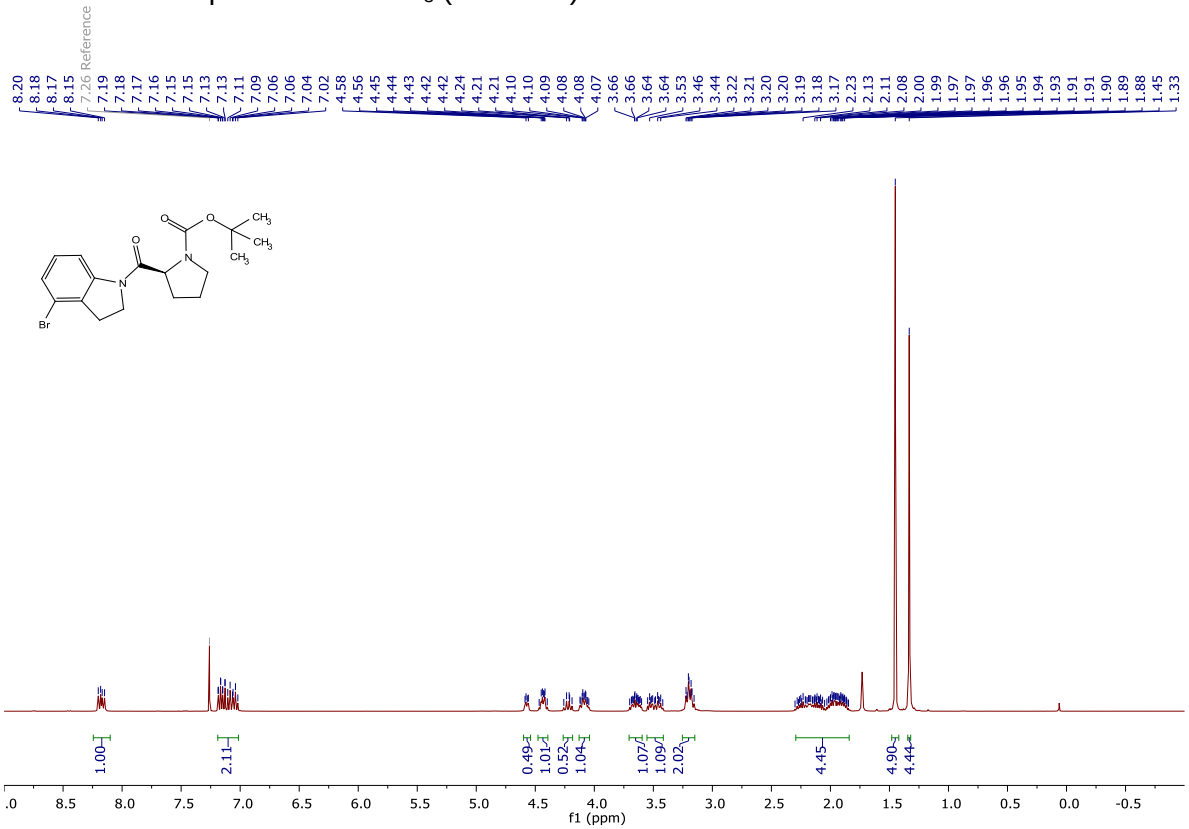
References

1. Broncel, M.; Serwa, R. A.; Ciepla, P.; Krause, E.; Dallman, M. J.; Magee, A. I.; Tate, E. W., Multifunctional reagents for quantitative proteome-wide analysis of protein modification in human cells and dynamic profiling of protein lipidation during vertebrate development. *Angew Chem Int Ed Engl* **2015,** *54* (20), 5948-5951.
2. Das, C.; Hoang, Q. Q.; Kreinbring, C. A.; Luchansky, S. J.; Meray, R. K.; Ray, S. S.; Lansbury, P. T.; Ringe, D.; Petsko, G. A., Structural basis for conformational plasticity of the Parkinson's disease-associated ubiquitin hydrolase UCH-L1. *Proc Natl Acad Sci U S A* **2006,** *103* (12), 4675-4680.
3. Strelow, J. M., A Perspective on the Kinetics of Covalent and Irreversible Inhibition. *SLAS Discov* **2017,** *22* (1), 3-20.
4. Kasperkiewicz, P.; Poreba, M.; Snipas, S. J.; Parker, H.; Winterbourn, C. C.; Salvesen, G. S.; Drag, M., Design of ultrasensitive probes for human neutrophil elastase through hybrid combinatorial substrate library profiling. *Proc Natl Acad Sci U S A* **2014,** *111* (7), 2518-2523.
5. Rappsilber, J.; Ishihama, Y.; Mann, M., Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* **2003,** *75* (3), 663-670.
6. Goya Grocin, A.; Serwa, R. A.; Morales Sanfrutos, J.; Ritzefeld, M.; Tate, E. W., Whole Proteome Profiling of N-Myristoyltransferase Activity and Inhibition Using Sortase A. *Mol Cell Proteomics* **2019,** *18* (1), 115-126.
7. Hernandez-Valladares, M.; Aasebo, E.; Mjaavatten, O.; Vaudel, M.; Bruserud, O.; Berven, F.; Selheim, F., Reliable FASP-based procedures for optimal quantitative proteomic

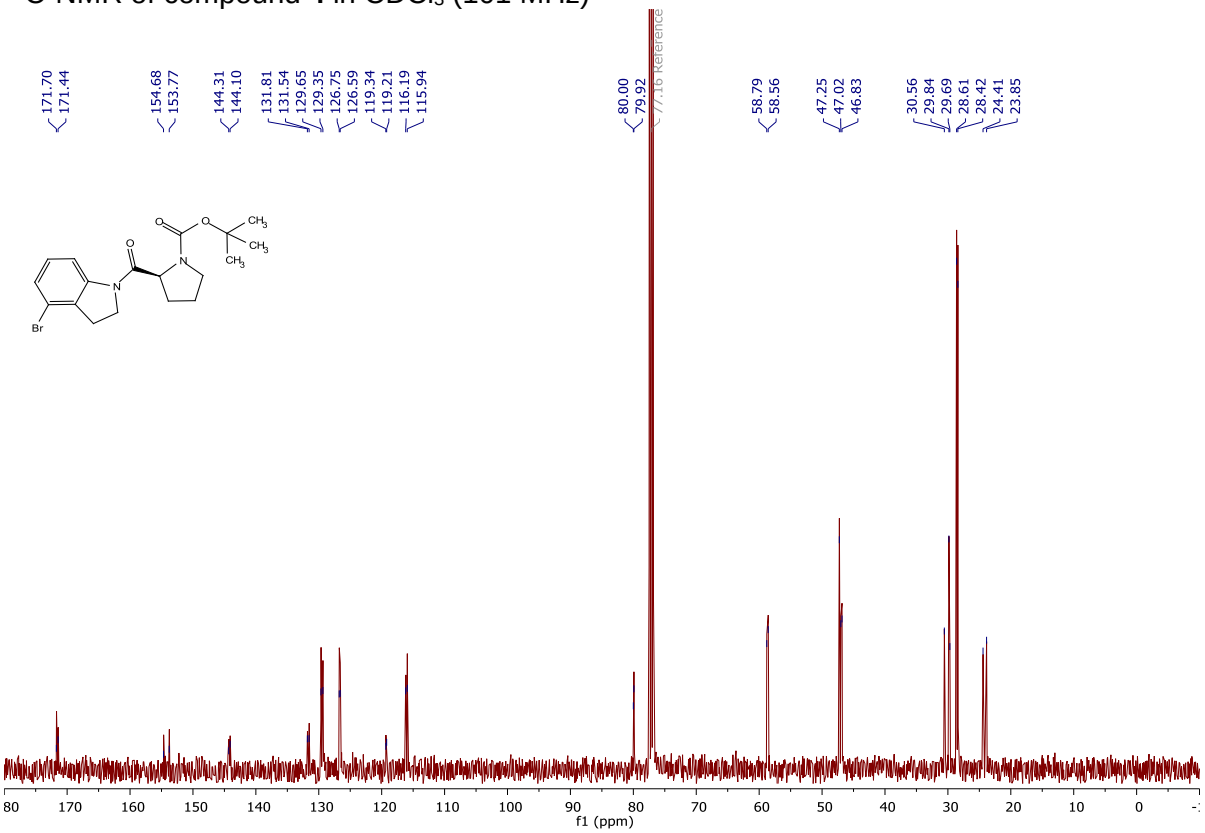
and phosphoproteomic analysis on samples from acute myeloid leukemia patients. *Biol Proced Online* **2016,** *18*, 13-22.
8. Storck, E. M.; Morales-Sanfrutos, J.; Serwa, R. A.; Panyain, N.; Lanyon-Hogg, T.; Tolmachova, T.; Ventimiglia, L. N.; Martin-Serrano, J.; Seabra, M. C.; Wojciak-Stothard, B.; Tate, E. W., Dual chemical probes enable quantitative system-wide analysis of protein prenylation and prenylation dynamics. *Nat Chem* **2019,** *11* (6), 552-561.
9. Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M., Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **2014,** *13* (9), 2513-2526.
10. Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M. Y.; Geiger, T.; Mann, M.; Cox, J., The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **2016,** *13* (9), 731-740.
11. Zhang, J.; Xin, L.; Shan, B.; Chen, W.; Xie, M.; Yuen, D.; Zhang, W.; Zhang, Z.; Lajoie, G. A.; Ma, B., PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol Cell Proteomics* **2012,** *11* (4), M111 010587.
12. Vizcaino, J. A.; Deutsch, E. W.; Wang, R.; Csordas, A.; Reisinger, F.; Rios, D.; Dianes, J. A.; Sun, Z.; Farrah, T.; Bandeira, N.; Binz, P. A.; Xenarios, I.; Eisenacher, M.; Mayer, G.; Gatto, L.; Campos, A.; Chalkley, R. J.; Kraus, H. J.; Albar, J. P.; Martinez-Bartolome, S.; Apweiler, R.; Omenn, G. S.; Martens, L.; Jones, A. R.; Hermjakob, H., ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol* **2014,** *32* (3), 223-226.
13. Kemp, M.; Stockley, M.; Jones, A. Cyanopyrrolidines as Dub Inhibitors for the Treatment of Cancer. WO2017009650 (A1), 2017/01/19/, 2017.
14. Fulmer, G. R.; Miller, A. J. M.; Sherden, N. H.; Gottlieb, H. E.; Nudelman, A.; Stoltz, B. M.; Bercaw, J. E.; Goldberg, K. I., NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist. *Organometallics* **2010,** *29* (9), 2176-2179.
15. Dufert, M. A.; Billingsley, K. L.; Buchwald, S. L., Suzuki-Miyaura Cross-Coupling of Unprotected, Nitrogen-Rich Heterocycles: Substrate Scope and Mechanistic Investigation. *J Am Chem Soc* **2013,** *135* (34), 12877-12885.
16. Billingsley, K. L.; Barder, T. E.; Buchwald, S. L., Palladium-Catalyzed Borylation of Aryl Chlorides: Scope, Applications, and Computational Studies. *Angewandte Chemie* **2007,** *119* (28), 5455-5459.
17. Chang, J. W.; Cognetta, A. B.; Niphakis, M. J.; Cravatt, B. F., Proteome-Wide Reactivity Profiling Identifies Diverse Carbamate Chemotypes Tuned for Serine Hydrolase Inhibition. *Acs Chemical Biology* **2013,** *8* (7), 1590-1599.

Appendix 1: ^{1}H and ^{13}C NMR Spectra

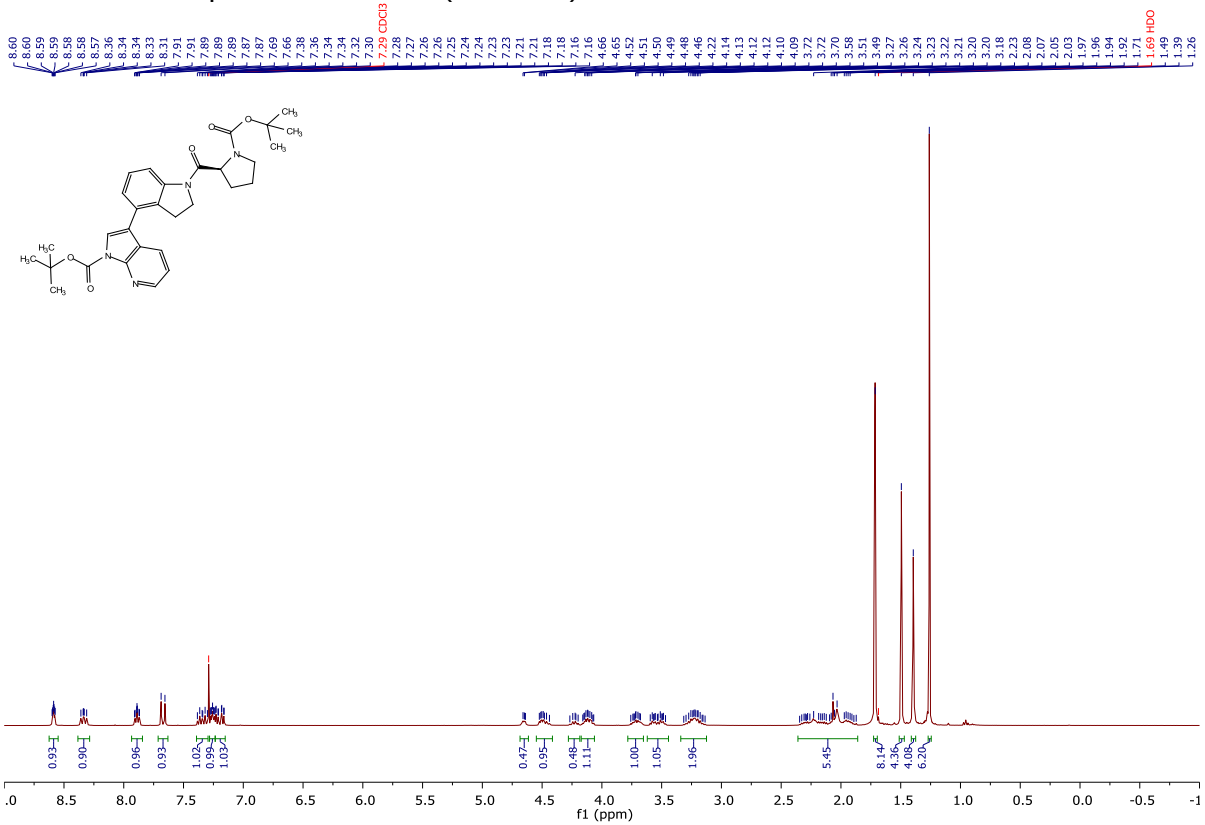
^{1}H-NMR of compound **4** in $CDCl_3$ (400 MHz)



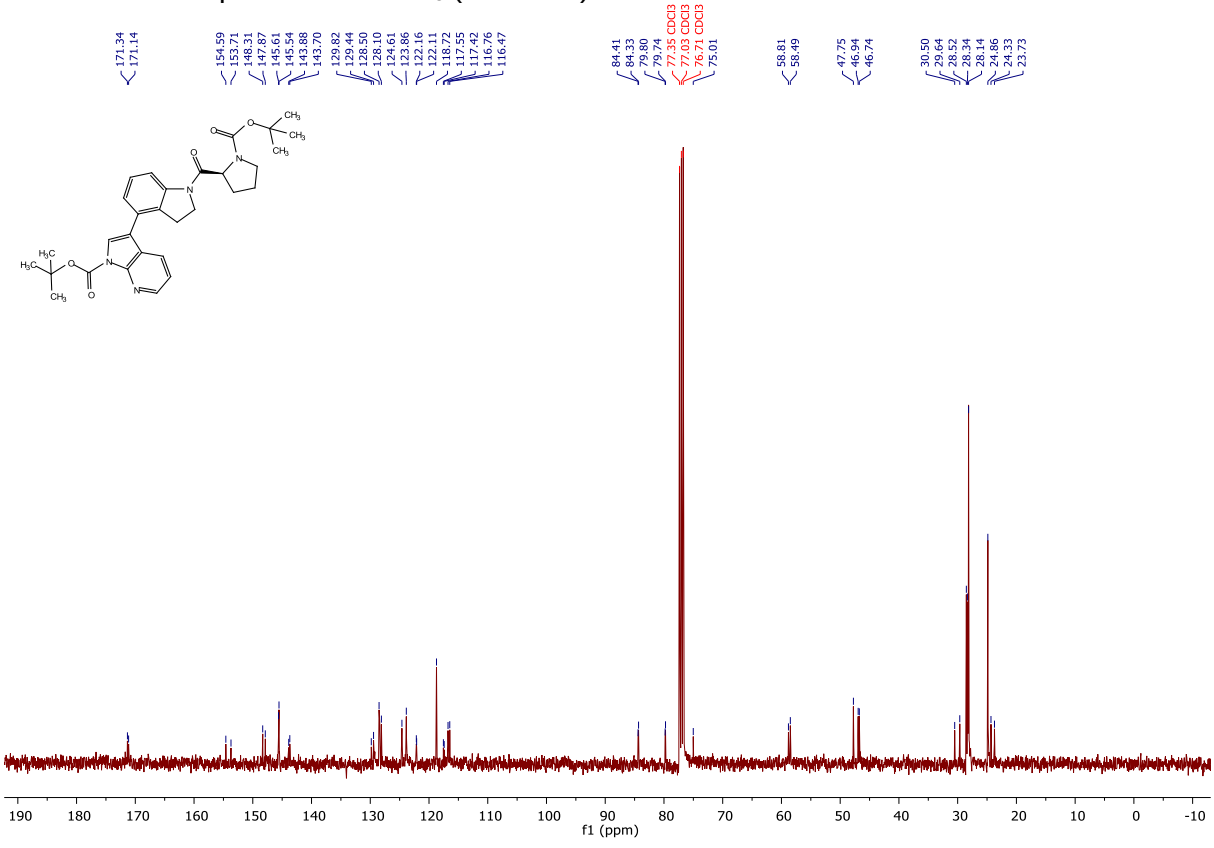
^{13}C-NMR of compound **4** in $CDCl_3$ (101 MHz)



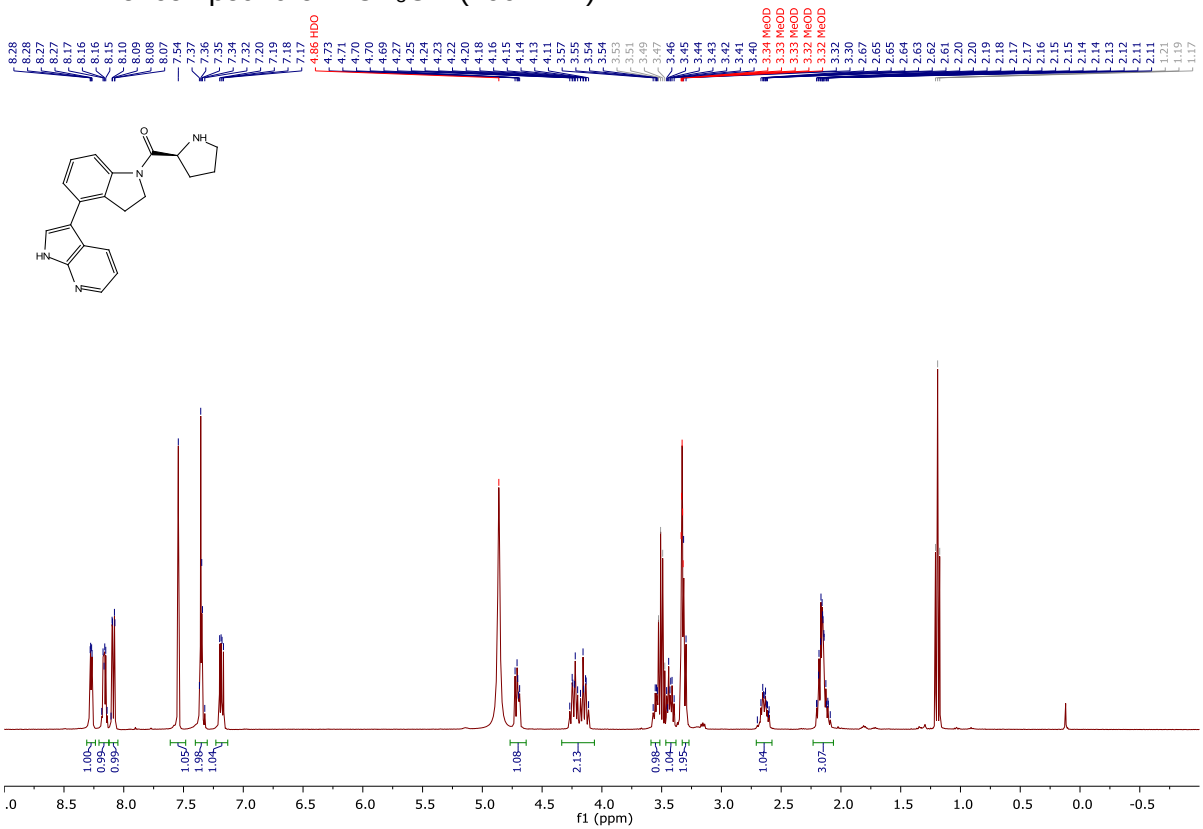
^{1}H-NMR of compound **5** in $CDCl_3$ (400 MHz)



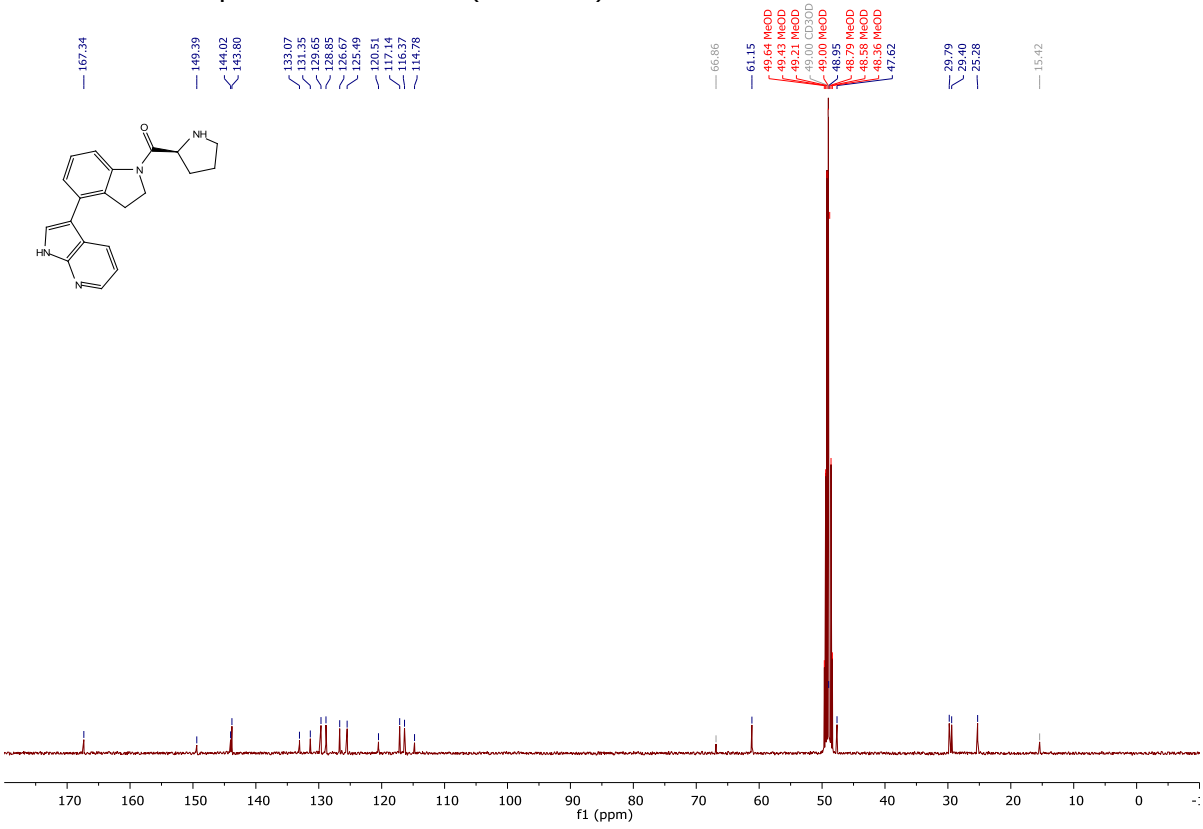
^{13}C-NMR of compound **5** in $CDCl_3$ (101 MHz)



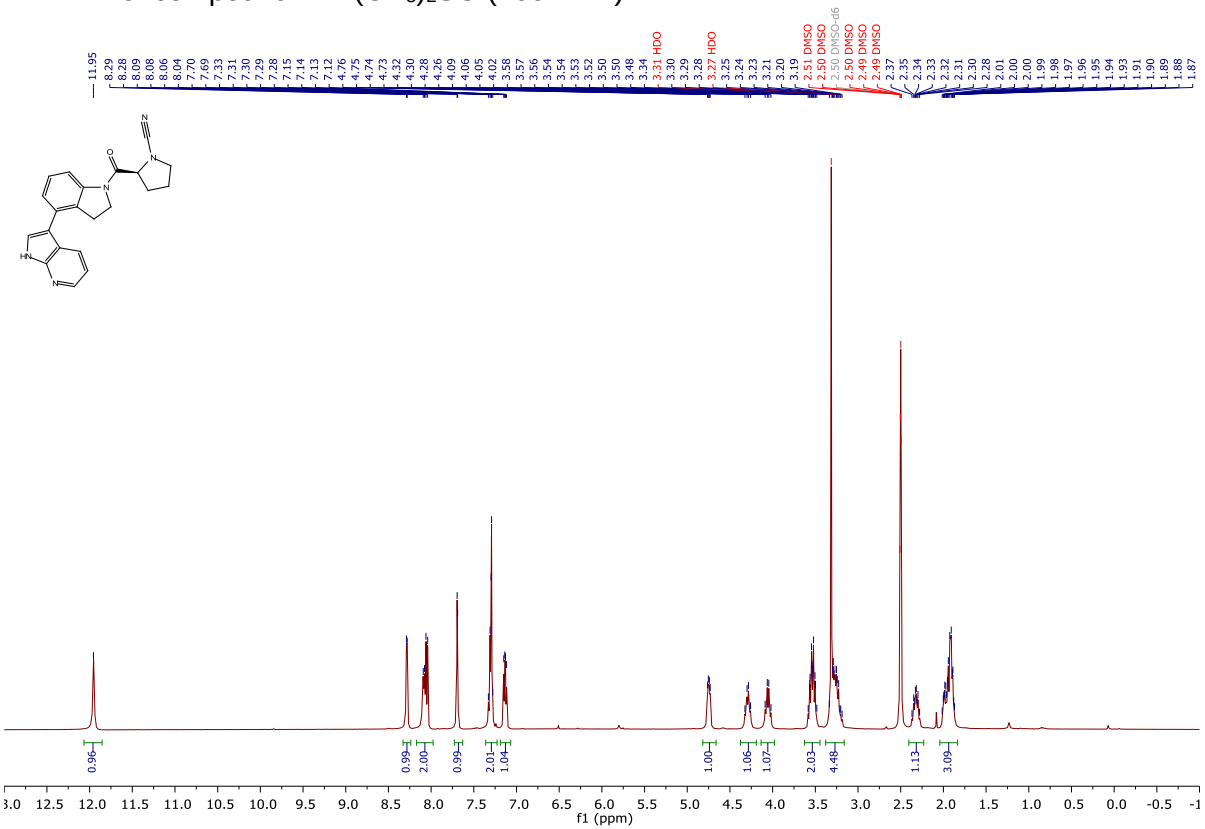
^{1}H-NMR of compound **6** in CD_3OD (400 MHz)



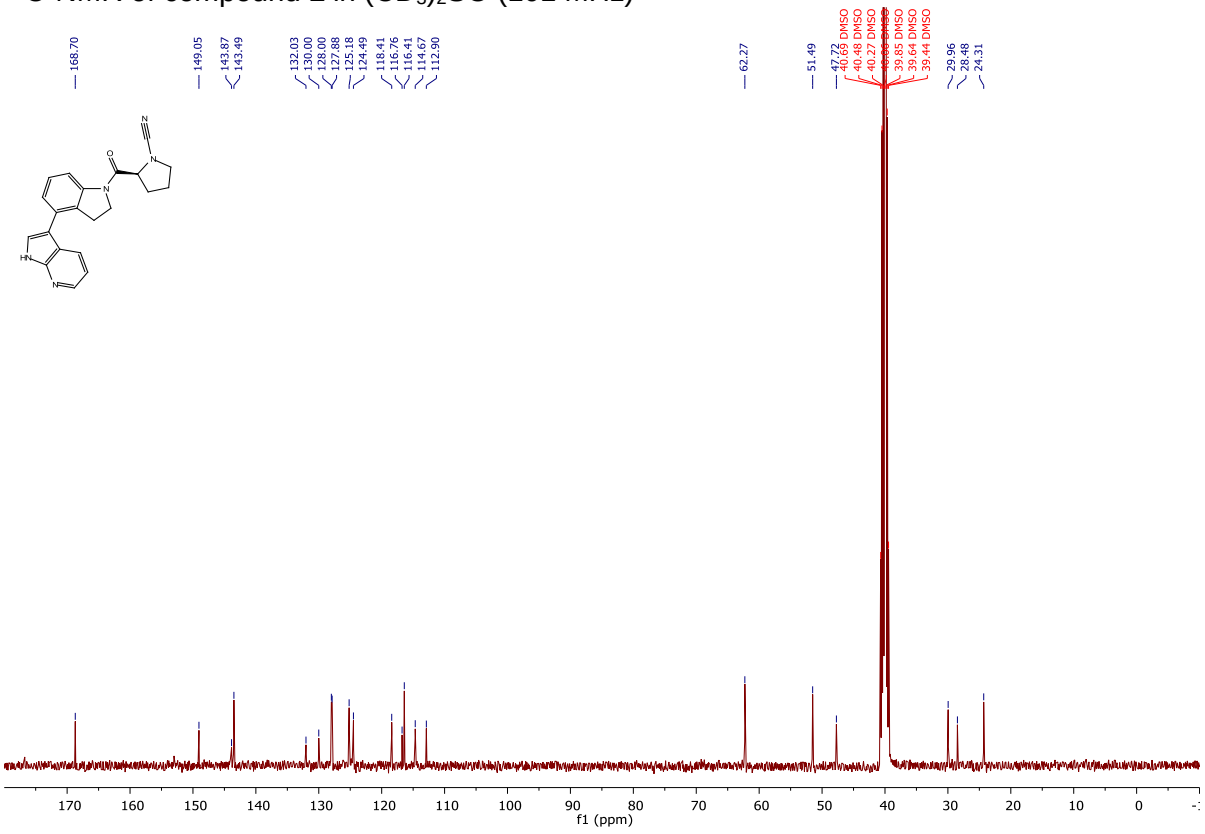
^{13}C-NMR of compound **6** in CD_3OD (101 MHz)



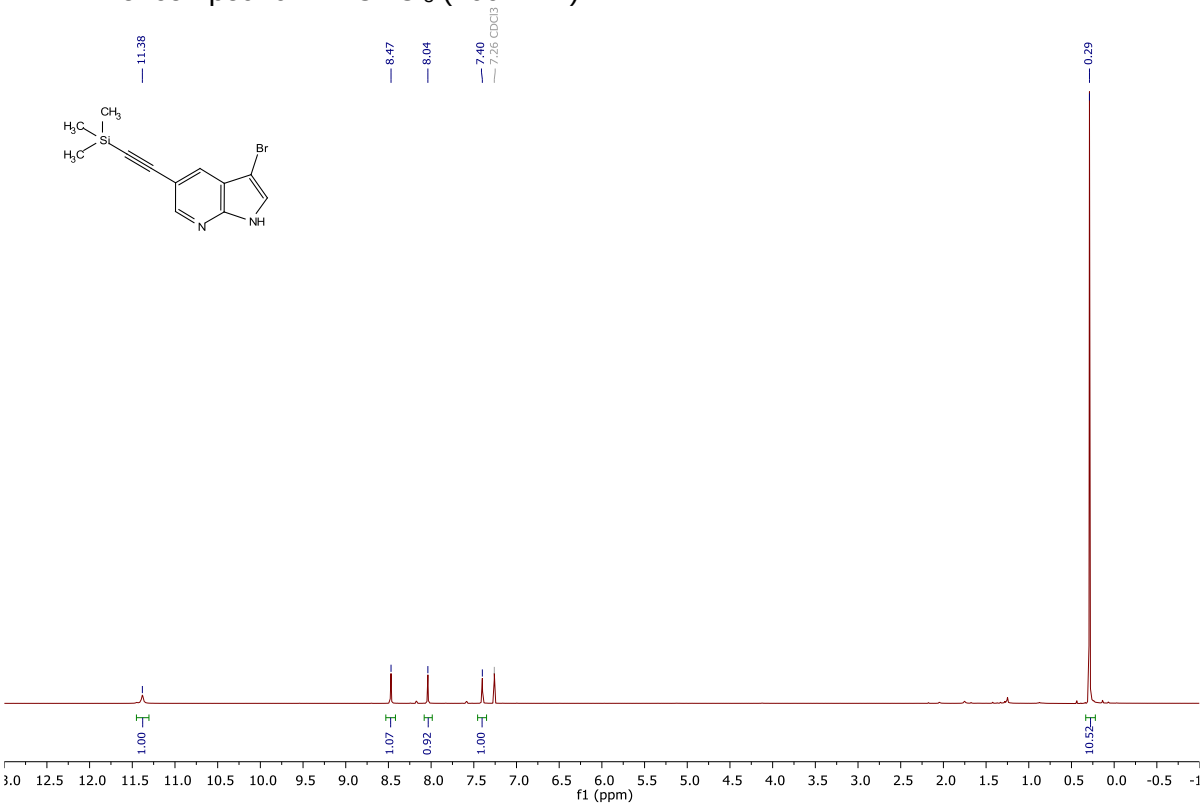
^{1}H-NMR of compound **1** in $(CD_3)_2SO$ (400 MHz)



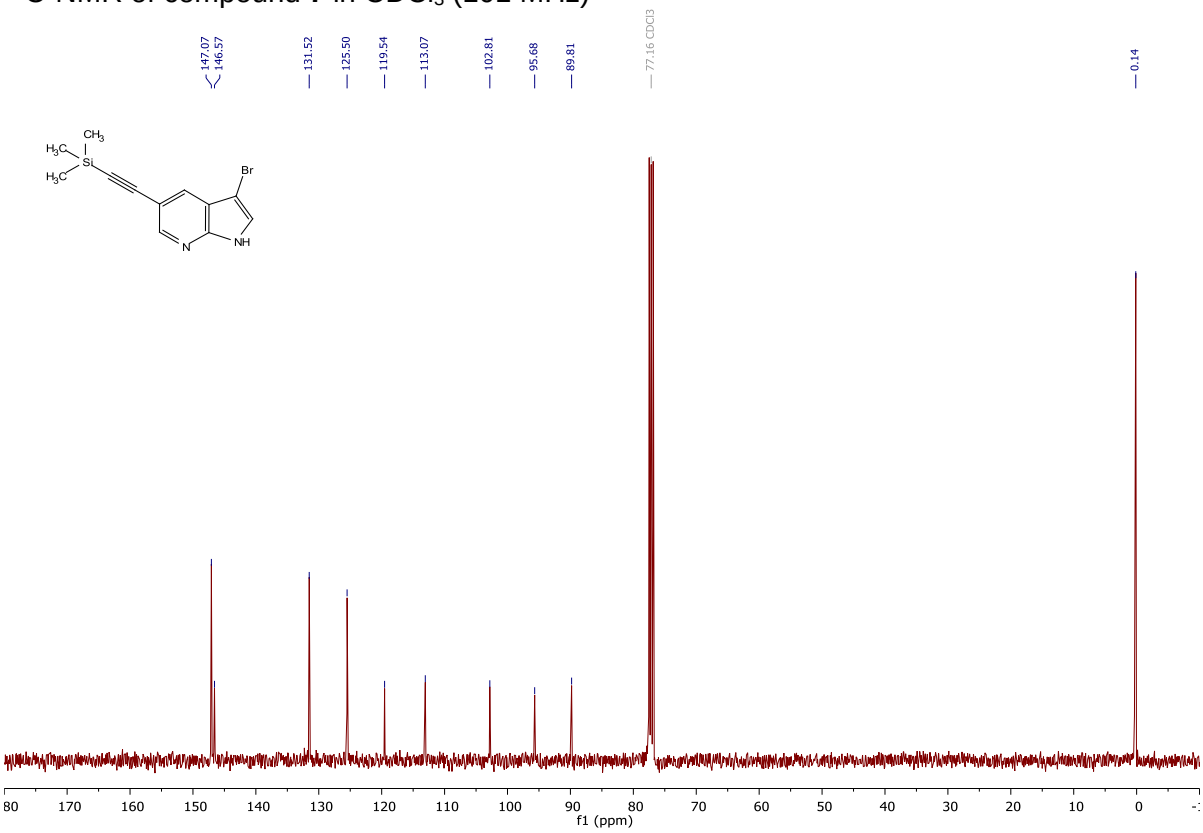
^{13}C-NMR of compound **1** in $(CD_3)_2SO$ (101 MHz)



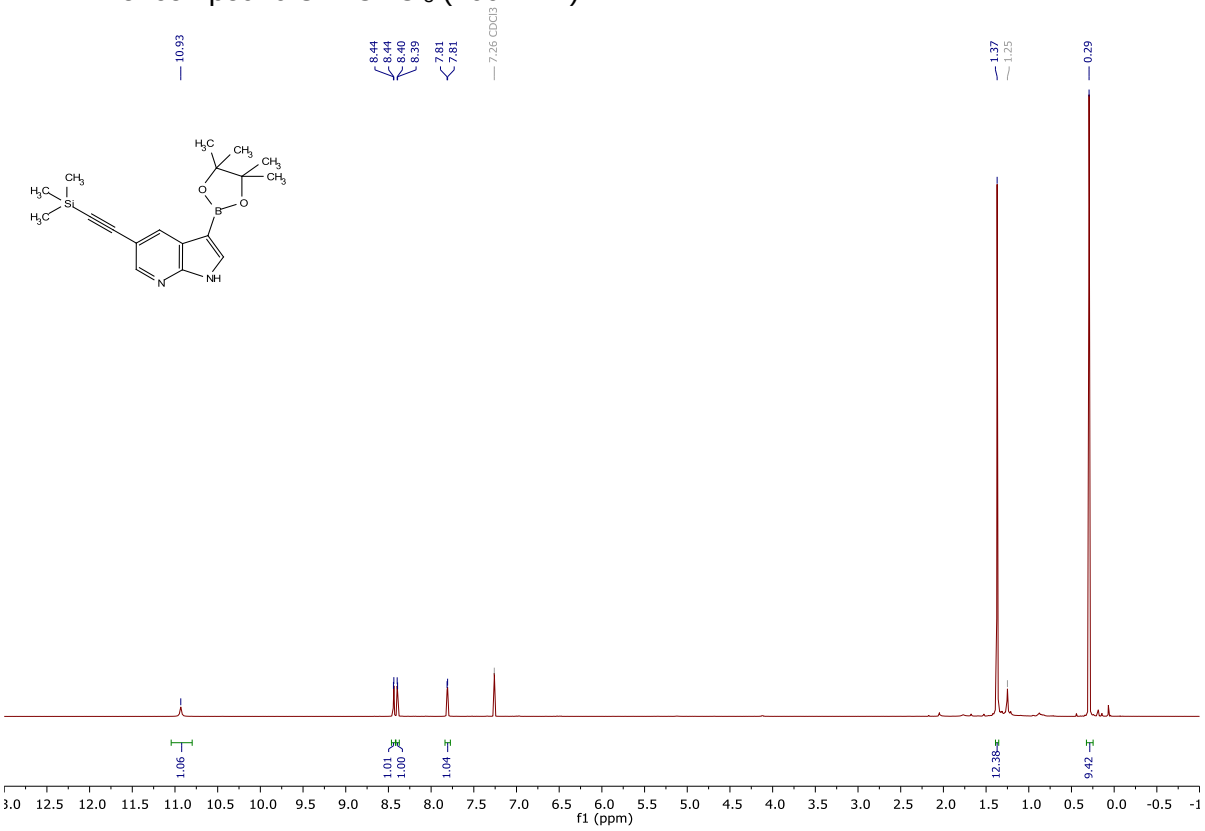
^{1}H-NMR of compound **7** in $CDCl_3$ (400 MHz)



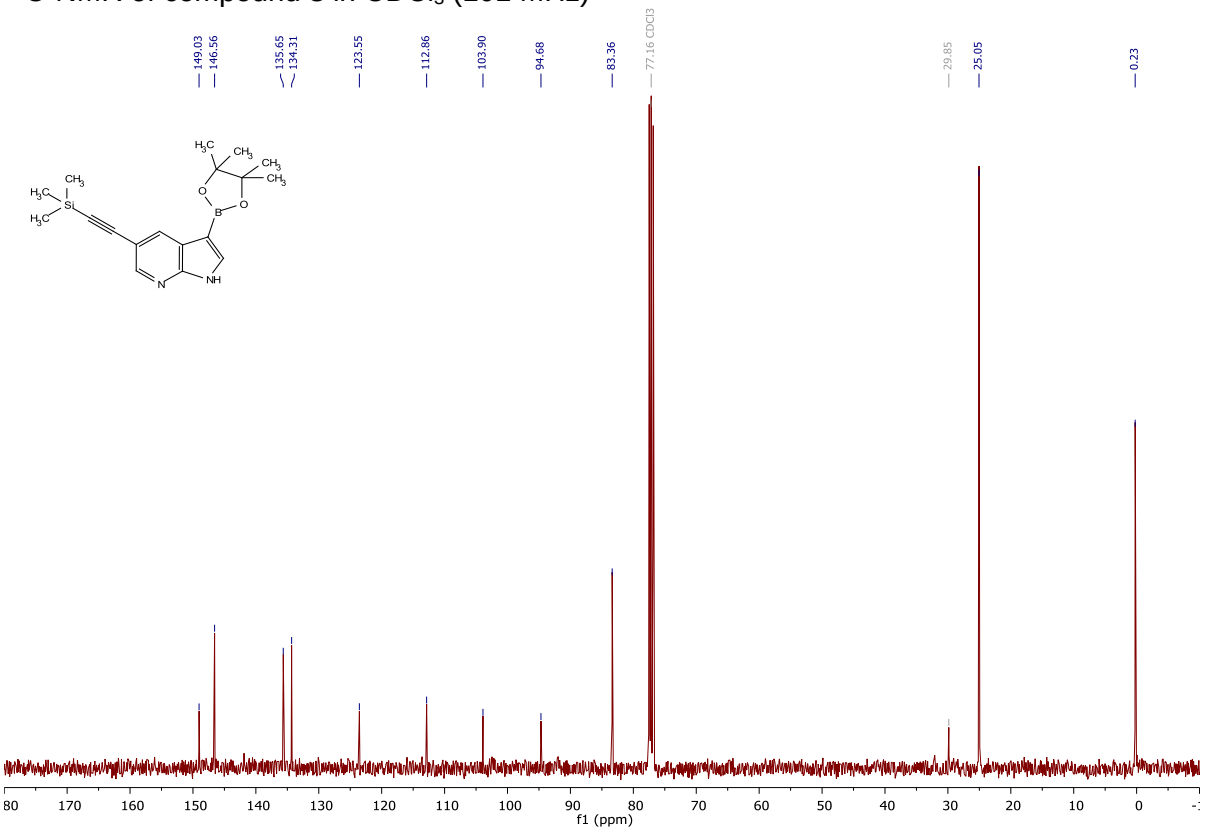
^{13}C-NMR of compound **7** in $CDCl_3$ (101 MHz)



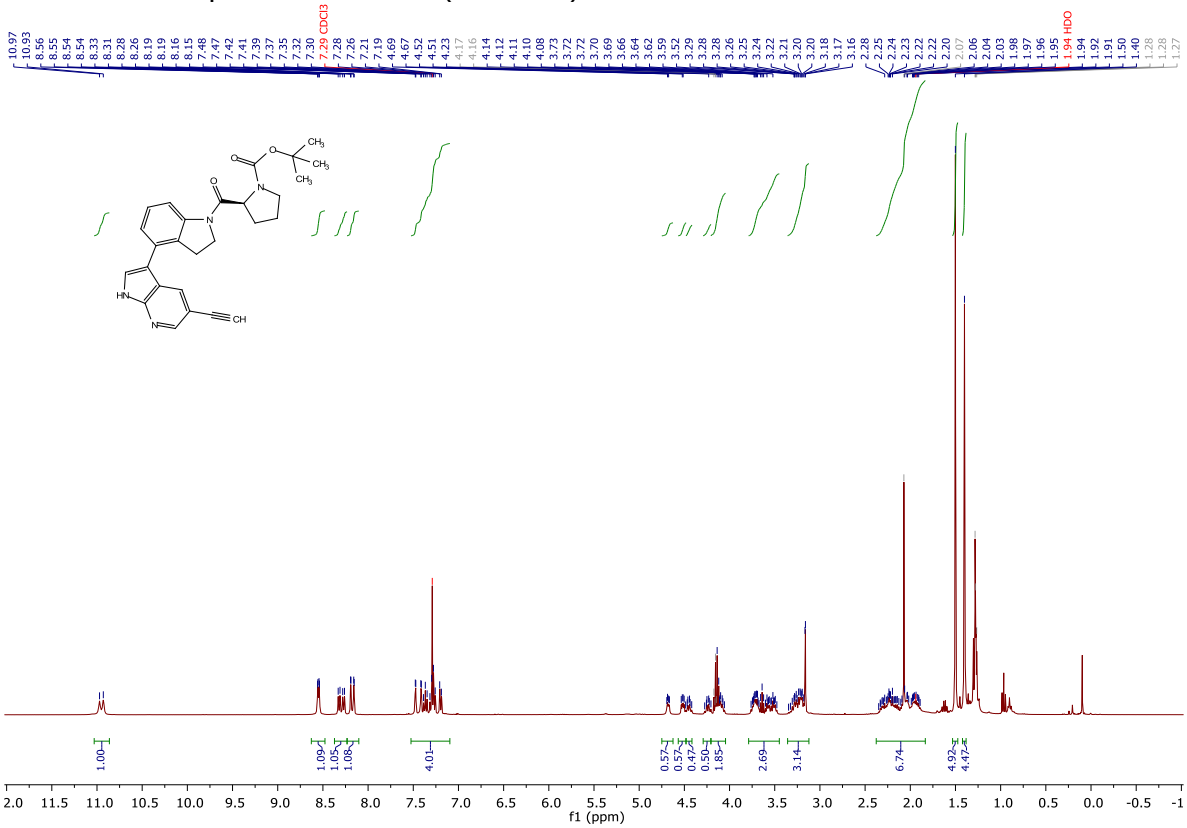
^{1}H-NMR of compound **8** in $CDCl_3$ (400 MHz)



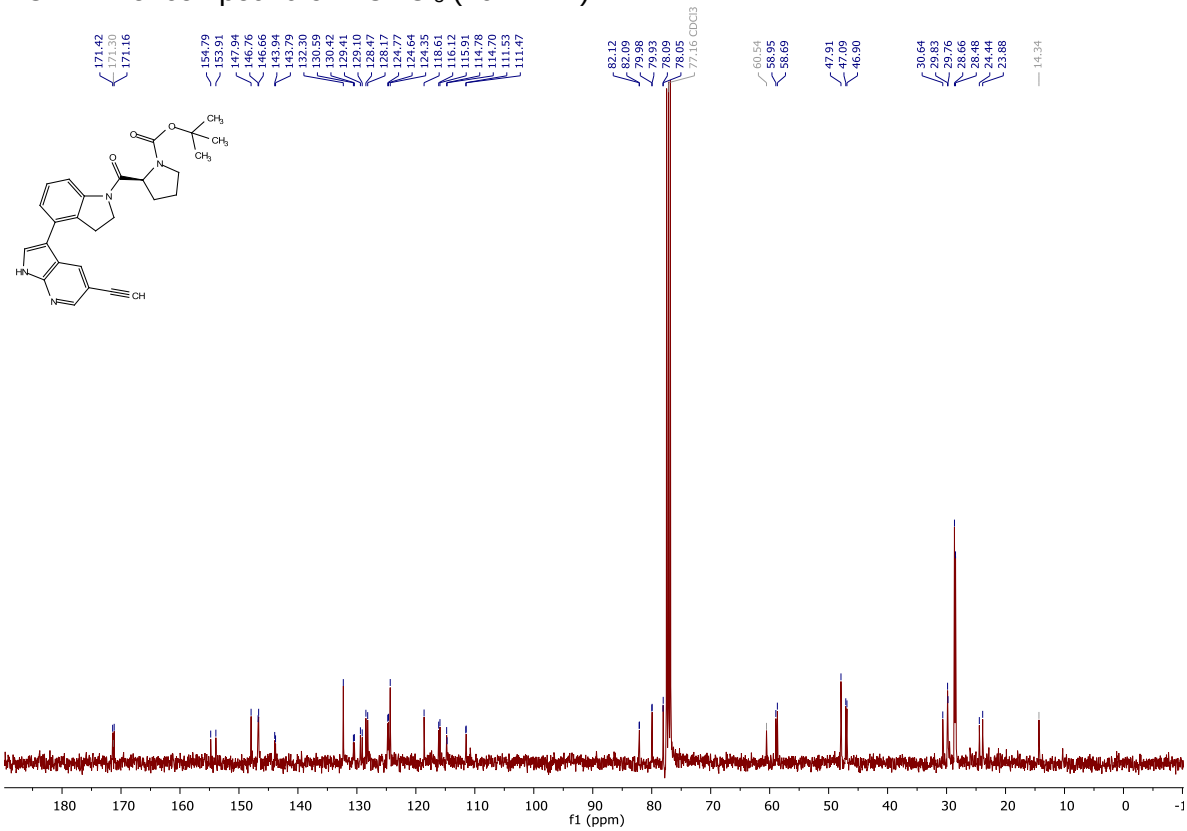
^{13}C-NMR of compound **8** in $CDCl_3$ (101 MHz)



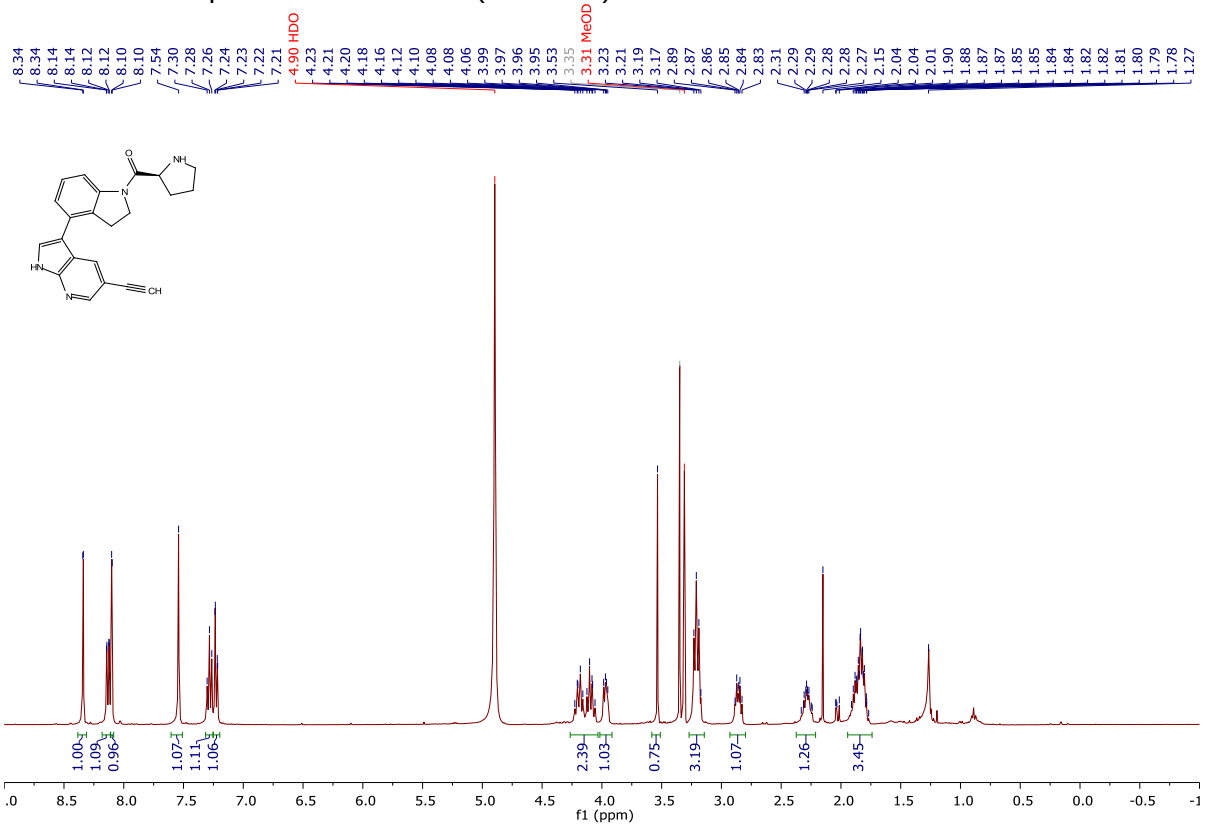
^{1}H-NMR of compound **9** in $CDCl_3$ (400 MHz)



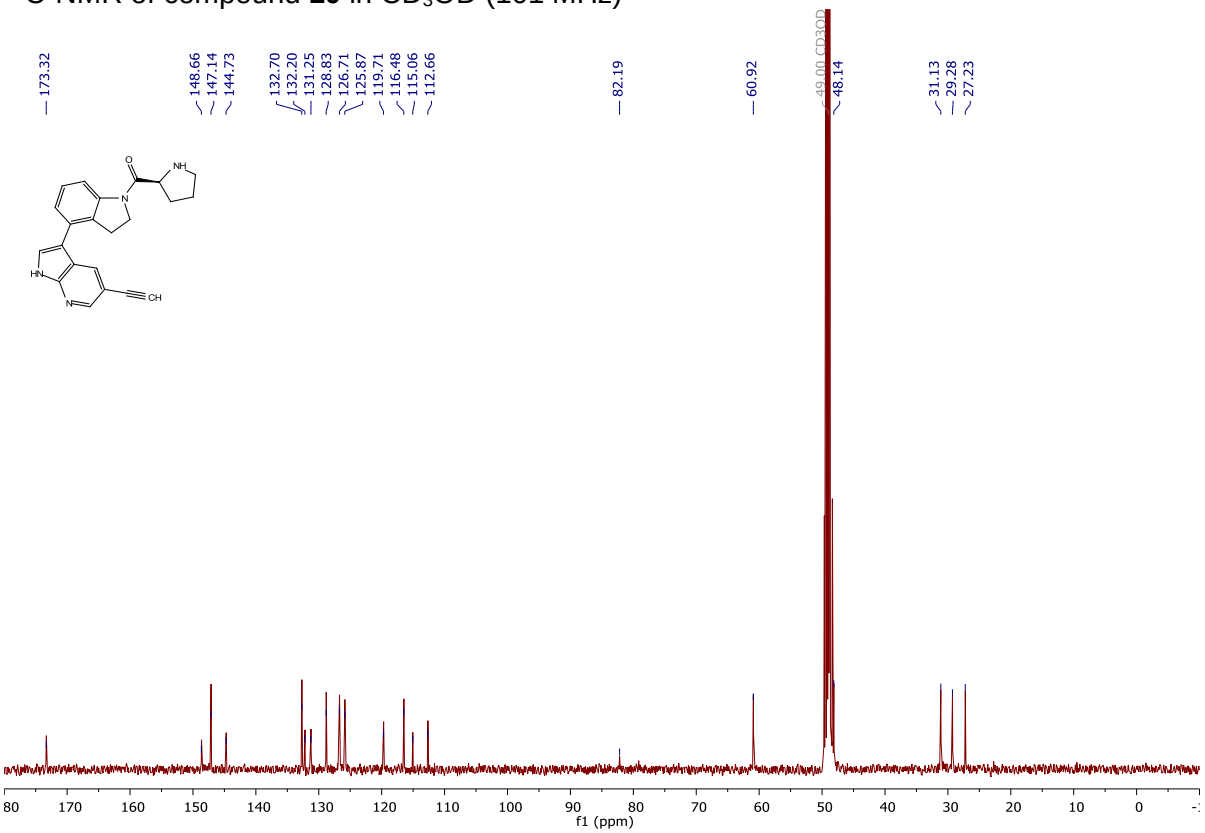
^{13}C-NMR of compound **9** in $CDCl_3$ (101 MHz)



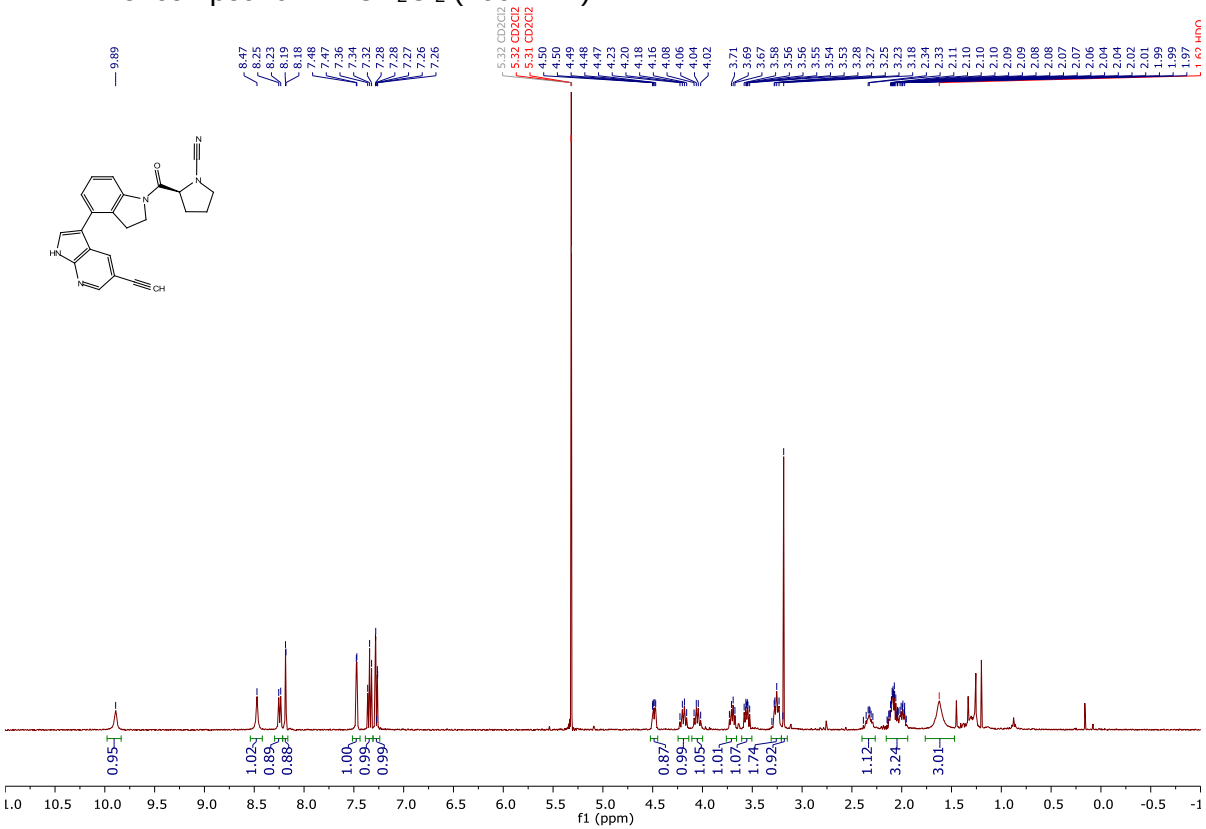
^{1}H-NMR of compound **10** in CD_3OD (400 MHz)



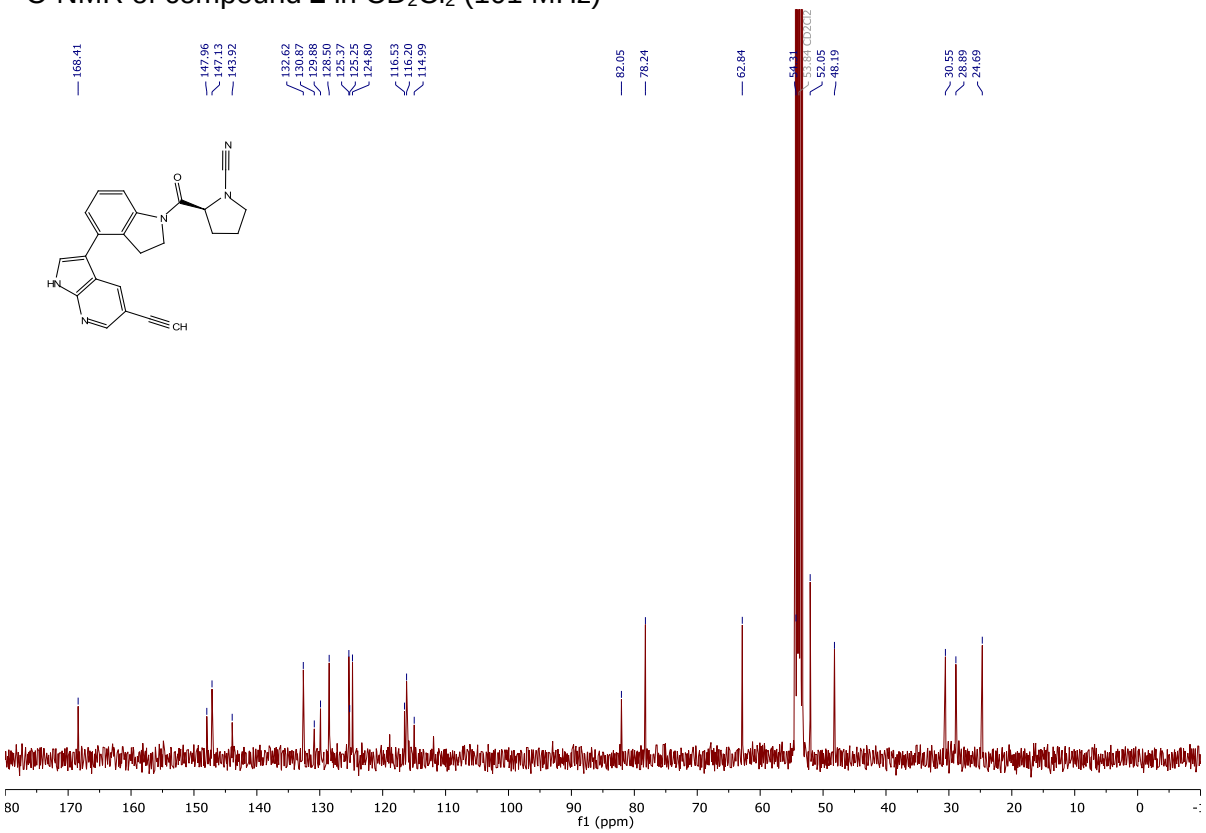
^{13}C-NMR of compound **10** in CD_3OD (101 MHz)



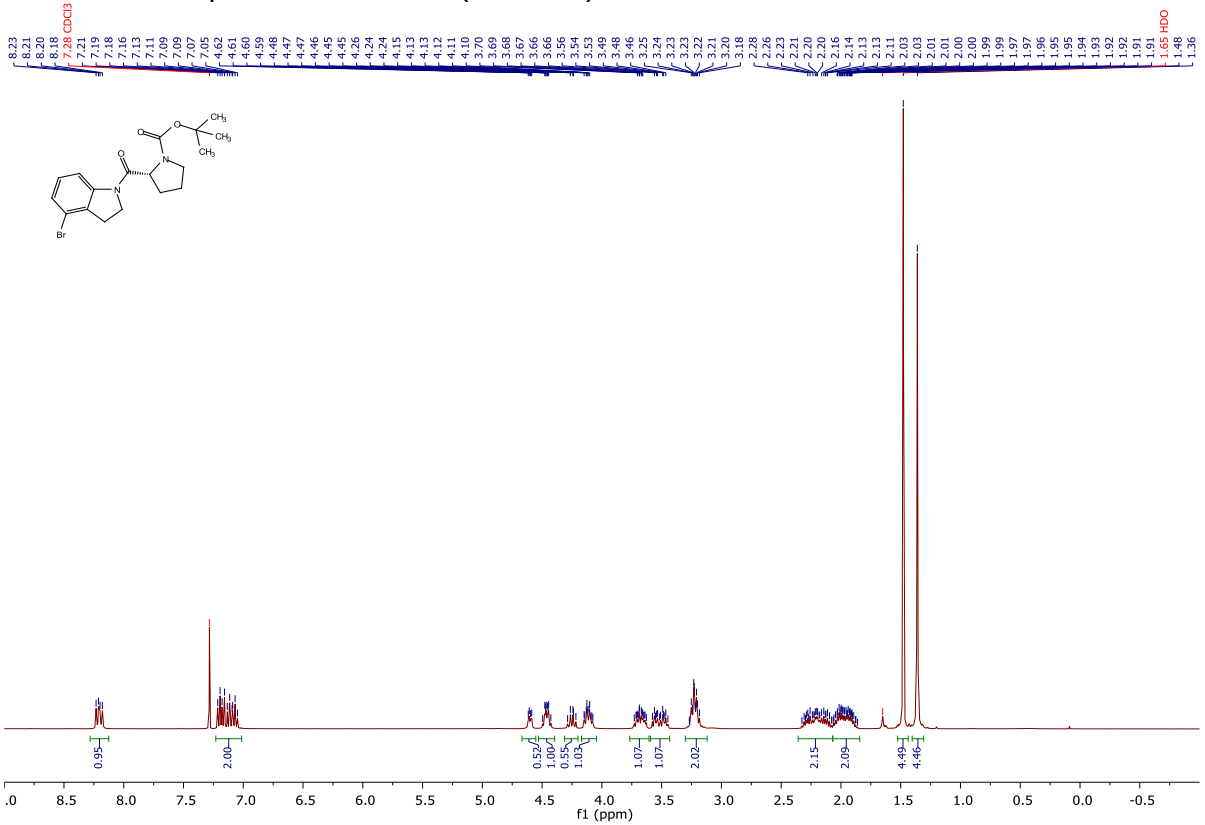
^{1}H-NMR of compound **2** in CD_2Cl_2 (400 MHz)



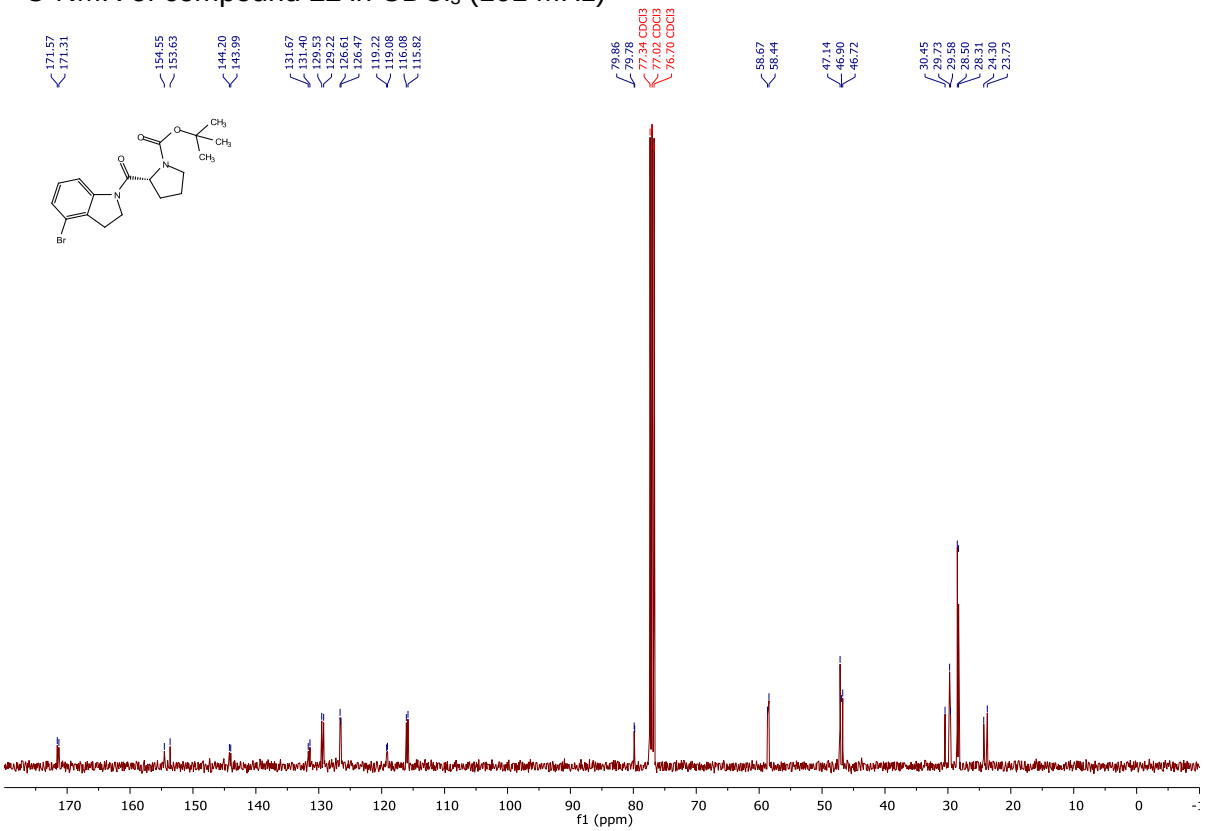
^{13}C-NMR of compound **2** in CD_2Cl_2 (101 MHz)



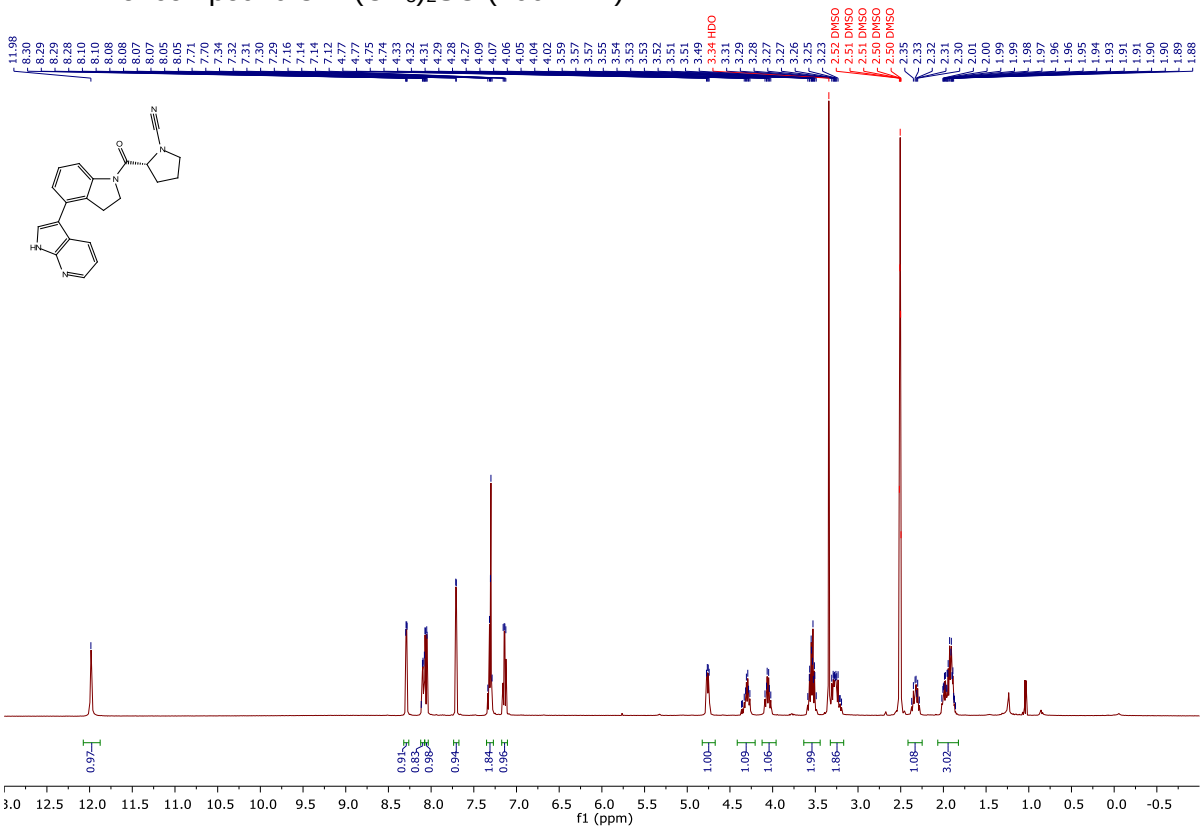
^{1}H-NMR of compound **11** in $CDCl_3$ (400 MHz)



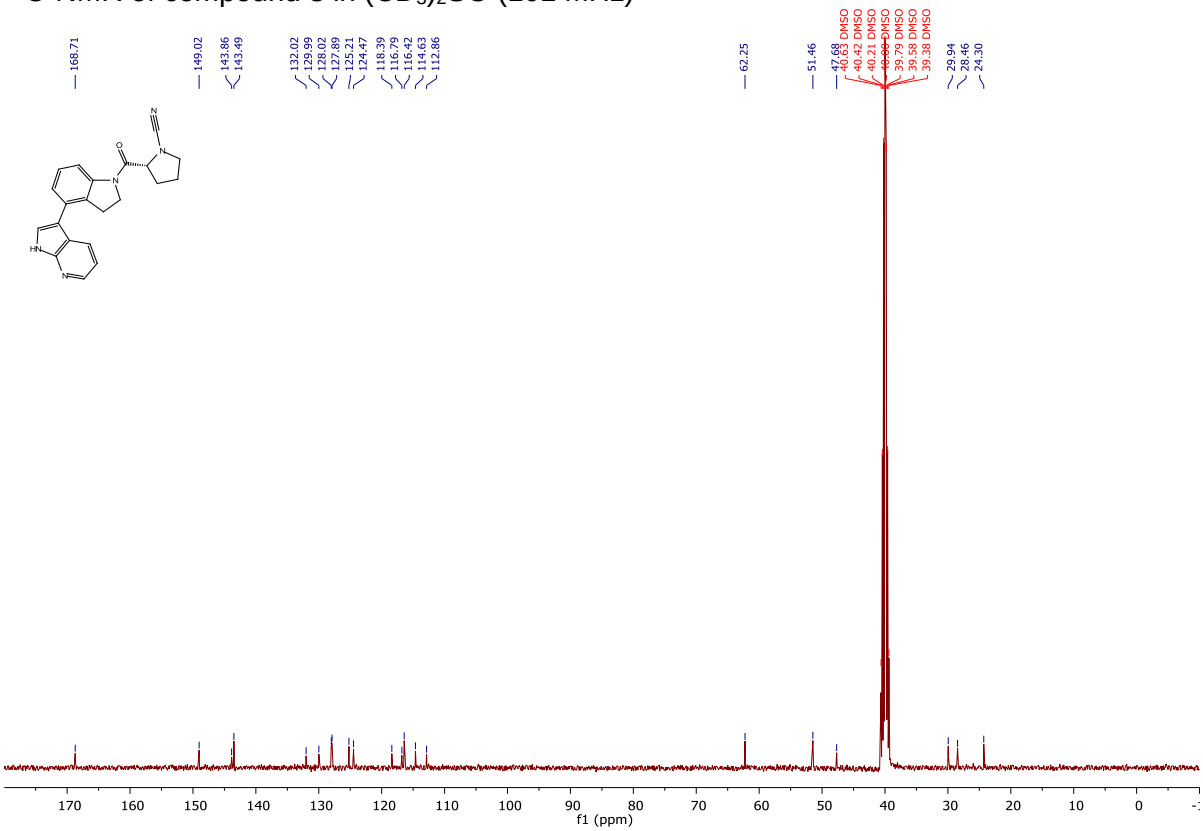
^{13}C-NMR of compound **11** in $CDCl_3$ (101 MHz)



^{1}H-NMR of compound **3** in $(CD_3)_2SO$ (400 MHz)



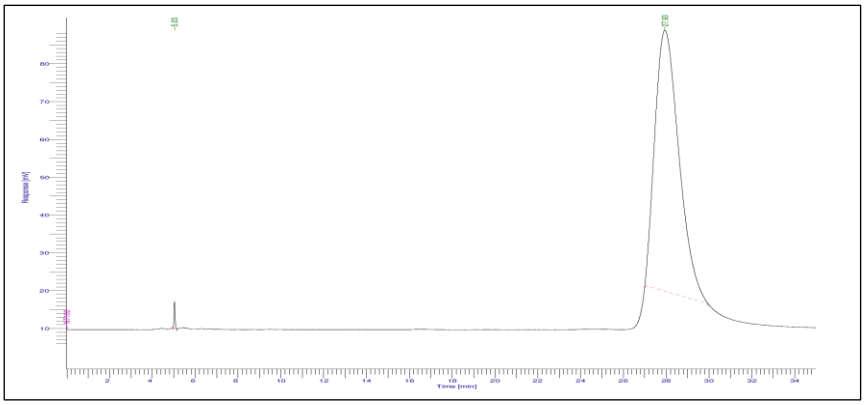
^{13}C-NMR of compound **3** in $(CD_3)_2SO$ (101 MHz)



Appendix 2: Chiral HPLC traces

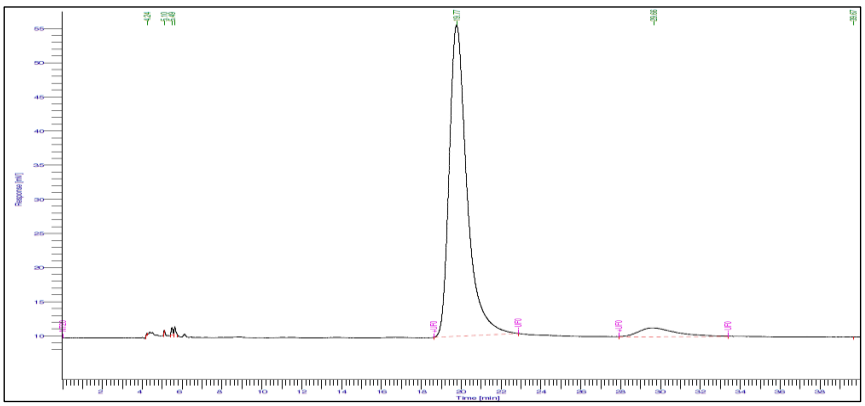
Compound 1

Peak #	Time (min)	Area (μV·s)	Height (μV)	Area (%)	Norm. Area (%)	Area/Height (s)
1	5.03	27139.7	7104.48	0.5	0.5	3.8201
2	**27.933**	**5367244**	**68966.5**	**99.5**	**99.5**	**77.824**
		5394383	76070.9	100	100	



Compound 2

Peak #	Time (min)	Area (μV·s)	Height (μV)	Area (%)	Norm. Area (%)	Area/Height (s)
1	4.242	329.07	39.53	0.01	0.01	8.3254
2	5.102	264.56	130.15	0.01	0.01	2.0327
3	5.490	6283.15	1082.68	0.21	0.21	5.8033
4	5.631	6514.30	1239.12	0.22	0.22	5.2572
5	**19.770**	**2809296.75**	**45592.13**	**93.94**	**93.94**	**61.6180**
6	29.663	167765.04	1294.69	5.61	5.61	129.5791
7	39.665	23.77	56.45	0.00	0.00	0.4211
		2990476.64	49434.74	100	100	



Compound 3

Peak #	Time (min)	Area (μV·s)	Height (μV)	Area (%)	Norm. Area (%)	Area/Height (s)
1	5.059	15806.51	3989.40	0.06	0.06	3.9621
2	**24.423**	**28216910.36**	**381202.79**	**99.94**	**99.94**	**74.0207**
		28232716.87	385192.19	100	100