Supplementary Information

Quantitative Chemoproteomic Profiling Reveals Multiple Target Interactions of Spongiolactone Derivatives in Leukemia Cells

M. H. Wright,[a, d] Y. Tao,[c] J. Drechsel,[a] J. Krysiak,[a] S. Chamni,[b,Δ] A. Weigert-Munoz,[a] N. L. Harvey,[b] D. Romo[c] * and S. A. Sieber[a] *

a. Center for integrated Protein Science (CIPSM), Department of Chemistry, Technical University of Munich, Lichtenbergstr. 4, D-85748, Garching, Germany.

b. Department of Chemistry, Texas A&M University, P.O. Box 30012, College Station, TX 77842 (USA).

c. Department of Chemistry & Biochemistry, Baylor University, One Bear Place 97348, Waco, TX 76798 (USA).

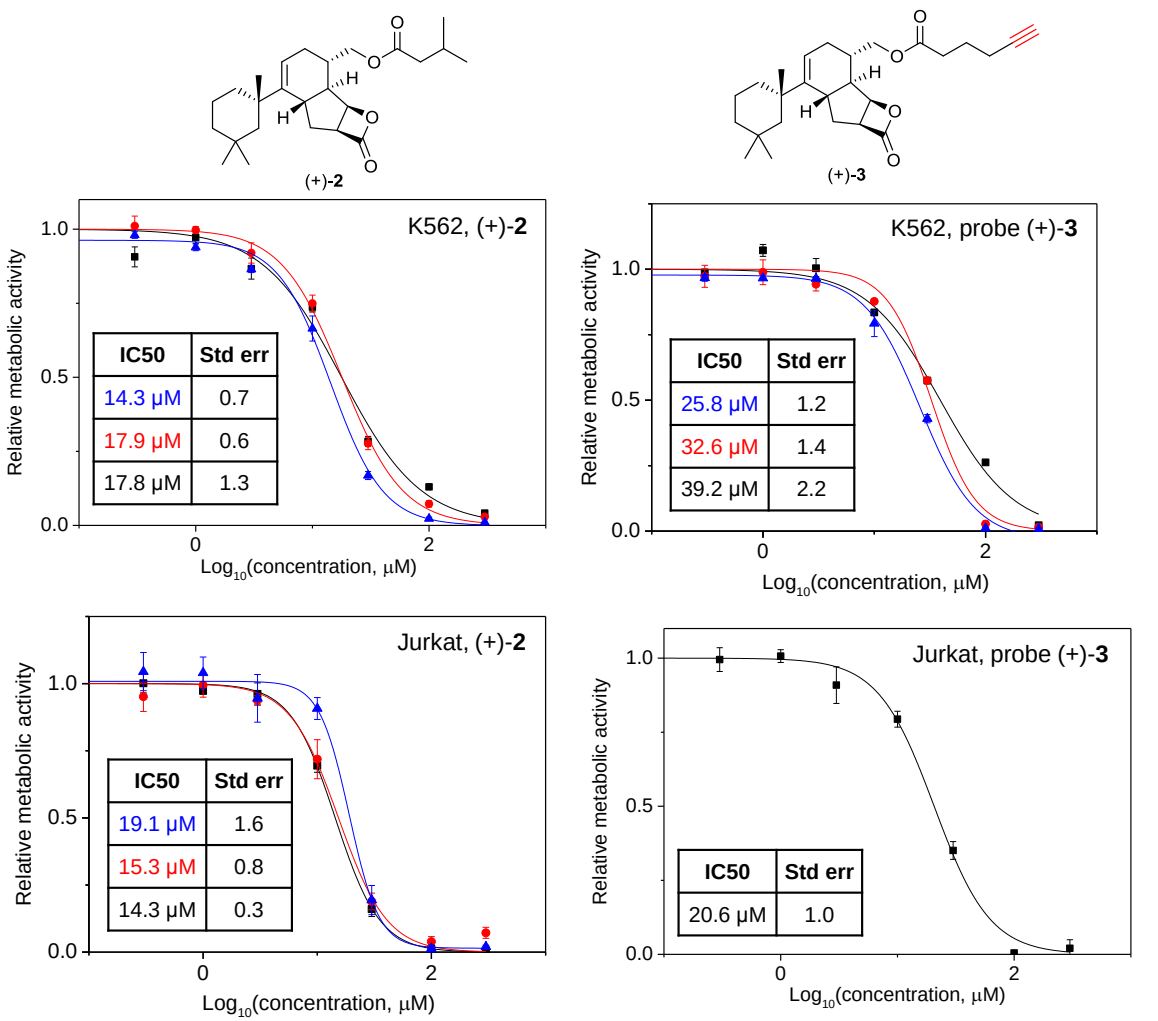
d. School of Chemistry, University of Leeds, Leeds, LS2 9JT, UK.

ΔCurrent address: Department of Pharmacognosy and Pharmaceutical Botany, Faculty of Pharmaceutical Science, Chulalangkorn University, Bangkok, Thailand.

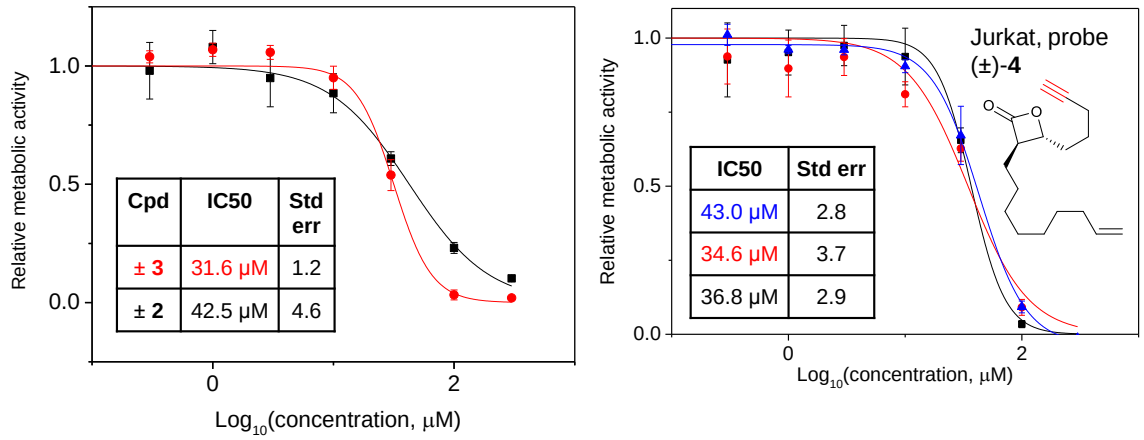
* Co-corresponding authors: stephan.sieber@mytum.de; Daniel_Romo@baylor.edu

Contents

1 Supporting Figures S1-S9



Racemates:



Activity on A549 cells:

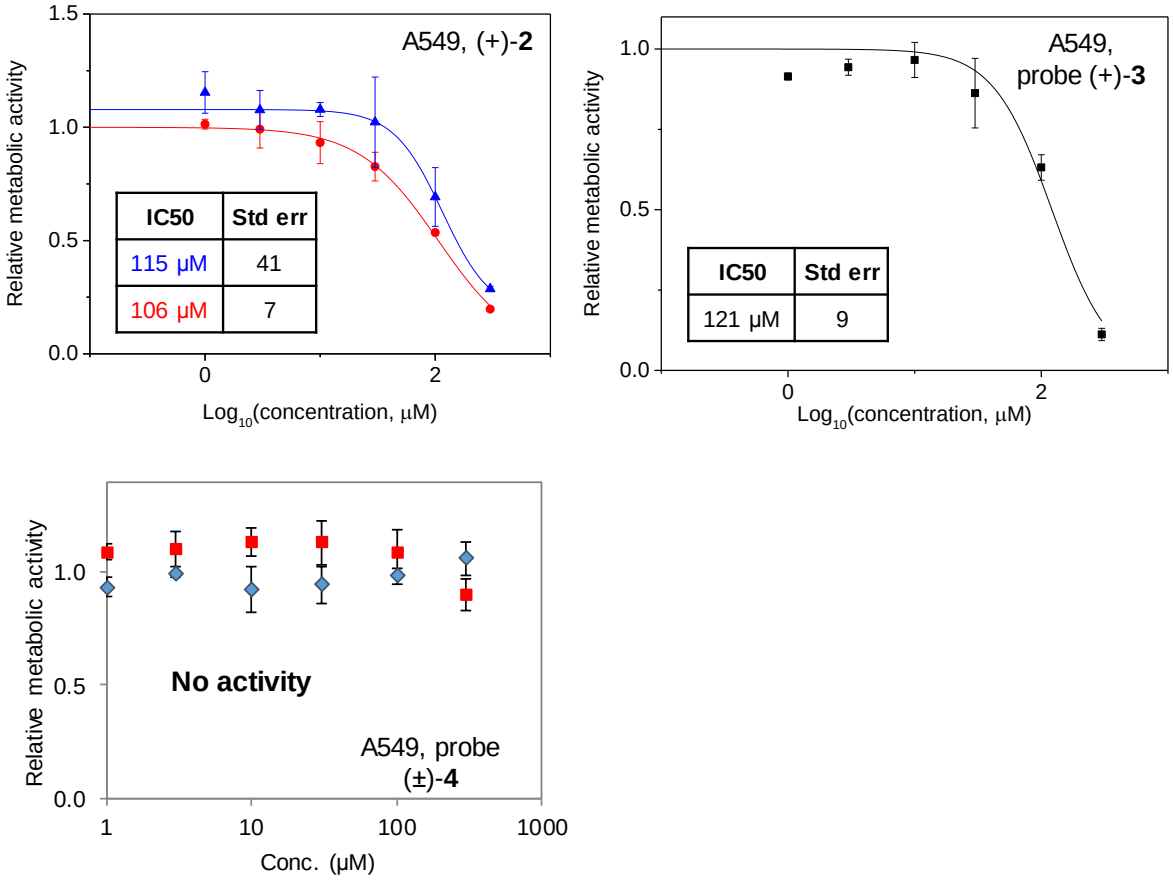


Figure S1. Cytotoxicity assay (MTT) data for compounds in different cell lines after 24 h. Each colour represents an independent biological replicate, performed in technical triplicate. Datapoints are at the mean with error bars indicating standard deviation.

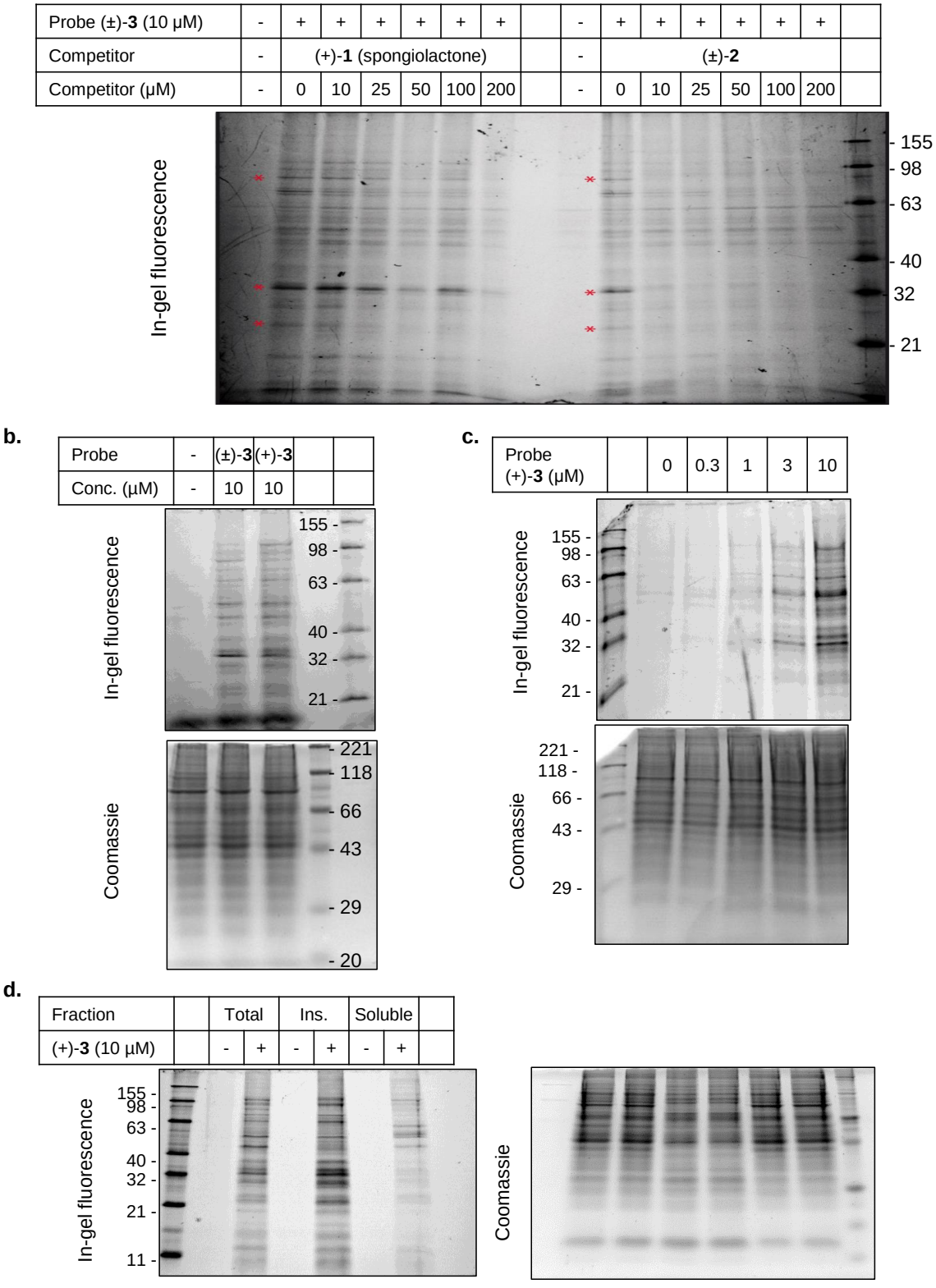
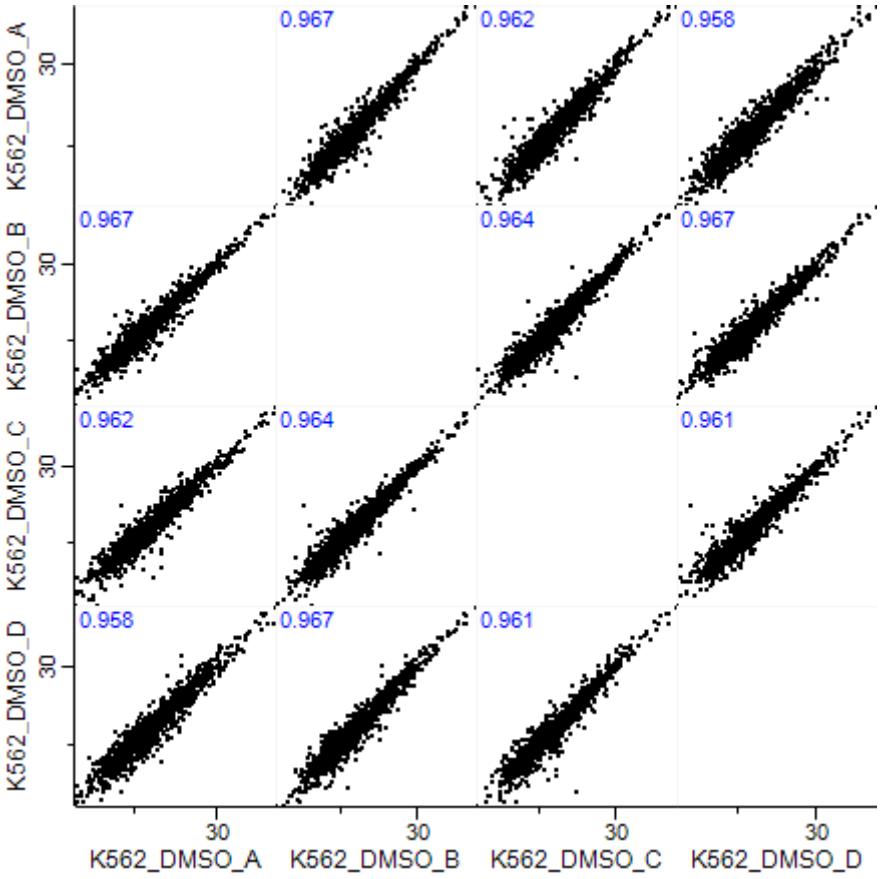
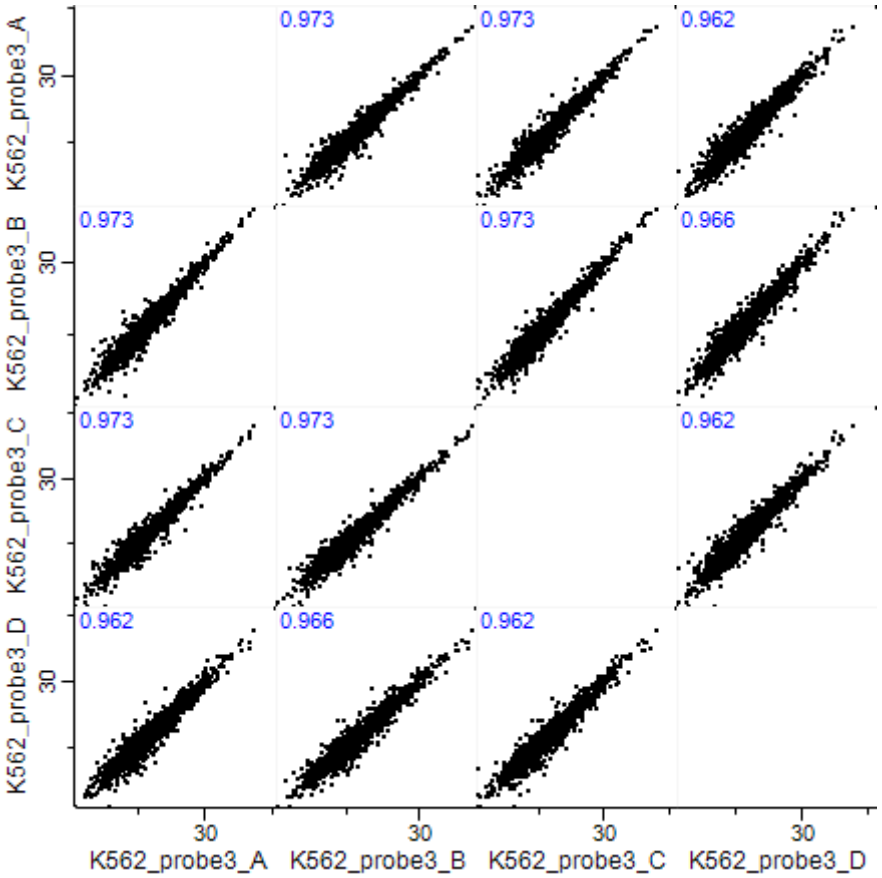


Figure S2. Gel-based data for spongiolactone-derived probes in K562 cells. **a.** Labelling and competition experiments with racemic (±)-**3**. Total lysate. Competition experiments were performed by pre-incubating cells in the presence of (+)-**1** or (±)-**2** as indicated for 30 min, before labelling with (±)-**3** for 2 h. * = bands clearly outcompeted by competitors. **b.** Comparison between racemic probe (±)-**3** and enantiopure probe (+)-**3** (used for all subsequent experiments). **c.** Concentration-dependent labelling of (+)-**3**. **d.** Fractionation of (+)-**3** labelled cells into insoluble (ins.) and soluble fractions. All probe incubations performed for 2 h.

a.


0.967
0.962
0.958
0.967
0.964
0.967
0.962
0.964
0.961
0.958
0.967
0.961
K562_DMSO_A
K562_DMSO_B
K562_DMSO_C
K562_DMSO_D
30


0.973
0.973
0.962
0.973
0.973
0.966
0.973
0.973
0.962
0.962
0.966
0.962
K562_probe3_A
K562_probe3_B
K562_probe3_C
K562_probe3_D
30

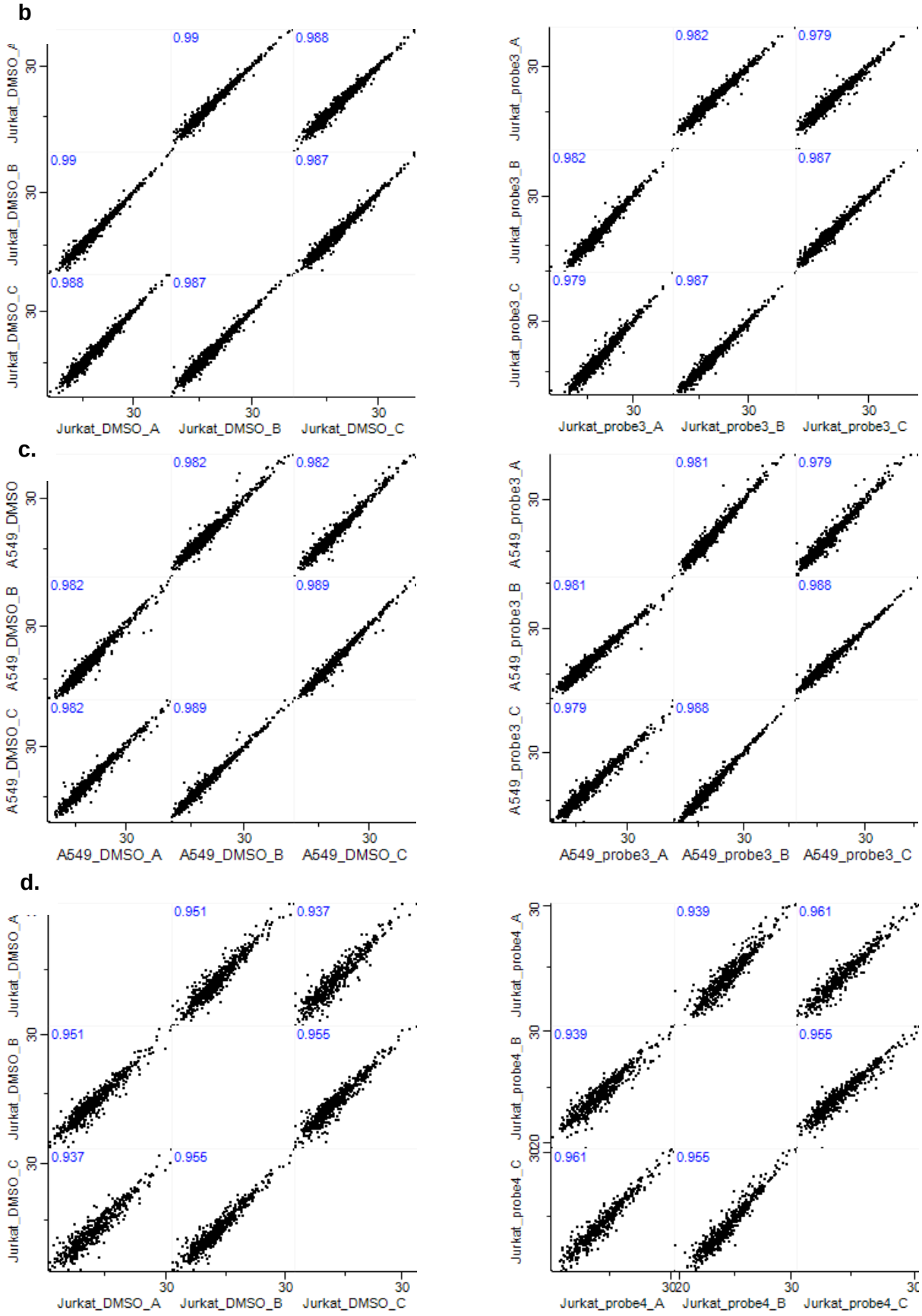


Figure S3. Scatter plots assessing the reproducibility of label-free proteomic data across biological replicates (10 µM probe, 2 h labelling). Pearson correlation shown in blue. **a.** (+)-**3** labelling in K562 cells (n=4 replicates, A, B, C, D). **b.** (+)-**3** labelling in Jurkat cells (n=3, A-C). **c.** (+)-**3** labelling in A549 cells (n=3, A-C). **d.** (±)-**4** labelling in Jurkat cells (n=3, A-C).

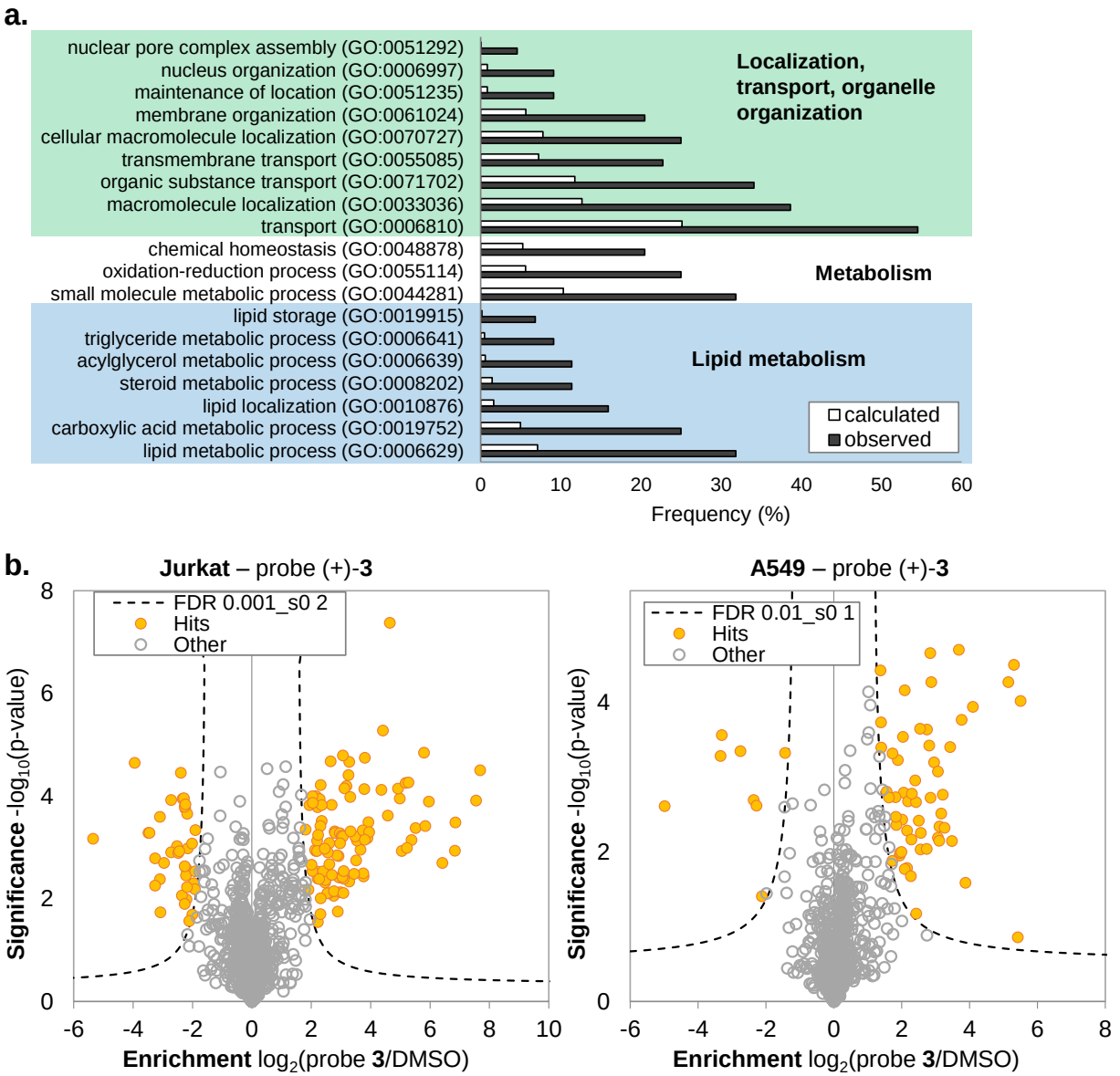


Figure S4. Supporting proteomic data for probe (+)-**3** (10 µM probe, 2 h labelling). **a.** Gene Ontology (GO) enrichment analysis of 48 hits common to K562 and Jurkat cells relative to the total human proteome. Calculated = frequency of GO term in total proteome; observed = frequency of GO term in sample. GO terms shown are statistically significantly enriched (by hypergeometric test; significance level 0.05; with Benjamini & Hochberg False Discovery Rate (FDR) correction). **b.** Volcano plots of proteomic data from enrichment experiments after labelling with (+)-**3** in Jurkat (left) and A549 (right) cells.

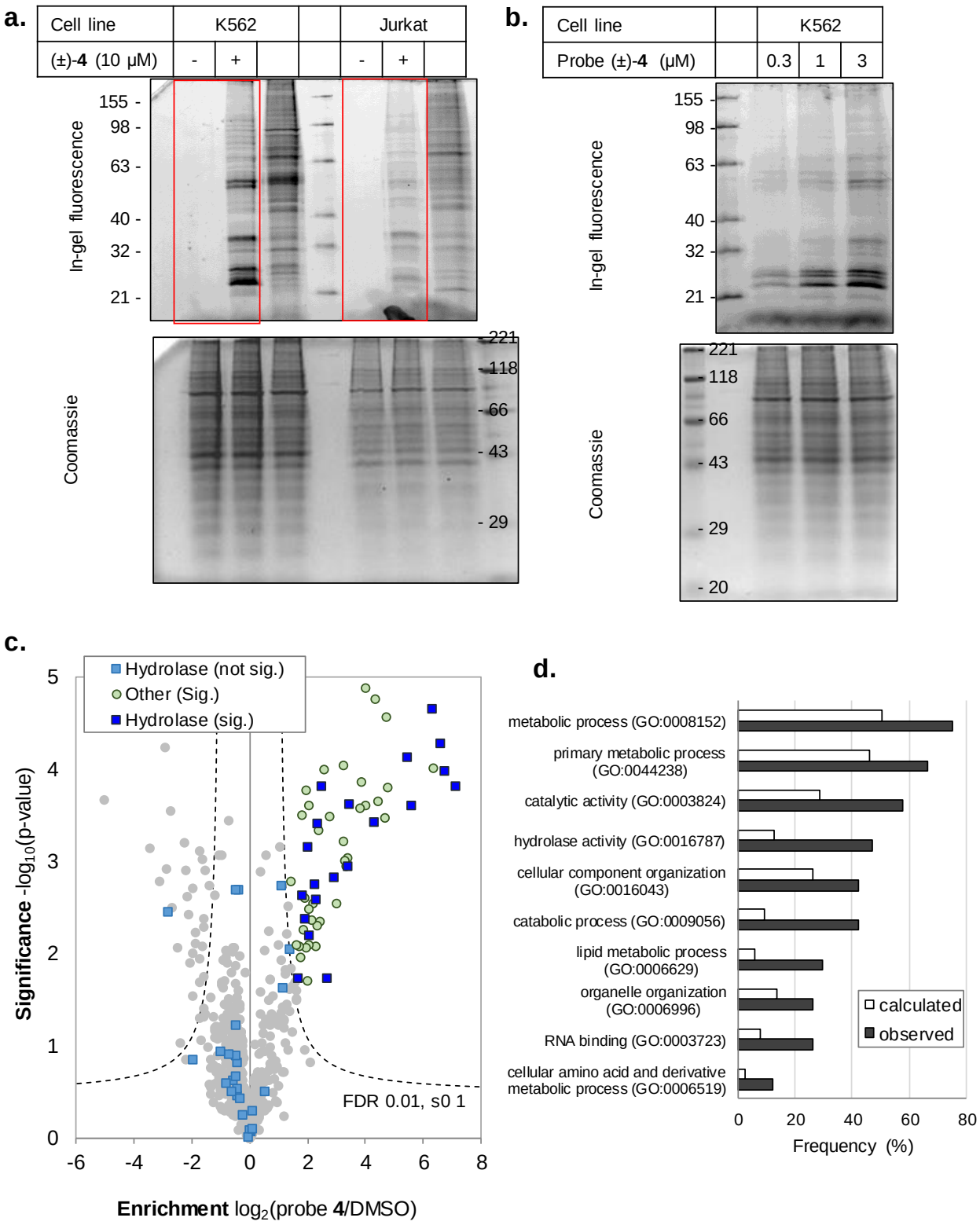


Figure S5. Supporting gel and proteomic data for probe (±)-**4. a.** Labelling with probe **4** in K562 and Jurkat cells (10 µM probe, 2 h labelling). Red rectangles indicate portions of gel shown in Figure 3 of main text**. b.** Concentration-dependent labelling with **4** in K562 cells. **c.** Volcano plot of proteomic data from enrichment experiments after labelling with **4** in Jurkat cells (10 µM, 2 h). Proteins with annotated hydrolase activity are coloured blue. **d.** Gene Ontology (GO) enrichment analysis of hits relative to the total human proteome. Calculated = frequency of GO term in total proteome; observed = frequency of GO term in sample. GO terms shown are statistically significantly enriched (by hypergeometric test; significance level 0.05; with Benjamini & Hochberg False Discovery Rate (FDR) correction).

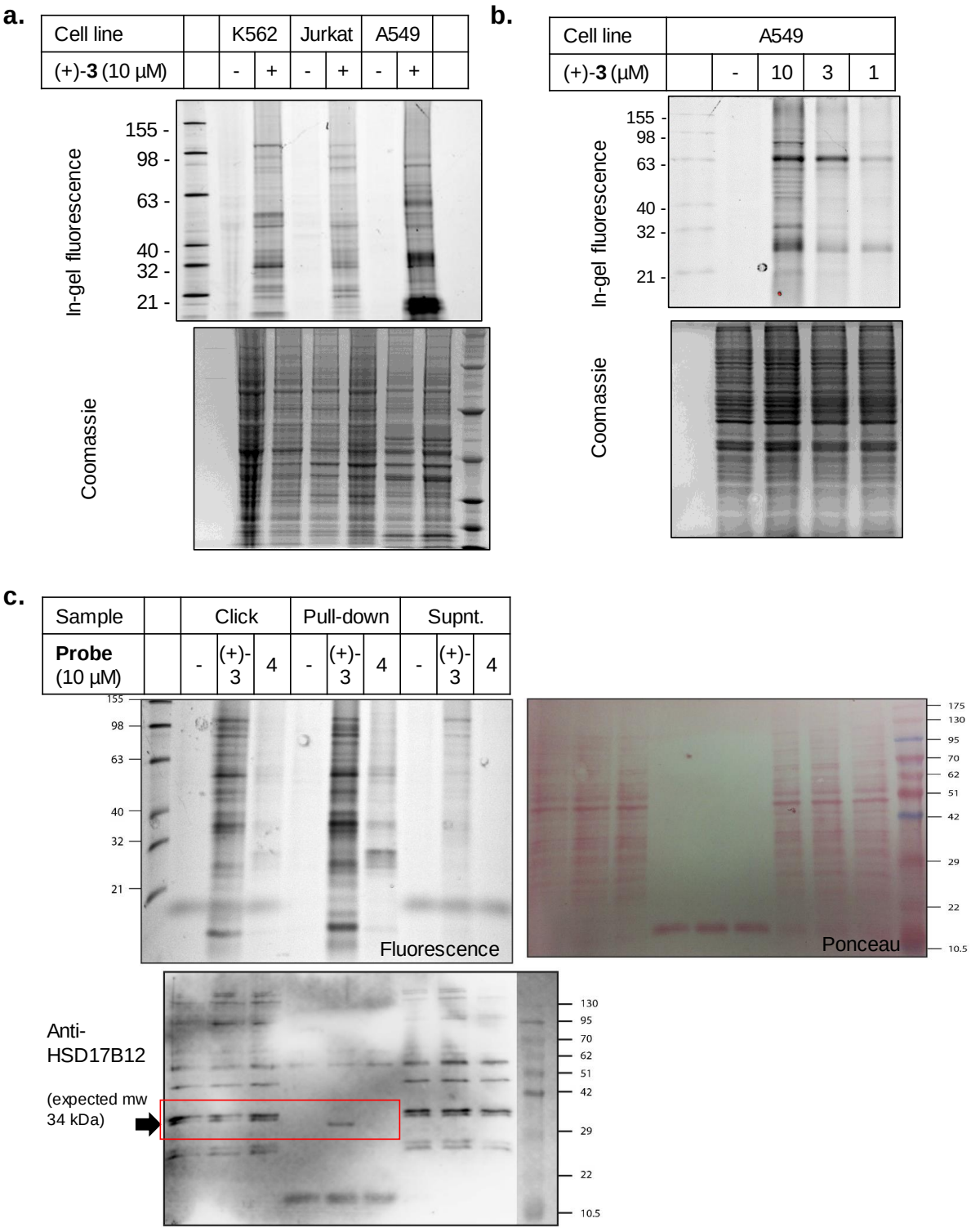


Figure S6. Gel-based data for probe (+)-**3** comparative labelling. **a.** Labelling (2 h) in three cell lines. Total fluorescence gel and Coomassie for gel shown in main text Figure 3. **b.** Concentration-dependent labelling with (+)-**3** in A549 cells (2 h). **c.** Whole gel and blot from Fig. 3c: validation of HSD17B12 as a (+)-3-specific target. K562 cells were incubated with DMSO (-), (+)-**3** or (±)-**4** for 2 h at 10 µM probe concentration, lysed and ligated to a trifunctional azide-biotin-tamra reagent. A portion (click) was removed, and the remainder enriched on avidin beads. Samples of the supernatant (Supnt) were also taken. The beads were boiled and the samples run on a gel. Top: in-gel fluorescence. Middle: Western blot for HSD17B12. Red boxed area shown in Fig. 3c. Right: Ponceau staining to show equal loading.

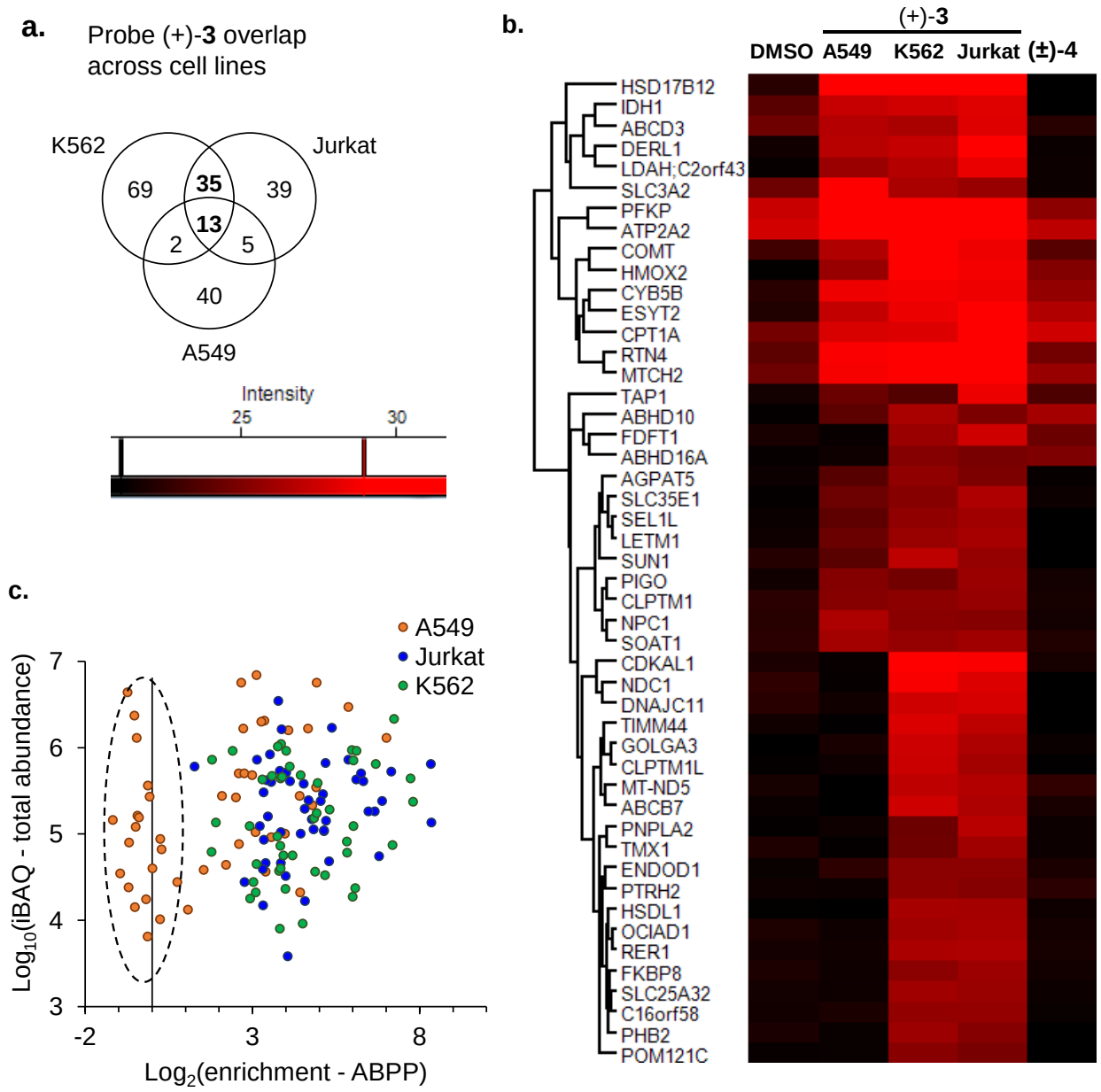
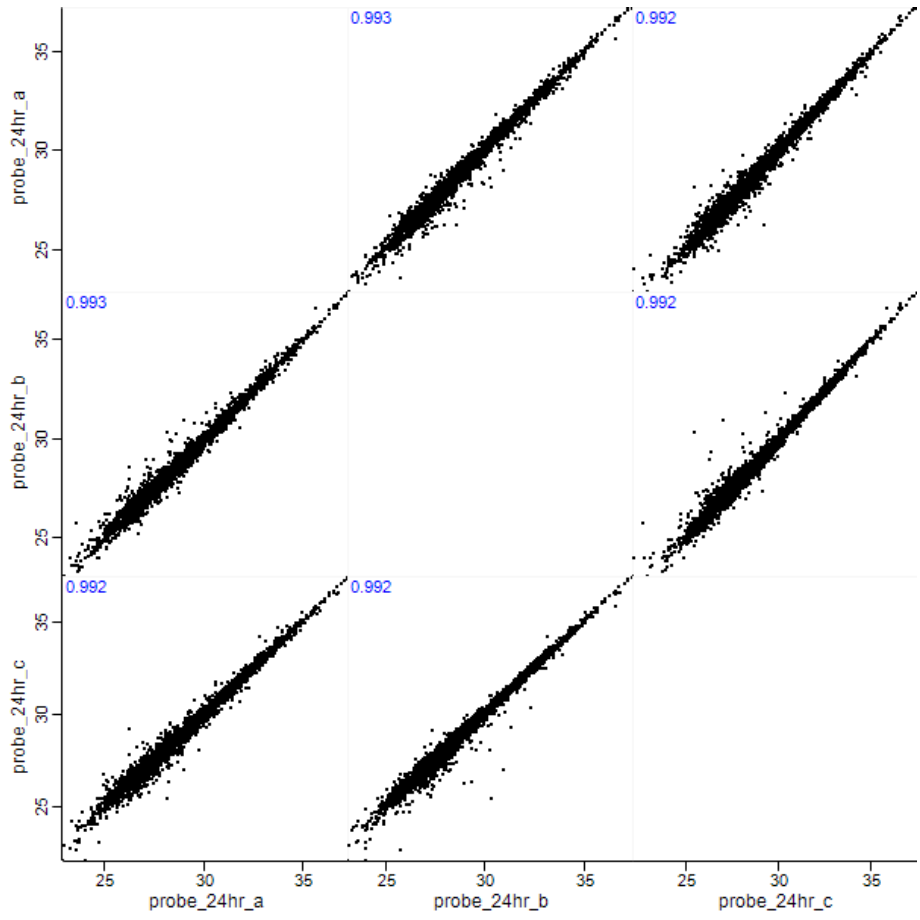
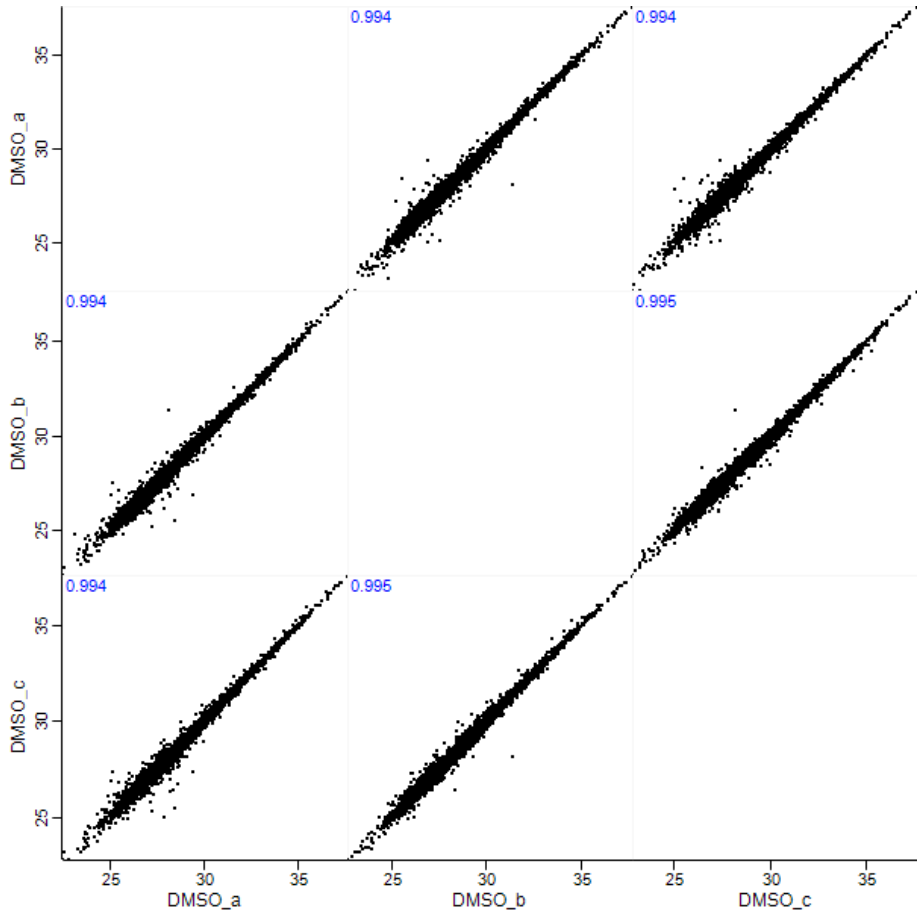


Figure S7. Comparative analysis of labelling with (+)-**3** (10 µM, 2 h) across the three different cell lines, and with (±)-**4**-specific hits (**4** in Jurkat cells). **a.** Venn overlap of (+)-**3** hits. **b.** Heat map showing hierarchical clustering of hits. Coloured according to mean LFQ intensity. A reduced version of this heatmap (with proteins showing intensity in DMSO controls removed) is shown in main text Figure 3. **c.** Plot of enrichment (probe (+)-**3** vs DMSO) versus abundance as given in Proteomics DB (iBAQ values used) for protein hits in leukemia cell lines (see Table S4). Circled datapoints: not hits in A549.


0.993
0.992
0.993
0.992
0.992
0.992
probe_24hr_a
probe_24hr_b
probe_24hr_c
25
30
35


0.994
0.994
0.994
0.995
0.994
0.995
DMSO_a
DMSO_b
DMSO_c
25
30
35

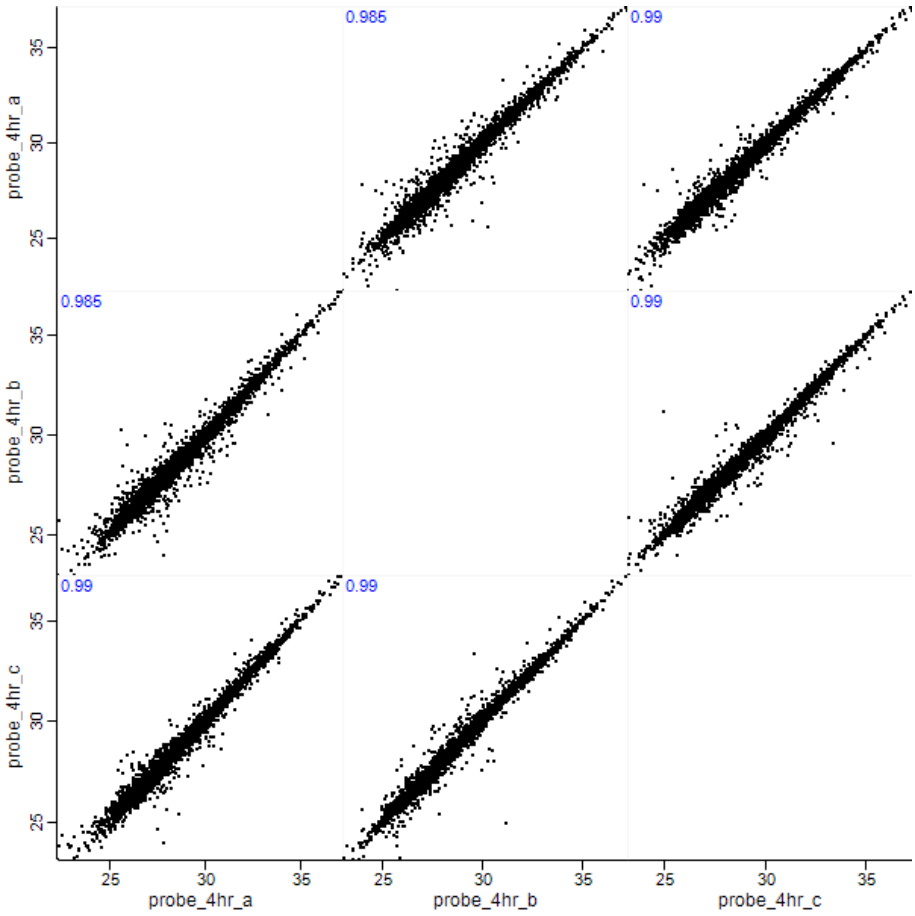


Figure S8. Scatter plots assessing the reproducibility of label-free total proteomic data (LFQ) across replicates (30 µM probe (+)-**3**, timing as indicated). Pearson correlation shown in blue. n=3, a-c replicates. This dataset corresponds to Table S5.

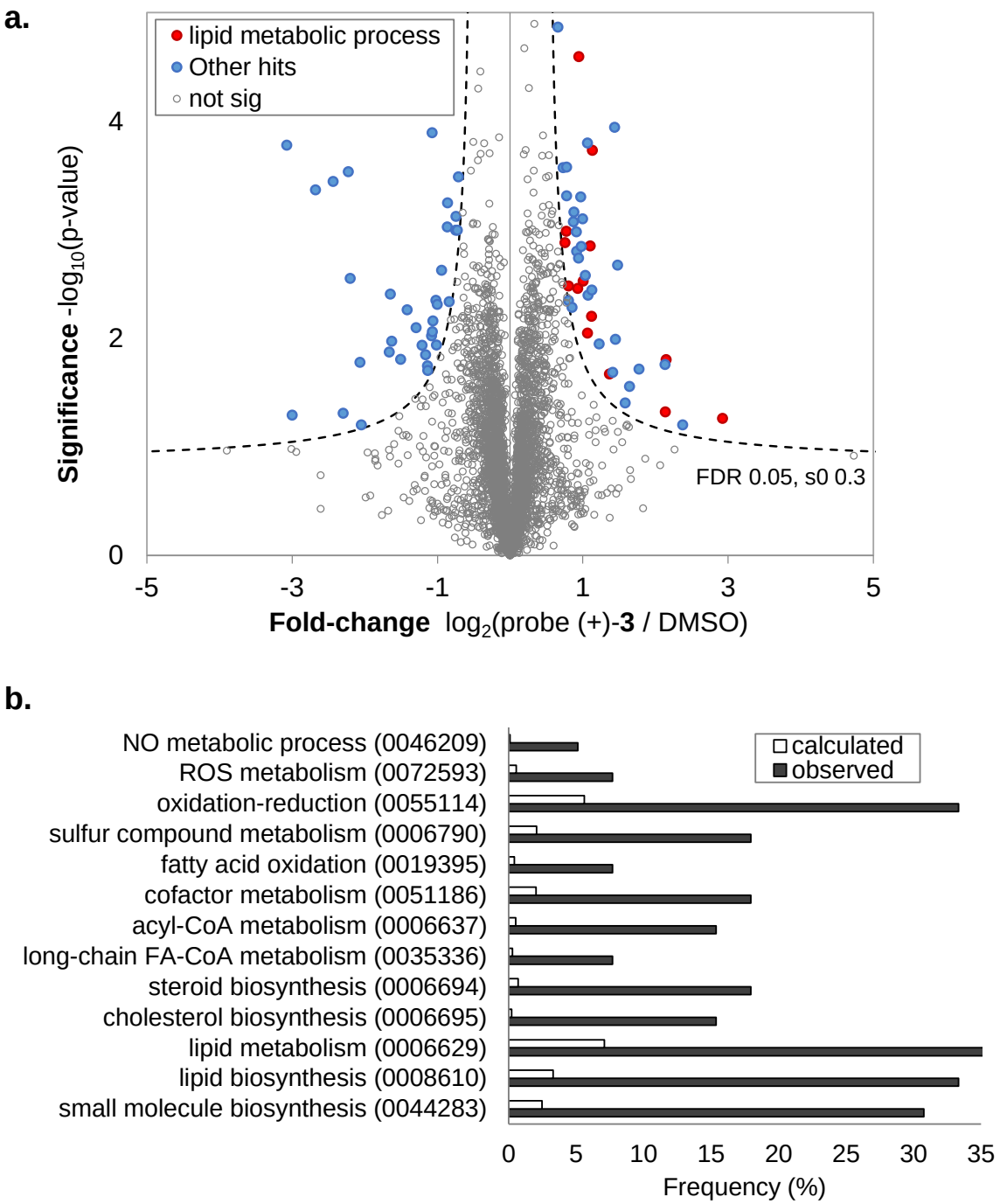


Figure S9. Total proteome data. **a.** Volcano plot of fold-change in abundance following treatment of K562 cells with 30 µM (+)-**3** probe for 24 h. Shown: the 4247 proteins quantified in triplicate (LFQ). 76 are significantly changed in response to the probe. Those proteins upregulated and annotated as involved in lipid metabolism are shown in red. **b.** GO enrichment analysis of upregulated hits (41). GO ID in brackets. Calculated = frequency of GO term in human proteome; observed = frequency of GO term in sample. GO terms shown are statistically significantly enriched (by hypergeometric test; significance level 0.05; with Benjamini & Hochberg False Discovery Rate (FDR) correction). The full analysis is shown in Table S5.

2 Supporting Tables S1-S5

See separate Excel files.

Table S1. Chemical proteomic identification of probe (+)-**3**-labelled proteins in K562 cells.

Table S2. Chemical proteomic identification of probe (+)-**3**-labelled proteins in Jurkat cells and comparison with K562 cells.

Table S3. Chemical proteomic identification of probe (±)-**4**-labelled proteins in Jurkat cells and comparison with **3**-labelled proteins in Jurkat cells.

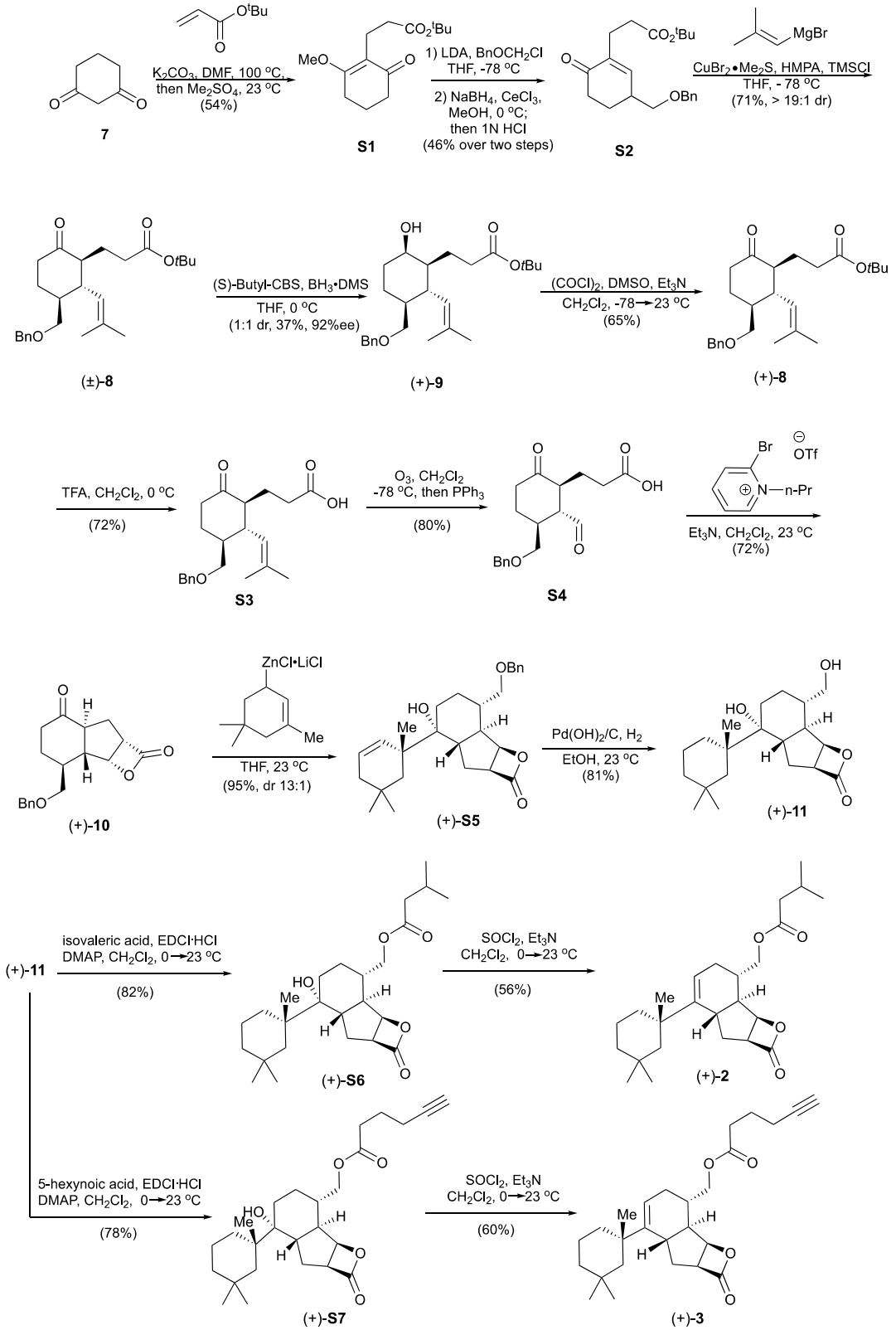
Table S4. Chemical proteomic identification of probe (+)-**3**-labelled proteins in A549 cells and comparison with data from other cell lines (Table S1 data). iBAQ total abundance levels of hits.

Table S5. Global proteomics analysis following treatment of K562 cells with 30 µM (+)-**3** for 24 or 4 h.

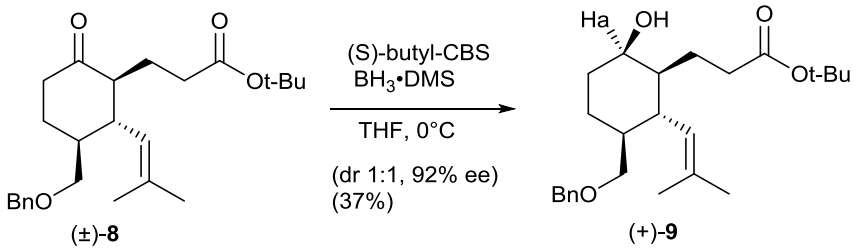
3 Synthetic chemistry

Probe (±)-**4** and reagent AzRB were prepared as previously described.[1, 2]

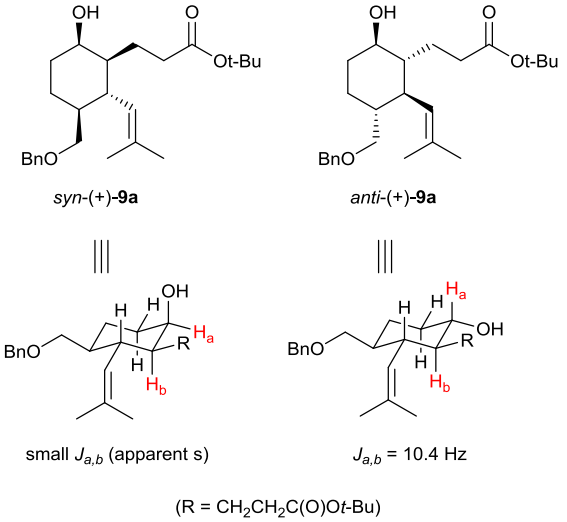
3.1 Synthesis of compound (+)-2 and probe (+)-3



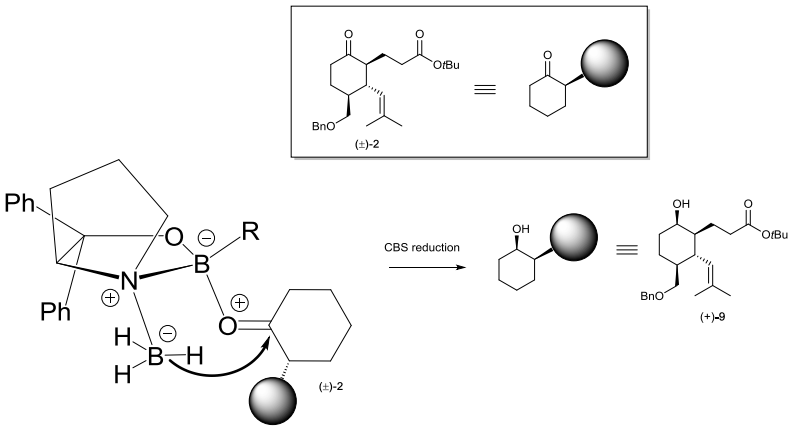
Scheme 1. Synthesis of compound (+)-**2** and probe (+)-**3**

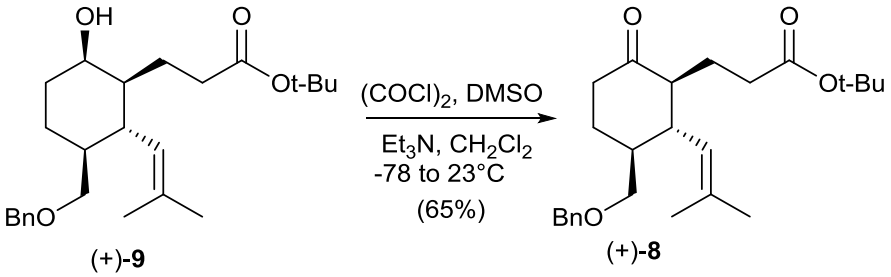


Alcohol syn-(+)-9. The reduction was performed according to the literature procedure.[3, 4] The (*S*)-butyl-CBS catalyst (1 M solution in toluene, 1.4 mL, 0.2 equiv) was dissolved in THF at 0 °C, into which borane dimethyl sulfide complex (0.7 mL, 1.0 equiv) was added and the mixture was stirred at that temperature for 30 minutes. The above solution was then added to a THF solution of ketone (±)-**8** (2.800 g, 1.0 equiv) and stirred at 0 °C for 4 h until complete reduction was determined by TLC. The reaction mixture was then carefully quenched, concentrated, and purified by flash chromatography (SiO_2, 0 → 25% EtOAc/Hexanes) to afford a pair of separable diastereomers (dr ~1:1). The *syn*-diastereomeric alcohol (+)-**9** had a slightly higher R_f compared to the *anti*-diastereomer and was obtained in 37% yield (1.021 g) and 92% ee (chiral HPLC, OD column) as a colorless oil. The relative stereochemistry of the *syn* vs *anti* diastereomeric alcohols **9** was determined by coupling constant analysis of the secondary alcohol methine proton (H_a). In the *syn* diastereomer (+)-**9**, H_a appears as an apparent singlet whereas in the *anti* diastereomer, this proton is an apparent dt ($J_{a,b}$ = 10.2 Hz). A racemic sample of the *syn*-diastereomer was prepared by $NaBH_4$ reduction of ketone (±)-**8** for comparison following separation of *syn*/*anti* diastereomers. Data for syn (+)-**9a**. $[\alpha]_D^{20}$ = +8.01 (*c* 1.00, $CHCl_3$); TLC (EtOAc:hexanes, 1:3 v/v): Rf = 0.48; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.36 – 7.28 (m, 5H), 4.74 (dq, *J* = 10.2, 1.4 Hz, 1H), 4.50 – 4.39 (m, 2H), 3.86 (s, 1H), 3.45 (dd, *J* = 9.0, 3.2 Hz, 1H), 3.17 (dd, *J* = 9.0, 8.0 Hz, 1H), 2.33 – 2.14 (m, 4H), 1.94 (app dq, *J* = 13.5, 3.1 Hz, 1H), 1.81 (dq, *J* = 12.7, 3.4 Hz, 1H), 1.75 – 1.58 (m, 6H), 1.56 (d, *J* = 1.4 Hz, 3H), 1.55 – 1.46 (m, 2H), 1.45 (s, 9H), 1.09 (app ddt, *J* = 11.1, 3.8, 2.3 Hz, 1H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 174.3, 138.9, 133.0, 128.2 (2), 128.1, 127.5 (2), 127.3, 80.4, 74.0, 73.0, 65.6, 46.3, 42.9, 38.2, 33.2, 32.3, 28.1 (3), 25.8, 23.9, 23.1, 18.7; HRMS (ESI+) $C_{25}H_{38}O_4Li$ $[M+Li]^+$: 409.2925. Found: 409.2938.

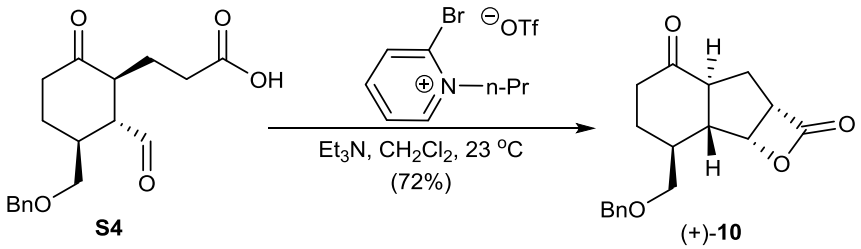


The absolute stereochemistry of alcohol syn-(+)-**9** was assigned by the CBS reduction nemonic,[3] which is most reliable when the steric difference around the ketone is significantly different, as in the case of (±)-**2**. One side of the ketone **2** is substituted while the other side is not, thus leading to differential steric interactions with the boron catalyst. This leads to facial selectivity predicted by the CBS reduction model as pictured above through hydride addition from the front face of the ketone as drawn – shown for the *syn* alcohol (+)-**9**.

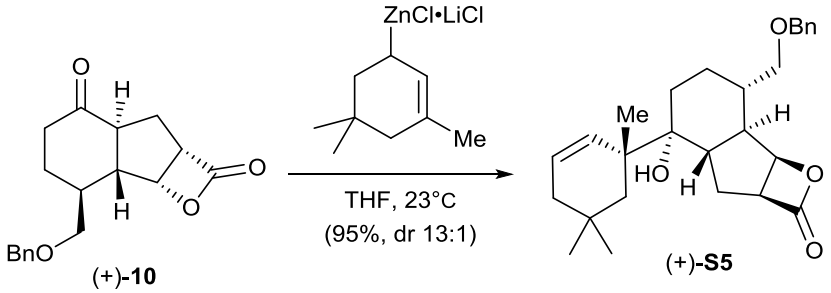




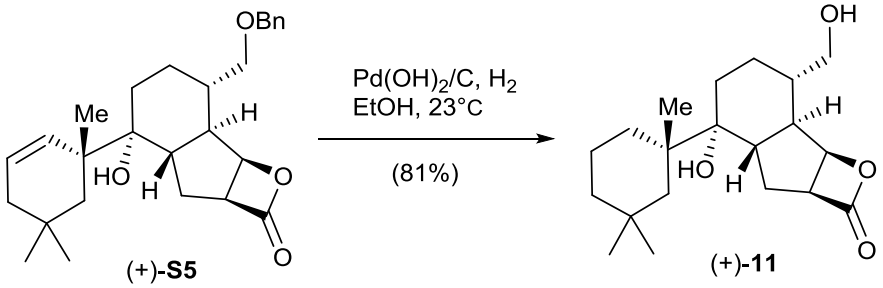
Ketone (+)-8. An oven-dried flask was charged with 50 mL of dry CH_2Cl_2 followed by $(COCl)_2$ (10 mL of 2 M solution in DCM, 5.0 equiv) and the solution was cooled to -78 °C. DMSO (2.94 mL, 10.0 equiv) was slowly added and the reaction mixture was stirred for 30 min at -78 °C. A solution of the alcohol (+)-**9** (1.605 g, 1.0 equiv) in 5 mL CH_2Cl_2 was then added over 30 min and stirred for an additional 2 h at -78 °C. The reaction was then quenched by addition of triethylamine (8.64 mL, 15.0 equiv), warmed to 23 °C temperature, and stirred overnight. The reaction mixture was concentrated and purified by silica gel flash chromatography (0 → 25% EtOAc/hexanes) to afford ketone (+)-**8** (1.015 g, 65%) as a light, yellow oil. $[\alpha]_D^{20}$ = +22.91 (*c* 0.87, $CHCl_3$); Other data matched that previously reported for the racemic ketone (±)-**8**.[5]



β-lactone (+)-10. Modified Mukaiyama's reagent (1.260 g, 3.0 equiv) was dissolved in 50 mL dry CH_2Cl_2. To the above solution was added triethylamine (1.0 mL, 5.0 equiv). A solution of **S4** (0.490 g, 1.0 equiv) in 5 mL CH_2Cl_2 was then added to the above solution at 23 °C over 1 h *via* syringe pump. The reaction mixture was then stirred at that temperature for 24 h. It was then concentrated and purified by silica gel flash chromatography (0 → 25% EtOAc/hexanes) to afford ketone (+)-**10** (0.353 g, 72%) as a pale yellow oil. $[\alpha]_D^{20}$ = +29.28 (*c* 0.20, $CHCl_3$); Other data matched that previously reported for the racemic ketone (±)-**10**.[5]

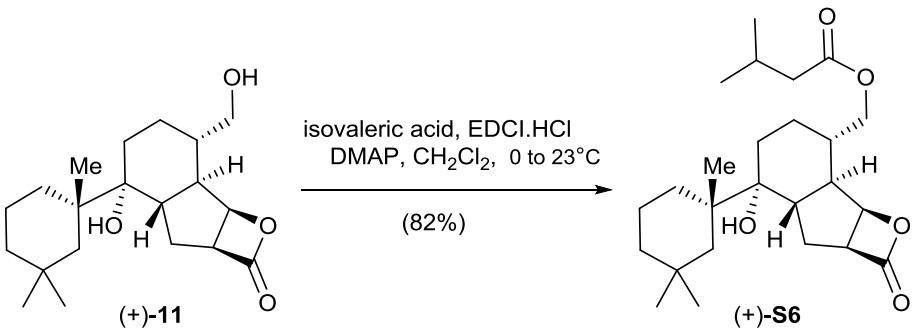


Homoallylic alcohol (+)-S5. A solution of β-lactone (+)-**10** (312.7 mg, 1.0 mmol) in 10 mL of dry THF was cooled to -78 ˚C and a solution of freshly prepared (3,5,5-trimethylcyclohex-2-en-1-yl)zinc(II) chloride[6] (9.5 mL, 1.0 mmol, 0.11 M in THF)[I] was added slowly. The reaction mixture was stirred at -78 ˚C for 1 h, quenched with pH 7 buffer solution (50 mL) at -78 ˚C, allowed to warm to 23 ºC for 30 min, and filtered through a pad of Celite with Et_2O. The layers were separated in a separatory flask and the aqueous layer was extracted with Et_2O (3 x 50 mL). The combined organic layers were washed with brine (50 mL), dried over anhydrous $MgSO_4$, and concentrated *in vacuo*. Purification by flash chromatography on SiO_2 eluting with 20% EtOAc/hexanes gave 421.3 mg (95%, dr 13:1) of homoallylic alcohol (+)-**5** as a colorless oil. $[\alpha]_D^{20}$ = +64.17 (*c* 0.10, $CHCl_3$). Other data matched that previously reported for the racemic homoallylic alcohol (±)-**S5**.[5]

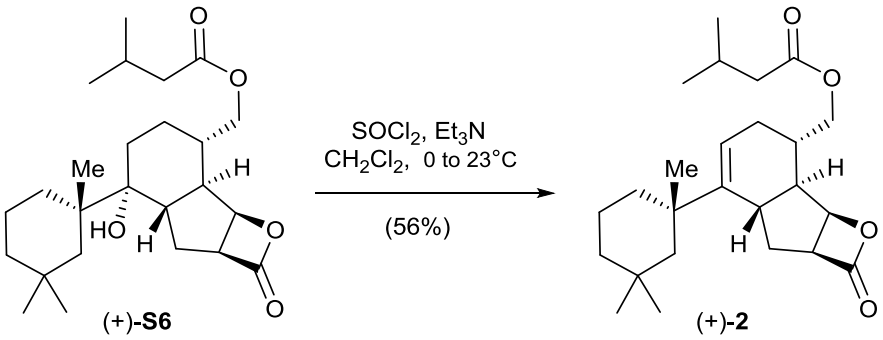


Diol (+)-11. To a solution of homoallylic alcohol (+)-**S5** (373.7 mg, 0.88 mmol) in 50 mL of EtOH was added $Pd(OH)_2$ on carbon (75.0 mg, 20% Pd). The reaction flask was charged with hydrogen gas and a balloon of H_2 was attached. The heterogeneous solution was stirred vigorously at ambient temperature (23 ˚C) for 24 h. The reaction was filtered through a pad of Celite with Et_2O, dried over anhydrous $MgSO_4$, and concentrated *in vacuo*. Purification by flash chromatography on SiO_2 eluting with 25% EtOAc/hexanes gave 240.1 mg (81%) of diol (+)-**11** as a colorless oil. $[\alpha]_D^{20}$ = +60.01 (*c* 0.05, $CHCl_3$). Other data matched that previously reported for the racemic homoallylic alcohol (±)-**11**.[5]

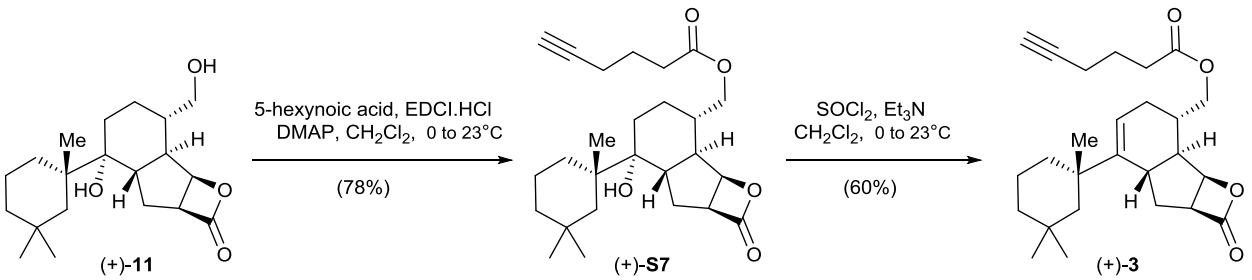
[I] The zinc chloride was prepared according to the procedure of Knochel (ref. 6. H. Ren, G. Dunet, P. Mayer and P. Knochel, *J Am Chem Soc*, 2007, **129**, 5376-5377.) as described previously (ref. 5. N. L. Harvey, J. Krysiak, S. Chamni, S. W. Cho, S. A. Sieber and D. Romo, *Chem Eur J*, 2015, **21**, 1425-1428.).



Isovaleric ester (+)-S6. To a solution of alcohol (+)-**11** (45 mg, 1.0 equiv) in 10 mL of dry CH_2Cl_2 was added DMAP (25 mg, 1.5 equiv) followed by EDCI·HCl (40.0 mg, 1.5 equiv). The reaction was stirred at ambient temperature (23 °C) for 10 min until all solids well dissolved and then isovaleric acid (16.4 mg, 1.2 equiv) was added. The mixture was stirred for 4 h and quenched with saturated aq. $NaHCO_3$ solution. The layers were separated in a separatory funnel and the aqueous layer was extracted with EtOAc (3 x 50 mL). The combined organic layers were washed with brine (50 mL), dried over anhydrous $MgSO_4$, and concentrated *in vacuo*. Purification by flash chromatography on SiO_2 eluting with 30% EtOAc/Hexanes gave 46 mg (82%) of ester (+)-**S6** as a colorless oil. $[\alpha]_D^{20}$ = +67.42 (*c* 0.14, $CHCl_3$). Other data matched that previously reported for the racemic homoallylic alcohol (±)-**S6**.[5]



Regio, bis-epi-spongiolactone (+)-2. To a solution of alcohol (+)-**S6** (46 mg, 1.0 equiv) in 2 mL of dry CH_2Cl_2 cooled to 0 °C, was added triethylamine (104 µL, 5.25 equiv) followed by slow addition of thionyl chloride ($SOCl_2$) (46 µL, 4.25 equiv). The reaction was allowed to warm slowly to ambient temperature (23 °C) and stirred for 12 h. The mixture was then diluted with 5 mL of dry CH_2Cl_2 and quenched with pH 7 buffer solution. The layers were separated in a separatory funnel and the aqueous layer was extracted with CH_2Cl_2 (3 x 10 mL). The combined organic layers were washed with brine (10 mL), dried over anhydrous $MgSO_4$, and concentrated *in vacuo*. Purification by silica gel chromatography using 10% EtOAc/hexanes gave 19.6 mg (56%) of regio, bis-epi-spongiolactone (+)-**2** as a colorless oil. $[\alpha]_D^{20}$ = +65.77 (*c* 1.96, $CHCl_3$). Other data matched that previously reported for the racemic regio, bis-epi-spongiolactone (±)-**2**.[5]

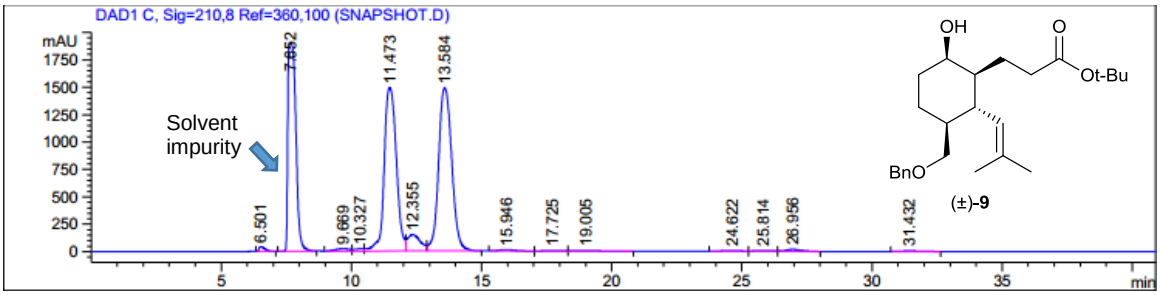


Alkyne probe (+)-3. To a solution of diol (+)-**11** (41 mg, 1.0 equiv) in 10 mL of dry CH_2Cl_2 was added DMAP (25 mg, 1.5 equiv) followed by EDCI·HCl (40 mg, 1.5 equiv). The reaction was stirred at ambient temperature (23 °C) for 10 min until all solids were dissolved and 5-hexynoic acid (16 mg, 1.2 equiv) was added. The mixture was stirred for 4 h and quenched with saturated aq. $NaHCO_3$ solution. The layers were separated in a separatory funnel and the aqueous layer was extracted with EtOAc (3 x 50 mL). The combined organic layers were washed with brine (50 mL), dried over anhydrous $MgSO_4$, and concentrated *in vacuo*. Purification of the crude product by flash chromatography on SiO_2 eluting with 30% EtOAc/Hexanes gave 39 mg (78%) of ester (+)-**S7** as a colorless oil, which was used immediately in the next reaction without further purification. To a solution of the tertiary alcohol (+)-**S7** (39 mg, 1.0 equiv) in 2 mL of dry CH_2Cl_2 was cooled at 0 °C was added triethylamine (66 µL, 5.25 equiv) followed by slow addition of thionyl chloride ($SOCl_2$) (29 µL, 4.25 equiv). The reaction was allowed to slowly warm to ambient temperature (23 °C) and stirred for 12 h. The mixture was then diluted with 5 mL of dry CH_2Cl_2 and quenched with pH 7 buffer solution. The layers were separated in a separatory funnel and the aqueous layer was extracted with CH_2Cl_2 (10 mL x 3). The combined organic layers were washed with brine (10 mL), dried over anhydrous $MgSO_4$, and concentrated *in vacuo*. Purification by silica gel chromatography using 10% EtOAc/Hexanes gave 17 mg (60%) of alkyne probe (+)-**3** as a colorless oil. $[\alpha]_D^{20}$ = +58.33 (*c* 1.02, $CHCl_3$); TLC (EtOAc:hexanes, 1:9 v/v): Rf = 0.35; ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.54 (dt, *J* = 5.2, 2.5 Hz, 1H), 4.95 (app t, *J* = 3.8 Hz, 1H), 4.26 (dd, *J* = 11.2, 4.8 Hz, 1H), 4.00 (dd, *J* = 11.1, 6.4 Hz, 1H), 3.95 (dd, *J* = 7.5, 4.1 Hz, 1H), 2.66 – 2.57 (m, 1H), 2.53 (dd, *J* = 12.5, 5.2 Hz, 1H), 2.47 (t, *J* = 7.4 Hz, 2H), 2.32 (dq, *J* = 6.1, 2.0 Hz, 1H), 2.27 (dt, *J* = 7.0, 2.6 Hz, 2H), 2.15 (app tt, *J* = 11.4, 5.6 Hz, 1H), 1.96 (app t, *J* = 3.5 Hz, 2H), 1.85 – 1.82 (m, 3H), 1.68 (dt, *J* = 12.7, 7.6 Hz, 1H), 1.61 – 1.57 (m, 2H), 1.54 – 1.43 (m, 2H), 1.26 – 1.14 (m, 4H), 1.04 (s, 3H), 0.89 (s, 3H), 0.81 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 173.1, 171.2, 145.3, 119.3, 83.2, 74.6, 69.2, 66.7, 55.0, 49.7, 49.2, 39.9, 39.8, 39.4, 38.0, 33.5, 32.8, 32.2, 31.8, 31.4, 30.4, 30.1, 29.32, 23.5, 19.8, 17.9; IR (thin film): 3285, 2120, 1786 cm^{-1}; HRMS (ESI+) $C_{26}H_{36}O_4Li$ $[M+Li]^+$: 419.2768. Found: 419.2760.

3.2 Chiral HPLC analysis of alcohol (+)-9

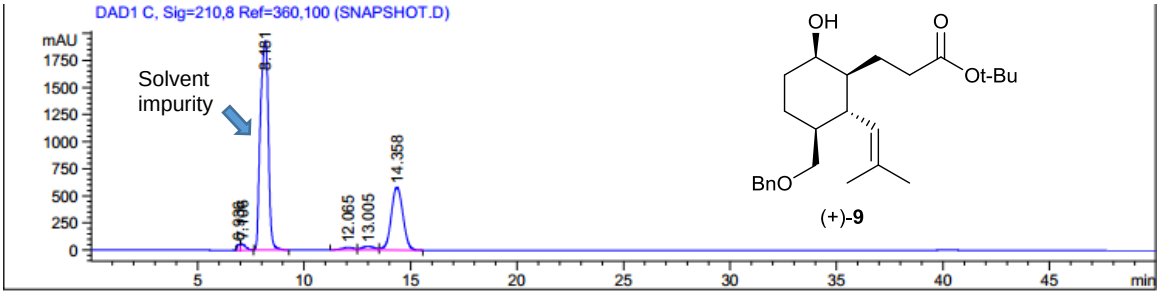
Chiral HPLC analysis was conducted using a chiral OD column (detection: DAD; wavelength 210 nm) 250 x 4.5 mm, L x I.D., solvent 95:5 hexanes/2-propanol, flow rate 0.5 mL/min.

HPLC profile of racemic *anti*-alcohol (±)-**9** (obtained from $NaBH_4$ reduction of racemic ketone (±)-**8**, dr ~1:1) following separation of diastereomers.



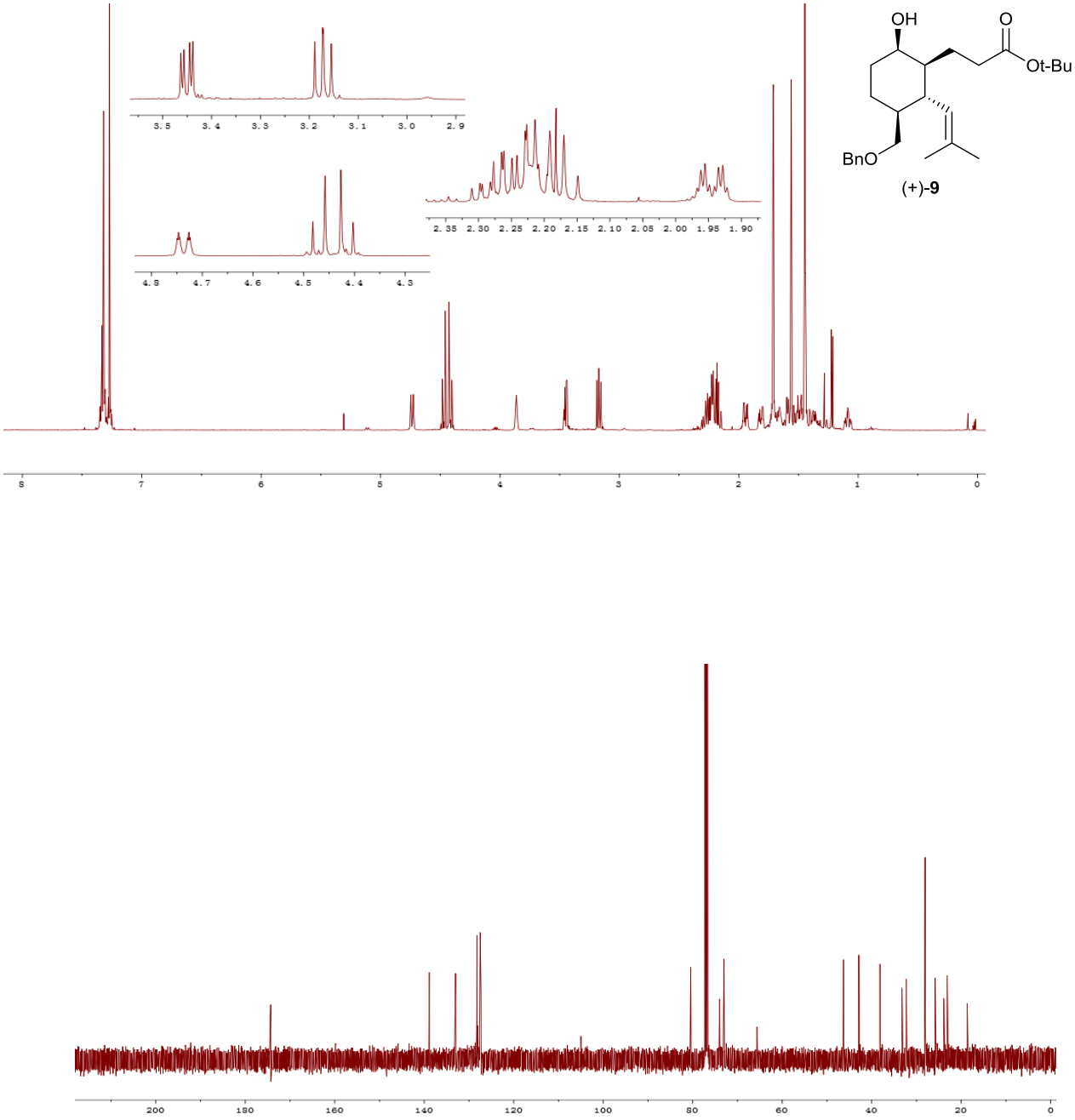

```
Signal 3: DAD1 C, Sig=210,8 Ref=360,100

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]    [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1  11.473 MM    0.5272 4.68434e4  1480.97534  49.3184
   2  13.584 MM    0.5566 4.81382e4  1441.54529  50.6816

Totals :                  9.49817e4  2922.52063
```

HPLC profile for alcohol (+)-**9**

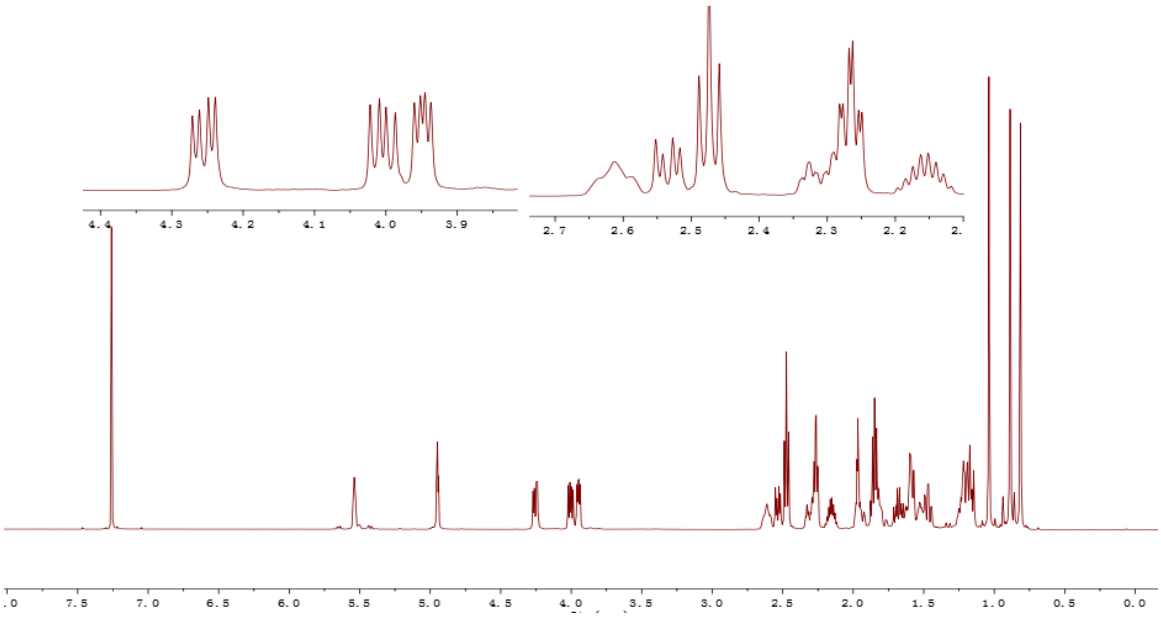


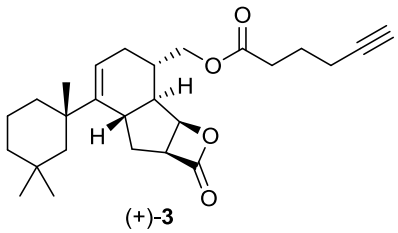

```
Signal 3: DAD1 C, Sig=210,8 Ref=360,100

Peak RetTime Type  Width     Area      Height     Area
  #   [min]        [min]    [mAU*s]     [mAU]       %
----|-------|----|-------|----------|----------|--------|
   1  12.066 MM    0.5748  690.99231   20.03674   3.1517
   2  14.358 MM    0.6195 2.12337e4   571.28802  96.8483

Totals :                  2.19247e4   591.32477
```

3.3 NMR spectra of key compounds

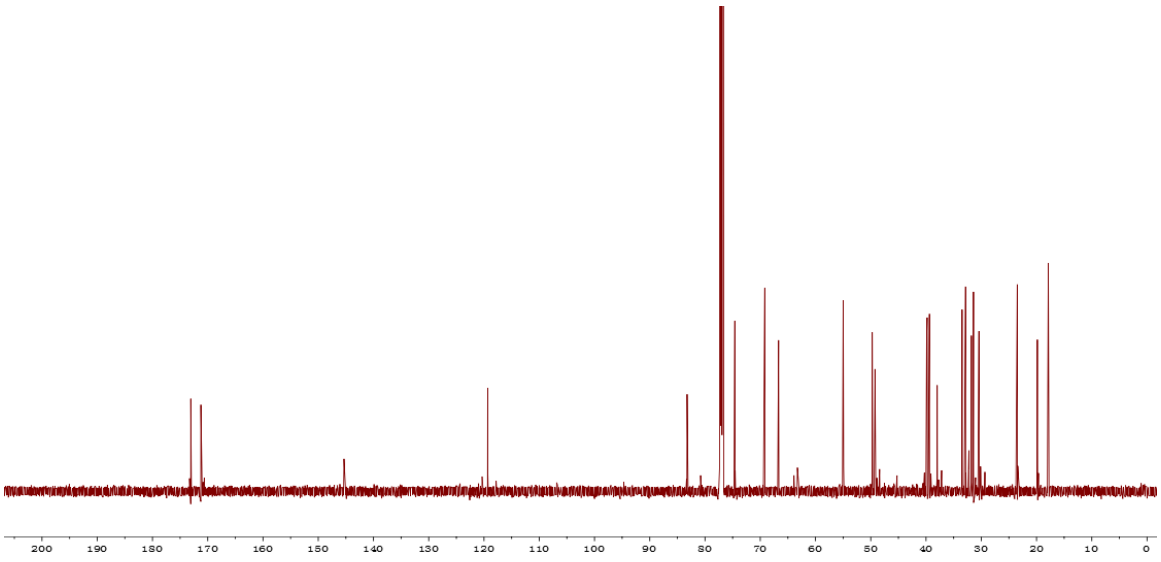


^{1}H NMR (500 MHz) and ^{13}C NMR (125 MHz) of alcohol (+)-**9** in $CDCl_3$





(+)-**3**



^{1}H NMR (500 MHz) and ^{13}C NMR (125 MHz) of alkyne probe (+)-**3** in $CDCl_3$

4 Biological and biochemical methods

4.1 Human cell culture

Cells were incubated at 37 °C in a humidified atmosphere containing 5 % CO_2.

K562 (chronic myelogenous leukemia) and Jurkat (acute T cell leukemia) suspension cell lines were cultured in RPMI-1640 medium (Sigma R7509) supplemented with 2 mM L-glutamine (PAA) and 10% heat-inactivated FCS (Sigma Life Science). Medium used for MTT assay did not contain Phenol Red. Cells were maintained at 1 x 10^5 to 1 x 10^6 viable cells/mL.

A549 (lung carcinoma) adherent cell line was cultured in Dulbecco's Modified Eagle Medium (DMEM high glucose, Sigma Life Science) supplemented with 10 % FCS and 2 mM L-glutamine. Cells were detached with trypsin-EDTA.

4.2 Cell cytotoxicity assays (MTT)

4.2.1 K562 and Jurkat MTT assay

Cells were diluted with a fresh culture medium to the concentration 1.6×10^5 cells/mL. 50 µL cell suspension were plated in a 96-well round-bottom transparent plate, so that cell density was 8000 cells per well. Blank wells not containing cells were included. Compounds were diluted 1:49 (v/v) from DMSO stocks (stored at -80° C) with pre-warmed culture medium. 50 µL aliquots of the compound-medium solution were added to cells in triplicate wells. Final volume in each well was 100 µL, cell density was 8 x10^4 cells/mL and final compound concentration ranged between 0.3 and 300 µM.

After 24 h exposure, 20 µL MTT solution (5 mg/mL in PBS, sterile filtered) [(3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) (Sigma M5655)] were added to each well and mixed well. Plates were incubated at 37 °C, 5% CO_2 for 4 h in the dark to allow MTT to be metabolized. Next, 50 µL 'triplex lysis solution' (10% SDS (w/v), 5% isobutanol (v/v), 0.012 M HCl in H_2O) was added to each well and mixed well. Plates were incubated overnight (14-15 h) at 37 °C, 5% CO_2. Optical density was measured at 570 nm by a TECAN Infinite® M200 Pro. A background control was read at 630 nm and subtracted from the measurement.

4.2.2 A549 MTT assay

Performed as above with the following modifications. Cells were plated at 7000 cells / well in 200 µL aliquots and incubated for 16 h to allow adherence. Compounds were diluted 1:99 (v/v) from DMSO stocks (stored at -80° C) with pre-warmed culture medium. The medium was removed from cells by gentle suction and 100 µL aliquots of the compound-medium solution were added to cells in triplicate wells. After incubation and MTT addition as described above, the medium was removed by gentle suction and 200 µL DMSO added to each well. The plates were incubated at RT, 300-500 rpm for 5 minutes to dissolve formazan before absorbance measurement as above.

4.2.3 MTT assay data processing

All biological experiments were performed with three technical replicates, and biological experiments were performed in triplicate unless otherwise noted. For each experiment, average blank absorbance (medium, MTT solution and lysis buffer) was subtracted from average absorbance of wells containing cells. Percentage inhibition of metabolic activity (cell viability) was calculated as a fraction of control (cells treated only with the DMSO (vehicle)). Where applicable,

IC_{50} values were calculated using sigmoidal dose response curve fitting algorithm in Origin 2016 (OriginLab Corporation). Data were plotted as mean ± standard deviation. Each experiment gave an IC_{50} and error value. A weighted mean and standard error was calculated to give one value for each probe and cell line (given in Fig. 2a), according to the following formulae:[7]

$$weighted\ mean = \frac{\left(\frac{A}{a^2}\right) + \left(\frac{B}{b^2}\right) + \left(\frac{C}{c^2}\right)}{\left(\frac{1}{a^2}\right) + \left(\frac{1}{b^2}\right) + \left(\frac{1}{c^2}\right)}$$

and

$$error = \frac{\left(\frac{1}{a}\right) + \left(\frac{1}{b}\right) + \left(\frac{1}{c}\right)}{\left(\frac{1}{a^2}\right) + \left(\frac{1}{b^2}\right) + \left(\frac{1}{c^2}\right)}$$

4.3 Probe labelling in cells and lysis

4.3.1 Labelling in K562 and Jurkat cells

Cells were washed 1× warm PBS and resuspended at 1×10^6 cells / mL in warm RPMI without FCS. ~2×10^6 cells were used for analytical labelling experiments, and 8×10^6 cells (K562) or 10×10^6 cells (Jurkat) for proteomic experiments. Labelling was carried out with 1×10^6 cells per well of a 6 well plate. Probe in DMSO or DMSO control (final DMSO concentration of 0.1%, 10 mM probe stock) was added to each well, and cells incubated at 37 °C, 5% CO_2 for 2 h. Cells were transferred (with gentle scraping due to adherence to the plate) to a tube, pelleted (500 ×g, 5 min) and washed 3× cold PBS. Cells were flash frozen in liquid N_2 and stored at -80 °C until lysis.

For lysis, cells were thawed on ice and resuspended in 10 mM sodium phosphate pH 7.5 with EDTA-free protease inhibitor cocktail (Roche). Solutions were sonicated 2 x 15 sec at 60% intensity with 30 sec rest on ice in between. NP40 was added to 1% and SDS to 0.5% from 10% stock solutions. The samples were left on ice for 20-30 min then centrifuged (21,000 ×g, 15 min, 4 °C), the supernatant transferred to a new tube and the pellet discarded.

4.3.2 Labelling in A549 cells

Cells were plated 24 hours before labelling to be ~80% confluent on the day of labelling. Cells were washed 2 × warm PBS, then medium containing probe in DMSO added. Following incubation as above, cells were washed 3 × PBS in the plate and lysis buffer (1% NP40, 10 mM sodium phosphate pH 7.5, EDTA-free protease inhibitor) added directly to the plate. Lysate was transferred to microcentrifuge tubes and SDS added to 0.2%. Solutions were sonicated 2 x 15 sec at 60% intensity with 30 sec rest on ice in between to shear nucleic acid. The samples were left on ice for 20-30 min then centrifuged (21,000 ×g, 15 min, 4 °C), the supernatant transferred to a new tube and the pellet discarded.

4.3.3 Competitive labelling in K562 cells

Analytical labelling was performed as above (section 4.3.1) but with pre-incubation of cells with competitor compounds for 30 min prior to addition of probe.

4.4 Cell lysis

Protein concentration was measured using the Roti®-Quant kit (Carl Roth). Lysate concentration was adjusted to ~1 mg/mL.

For fractionation of soluble and insoluble proteins, cells were resuspended in 10 mM sodium phosphate pH 7.5 with EDTA-free protease inhibitor cocktail. Solutions were sonicated 2 x 10 sec at 10% intensity with 30 sec rest on ice in between. Samples were centrifuged at 2000 ×g, 10 min, 4°C to pellet nuclei. The supernatant was transferred to a new tube and centrifuged at 21000 ×g, 1 h, 4 °C to pellet the microsomal fraction. The microsomal (membrane fraction) was resuspended in PBS with EDTA-free protease inhibitor cocktail. 0.5% final SDS concentration was added to samples before CuAAC.

Lysates were flash frozen and stored at -80 °C until use.

4.5 CuAAC and gel-based fluorescence imaging

Click reagents were premixed in the following order and added to the lysate at the indicated final concentrations: 100 µM azide reagent (10 mM stock in DMSO, see below), 1 mM $CuSO_4$ (50 mM stock in water), 1 mM TCEP (50 mM stock in water) and 100 µM TBTA (10 mM stock in DMSO). Samples were incubated for 1 h with gentle agitation, RT. The azide reagent used for analytical CuAAC and gel-based analysis was RhN_3: tetramethylrhodamine 5-Carboxamido-(6-Azidohexanyl), 5-isomer (Life Technologies, T10182). Following CuAAC, proteins were precipitated with ice-cold acetone (4 volumes) at -20 °C for 1 h, centrifuged at 21,000 ×g, 10 min, 4 °C and the pellet was then washed 2× with ice-cold MeOH. Proteins were resuspended in 1× sample loading buffer (SLB: 32 mM Tris pH 6.8, 5% glycerol, 1% SDS, 0.0013% bromophenol blue, 2.5% 2-mercaptoethanol) by vortexing and heating to 40 °C.

Proteins were separated by SDS-PAGE (150 V, 2 h) with 4% stacking gels and 12.5% separating gels on a PEQLAB Biotechnologie GmbH, Erlangen, PerfectBlue Dual Gel System with 8 µL fluorescent marker (BenchMark™ Fluorescent Protein Standard by Invitrogen) and 12 µL Roti®-Mark STANDARD (Carl Roth) for protein detection by Coomassie staining. Fluorescence was measured with a LAS-4000 gel scanning station using a Cy3 filter. Gels were the Coomassie stained by overnight incubation in staining solution (0.25% w/v Coomassie Brilliant blue R-250, 9.2% AcOH, 45.4% EtOH), followed by destaining with 10% AcOH, 40% EtOH.

4.6 Pull-down of labelled proteins

To test the pull-down and enrichment of labelled proteins, CuAAC was performed as above but with a trifunctional azido-biotin-rhodamine reagent in place of RhN_3. Following, CuAAC and protein precipitation, proteins were resuspended at 10 mg/mL in 2 % SDS, 10 mM EDTA in PBS, and then diluted to 1 mg/mL with PBS. Affinity enrichment were performed with Avidin agarose resin (A9207 Sigma, pre-washed 3 × 0.2 % SDS in PBS; 2000 rpm 2-3 min was used to pellet beads; typically 50-100 µL of bead slurry was used for 0.5 mg of lysate). Samples were added to the resin in LoBind Eppendorf tubes and incubated with agitation for 2 h at RT. Beads were washed following pull-down (0.5 mL each time) 4 × 1 % SDS in PBS. Sample loading buffer was added and incubated at 90 °C 10 min to elute bound proteins.

4.7 Western blot analysis of HSD17B12

For western blot analysis, samples were separated by SDS-PAGE as above. For one gel, two filter papers (extra thick blot paper, Bio-Rad, USA) were soaked in blotting buffer (48 mM Trizma, 39 mM Glycine, 0.04 % SDS, 20 % MeOH) and one was put on top of the anode of a semi-dry western blot cell (Trans-Blot SD, Bio-Rad, Germany). A PVDF membrane (Immun-blot® PVDF membrane, Bio-Rad, USA) was cut to the size of the gel and submerged for at least 5 min in 100% methanol. Membrane, gel and the second filter paper were stacked in the station and the blotting was carried out at 10 V for 1 h. Next, the membrane was placed in a solution of 5% milk powder in PBS-T (PBS + 0.5 % Tween-20) at room temperature for 45 min and then washed 5 min in PBS-T. The HSD17B12 antibody (polyclonal, mouse anti-human, Abnova) was diluted (1:1000 in 5% milk powder in PBS-T) and incubated overnight at 4 °C with constant shaking. The next day, the membrane was washed three times with PBS-T for 10 min and incubated with secondary goat anti-Mouse IgG (polyclonal, HRP-conjugate, ThermoFisher, diluted 1:10000 in 5% milk powder in PBS-T) for 1 h at RT. The membrane was then washed three times in PBS-T for 10 min, incubated for 30 min with 1 mL (1:1) of a freshly prepared ECL solution (ECL Prime Western Blotting Detection Reagent, GE Healthcare, UK) and scanned in a LAS-4000 gel scanning station (Fujifilm, Japan).

5 Proteomics methods

5.1 CuAAC, enrichment, reduction, alkylation and digest for proteomics

0.5 mg lysate per sample was prepared for proteomic analysis. Proteins were ligated via CuAAC as before with the following modifications: reaction was carried out with Biotin-PEG_3-N_3 (Jena Bioscience, CLK-AZ104P4-100) for probe (±)-**4** experiments and trypsin cleavable biotin-azide (AzRB)[1] for all probe (+)-**3** experiments. Following CuAAC, EDTA was added to 10 mM final (from 500 mM stock) and proteins precipitated with ice-cold acetone (4 volumes) at -20 °C overnight, centrifuged at 21,000 xg, 10 min, 4 °C and the pellet was then washed 2× with ice-cold MeOH with resuspension in a sonication bath. Proteins were finally resuspended at 10 mg/mL in 2 % SDS, 10 mM EDTA in PBS, and then diluted to 1 mg/mL with PBS. DTT (from a fresh 100 × stock in water) was added to give a final concentration of 1 mM to facilitate resuspension. Samples were centrifuged at 21,000 xg, 10 min, RT to pellet any particulates. Eppendorf LoBind® tubes were used throughout proteomic preparation.

Affinity enrichment were performed with Avidin agarose resin (A9207 Sigma, pre-washed 3 × 0.2 % SDS in PBS; 2000 rpm 2-3 min was used to pellet beads; typically 50-100 µL of bead slurry was used for 0.5 mg of lysate). Samples were added to the resin in LoBind Eppendorf tubes and incubated with agitation for 2 h at RT. Beads were stringently washed following pull-down (0.5 mL each time): 2 × 1 % SDS in PBS, 3 × 4M Urea in PBS, 3 × TEAB (50 mM triethylammonium bicarbonate, Sigma Aldrich). For a 50-100 µL bed of beads resuspended in 100 µL TEAB, samples were reduced (4 µL of 250 mM DTT; final concentration 10 mM) at 55 °C for 30 minutes and allowed to cool to room temperature. The beads were washed 1 x TEAB (0.5 mL). Cysteines were alkylated (4 µL of 500 mM iodoacetamide; final concentration 20 mM) at RT for 30 min in the dark. The beads were washed 2 x TEAB and resuspended in 100 µL TEAB. Trypsin (1 µg, 2 µL 0.5 µg/µL in 50 mM acetic acid) was added to the beads and samples digested overnight at 37 °C with shaking at 1050 rpm. The samples were centrifuged and the supernatant was transferred into clean tubes and the pH adjusted to 2-3 with 0.5 µL FA. The beads were washed twice with 0.1% FA in H_2O, and these washes were combined with the first supernatant.

The solutions were stage-tipped according to a published protocol.[8] Briefly, tips were prepared with 3 layers of C18 material (SDC-XC from 3M) and washed with 80 µL MeOH, 80 µL 80% ACN, 0.5% FA and 2 x 80 µL 0.5% FA in H_2O (using centrifugation at 500 xg, 1-2 min, RT). Samples were loaded and desalted with 0.1% FA (150 µL). Elution from the sorbent with 100 µL 80% ACN, 0.5% FA in H_2O was followed by speed-vac-assisted solvent removal. Peptides were stored dry at -20 °C until needed.

5.2 Preparation of samples for global proteome analysis

K562 cells were incubated with probe (+)-**3** or DMSO for 24 h, or with probe (+)-**3** for 4 h, and lysed as described above. Proteins were precipitated with ice-cold acetone (4 volumes) at -20 °C overnight, centrifuged at 21,000 xg, 10 min, 4 °C and the pellet was then reconstituted in 600 µL 7 M urea, 2 M thiourea in 20mM HEPES buffer pH 7.5 using sonication. Cysteines were reduced (1 mM DTT, 45 min, RT, 450 rpm), alkylated (5.5 mM iodoacetamide, 30 min, RT, 450 rpm in the dark) and the reaction was stopped (4 mM DTT, 30 min, RT, 450 rpm). Tryptic digest was carried out as indicated above. The digest was stopped by adjustment of the pH to 2-3 with FA. The solution was desalted using 50 mg Sep-Pak® C18 columns (Waters). The Columns were equilibrated with 1 mL ACN, 0.5 mL elution buffer (80% ACN, 0.5% FA), and 3 x 1 mL 0.1 % FA in H_2O. Samples were loaded, and the peptides washed with 3 x 1 mL 0.1 % FA in H_2O and 250 µL 0.5 % FA in H_2O. Peptides were eluted using 3 x 250 µL elution buffer, followed by speed-vac-assisted solvent removal. Peptides were stored dry at -80 °C until needed.

5.3 LC-MS/MS

Peptides were reconstituted in 1 % FA in H_2O in a sonication bath and with vortexing, filtered using centrifugal filters (modified Nylon, 0.45 µm, low protein binding, VWR International, LLC) and transferred into LC-MS sample vials.

5.3.1 Chemical proteomics samples

Nanoflow LC-MS/MS analysis was performed with an UltiMate 3000 Nano HPLC system (Thermo Scientific) coupled to an Orbitrap Fusion (Thermo Scientific). Peptides were loaded on a trap column (Acclaim C18 PepMap100 75 µm ID x 2 cm) and washed for 10 min with 0.1 % FA and 5% DMSO (10 µL/min flow rate), then transferred to an analytical column (Acclaim C18 PepMap RSLC, 75 µM ID x 15 cm) and separated using a 125 min gradient from 3 % to 40 % (120 min from 3 % to 25 % and 5 min to 40 %) MeCN in 0.1 % FA and 5 % DMSO at a flow rate of 200 nL/min. Peptides were ionised using a nanospray source at 1.9 kV and a capillary temperature of 275 °C. Orbitrap Fusion was operated in a top speed data dependent mode with a cycle time of 3 s. Full scan acquisition (scan range of 300 – 1700 m/z) was performed in the orbitrap at a resolution of 120000 (at m/z 200) and with an automatic gain control ion target value of 4e5. Monoisotopic precursor selection as well as dynamic exclusion of 60 s were enabled. Internal calibration was performed using the ion signal of fluoranthene cations (EASY-ETD/IC source). Most intense precursors with charge states of 2 - 7 and intensities greater than 5e3 were selected for fragmentation. Isolation was performed in the quadrupole using a window of 1.6 m/z. Ions were collected to a target of 1e2 for a maximum injection time of 250 ms with “inject ions for all available parallelizable time” enabled (“Universal” method, Eliuk *et al.*, Thermo Scientific Poster Note PN40914). Fragments were generated using higher-energy collisional dissociation (HCD) and detected in the ion trap at a rapid scan rate.

5.3.2 Total proteome samples

Samples were analyzed with an UltiMate 3000 nano HPLC system (Dionex) using Acclaim C18 PepMap100 75 µm ID x 2 cm trap and Acclaim Pepmap RSLC C18 (75 µm ID x 50 cm) separation columns in an EASY-spray setting coupled to a Q Exactive Plus (Thermo Fisher). Samples were loaded on the trap and washed with 0.1% TFA, then transferred to the analytical column (buffer A: H_2O with 0.1% FA, buffer B: MeCN with 0.1% FA, flow 300 nL/min, gradient 5 to 22% buffer B in 115 min, then to 32% buffer B in 10 min, then to 90% buffer B in 10 min and hold 90% buffer B for 10 min, then to 5% buffer B in 0.1 min and hold 5% buffer B for 9.9 min). Q Exactive Plus was operated in a TOP12 data dependent mode. Full scan acquisition was performed in the Orbitrap at a resolution of 70000 and an AGC target of 3e6 in a scan range of 300−1500 m/z. Monoisotopic precursor selection as well as dynamic exclusion (exclusion duration: 60 s) was enabled. Precursors with charge states of >1 and intensities greater than 1e5 were selected for fragmentation. Isolation was performed in the quadrupole using a window of 1.6 m/z. Precursors were collected to an AGC target of 5e4 for a maximum injection time of 50 ms. Fragments were generated using higher-energy collisional dissociation and detected in the ion trap at a rapid scan rate.

5.4 Data processing: general comments

The data were processed with MaxQuant versions 1.5.3.8 or 1.5.5.1, and peptides were identified from the MS/MS spectra searched against the Uniprot human reference proteome (canonical version, without isoforms, downloaded 13/04/2016) using the Andromeda search engine.[9] Cysteine carbamidomethylation was used as a fixed modification, and methionine oxidation and N-terminal acetylation as variable modifications. The false discovery rate was set to 0.01 for peptides, proteins and sites. Other parameters were used as pre-set in the software. “Unique and razor peptides” mode was selected to allow for protein grouping; this calculates ratios from unique and razor peptides (razor peptides are uniquely assigned to protein groups and not to individual proteins). Match between runs was enabled between replicates (i.e. between DMSO replicates A, B, C etc., and between probe replicates, but not between probe and DMSO samples). LFQ experiments in MaxQuant were performed using the built-in label-free quantification algorithm (MaxLFQ).[10] Data were elaborated using Perseus versions 1.5.0.31 and 1.5.5.3, and Excel. Data have been deposited to ProteomeXchange[11] via the PRIDE[12] partner repository with the data set identifier PXD006811.

5.5 Statistical analysis of proteomics data

5.5.1 General comments

ProteinGroups files were used for analysis using Perseus (http://www.coxdocs.org/doku.php?id=perseus:start). For all files, proteins falling into the following categories were removed: identified only by site, reverse, contaminant. Data were matched to annotation data downloaded from Uniprot (http://www.uniprot.org/), including GO annotation. LFQ intensities were Log_2 transformed, and filtered to retain only proteins quantified in at least one sample. Proteins lacking >1 razor+unique peptides were also removed.

5.5.2 Chemical proteomics with (+)-3 in K562 cells

Samples searched: DMSO (replicates A, B, C, D) and probe (+)-**3** (replicates A, B, C, D) in K562 cells.

The dataset was filtered to retain proteins that satisfied the razor+unique peptides requirement in either DMSO or probe (+)-**3** (≥2 in each replicate); then filtered for LFQ quantification in three out of four replicates and >2 razor+unique peptides overall. Missing values were imputed to mimic low abundance values (width 0.3, downshift 1.8, total matrix). *t*-tests were then applied to analyse the data: two-sample two-sided *t*-test with permutation-based statistics (250 permutations); FDR = 0.001, s0 = 0.5 (probe (+)-**3**/DMSO).

5.5.3 Chemical proteomics with 3 in Jurkat cells and comparison with K562 cells

Samples searched: DMSO (replicates A, B, C, D) and probe (+)-**3** (replicates A, B, C, D) in K562 cells; DMSO (replicates A, B, C) and probe (+)-**3** (replicates A, B, C) in Jurkat cells; DMSO (replicates A, B, C) and probe (+)-**3** (replicates A, B, C) in A549 cells.

T-test in Jurkat cells. The dataset was filtered to retain proteins that satisfied the razor+unique peptides requirement in either DMSO or probe (+)-**3** (≥2 in each replicate) for Jurkat cells; then filtered for LFQ quantification in three out of three replicates. Missing values were imputed to mimic low abundance values (width 0.3, downshift 1.8, total matrix). *t*-tests were then applied to analyse the data: two-sample two-sided *t*-test with permutation-based statistics (250 permutations); FDR = 0.001, s0 = 2 (probe (+)-**3**/DMSO).

T-test in K562 cells. The dataset was filtered to retain proteins that satisfied the razor+unique peptides requirement in either DMSO or probe (+)-**3** (≥2 in each replicate) for K562 cells; then filtered for LFQ quantification in four out of four replicates. Missing values were imputed to mimic low abundance values (width 0.3, downshift 1.8, total matrix). *t*-tests were then applied to analyse the data: two-sample two-sided *t*-test with permutation-based statistics (250 permutations); FDR = 0.001, s0 = 2 (probe (+)-**3**/DMSO).

Comparison. The total dataset was filtered for 3 valid values in at least one group (DMSO or probe, either cell line). Missing values were imputed to mimic low abundance values (width 0.3, downshift 1.8, total matrix). Data were cross-referenced with the *t*-tests described above and hits selected on the basis of significance and a $\log_2$(probe/DMSO) enrichment of >2.

5.5.4 Chemical proteomics with 4 in Jurkat cells

Samples searched: DMSO (replicates A, B, C) and probe (±)-**4** (replicates A, B, C) in Jurkat cells.

The dataset was filtered to retain proteins that satisfied the razor+unique peptides requirement in either DMSO or probe (±)-**4** (≥2 in each replicate); then filtered for LFQ quantification in three out of three replicates. Missing values were imputed to mimic low abundance values (width 0.3, downshift 1.8, total matrix). *t*-tests were then applied to analyse the data: two-sample two-sided *t*-test with permutation-based statistics (250 permutations); FDR = 0.01, s0 = 1 (probe (±)-**4**/DMSO).

5.5.5 Comparative analysis of 3 and 4 targets in Jurkat cells

Samples searched: DMSO (replicates A, B, C) and probe (±)-**4** (replicates A, B, C) in Jurkat cells; DMSO (replicates A, B, C) and probe (+)-**3** (replicates A, B, C) in Jurkat cells.

T-test probe (+)-3. The dataset was filtered for LFQ quantification in two out of three replicates. Missing values were imputed to mimic low abundance values (width 0.3, downshift 1.8, separately each column). *t*-tests were then applied to analyse the data: two-sample two-sided *t*-test with permutation-based statistics (250 permutations); FDR = 0.001, s0 = 0.5 (probe (+)-**3**/DMSO).

T-test probe (±)-4. The dataset was filtered for LFQ quantification in two out of three replicates. Missing values were imputed to mimic low abundance values (width 0.3, downshift 1.8, separately each column). *t*-tests were then applied to analyse the data: two-sample two-sided *t*-test with permutation-based statistics (250 permutations); FDR = 0.01, s0 = 0.5 (probe (±)-**4**/DMSO).

Comparison. The total dataset was filtered for 3 valid values in at least one group (DMSO or probe, either cell line). Missing values were imputed to mimic low abundance values (width 0.3, downshift 1.8, separately each column). Data were cross-referenced with the *t*-tests described above and a global difference in LFQ intensity calculated for plotting in Fig. 3b. Hits were selected on the basis of significance and a $\log_2$(probe/DMSO) enrichment of >2.

5.5.6 Chemical proteomics with (+)-3 in A549 cells

Samples searched: DMSO (replicates A, B, C, D) and probe (+)-**3** (replicates A, B, C, D) in K562 cells; DMSO (replicates A, B, C) and probe (+)-**3** (replicates A, B, C) in Jurkat cells; DMSO (replicates A, B, C) and probe (+)-**3** (replicates A, B, C) in A549 cells. (Same search as for datasets analysed as described in section 5.5.3).

T-test in A549 cells. The dataset was filtered to retain proteins that satisfied the razor+unique peptides requirement in either DMSO or probe (+)-**3** (≥2 in each replicate) for A549 cells; then filtered for LFQ quantification in three out of three replicates. Missing values were imputed to mimic low abundance values (width 0.3, downshift 1.8, total matrix). *t*-tests were then applied to analyse the data: two-sample two-sided *t*-test with permutation-based statistics (250 permutations); FDR = 0.01, s0 = 1 (probe (+)-**3**/DMSO).

5.5.7 Comparative analysis of all datasets

Total dataset (all cell lines) before imputation was cross-referenced to probe (±)-**4** dataset. Missing values were globally imputed (width 0.3, downshift 1.8, across total dataset). Leukemia hits were selected (48 hits - defined according to analysis in Table S2). LFQ intensities across replicates and across all DMSO samples were averaged (mean). Hierarchical clustering was performed on this data in Perseus to generate heatmaps (See Supp. Fig. S7 and Fig. 3d).

5.5.8 Total proteomics

Samples searched: DMSO (replicates a, b, c) at 24 h, probe (+)-**3** (replicates a, b, c) at 24 h, probe (+)-**3** (replicates a, b, c) at 4 h; K562 cells.

The dataset was filtered to retain proteins with ≥2 razor+unique peptides, then filtered for LFQ quantification in triplicate in either DMSO_24h, probe_24h or probe_4h datasets. Missing values were imputed from a normal distribution to mimic low abundance values (width 0.3, downshift 1.8, across total dataset) and a two-sample *t*-test performed to compare DMSO and 3-treated samples: two sided, permutation-corrected (250 permutations), FDR 0.05, s0 0.3 (24 h samples; probe (+)-**3**/DMSO).

5.5.9 Gene Ontology enrichment analysis

Gene Ontology analysis was performed in Cytoscape using the BiNGO app.[13] Human gene association (.gaf) and ontology (.obo) files were downloaded from http://www.geneontology.org/ (04/2017). Biological process name space was used to select terms. A hypergeometric test was performed, with selected significance level of 0.05 and Benjamini & Hochberg False Discovery Rate (FDR) correction.

6 References

1. M. H. Wright, D. Paape, E. M. Storck, R. A. Serwa, D. F. Smith and E. W. Tate, *Chem Biol*, 2015, **22**, 342-354.
2. T. Bottcher and S. A. Sieber, *J Am Chem Soc*, 2008, **130**, 14400-14401.
3. E. J. Corey and C. J. Helal, *Angew Chem Int Ed*, 1998, **37**, 1987-2012.
4. E. J. Corey, R. K. Bakshi and S. Shibata, *J Am Chem Soc*, 1987, **109**, 5551-5553.
5. N. L. Harvey, J. Krysiak, S. Chamni, S. W. Cho, S. A. Sieber and D. Romo, *Chem Eur J*, 2015, **21**, 1425-1428.
6. H. Ren, G. Dunet, P. Mayer and P. Knochel, *J Am Chem Soc*, 2007, **129**, 5376-5377.
7. D. C. Jones, I. Hallyburton, L. Stojanovski, K. D. Read, J. A. Frearson and A. H. Fairlamb, *Biochem Pharmacol*, 2010, **80**, 1478-1486.
8. J. Rappsilber, Y. Ishihama and M. Mann, *Anal Chem*, 2003, **75**, 663-670.
9. J. Cox, N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen and M. Mann, *J Proteome Res*, 2011, **10**, 1794-1805.
10. J. Cox, M. Y. Hein, C. A. Luber, I. Paron, N. Nagaraj and M. Mann, *Mol Cell Proteomics*, 2014, **13**, 2513-2526.
11. J. A. Vizcaino, E. W. Deutsch, R. Wang, A. Csordas, F. Reisinger, D. Rios, J. A. Dianes, Z. Sun, T. Farrah, N. Bandeira, P. A. Binz, I. Xenarios, M. Eisenacher, G. Mayer, L. Gatto, A. Campos, R. J. Chalkley, H. J. Kraus, J. P. Albar, S. Martinez-Bartolome, R. Apweiler, G. S. Omenn, L. Martens, A. R. Jones and H. Hermjakob, *Nat Biotechnol*, 2014, **32**, 223-226.
12. J. A. Vizcaino, A. Csordas, N. Del-Toro, J. A. Dianes, J. Griss, I. Lavidas, G. Mayer, Y. Perez-Riverol, F. Reisinger, T. Ternent, Q. W. Xu, R. Wang and H. Hermjakob, *Nucleic Acids Res*, 2016, **44**, 11033.
13. S. Maere, K. Heymans and M. Kuiper, *Bioinformatics*, 2005, **21**, 3448-3449.