

# Direct mapping of ligandable tyrosines and lysines in cells with chiral sulfonyl fluoride probes

Received: 23 November 2022

Accepted: 23 June 2023

 Check for updates

Ying Chen<sup>1</sup>, Gregory B. Craven<sup>1</sup>, Roarke A. Kamber<sup>2</sup>, Adolfo Cuesta<sup>1</sup>, Serhii Zhersh<sup>3</sup>, Yurii S. Moroz<sup>4,5</sup>, Michael C. Bassik<sup>2,6,7</sup> & Jack Taunton<sup>1</sup>✉

Advances in chemoproteomic technology have revealed covalent interactions between small molecules and protein nucleophiles, primarily cysteine, on a proteome-wide scale. Most chemoproteomic screening approaches are indirect, relying on competition between electrophilic fragments and a minimalist electrophilic probe with inherently limited proteome coverage. Here we develop a chemoproteomic platform for direct electrophile-site identification based on enantiomeric pairs of clickable arylsulfonyl fluoride probes. Using stereoselective site modification as a proxy for ligandability in intact cells, we identify 634 tyrosines and lysines within functionally diverse protein sites, liganded by structurally diverse probes. Among multiple validated sites, we discover a chiral probe that modifies Y228 in the MYC binding site of the epigenetic regulator WDR5, as revealed by a high-resolution crystal structure. A distinct chiral probe stimulates tumour cell phagocytosis by covalently modifying Y387 in the recently discovered immuno-oncology target APMAP. Our work provides a deep resource of ligandable tyrosines and lysines for the development of covalent chemical probes.

Electrophilic probes are powerful tools for investigating protein function and quantifying target engagement. Cysteine, which is the most intrinsically reactive amino acid, is the primary focus of targeted covalent drugs, as exemplified by multiple cancer therapeutics, including osimertinib<sup>1</sup>, ibrutinib<sup>2</sup> and sotorasib<sup>3</sup>. However, many ligand-binding sites lack a proximal cysteine, limiting the generalizability of this approach. Lysine (Lys) and tyrosine (Tyr), which are frequently located near ligand-binding sites, are potential alternative nucleophiles for covalent targeting. Despite their lower intrinsic nucleophilicity, recent efforts have led to the development of multiple Lys-targeted inhibitors (for example, Hsp90<sup>4</sup>, eIF4E<sup>5</sup>, PI3K<sup>6</sup> and XIAP<sup>7</sup>) and Tyr-targeted inhibitors (for example, Ral<sup>8</sup>, BCL6<sup>9</sup>, SRPK1<sup>10</sup>, Bcl-xL<sup>11</sup> and A1-AR<sup>12</sup>). Nevertheless, covalent probes are not available for the vast majority of human proteins.

Fragment-based ligand discovery coupled with quantitative chemical proteomics has been utilized to identify 'ligandable' sites in native biological systems<sup>13–17</sup>. Typically, the ligandability of a nucleophilic site is assessed indirectly by measuring competition ratios, in which cell lysates or intact cells are first treated with structurally diverse electrophilic fragments, followed by treatment with a minimalist electrophilic probe, which is designed to react broadly with either Cys (for example, iodoacetamide), Lys (for example, active ester) or Tyr (for example, sulfonyl triazole).

Enriched peptides bearing the minimalist probe are identified and competition ratios quantified ( $\pm$  electrophilic fragment competitor) by data-dependent acquisition mass spectrometry (DDA MS). For example, 818 Lys were liganded (out of 14,000 Lys quantified) after screening ~180 electrophilic fragments in cell lysates for their ability to

<sup>1</sup>Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA, USA. <sup>2</sup>Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA. <sup>3</sup>Enamine Ltd, Kyiv, Ukraine. <sup>4</sup>National Taras Shevchenko University of Kyiv, Kyiv, Ukraine. <sup>5</sup>Chemspace LLC, Kyiv, Ukraine. <sup>6</sup>Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA, USA. <sup>7</sup>Program in Chemistry, Engineering, and Medicine for Human Health (ChEM-H), Stanford University, Stanford, CA, USA. ✉e-mail: [jack.taunton@ucsf.edu](mailto:jack.taunton@ucsf.edu)

block subsequent labelling with a clickable activated ester<sup>15</sup>. Although hundreds of liganded Lys and Tyr have been reported in these studies, a limitation to the indirect competition approach is that only a tiny fraction (<1%) of the 650,000 Lys and 300,000 Tyr in the human proteome can be identified in a typical DDA MS dataset. This low site coverage arises from two factors: (1) the minimalist probe may not react sufficiently with all potentially ligandable Lys and Tyr; (2) the stochastic nature of DDA MS, coupled with variations in protein abundance and site reactivity, severely limit the absolute number of sites that can be reproducibly identified. We envisioned that direct identification of probe-modified peptides<sup>18</sup>, achieved by incorporating a clickable alkyne handle into a small library of structurally distinct electrophilic probes, could increase the likelihood of detecting ligandable nucleophiles that would be missed by indirect methods based on competition with 'minimalist' electrophilic probes.

Stereochemistry is a critical determinant of ligand–protein interactions and has been widely exploited in the discovery of electrophilic probes. Recently, enantiomeric pairs of chiral fragments bearing a diazine moiety were employed in whole-proteome photo-crosslinking experiments<sup>19</sup>. Although this innovative approach led to the identification of >170 stereoselectively liganded proteins, direct identification of the actual photo-crosslinking site was only achieved for a small fraction of the targets. Moreover, the photo-crosslinking approach does not immediately lend itself to the design of site-selective electrophilic probes. Although chiral acrylamides have recently been used to map ligandable cysteines, these studies still relied on an indirect competition approach<sup>20,21</sup>.

In this Article we report the design of a small library of alkyne-functionalized arylsulfonyl fluoride probes comprising a chiral 2-methylpiperazine amide linked to a heteroaryl diversity element. We directly quantify the ability of each pair of enantiomeric probes to covalently modify tyrosines and lysines in live cells, leading to the identification of hundreds of stereoselectively modified sites. Ligandable hotspots are identified in substrate binding pockets and protein–protein interfaces, many of which lack existing chemical probes or drug leads. We validate stereoselective modification of the epigenetic regulator WDR5 and reveal the structural basis of this interaction by X-ray crystallography. Finally, we discover a chiral sulfonyl fluoride probe that covalently modifies and functionally inactivates APMAP, a recently characterized cancer immunotherapy target.

## Results and discussion

### Chemoproteomic platform based on chiral sulfonyl fluorides

To directly identify ligandable Tyr and Lys in intact cells, we designed a set of 20 chiral sulfonyl fluoride probes—ten enantiomeric pairs—with the following features: (1) arylsulfonyl fluoride to react with proximal Tyr and Lys (Fig. 1a); (2) an alkyne handle for affinity enrichment and direct identification of covalently modified protein sites; (3) an (*S*)- or (*R*)-2-methylpiperazine amide linker, capable of providing stereoselective discrimination at the level of noncovalent and covalent binding<sup>4</sup>; (4) a diversity element comprising ten heteroaryl fragments (Fig. 1b), each with at least one *sp*<sup>3</sup> carbon and two rotatable bonds (*M*<sub>w</sub> of 159–265 Da).

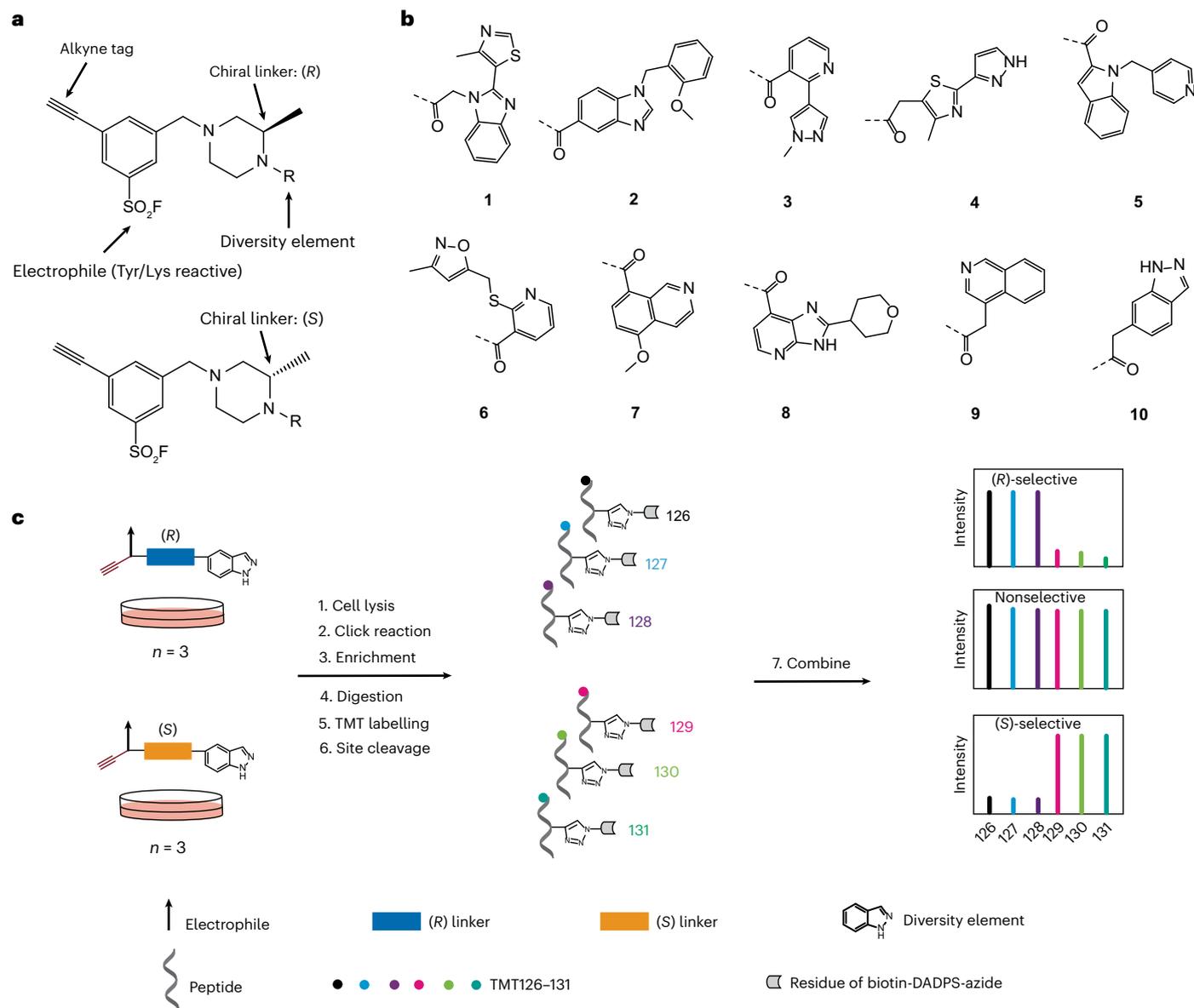
The fully elaborated probes have molecular weights and calculated lipophilicities (*M*<sub>w</sub>, 454–560 Da; *c log P*, 2.9–5.2; *pK<sub>a</sub>*, -8) within the range of approved drugs and are expected to be cell-permeable. Although the probes share a common arylsulfonyl fluoride and piperazine linker, we reasoned that the chirality of the linker (the axial methyl group enforced by the piperazine amide bond), along with the structural diversity and hydrogen-bonding potential of the heteroaryl fragments (Fig. 1b), would enable covalent modification of proximal Tyr and Lys residues within distinct binding sites. Instead of using competition ratios, we performed a head-to-head comparison of matched (*R*) and (*S*) enantiomeric probes (bearing the same heteroaryl fragment, Fig. 1b) for their ability to modify endogenous proteins in living cells.

We developed a multiplexed chemoproteomics workflow based on isobaric tandem mass tags (TMTs) to directly identify probe-modified sites and quantify the enantiomer modification ratios. Jurkat cells were separately treated for 1 h with 10 μM of each enantiomer of a given probe pair (three biological replicates, six samples total; Fig. 1c). Cell lysates were prepared, and biotin–DADPS (dialkoxydiphenylsilane)–azide was conjugated to probe-modified proteins under denaturing conditions via Cu(I)-promoted click chemistry. Biotin–DADPS–azide, which contains a formic acid-cleavable dialkoxydiphenylsilane linker<sup>22</sup>, has been shown to perform optimally in chemoproteomics studies involving modified site identification<sup>23,24</sup>. After neutravidin enrichment of biotinylated probe-modified proteins and on-bead digestion to remove unmodified tryptic peptides, each sample was separately labelled on-bead with 6-plex TMT reagents (Fig. 1c). Finally, probe-modified peptides were treated with 2% formic acid to cleave the DADPS linker, eluted from the neutravidin beads, and combined for liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis. To determine the enantiomer modification ratio of a given probe-modified peptide identified by LC-MS/MS analysis, we compared the relative intensities of the six TMT reporter ions derived from each set of triplicate samples (Fig. 1c). We used the enantiomer modification ratio (e.m.r.) as our primary metric for determining ligandability, such that probe-modified sites were defined as 'liganded' if they had a mean enantiomer modification ratio of at least twofold (e.m.r. ≥ 2, *P* ≤ 0.05, *n* = 3, Student's *t*-test).

### MS quantitation of stereoselectively modified sites

MS2-based quantitation of TMT reporter ions often suffers from ratio distortion due to co-isolation and fragmentation of interfering precursor ions, a problem that is partially addressed by MS3-based quantitation. Recently, real-time database search (RTS) has been developed to augment the speed of MS3 quantitation<sup>17,25,26</sup>. We separately analysed each 6-plex sample using MS2 and RTS-MS3 acquisition methods (two LC-MS runs each, four total runs per 6-plex TMT sample) to maximize the number of quantified peptides (MS2), while also maximizing quantification accuracy (RTS-MS3). Although there are isolated examples of arylsulfonyl fluorides reacting with other protein nucleophiles (for example, Ser, Thr, Cys and His)<sup>27–30</sup>, we focused our search on tyrosines and lysines, which are the predominant targets of arylsulfonyl fluorides in unbiased chemoproteomic studies<sup>31</sup>.

Using the criteria defined above for stereoselective site modification, we identified 41 (*R*)-selective and 40 (*S*)-selective sites for probe 1, whereas the majority of sites (540) were modified to a similar extent by both enantiomers (Fig. 2a). Similar trends were observed for the other enantiomeric probe pairs 2–10 (Extended Data Fig. 1). Across all probes 1–10, the RTS-MS3 method identified more stereoselectively modified sites than the MS2 method, despite identifying fewer total sites (Fig. 2b–e and Supplementary Data 1). Probes 1–10 exhibited a strong preference for modifying tyrosines over lysines (Fig. 2b–f and Supplementary Data 1), consistent with previous studies of arylsulfonyl fluoride reactivity<sup>18,31–33</sup>. From a total of 40 LC-MS runs (probe pairs 1–10, four runs each), we identified 513 stereoselectively modified Tyr (30% of 1,728 total modified Tyr) and 121 stereoselectively modified Lys (20% of 615 total modified Lys) derived from 483 protein targets (37% of 1,312 modified proteins) (Fig. 2f). Of the 513 Tyr and 121 Lys that were stereoselectively modified by probes 1–10, only 7% and 16%, respectively, were found to be liganded in two previous chemoproteomic studies<sup>14,15</sup> (Extended Data Fig. 2a–f), suggesting that our direct method is complementary to indirect, competition-based methods for assessing proteome-wide ligandability. Finally, we identified an additional subset of potentially ligandable sites by analysing fragment selectivity (Fig. 1b). With the caveat that certain probe-modified peptides may have escaped detection during all four LC-MS runs (for example, due to a relatively low abundance and/or ionization efficiency), 58% of the 2,343 sites in our dataset were identified by only one or two probe pairs



**Fig. 1 | Design of ten enantiomeric pairs of chiral sulfonamide probes to identify ligandable tyrosines and lysines in living cells.** **a**, Design of chiral sulfonamide probes. **b**, Chemical structures of diversity fragments 1–10. **c**, TMT6 chemoproteomic workflow to identify stereoselectively modified sites. Jurkat cells were treated separately with 10  $\mu\text{M}$  (*R*) and (*S*) probes for 1 h at 37 °C ( $n = 3$  for each enantiomeric probe). Cell lysates were conjugated with biotin–

DADPS–azide via click chemistry. After enrichment and trypsin digestion, probe-modified peptides from each sample were separately labelled with TMT126–TMT128 (*R*) and TMT129–TMT131 (*S*), respectively. Finally, probe-modified peptides from each sample were eluted with 2% formic acid and combined. After data acquisition and analysis, stereoselectively modified sites were identified by comparing the relative intensities of TMT reporter ions.

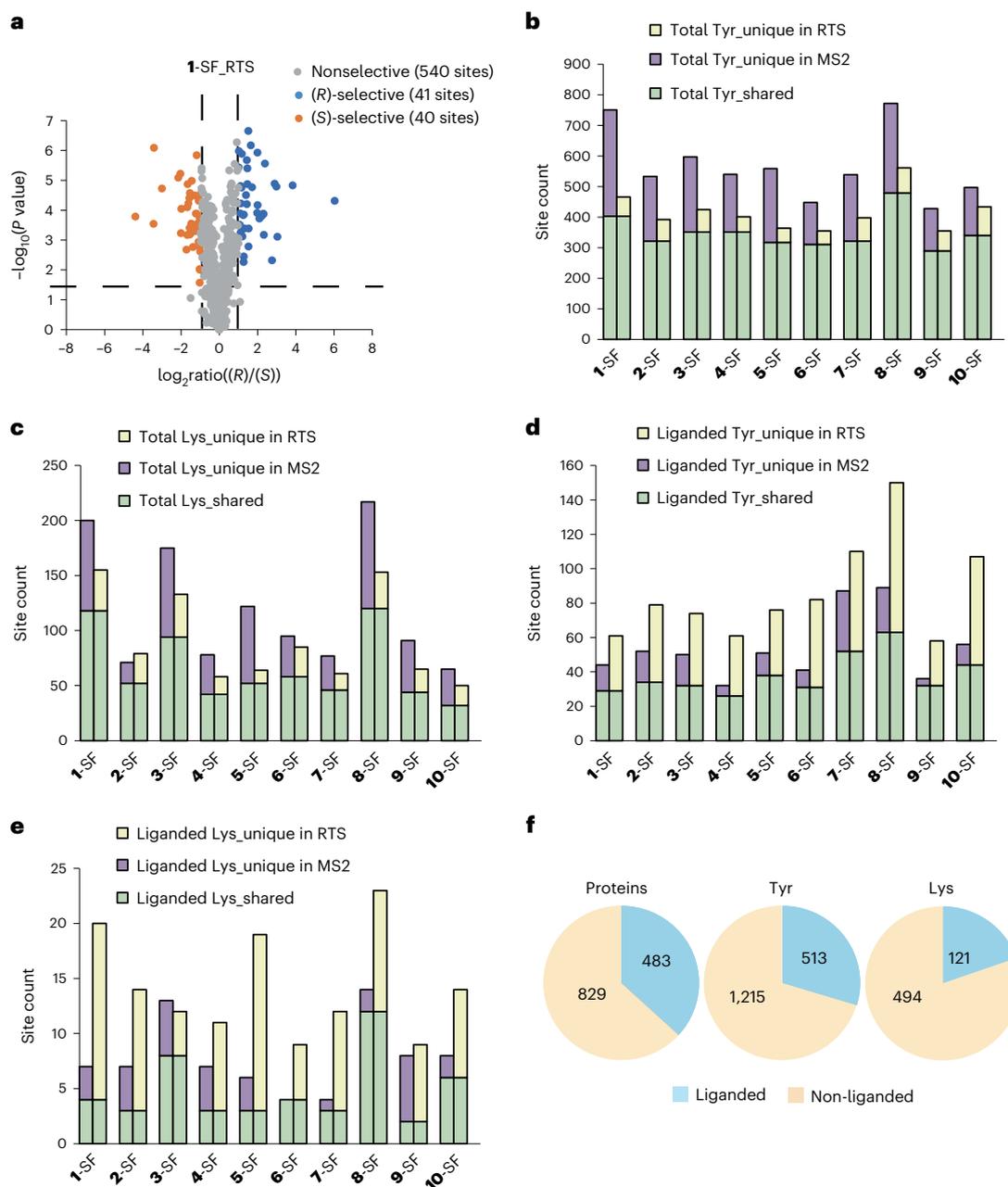
(Extended Data Fig. 2g), including 1,127 sites that did not meet our stereoselective modification criteria (Supplementary Data 1). We hypothesize that many of these sites were modified in a fragment-selective manner and are in fact ligandable (for example, APMAP; see below).

### Chiral sulfonamide (SF) probes target diverse protein sites

We queried all 2,343 probe-modified sites against the Protein Data Bank (PDB) using the NCBI BLAST database (downloaded from <https://ftp.ncbi.nlm.nih.gov/blast/db/>) and found that ~40% are represented in experimental protein structures (resolution cutoff of  $\leq 3.5$  Å). Nevertheless, only ~10% of probe-modified sites are within 10 Å of a bound small molecule (Supplementary Data 1). These data suggest that our direct site-specific chemoproteomic method can potentially identify ligandable sites in previously unexplored regions of the proteome. Structure-based analysis of selected sites from our dataset (Fig. 3a

and Supplementary Data 1) reveals a rich diversity of biochemical functions, including active sites of post-translational modification enzymes (for example, PARP1 and LCMT1; Fig. 3b,c), a phosphoinositide binding domain (ABR; Fig. 3d), protein–protein interaction domains in epigenetic regulators (for example, WDR5 and WDR82; Fig. 3e,f), RNA-binding domains (for example, FUS and YBX1; Extended Data Fig. 3a,b) and a druggable allosteric site (KIF11; Extended Data Fig. 3c). Functional diversity is also evident in fragment-selective sites that were identified by only one or two probes (Fig. 3a), but which did not meet our criteria for stereoselective modification (for example, APMAP, Fig. 3g; KDM1A, PRMT5 and SIRT2, Extended Data Fig. 3d–f).

With the goal of highlighting functional diversity, we selected six probe-modified sites for validation studies (Fig. 3b–g). Of these proteins, small-molecule inhibitors have only been reported for the poly-adenosine diphosphate (ADP)-ribose polymerase PARP1 and



**Fig. 2 | Global analysis of tyrosines and lysines modified by chiral sulfonyl fluoride probes.** **a**, Volcano plot showing modified sites for (R) and (S) 1-SF with  $\log_2 \text{ratio}((R)/(S))$  on the x axis and  $-\log_{10}(P \text{ value})$  on the y axis (RTS-MS3 acquisition). P values were determined by Student's *t*-test (two-tailed, two-sample equal variance). Modified sites are indicated as (R)-selective (blue), (S)-selective (orange) and nonselective (grey). **b,c**, Total number of probe-modified

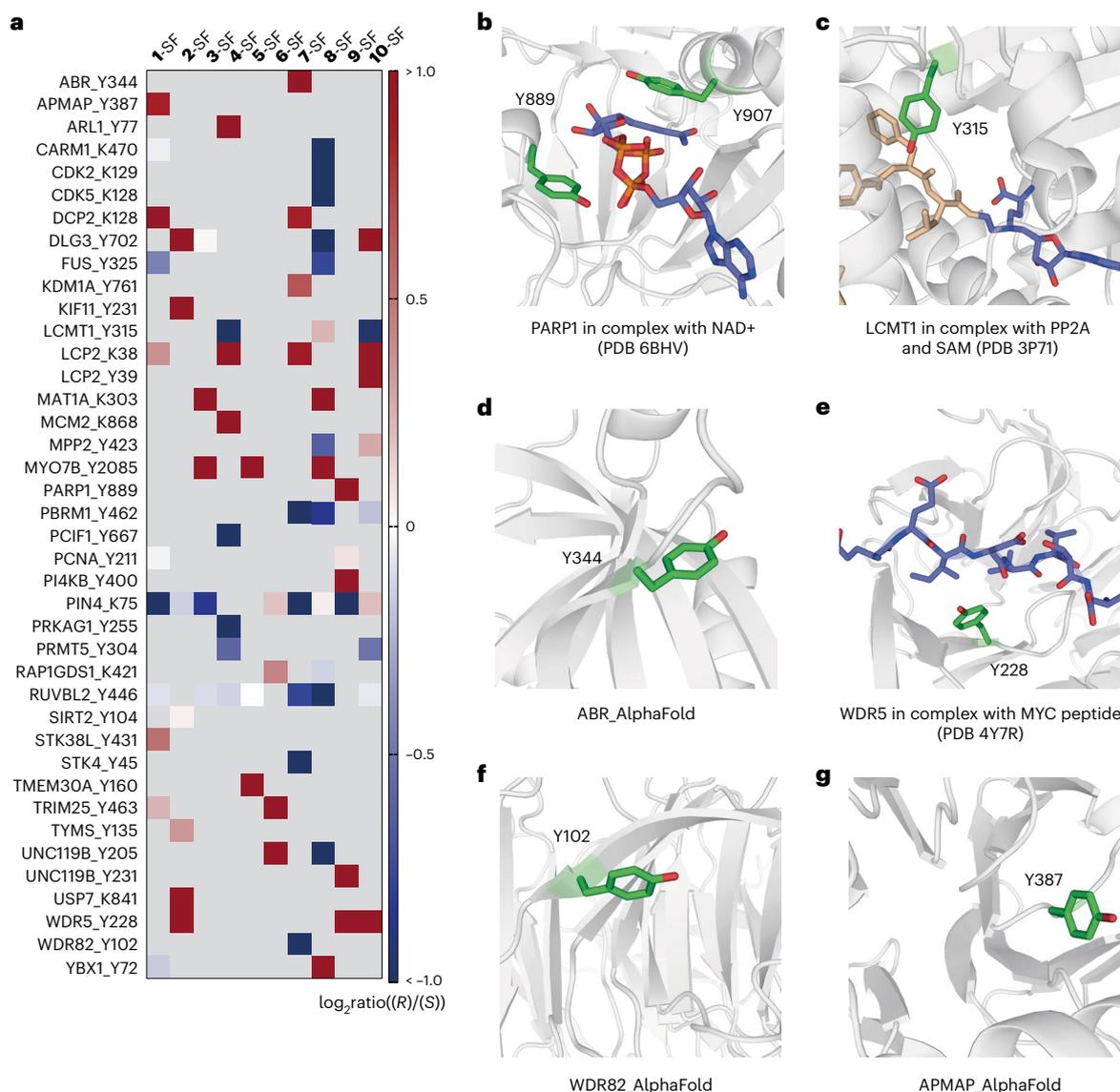
Tyr (**b**) and Lys (**c**) from MS2 and RTS-MS3 acquisition modes. **d,e**, Number of stereoselectively modified (liganded) Tyr (**d**) and Lys (**e**) from MS2 and RTS-MS3 acquisition modes. **f**, Total number of modified proteins, Tyr and Lys that were stereoselectively modified (liganded), or not, by one or more sulfonyl fluoride probes.

the epigenetic scaffolding protein WDR5. To evaluate stereoselective and/or fragment-selective modification, HEK293T cells were first transfected with a Flag epitope-tagged expression construct encoding either the wild-type (WT) full-length protein or a mutant in which the putatively reactive Tyr was changed to Phe.

Transfected cells were treated with chiral SF probes for 1 h, and, after cell lysis, Flag-tagged proteins were immunoprecipitated, eluted with excess Flag peptide, conjugated to TAMRA-azide via click chemistry and analysed by in-gel fluorescence.

PARP1 catalyses poly-ADP ribosylation of protein substrates and is a validated drug target for breast and ovarian cancer. Y889 interacts

with the NAD<sup>+</sup> cofactor (Fig. 3b) and resides on the opposite side of the PARP1 active site from Y907, which was previously identified as the target of a sulfamidimidoyl fluoride probe<sup>34</sup>. Consistent with our chemoproteomics data (Fig. 3a and Supplementary Data 1), PARP1 was modified by (R)-9-SF with high fragment selectivity and stereoselectivity (Fig. 4a,b). Importantly, the Y889F mutation abrogated labelling by (R)-9-SF (Fig. 4b), confirming that Y889 is the dominant modification site on PARP1, which has 32 Tyr. Crystal structures reveal that Y889 directly contacts the clinically approved PARP1 inhibitors rucaparib (Fig. 4c) and olaparib. As expected with an active-site-directed probe, covalent modification by (R)-9-SF was reduced by pretreatment of cells



**Fig. 3 | Chiral SF probes target functionally diverse protein sites.** **a**, Heatmap of selected sites. Red, (*R*)-selective; blue, (*S*)-selective; white, nonselective; grey, not detected. **b–g**, Crystal structures and AlphaFold models indicating modified sites on PARP1 (**b**, PDB 6BHV), LCMT1 (**c**, PDB 3P71), ABR (**d**, AlphaFold), WDR5 (**e**, PDB 4Y7R), WDR82 (**f**, AlphaFold) and APMAP (**g**, AlphaFold).

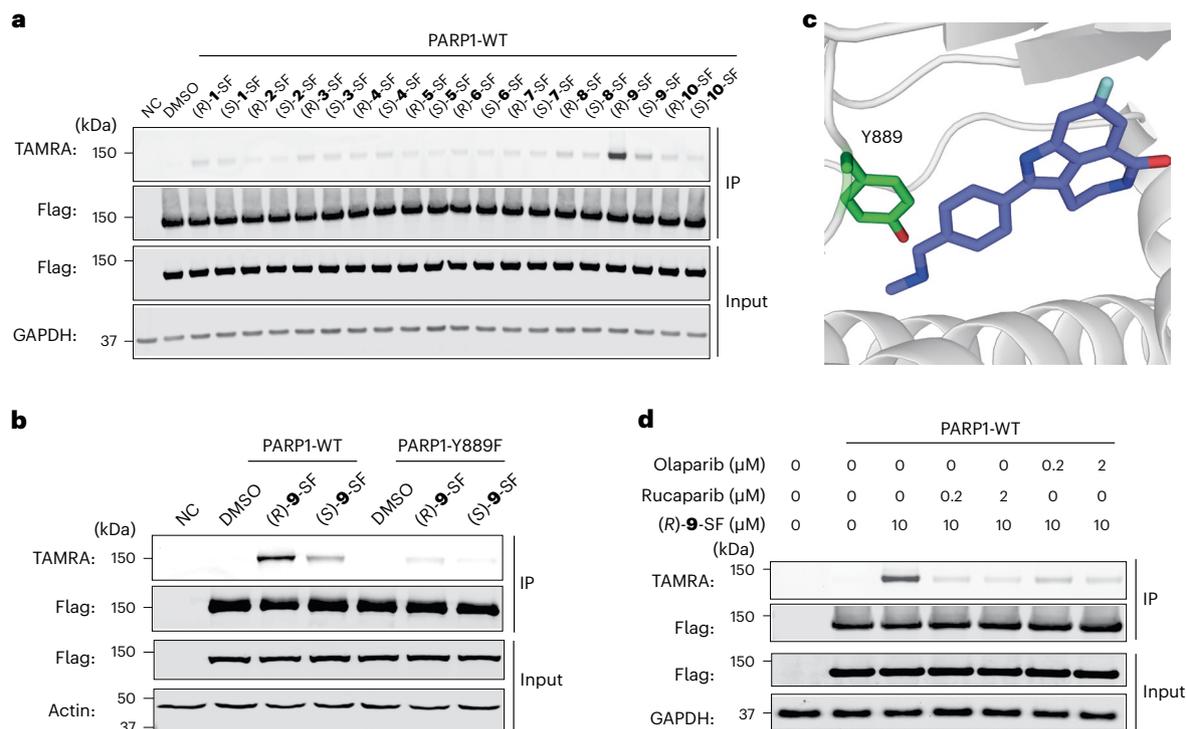
with either rucaparib or olaparib (Fig. 4d). Although (*R*)-9-SF lacks sufficient potency and selectivity for interrogating PARP1 biology, our data demonstrate that this clickable SF has utility as a cellular occupancy probe for assessing PARP1 inhibitors.

We confirmed stereoselective and site-selective modification of full-length LCMT1, ABR and WDR82 by (*S*)-4-SF, (*R*)-7-SF and (*S*)-7-SF, respectively; we noted that fragment selectivity was not uniformly high across all three proteins, and some degree of nonspecific labelling was observed in each case (Extended Data Figs. 4–6). Importantly, unlike with PARP1, there are no reported small molecules that bind to these sites. Based on structural analysis<sup>35</sup>, probes that modify Y315 of the leucine carboxy-terminal methyltransferase LCMT1 are expected to disrupt interactions with its protein substrates (Ser/Thr phosphatase catalytic subunits PP2A, PP4 and PP6), methylation of which is essential for function<sup>36,37</sup>. ABR Y344 is in the PH (pleckstrin homology) domain, which probably binds the phosphoinositide PIP2 and mediates membrane recruitment based on studies of the related protein BCR<sup>38</sup>. WDR82 is a scaffolding protein that binds transcriptional regulators in large multi-protein complexes, including the histone

methyltransferase SET1<sup>39,40</sup>. Although the structural details of WDR82–protein interactions remain unknown, an AlphaFold model reveals that Y102 is adjacent to a likely protein–protein interaction site (Extended Data Fig. 6d), inferred by analogy to the distantly related scaffolding protein WDR5. Hence, compounds that covalently modify WDR82 Y102, such as (*S*)-7-SF, may directly or allosterically disrupt scaffolding interactions with transcriptional regulators.

#### (*R*)-2-SF blocks WDR5 binding to the MYC oncoprotein

WDR5 regulates gene expression in cancer cells, in part through direct interactions with histone methyltransferases and the oncogenic transcription factor MYC<sup>41</sup>. As such, the MYC binding site on WDR5—a shallow groove on the opposite side of the beta-propeller domain from the histone methyltransferase binding site—has emerged as a potential therapeutic target for MYC-driven cancers. Despite the challenges of targeting a protein–protein interface, recent work has led to the development of high-affinity ligands ( $K_d \approx 100$  nM) that disrupt the WDR5/MYC complex at micromolar concentrations in cells<sup>42,43</sup>, including a lysine-targeted covalent inhibitor<sup>44</sup>.



**Fig. 4 | (R)-9-SF modifies PARP1 Y889 in cells. a**, HEK293T cells were transfected with Flag-PARP1 and treated with SF probes (10  $\mu\text{M}$ , 1 h). After cell lysis, anti-Flag enrichment, Flag peptide elution, and click reaction with TAMRA-azide, the samples were analysed by in-gel fluorescence and western blotting. Data are representative of two independent experiments. NC, nontransfected cells. **b**, HEK293T cells were transfected with WT and Y889F Flag-PARP1 and treated

with 9-SF probes (10  $\mu\text{M}$ , 1 h). Cells were lysed and processed as described above. Data are representative of two independent experiments. **c**, Crystal structure of PARP1 bound to rucaparib (PDB 4RV6). **d**, HEK293T cells were transfected with Flag-PARP1 and treated with rucaparib or olaparib for 1 h followed by treatment with (R)-9-SF for 1 h. The cells were lysed and processed as described above. Data are representative of two independent experiments.

Our chemoproteomic analysis revealed that (R)-2-SF, (R)-9-SF and (R)-10-SF stereoselectively modify WDR5 Y228 (Fig. 3a), which forms part of the MYC binding groove. We directly compared all 20 probes in cells overexpressing Flag-WDR5 and assessed covalent modification using the experimental workflow described above. This experiment revealed stronger labelling of WDR5 by (R)-2-SF and (R)-9-SF than the other 18 probes, including (R)-10-SF (Fig. 5a). Within each probe pair, the (R) enantiomer was consistently superior, suggesting that the chiral 2-methylpiperazine amide plays a dominant role in molecular recognition (Fig. 5a,b). Finally, the Y228F mutation led to decreased labelling, both in cells (Fig. 5b) and with purified WDR5 (Fig. 5c), confirming that Y228 is the major modification site.

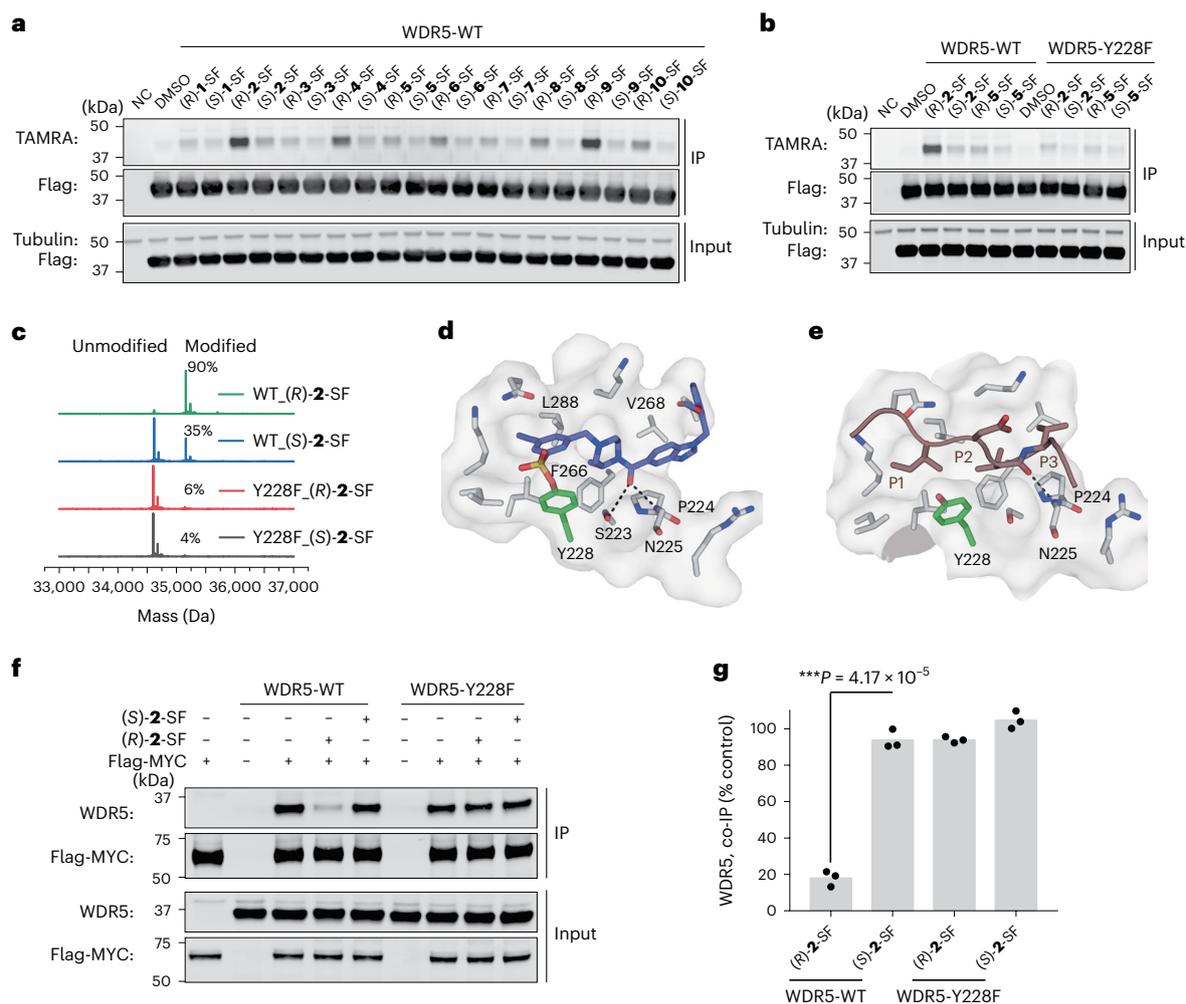
To elucidate the molecular details of WDR5 modification by (R)-2-SF, we solved the structure of the covalent complex by X-ray crystallography. The 2.3 Å resolution structure unambiguously defined the ligand-binding pose and the covalent bond to Y228 (Fig. 5d and Extended Data Figs. 7 and 8). MYC engages WDR5 via a conserved IDVV tetrapeptide motif, in which the Ile and Val side chains occupy distinct hydrophobic pockets P1–P3 (Fig. 5e). In the case of (R)-2-SF, the ethynyl phenylsulfonyl moiety occupies the P1 pocket in a manner compatible with covalent-bond formation with Y228 (Fig. 5d). The piperazine linker fills the P2 pocket such that the axial (R)-methyl group slots in between F266 and L288. In addition, the (R)-stereocentre allows the adjacent amide carbonyl to accept hydrogen bonds from both the side-chain hydroxyl of S223 and the backbone amide of N225 (Fig. 5d). The latter hydrogen bond mimics a backbone carbonyl–NH interaction between MYC and WDR5 (Fig. 5e). Finally, the benzimidazole fragment extends into the P3 site, packing between P224 and V268 (Fig. 5d). Our co-crystal structure therefore predicts that WDR5 modification by (R)-2-SF should inhibit MYC binding. To test this, we immunoprecipitated full-length, Flag-MYC from cell lysates containing exogenous

recombinant WDR5. Pretreatment of WT WDR5 with (R)-2-SF, but not the (S) enantiomer, abrogated binding to Flag-MYC (Fig. 5f,g). By contrast, (R)-2-SF had no effect on the interaction between Y228F WDR5 and Flag-MYC. Although further work is required to improve cellular potency and proteome-wide selectivity, targeting Y228 with an arylsulfonyl fluoride linked to a chiral 2-methylpiperazine amide represents a distinct approach to disrupt the WDR5–MYC complex and provides a clickable probe for assessing WDR5 occupancy in cells by inhibitors that compete with MYC.

### Modification of APMAP Y387 by (R)-1-SF promotes phagocytosis

Adipocyte plasma membrane-associated protein (APMAP) is a putative esterase comprising a single transmembrane segment and an ectodomain with predicted structural homology to the calcium-dependent lactonase PON1. Direct evidence of APMAP enzymatic activity is lacking, and its biological functions remain poorly understood. Recently, genetic deletion of APMAP from cancer cells was found to promote antibody-dependent phagocytosis by associated macrophages<sup>45</sup>. APMAP thus represents a compelling immuno-oncology target. To the best of our knowledge, no small-molecule inhibitors or endogenous substrates of APMAP have been reported, motivating the development of new chemical probes.

Our chemoproteomic data revealed that (R)-1-SF modifies APMAP Y387 with exceptional fragment selectivity (Fig. 3a), which we confirmed in cells that stably overexpress Flag-APMAP using TAMRA-azide click chemistry and in-gel fluorescence (Fig. 6a). Treatment of cells with (R)-1-SF resulted in saturable, dose-dependent labelling with a half-maximum effective concentration ( $\text{EC}_{50}$ ) of ~8  $\mu\text{M}$ , whereas APMAP labelling by (S)-1-SF was barely detected at 60  $\mu\text{M}$  (Fig. 6b). These results indicate that the stereoselectivity of APMAP modification is



**Fig. 5 | (R)-2-SF blocks WDR5 binding to MYC. a**, HEK293T cells were transfected with Flag-WDR5 WT and then treated with the indicated SF probes (10  $\mu$ M, 1 h). After cell lysis, anti-Flag enrichment, Flag peptide elution, and click reaction with TAMRA-azide, samples were analysed by in-gel fluorescence and western blotting. Data are representative of two independent experiments. NC, nontransfected cells. **b**, HEK293T cells were transfected with WT or Y228F Flag-WDR5 and treated with 2-SF or 5-SF probes (10  $\mu$ M, 1 h). The cells were lysed and processed as described above. Data are representative of two independent experiments. **c**, Recombinant WT and Y228F WDR5 (2  $\mu$ M) were treated with 2-SF probes (40  $\mu$ M, 37  $^{\circ}$ C, 6 h) and analysed by intact-protein MS. Data are

representative of two independent experiments. **d**, Co-crystal structure of WDR5 covalently modified by (R)-2-SF. **e**, Co-crystal structure of MYC peptide bound to WDR5 near Y228 (PDB 4Y7R). **f**, Recombinant WT and Y228F WDR5 (0.5 mg ml<sup>-1</sup>) were treated with 2-SF probes (70  $\mu$ M, 37  $^{\circ}$ C, 7 h) and then added to lysates prepared from HEK293T cells expressing Flag-MYC. After anti-Flag enrichment and Flag peptide elution, samples were analysed by western blotting. **g**, Co-immunoprecipitated WDR5 was quantified from three biological replicates of the experiment in **f**. The *P* value was determined by Student's *t*-test (two-tailed, two-sample equal variance).

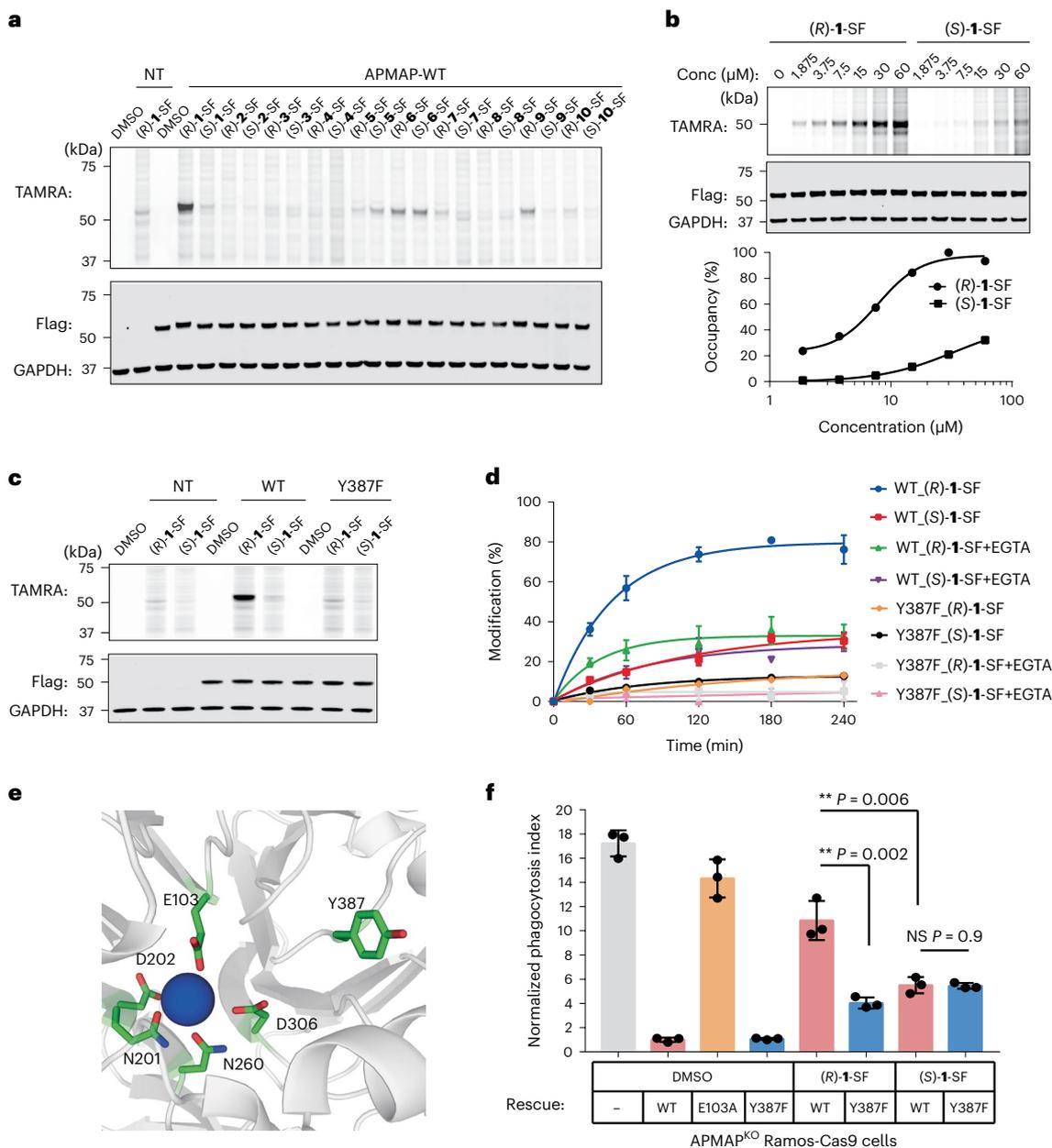
actually higher than suggested by the initial chemoproteomic data, probably due to MS2-based TMT ratio compression (MS2 e.m.r., 1.82;  $P = 0.001$ ; RTS-MS3, not detected; Supplementary Data 1). Modification of the Y387F mutant was greatly reduced compared to WT APMAP, both in cells and with the purified recombinant ectodomain (Fig. 6c,d). Comparison of the AlphaFold model of APMAP with the crystal structure of PON1 reveals a potential calcium-binding motif proximal to Y387 (Fig. 6e), and mutation of the putative calcium ligand E103 was previously shown to impair its ability to suppress cancer-cell phagocytosis<sup>45</sup>. Consistent with a structural and/or functional role for calcium, treatment of recombinant APMAP with EGTA-reduced covalent modification by (R)-1-SF (Fig. 6d).

The apparently saturable labelling of Flag-APMAP in cells treated with (R)-1-SF (Fig. 6b) motivated us to evaluate our multiplexed chemoproteomics platform in dose–response mode.

We performed a TMT 11-plex experiment by treating cells that stably express Flag-APMAP with five concentrations of (R)-1-SF and (S)-1-SF (3.75, 7.5, 15, 30 and 60  $\mu$ M), along with a dimethylsulfoxide (DMSO)

control (Extended Data Fig. 9a and Supplementary Data 2). Analysis of TMT intensities derived from probe-modified peptides (using both MS2 and RTS-based TMT quantification) confirmed stereoselective modification of APMAP Y387, with dose–response curves that broadly recapitulate the in-gel fluorescence results (Extended Data Fig. 9b). By contrast, modification of APMAP K347, presumably a minor modification site, was not stereoselective and produced a linear dose–response profile up to probe concentrations of 60  $\mu$ M (Extended Data Fig. 9c). The dose–response experiment also revealed stereoselective and saturable modification of MTHFD1 Y52, which lies within the active site of cytosolic methylenetetrahydrofolate dehydrogenase and directly interacts with a potent inhibitor (Extended Data Fig. 9d,e).

To test whether covalent modification of APMAP Y387 has a functional impact on cancer cells, we studied the effect of (R)-1-SF pretreatment on antibody-dependent phagocytosis by activated macrophages. Ramos lymphoma cells lacking APMAP (APMAP<sup>KO</sup> Ramos-Cas9) were genetically rescued by transduction with either GFP (negative control), WT APMAP, or the APMAP mutants E103A or Y387F (Extended Data



**Fig. 6 | Covalent modification of APMAP Y387 by (R)-1-SF induces cancer-cell phagocytosis. a**, HEK293T cells stably overexpressing Flag-APMAP were treated with the indicated SF probes (10 μM, 1 h). After cell lysis and click reaction with TAMRA-azide, samples were analysed by in-gel fluorescence and western blotting. Data are representative of two independent experiments. NT, nontransfected cells. **b**, HEK293T cells stably overexpressing Flag-APMAP were treated with 1-SF probes (1.875, 3.75, 7.5, 15, 30 and 60 μM, 1.5 h), collected and lysed. Cell lysates were processed as described above. Data are representative of three independent experiments. **c**, HEK293T cells stably overexpressing WT or Y387F Flag-APMAP were treated with 1-SF probes (10 μM, 1 h), collected and lysed. Cell lysates were processed as described above. Data are representative of two independent experiments. **d**, Recombinant WT and Y387F APMAP ectodomains (1 μM) were treated with 1-SF probes (80 μM, 37 °C) in the presence

of 2 mM CaCl<sub>2</sub> or 10 mM EGTA and analysed by intact-protein MS. The experiment was performed with three biological replicates and the results are plotted as mean values ± s.d. **e**, AlphaFold model of the APMAP calcium binding site. **f**, APMAP<sup>KO</sup> Ramos-Cas9 cells (rescued with green fluorescent protein (GFP) or APMAP WT, E103A or Y387F) were treated with 1-SF probes (30 μM, 1.5 h) and stained with calcein AM and pHrodo-Red succinimidyl ester dyes. Cells were pelleted and resuspended in medium containing 500 ng ml<sup>-1</sup> anti-CD20. Phagocytosis was monitored by time-lapse imaging after cells were added to LPS (lipopolysaccharide)-treated J774 macrophages. The experiment was performed with three biological replicates and the phagocytosis index (Methods) is plotted as mean values ± s.d. *P* values were calculated using Student's *t*-test (two-tailed, two-sample equal variance).

Fig. 10a). The resulting cell lines were treated with (R)-1-SF or (S)-1-SF (30 μM, 1.5 h) and incubated with pHrodo-Red succinimidyl ester dye and calcein AM (acetoxymethyl ester) dye to report on phagocytosis events and control for cell viability, respectively. Phagocytosis was initiated by adding activated J774 macrophages and the anti-CD20 antibody, rituximab. Cells were monitored by live imaging, and the

phagocytosis index was calculated as described in the Methods (Extended Data Fig. 10b).

Consistent with our previous report<sup>45</sup>, APMAP knockout cells showed substantial levels of phagocytosis, which was strongly suppressed by genetic rescue with WT, but not E103A APMAP (Fig. 6f). Remarkably, (R)-1-SF treatment promoted phagocytosis of APMAP<sup>KO</sup>

cells rescued with WT APMAP, similar in magnitude to untreated cells rescued with the inactive E103A mutant (Fig. 6f). Rescue of knock-out cells with Y387F APMAP suppressed phagocytosis, similar to WT. However, the Y387F APMAP mutant, which lacks the site of covalent modification, was significantly resistant to the phagocytosis-inducing effects of (*R*)-**1-SF** (Fig. 6f). This result provides strong genetic evidence for an on-target effect of (*R*)-**1-SF** and argues that induction of Ramos cell phagocytosis is mediated by covalent modification of APMAP Y387.

The enantiomeric probe (*S*)-**1-SF**, which weakly modifies APMAP (Fig. 6c), induced phagocytosis to a lesser extent than (*R*)-**1-SF**, in a manner that was unaffected by the Y387F mutation. Thus, by screening only 20 chiral SFs with our chemoproteomic platform, we have discovered a small molecule that functionally modulates APMAP. (*R*)-**1-SF** covalently modifies APMAP Y387 to block production of an as yet uncharacterized ‘don’t eat me’ signal in cancer cells, thereby promoting antibody-dependent phagocytosis and cell death.

## Conclusions

Bottom-up chemoproteomic methods have the potential to reveal new ligandable sites, and at the same time, new covalent probes that target these sites. Competitive chemoproteomic profiling is a powerful and widely used approach, especially for mapping ligandable cysteines.

Nevertheless, this indirect approach for identifying liganded sites is challenged by the following requirements: (1) a minimalist electrophilic probe that efficiently reacts with every nucleophilic site of potential interest throughout the proteome (that is, every ligandable Cys, Tyr or Lys) and (2) MS data acquisition and analysis methods that can reproducibly identify and quantify all of these sites. Although the total number of ligandable Cys, Tyr and Lys is unknown, it is likely that many, if not most, of these sites remain unexplored by current chemoproteomic approaches.

To address these challenges for Tyr and Lys, we developed a chemoproteomic platform that leverages (1) a bespoke library of clickable, chiral SF probes and (2) direct MS identification and quantification of probe-modified peptides. Using enantiomer modification ratios as a proxy for ligandability, we identified 634 stereoselectively modified sites (513 Tyr, 121 Lys)—along with matched clickable covalent probes—by screening only ten enantiomeric pairs of SFs in living cells. Out of 2,343 total modified sites, many were preferentially identified by a subset of the 20 tested probes (for example, APMAP). Such sites may also reflect specific molecular recognition, despite not meeting our criteria for stereoselective modification via TMT quantification.

Our study provides a rich resource of liganded sites and the first reported clickable covalent probes for most of these sites. We envision this resource being used by chemists and chemical biologists in at least three ways: (1) newly identified sites in potential therapeutic targets can become the focus of small-molecule screens; (2) SFs **1–10** can serve as chemical starting points for structure-based design of ligands with improved potency and selectivity toward selected sites; (3) SFs **1–10** can serve as clickable occupancy probes for cellular target engagement assays, as we have shown for (*R*)-**9-SF** and PARP1 (Fig. 4d).

Despite the broad utility of this resource, our chemoproteomic platform has certain limitations. First, unlike indirect competition-based methods, in which all sites are modified by the same minimalist probe, our method does not enable multiplexed quantitative comparison of chemically distinct probes due to differences in the mass, retention time and ionization efficiency of the resulting probe-modified peptides. Second, although direct identification of stereoselectively modified sites is advantageous, our chemoproteomic workflow does not reveal the fractional occupancy of a given site, unlike competition-based approaches. Most of the liganded sites in our dataset are probably modified in cells to a relatively low extent under the conditions of our experiments, with APMAP Y387 being a clear exception. This limitation could be addressed by performing the chemoproteomic analysis in dose–response mode, as we did

for (*R*)-**1-SF** and (*S*)-**1-SF** (Extended Data Fig. 9). Third, arylsulfonyl fluorides are metabolically unstable<sup>52,7</sup>, making it difficult to develop such probes for preclinical target validation studies. It may be possible to target ligandable Tyr with alternative electrophiles, such as sulfonyl triazoles<sup>14,46</sup>, fluorosulfates<sup>47–49</sup> or sulfuramidimidoyl fluorides<sup>34</sup>. Alternatively, because many reactive tyrosines<sup>50</sup> and ligand-binding sites<sup>51</sup> are proximal to a Lys (for example, WDR5 Y228, which is 9.6 Å from K250), it may be feasible to covalently and reversibly engage these sites using salicylaldehydes, which have been shown to retain activity in preclinical animal models<sup>52</sup> and humans<sup>53</sup>. The above limitations notwithstanding, our direct chemoproteomic approach powerfully complements indirect competition-based approaches for mapping the ligandable proteome.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41557-023-01281-3>.

## References

1. Cross, D. A. et al. AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. *Cancer Discov.* **4**, 1046–1061 (2014).
2. Byrd, J. C. et al. Targeting BTK with ibrutinib in relapsed chronic lymphocytic leukemia. *N. Engl. J. Med.* **369**, 32–42 (2013).
3. Hong, D. S. et al. KRAS(G12C) inhibition with sotorasib in advanced solid tumors. *N. Engl. J. Med.* **383**, 1207–1217 (2020).
4. Cuesta, A., Wan, X., Burlingame, A. L. & Taunton, J. Ligand conformational bias drives enantioselective modification of a surface-exposed lysine on Hsp90. *J. Am. Chem. Soc.* **142**, 3392–3400 (2020).
5. Wan, X. et al. Discovery of lysine-targeted eIF4E inhibitors through covalent docking. *J. Am. Chem. Soc.* **142**, 4960–4964 (2020).
6. Dalton, S. E. et al. Selectively targeting the kinase-conserved lysine of PI3K $\delta$  as a general approach to covalent kinase inhibition. *J. Am. Chem. Soc.* **140**, 932–939 (2018).
7. Baggio, C. et al. Design of potent pan-IAP and Lys-covalent XIAP selective inhibitors using a thermodynamics driven approach. *J. Med. Chem.* **61**, 6350–6363 (2018).
8. Bum-Erdene, K. et al. Small-molecule covalent bond formation at tyrosine creates a binding site and inhibits activation of Ras GTPases. *Proc. Natl Acad. Sci. USA* **117**, 7131–7139 (2020).
9. Teng, M. et al. Rationally designed covalent BCL6 inhibitor that targets a tyrosine residue in the homodimer interface. *ACS Med. Chem. Lett.* **11**, 1269–1273 (2020).
10. Hatcher, J. M. et al. SRPKIN-1: a covalent SRPK1/2 inhibitor that potently converts VEGF from pro-angiogenic to anti-angiogenic isoform. *Cell Chem. Biol.* **25**, 460–470 (2018).
11. Mukherjee, H. et al. Discovery and optimization of covalent Bcl-xL antagonists. *Bioorg. Med. Chem. Lett.* **29**, 126682 (2019).
12. Glukhova, A. et al. Structure of the adenosine A(1) receptor reveals the basis for subtype selectivity. *Cell* **168**, 867–877 (2017).
13. Backus, K. M. et al. Proteome-wide covalent ligand discovery in native biological systems. *Nature* **534**, 570–574 (2016).
14. Brulet, J. W., Borne, A. L., Yuan, K., Libby, A. H. & Hsu, K.-L. Liganding functional tyrosine sites on proteins using sulfur-triazole exchange chemistry. *J. Am. Chem. Soc.* **142**, 8270–8280 (2020).
15. Abbasov, M. E. et al. A proteome-wide atlas of lysine-reactive chemistry. *Nat. Chem.* **13**, 1081–1092 (2021).
16. Hacker, S. M. et al. Global profiling of lysine reactivity and ligandability in the human proteome. *Nat. Chem.* **9**, 1181–1190 (2017).

17. Kuljanin, M. et al. Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nat. Biotechnol.* **39**, 630–641 (2021).
18. McCloud, R. L. et al. Direct target site identification of a sulfonyl-triazole covalent kinase probe by LC-MS chemical proteomics. *Anal. Chem.* **93**, 11946–11955 (2021).
19. Wang, Y. et al. Expedited mapping of the ligandable proteome using fully functionalized enantiomeric probe pairs. *Nat. Chem.* **11**, 1113–1123 (2019).
20. Vinogradova, E. V. et al. An activity-guided map of electrophile-cysteine interactions in primary human T cells. *Cell* **182**, 1009–1026 (2020).
21. Tao, Y. et al. Targeted protein degradation by electrophilic PROTACs that stereoselectively and site-specifically engage DCAF1. *J. Am. Chem. Soc.* **144**, 18688–18699 (2022).
22. Szychowski, J. et al. Cleavable biotin probes for labeling of biomolecules via azide-alkyne cycloaddition. *J. Am. Chem. Soc.* **132**, 18351–18360 (2010).
23. Rabalski, A. J., Bogdan, A. R. & Baranczak, A. Evaluation of chemically-cleavable linkers for quantitative mapping of small molecule-cysteine reactivity. *ACS Chem. Biol.* **14**, 1940–1950 (2019).
24. Li, Z., Liu, K., Xu, P. & Yang, J. Benchmarking cleavable biotin tags for peptide-centric chemoproteomics. *J. Proteome Res.* **21**, 1349–1358 (2022).
25. Erickson, B. K. et al. Active instrument engagement combined with a real-time database search for improved performance of sample multiplexing workflows. *J. Proteome Res.* **18**, 1299–1306 (2019).
26. Schweppe, D. K. et al. Full-featured, real-time database searching platform enables fast and accurate multiplexed quantitative proteomics. *J. Proteome Res.* **19**, 2026–2034 (2020).
27. Cruite, J. T. et al. Cereblon covalent modulation through structure-based design of histidine targeting chemical probes. *RSC Chem. Biol.* **3**, 1105–1110 (2022).
28. Tam, J. et al. An activity-based probe for high-throughput measurements of triacylglycerol lipases. *Anal. Biochem.* **414**, 254–260 (2011).
29. Gambini, L. et al. Covalent inhibitors of protein-protein interactions targeting lysine, tyrosine or histidine residues. *J. Med. Chem.* **62**, 5616–5627 (2019).
30. Narayanan, A. & Jones, L. H. Sulfonyl fluorides as privileged warheads in chemical biology. *Chem. Sci.* **6**, 2650–2659 (2015).
31. Zanon, P. R. A. et al. Profiling the proteome-wide selectivity of diverse electrophiles. Preprint at <https://chemrxiv.org/engage/chemrxiv/article-details/60c755f2bb8c1a7d393dc505> (2021).
32. Gu, C. et al. Chemical proteomics with sulfonyl fluoride probes reveals selective labeling of functional tyrosines in glutathione transferases. *Chem. Biol.* **20**, 541–548 (2013).
33. Mukherjee, H. et al. A study of the reactivity of S((VI))-F containing warheads with nucleophilic amino-acid side chains under physiological conditions. *Org. Biomol. Chem.* **15**, 9685–9695 (2017).
34. Brighty, G. J. et al. Using sulfuramidimidoyl fluorides that undergo sulfur(VI) fluoride exchange for inverse drug discovery. *Nat. Chem.* **12**, 906–913 (2020).
35. Stanevich, V. et al. The structural basis for tight control of PP2A methylation and function by LCMT-1. *Mol. Cell* **41**, 331–342 (2011).
36. Hwang, J., Lee, J. A. & Pallas, D. C. Leucine carboxyl methyltransferase 1 (LCMT-1) methylates protein phosphatase 4 (PP4) and protein phosphatase 6 (PP6) and differentially regulates the stable formation of different PP4 holoenzymes. *J. Biol. Chem.* **291**, 21008–21019 (2016).
37. De Baere, I. et al. Purification of porcine brain protein phosphatase 2A leucine carboxyl methyltransferase and cloning of the human homologue. *Biochemistry* **38**, 16539–16547 (1999).
38. Reckel, S. et al. Structural and functional dissection of the DH and PH domains of oncogenic Bcr-Abl tyrosine kinase. *Nat. Commun.* **8**, 2101 (2017).
39. Lee, J. H. & Skalniak, D. G. Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. *Mol. Cell. Biol.* **28**, 609–618 (2008).
40. Wu, M. et al. Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Mol. Cell. Biol.* **28**, 7337–7344 (2008).
41. Thomas, L. R. et al. Interaction with WDR5 promotes target gene recognition and tumorigenesis by MYC. *Mol. Cell* **58**, 440–452 (2015).
42. Chacón Simon, S. et al. Discovery of WD repeat-containing protein 5 (WDR5)-MYC inhibitors using fragment-based methods and structure-based design. *J. Med. Chem.* **63**, 4315–4333 (2020).
43. Tian, J. et al. Discovery and structure-based optimization of potent and selective WD repeat domain 5 (WDR5) inhibitors containing a dihydroisoquinolinone bicyclic core. *J. Med. Chem.* **63**, 656–675 (2020).
44. Ding, J. et al. Discovery of potent small-molecule inhibitors of WDR5-MYC interaction. *ACS Chem. Biol.* **18**, 34–40 (2023).
45. Kamber, R. A. et al. Inter-cellular CRISPR screens reveal regulators of cancer cell phagocytosis. *Nature* **597**, 549–554 (2021).
46. Hahm, H. S. et al. Global targeting of functional tyrosines using sulfur-triazole exchange chemistry. *Nat. Chem. Biol.* **16**, 150–159 (2020).
47. Chen, W. et al. Arylfluorosulfates inactivate intracellular lipid binding protein(s) through chemoselective SuFEx reaction with a binding site Tyr residue. *J. Am. Chem. Soc.* **138**, 7353–7364 (2016).
48. Mortenson, D. E. et al. Inverse drug discovery strategy to identify proteins that are targeted by latent electrophiles as exemplified by aryl fluorosulfates. *J. Am. Chem. Soc.* **140**, 200–210 (2018).
49. Wang, N. et al. Genetically encoding fluorosulfate-L-tyrosine to react with lysine, histidine and tyrosine via SuFEx in proteins in vivo. *J. Am. Chem. Soc.* **140**, 4995–4999 (2018).
50. Hett, E. C. et al. Rational targeting of active-site tyrosine residues using sulfonyl fluoride probes. *ACS Chem. Biol.* **10**, 1094–1098 (2015).
51. Cuesta, A. & Taunton, J. Lysine-targeted inhibitors and chemoproteomic probes. *Annu. Rev. Biochem.* **88**, 365–381 (2019).
52. Yang, T. et al. Reversible lysine-targeted probes reveal residence time-based kinase selectivity. *Nat. Chem. Biol.* **18**, 934–941 (2022).
53. Metcalf, B. et al. Discovery of GBT440, an orally bioavailable R-state stabilizer of sickle cell hemoglobin. *ACS Med. Chem. Lett.* **8**, 321–326 (2017).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023

## Methods

A detailed Methods section is provided in the Supplementary Information.

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this Article.

## Data availability

Uncropped gel images are provided as Source Data files accompanying this paper. The reported crystal structure has been deposited to the Protein Data Bank (PDB) under accession no. [8F93](#). All proteomic raw data have been deposited to MassIVE (<http://massive.ucsd.edu>) with accession no. [MSV000090778](#), as well as in ProteomeXchange (<http://www.proteomexchange.org>) with accession no. [PXD042307](#). Source data are provided with this paper.

## Code availability

The script used for identifying modified sites and proximal ligands in the PDB has been deposited to GitHub (<https://github.com/aacuesta/Ying2023NatChem>).

## Acknowledgements

Funding for this study was provided by the Ono Pharma Foundation (J.T.), Pfizer and the Tobacco-Related Disease Research Program Postdoctoral Fellowship Award (T32FT4880 to G.B.C.). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank B. Küster for guidance on MS data acquisition and analysis.

## Author contributions

Y.C. and J.T. conceived the project, designed the experiments and analysed the data. S.Z. and Y.S.M. synthesized the SFs. Y.C. performed

the chemoproteomic and validation experiments. G.B.C. solved the WDR5 X-ray structure and performed biochemical experiments with recombinant APMAP. A.C. wrote the script for identifying modified sites in the PDB. R.A.K. performed the phagocytosis assays, with input from M.C.B. Y.C. and J.T. wrote the manuscript, with input from all of the authors.

## Competing interests

J.T. is a cofounder of Kezar Life Sciences and Terremoto Biosciences, and a scientific advisor to Entos. M.C.B. and R.A.K. have outside interests in DEM Biopharma. S.Z. and Y.S.M. are employees of Enamine Ltd and Chemspace LLC, respectively. The remaining authors declare no competing interests.

## Additional information

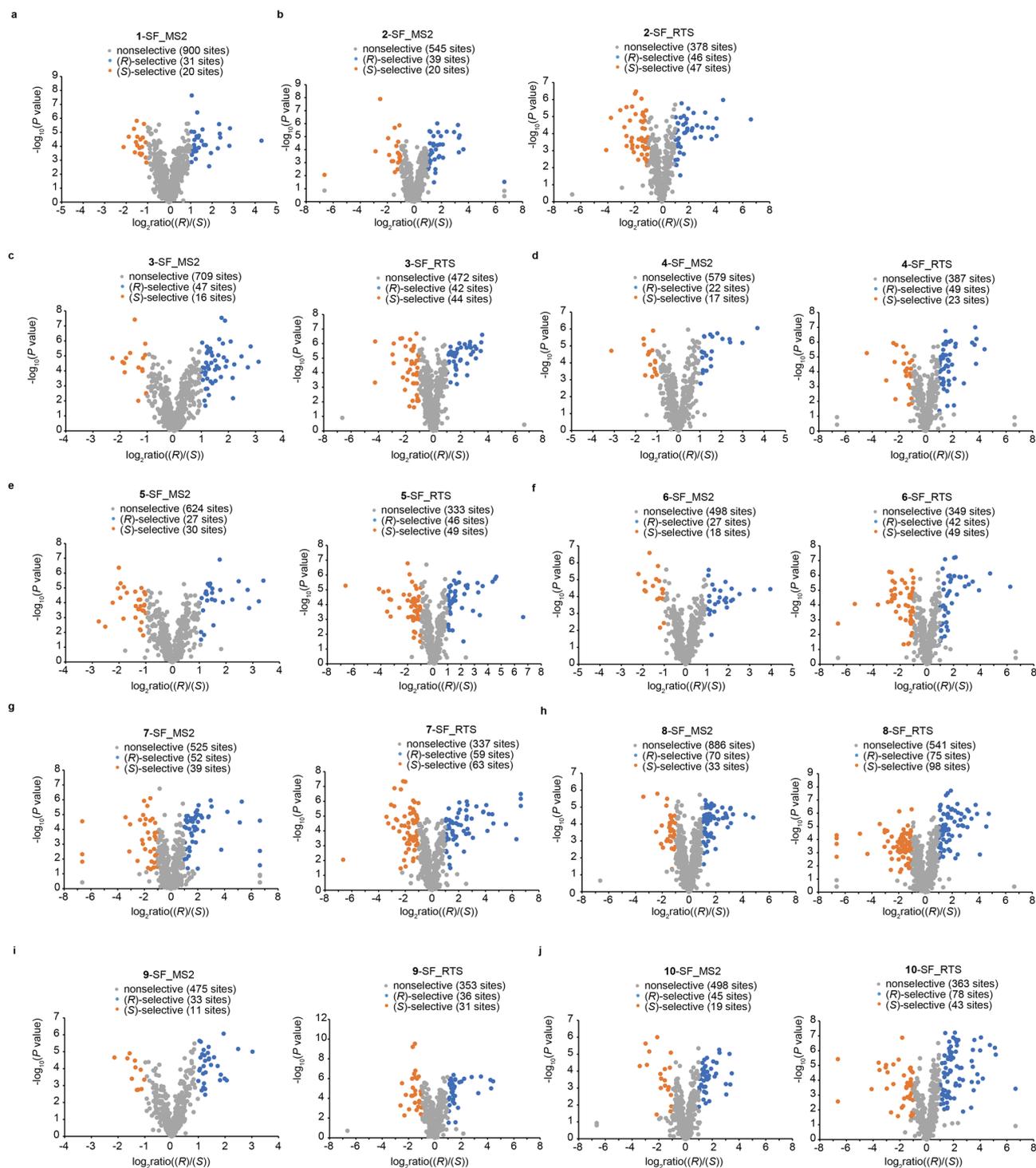
**Extended data** is available for this paper at <https://doi.org/10.1038/s41557-023-01281-3>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41557-023-01281-3>.

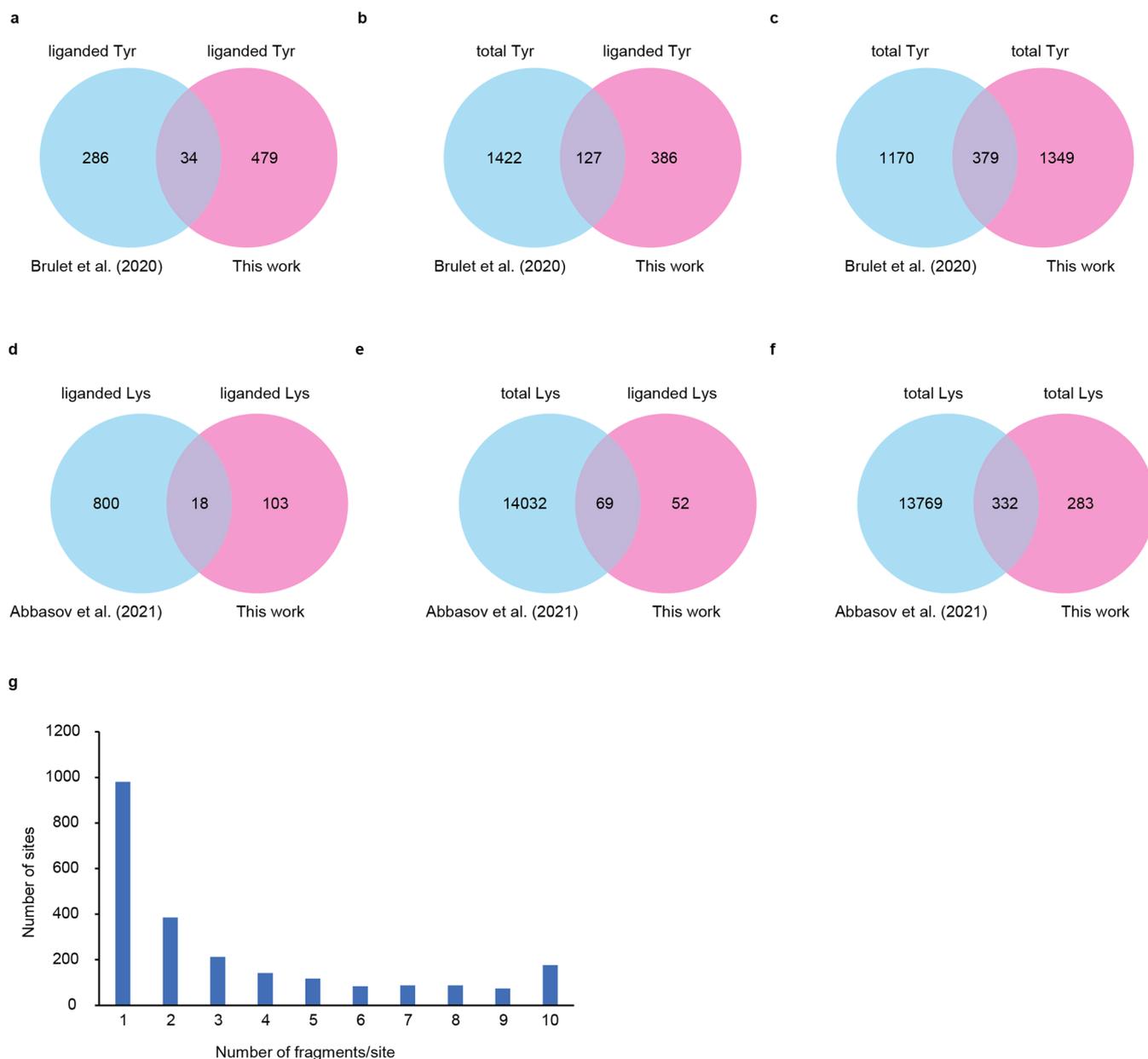
**Correspondence and requests for materials** should be addressed to Jack Taunton.

**Peer review information** *Nature Chemistry* thanks Benjamin Cravatt and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

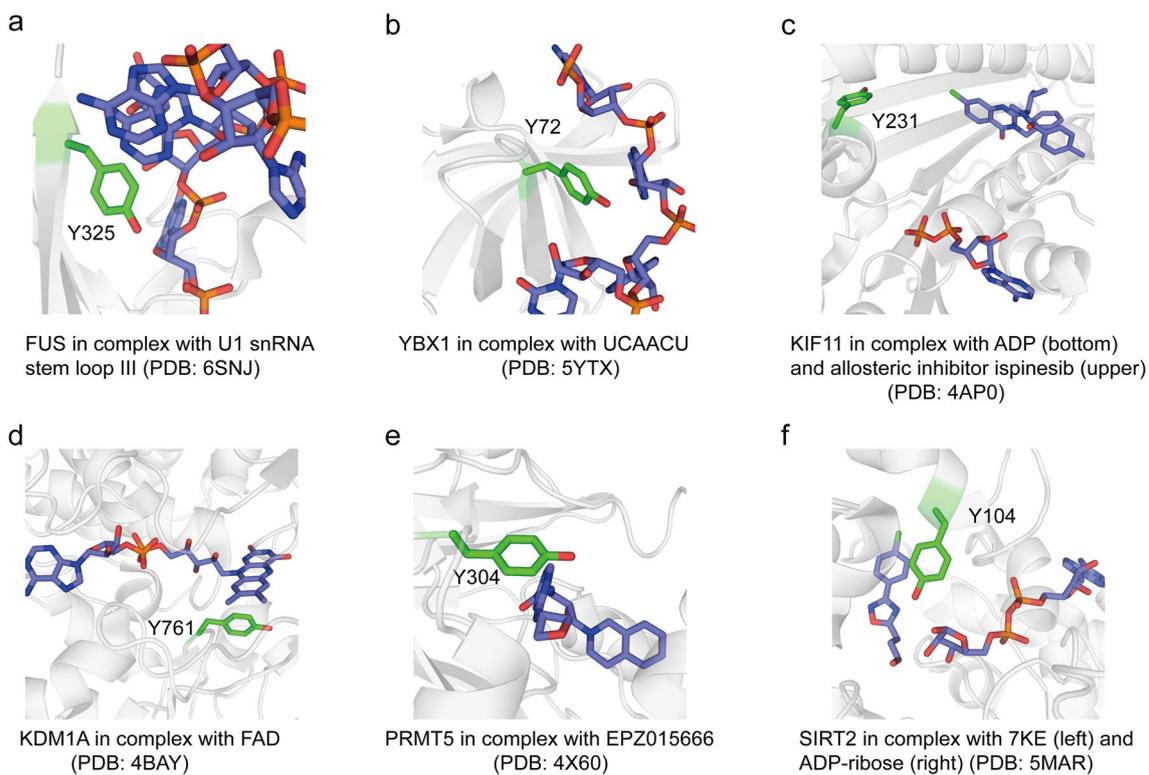


**Extended Data Fig. 1 | Volcano plots showing modified sites (MS2 and RTS-MS3 data) for probes 1–10. a–j,**  $\log_2\text{ratio}((R)/(S))$  is on the x-axis and  $-\log_{10}(P\text{ value})$  is on the y-axis. Modified sites are indicated as (R)-selective (blue), (S)-selective (orange), and nonselective (grey). *P* values were determined by Student's *t*-test (two-tailed, two-sample equal variance).

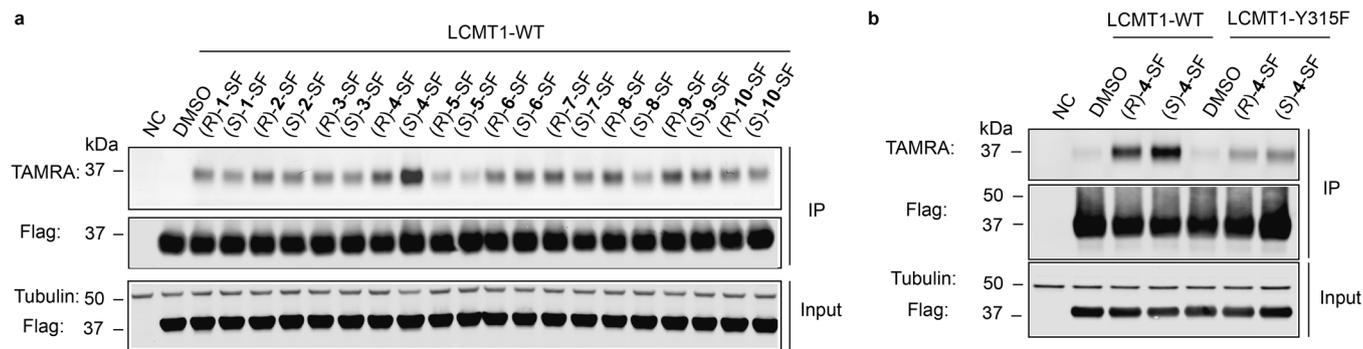


**Extended Data Fig. 2 | Analysis of tyrosines and lysines modified by chiral sulfonyl fluoride probes. a-f**, Number of liganded and total Tyr and Lys identified in this study compared with the indicated previous studies<sup>14,15</sup>.

**g**, The number of enantiomeric probe pairs (based on fragments 1-10) per site (x-axis) was plotted against the total number of sites modified by the indicated number of enantiomeric probe pairs (y-axis).

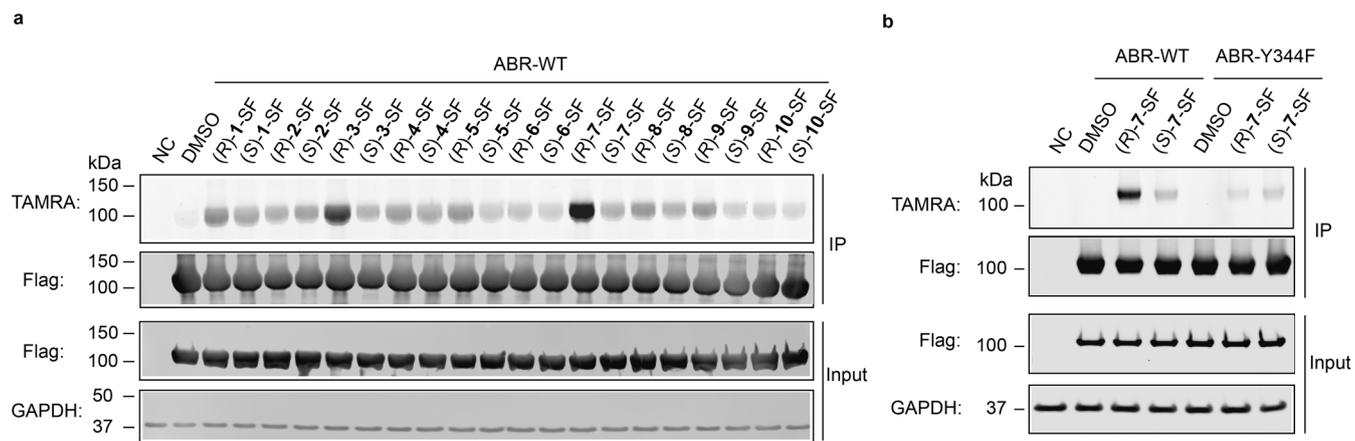


**Extended Data Fig. 3 | Functionally diverse protein sites targeted by chiral sulfonyl fluoride probes. a-f,** Structures of FUS, YBX1, KIF11, KDM1A, PRMT5, and SIRT2, highlighting the probe-modified Tyr proximal to bound ligands.



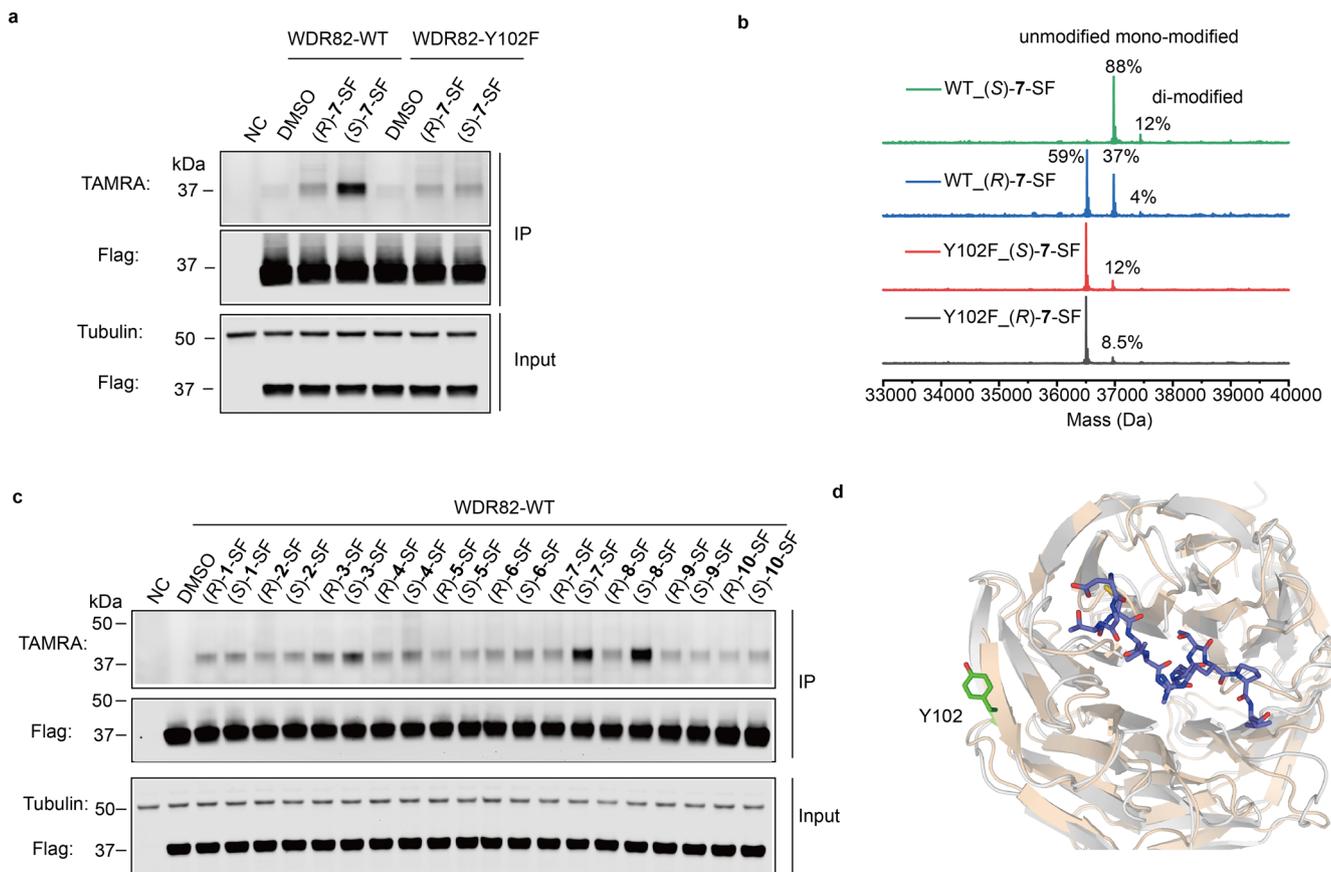
**Extended Data Fig. 4 | Probe modification of LCMT1. a**, HEK293T cells were transfected with Flag-LCMT1 and treated with SF probes (10  $\mu$ M, 1 h). After cell lysis, anti-Flag enrichment, 3xFlag peptide elution and click reaction with TAMRA-azide, samples were analyzed by in-gel fluorescence and western blotting. NC indicates nontransfected cells. Data are representative of two

independent experiments. **b**, HEK293T cells were transfected with WT or Y315F Flag-LCMT1 and treated with 4-SF probes (10  $\mu$ M, 1 h). The cells were lysed and processed as described above. Data are representative of two independent experiments.



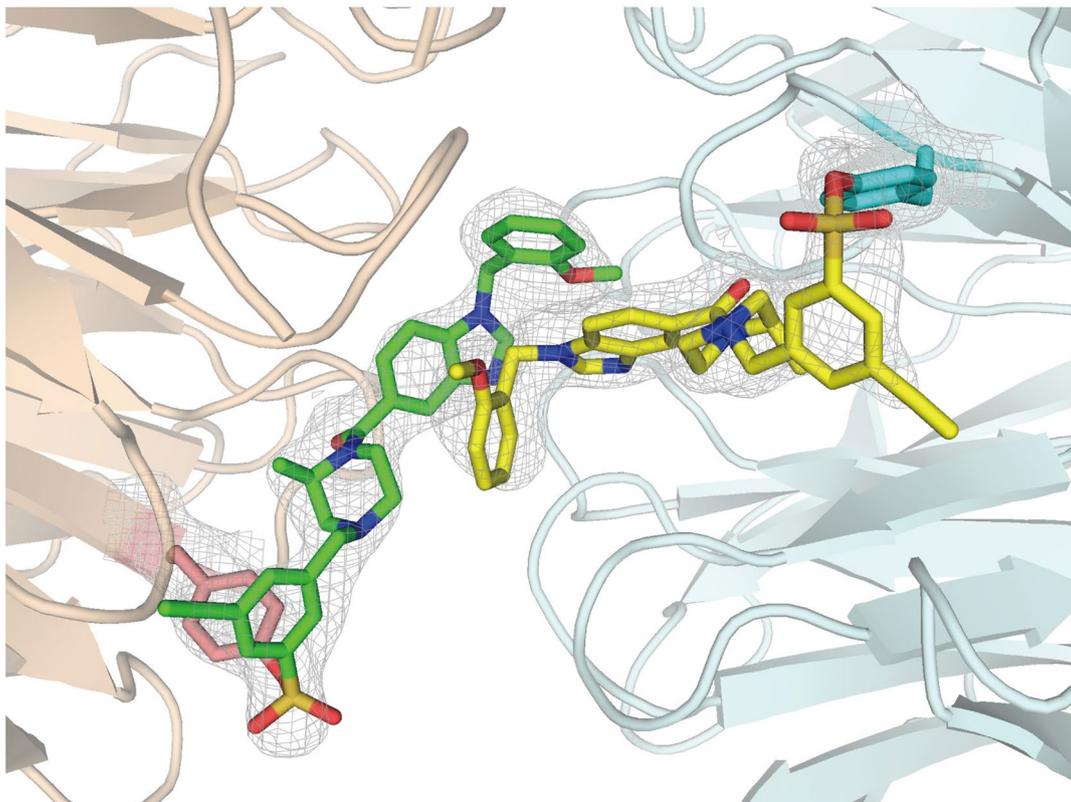
**Extended Data Fig. 5 | Probe modification of ABR.** **a**, HEK293T cells were transfected with Flag-ABR and treated with SF probes (10  $\mu$ M, 1 h). After cell lysis, anti-Flag enrichment, 3xFlag peptide elution and click reaction with TAMRA-azide, samples were analyzed by in-gel fluorescence and western blotting. NC

indicates nontransfected cells. Data are representative of two independent experiments. **b**, HEK293T cells were transfected with WT or Y344F Flag-ABR and treated with 7-SF probes (10  $\mu$ M, 1 h). The cells were lysed and processed as described above. Data are representative of two independent experiments.



**Extended Data Fig. 6 | Probe modification of WDR82.** **a**, HEK293T cells were transfected with WT or Y102F Flag-WDR82 and treated with 7-SF probes (50  $\mu$ M, 1 h). After cell lysis, anti-Flag enrichment, 3xFlag peptide elution and click reaction with TAMRA-azide, samples were analyzed by in-gel fluorescence and western blotting. Data are representative of two independent experiments. NC indicates nontransfected cells. **b**, Recombinant WT and Y102F WDR82 (0.44  $\mu$ M) were incubated with 7-SF probes (40  $\mu$ M, 37  $^{\circ}$ C, overnight) and analyzed by

intact-protein mass spectrometry. Data are representative of two independent experiments. **c**, HEK293T cells were transiently transfected with Flag-WDR82 and treated with the indicated SF probes (50  $\mu$ M, 1 h). The cells were lysed and processed as described above. Data are representative of two independent experiments. **d**, Overlay of the AlphaFold structure of WDR82 with the crystal structure of WDR5 bound to KANSL1 peptide (PDB: 4CY2).

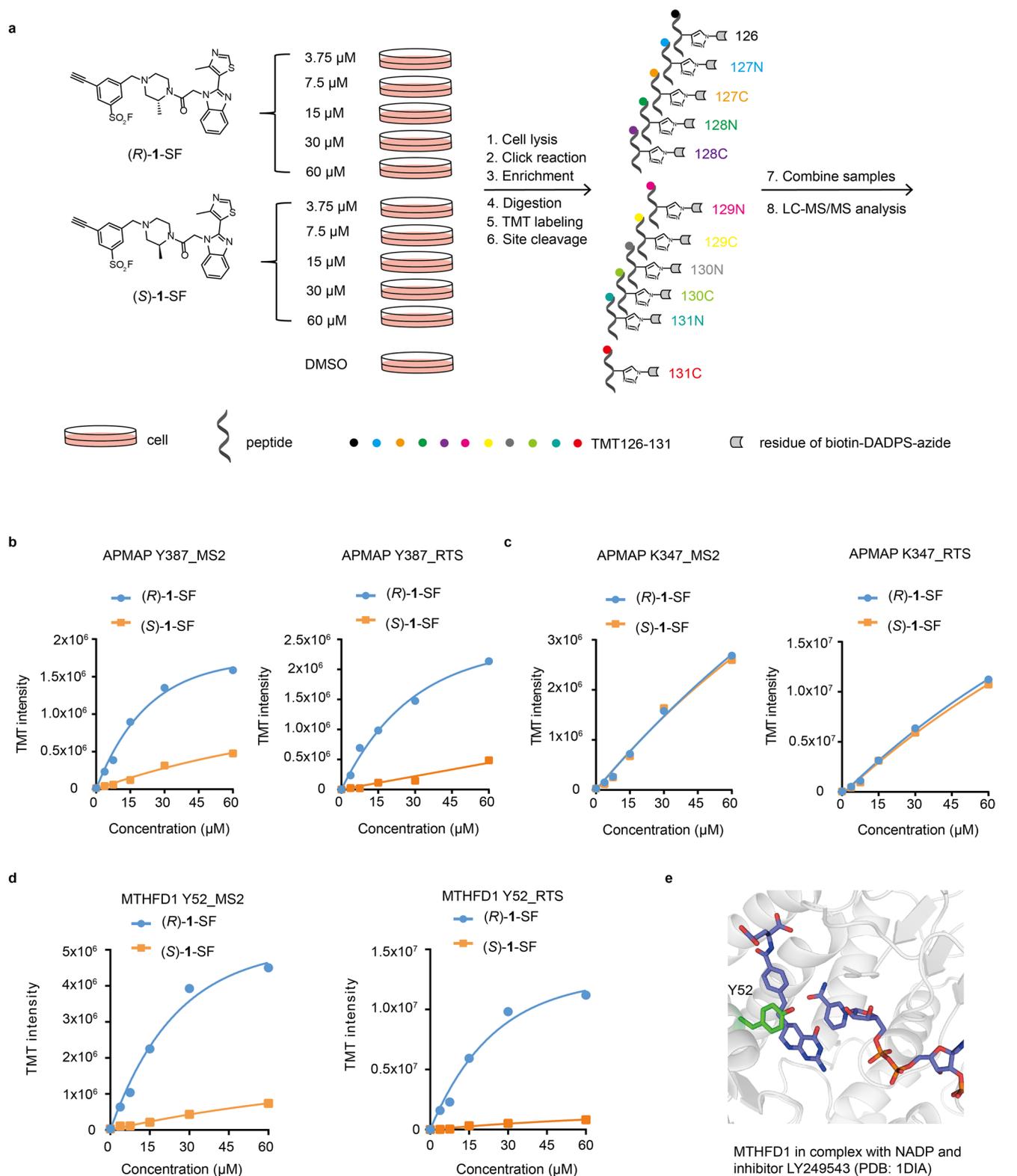


**Extended Data Fig. 7 | Electron density map of (R)-2-SF.** Electron density (2Fo-Fc) map of (R)-2-SF is shown at a contour level of 1 $\sigma$ .

WDR5-(R)-2-SF complex	
PDB Code	8F93
Data collection	
Space group	C121
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	92.78, 103.02, 81.01
$\alpha$ , $\beta$ , $\gamma$ (°)	90.0, 101.5, 90
Resolution (Å)	32.26 -2.30 (2.38 -2.30) *
<i>R</i> merge	0.077 (0.788)
<i>I</i> / $\sigma$ <i>I</i>	9.20 (1.61)
Completeness (%)	99.0 (99.4)
Redundancy	3.3 (3.3)
Refinement	
Resolution (Å)	32.26 -2.30
No. reflections	32853
<i>R</i> work / <i>R</i> free	0.2029 / 0.2527
No. atoms	
Protein	4761
Ligand/ion	122
Water	97
<i>B</i> -factors (Å <sup>2</sup> )	
Protein	54.58
Ligand/ion	61.05
Water	51.21
R.m.s. deviations	
Bond lengths (Å)	0.012
Bond angles (°)	1.25

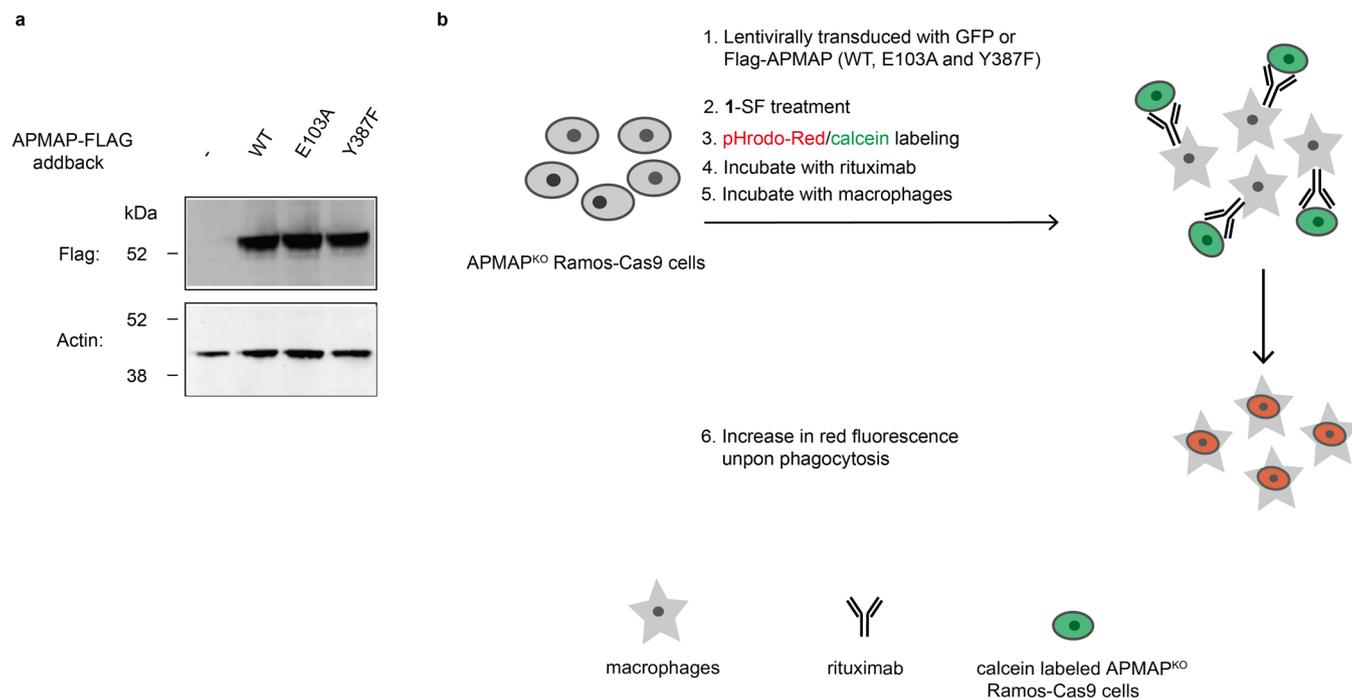
\*Highest-resolution shell is shown in parentheses.

**Extended Data Fig. 8 | Data collection and refinement statistics.** Data collection and refinement statistics (molecular replacement) for WDR5-(R)-2-SF complex.



**Extended Data Fig. 9 | Chemoproteomic analysis of 1-SF probes in dose-response mode.** **a**, HEK293T cells that stably overexpress Flag-APMAP were treated separately with 0, 3.75, 7.5, 15, 30 and 60  $\mu\text{M}$  of (*R*)-1-SF and (*S*)-1-SF for 1.5 h at 37  $^{\circ}\text{C}$ . Cell lysates were conjugated with biotin-DADPS-azide via click chemistry. After enrichment and trypsin digestion, probe-modified peptides from each sample were separately labeled with TMT 11-plex reagents.

Probe-modified peptides from each sample were eluted with 2% formic acid, combined, and analyzed by LC-MS/MS. **b-d**, TMT reporter ion intensities were plotted as a function of probe concentration for peptides quantified by MS2 and RTS: APMAP Y387 (**b**), APMAP K347 (**c**), and MTHFD1 Y52 (**d**). **e**, Crystal structure (PDB: 1DIA) depicting MTHFD1 Y52 proximal to the inhibitor LY249543.



**Extended Data Fig. 10 | Antibody-dependent phagocytosis assay. a,** APMAP<sup>KO</sup> Ramos-Cas9 cells (rescued with Flag-APMAP WT, E103A and Y387F) were resuspended in SDS loading buffer containing 100 mM DTT, sonicated and

heated (70 °C, 10 min). Samples were analyzed by western blotting. Data are representative of two independent experiments. **b,** Scheme of phagocytosis assay.

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

TAMRA fluorescence data were collected on Typhoon Imaging System (Molecular Dynamics). Western blots were scanned on an Odyssey infrared imager (LI-COR Biosciences). The LC-MS/MS data were collected on a Orbitrap Eclipse Tribrid Mass Spectrometer (Thermo) connected to an Ultimate 3000 RSLnano system. X-ray diffraction data were collected at beamline 2.0.1 of the Advanced Light Source (Lawrence Berkeley National Laboratory). NMR data were collected on a Bruker 400 MHz or Varian 500 MHz spectrometer. HRMS characterization data were obtained on a Waters LC system coupled with Xevo G2-XS Q-ToF Mass Spectrometer.

Data analysis

All in-gel fluorescence images were processed by Windows ImageJ software V.1.48 (National Institutes of Health), and contrast was adjusted appropriately. Western blots were analyzed with Image Studio Lite software V.5.2 (LI-COR Biosciences). Data displayed in Fig. 3a, 5g, 6b, 6d, 6f, Extended Data Fig. 9b, 9c, 9d were generated and analyzed using Graphpad Prism V7.00. EC50 in Fig. 6b was fit by GraphPad Prism V7.00. The LC-MS/MS data were searched using MaxQuant (V.1.6.7.0). The crystal structure was determined using iMosflm V.7.4.0, Scala V.3.3.18, Phaser V.2.7.0, Coot V.0.9.8.5 and PHENIX V.1.20.1. Protein structures were depicted using Pymol V2.6.0a0. (Schrödinger). Intact-protein MS data were depicted using Origin 2019 (9.60). NMR data were analyzed using MestReNova V.6.0.2. Phagocytosis was monitored using an Incucyte S3 system (Sartorius) with the 2021B software update.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Uncropped gel images were provided as Source Data files accompanying this paper. The reported crystal structure has been deposited to the Protein Data Bank (PDB) with the accession number of 8F93. All proteomic raw data have been deposited to MassIVE (<http://massive.ucsd.edu>) with the accession number of MSV000090778, as well as in ProteomeXchange (<http://www.proteomexchange.org>) with accession number PXD042307. The script used for identifying modified sites and proximal ligands in the PDB has been deposited to Github (<https://github.com/aacuesta/Ying2023NatChem>).

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was determined based on our experimental observations and experience, in which a particular sample size was found to provide reliable and reproducible results.
Data exclusions	No data were excluded.
Replication	In-gel fluorescence, western blotting, intact-protein MS, and phagocytosis experiments were performed independently at least twice. All data showing error bars were performed in triplicate. Information on the number of replicates performed for each measurement is provided in the figure legends. Three biological replicates were used for each condition in the TMT 6-plex chemoproteomic experiments. For the TMT 6-plex and TMT 11-plex chemoproteomic experiments, the final multiplexed samples were analyzed by two MS methods (MS2 and RTS TMT quantification) with two injections each (technical replicates).
Randomization	N/A. All experiments used identically grown cells or purified proteins and were therefore not randomized.
Blinding	Blinding is not relevant in this study, as cell lines used in experiments were grown under identical conditions, samples were treated uniformly, and the same data analysis procedure was applied to all samples of the same experiment.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

Antibodies used

Anti-Flag, Cell Signaling Technology, 14793S; anti-beta-actin, Cell Signaling Technology, 3700; anti-GAPDH, Santa Cruz Biotechnology, sc-32233; anti-beta-tubulin, Cell Signaling Technology, 2146S; IRDye 800CW goat anti-mouse IgG, LI-COR, 926-32210; IRDye 800CW goat anti-rabbit IgG, LI-COR, 926-32211; anti-WDR5, Cell Signaling Technology, 13105S; anti-Flag, clone M2, Sigma, F3165; anti-actin,

Abcam, ab8227; IRDye 680LT donkey anti-mouse IgG, Fisher, NC9030091; IRDye 800CW goat anti-rabbit IgG, Fisher, NC9401841; anti-CD20, Invivogen, hcd20-mab1.

## Validation

The antibodies used in this study have been validated by the manufacturers (see website links below).

Anti-Flag, Cell Signaling Technology, 14793S: <https://www.cellsignal.com/products/primary-antibodies/dykdddk-tag-d6w5b-rabbit-mab-binds-to-same-epitope-as-sigma-s-anti-flag-m2-antibody/14793>

Anti-beta-actin, Cell Signaling Technology, 3700: <https://www.cellsignal.com/products/primary-antibodies/b-actin-8h10d10-mouse-mab/3700>

Anti-Flag, Sigma, F3165: <https://www.sigmaaldrich.com/US/en/product/sigma/f3165>

Anti-GAPDH, Santa Cruz Biotechnology, sc-32233: <https://www.scbt.com/p/gapdh-antibody-6c5>

Anti-beta-tubulin, Cell Signaling Technology, 2146S: <https://www.cellsignal.com/products/primary-antibodies/b-tubulin-antibody/2146>

Anti-WDR5, Cell Signaling Technology, 13105S: <https://www.cellsignal.com/products/primary-antibodies/wdr5-d9e1i-rabbit-mab/13105>

Anti-actin, Abcam, ab8227: <https://www.abcam.com/beta-actin-antibody-ab8227.html>

Anti-CD20, Invivogen, hcd20-mab1: <https://www.invivogen.com/anti-hcd20-higg1>

## Eukaryotic cell lines

### Policy information about [cell lines](#)

#### Cell line source(s)

Jurkat, HEK293T, Ramos and J774A.1 cells were purchased from American Type Culture Collection (ATCC). Sf9 insect cells were purchased from Expression Systems.

#### Authentication

No additional authentication was performed.

#### Mycoplasma contamination

All cell lines were routinely tested for mycoplasma, showing mycoplasma-negative results.

#### Commonly misidentified lines (See [ICLAC](#) register)

No commonly misidentified cell lines were used.