nature chemical biology

ARTICLES

https://doi.org/10.1038/s41589-022-01019-1

Reversible lysine-targeted probes reveal residence time-based kinase selectivity

Tangpo Yang[1], Adolfo Cuesta[1], Xiaobo Wan[1,2], Gregory B. Craven[1], Brad Hirakawa[3], Penney Khamphavong[3], Jeffrey R. May[3], John C. Kath[3], John D. Lapek Jr.[3], Sherry Niessen[3], Alma L. Burlingame[2], Jordan D. Carelli[3] and Jack Taunton[1]✉

The expansion of the target landscape of covalent inhibitors requires the engagement of nucleophiles beyond cysteine. Although the conserved catalytic lysine in protein kinases is an attractive candidate for a covalent approach, selectivity remains an obvious challenge. Moreover, few covalent inhibitors have been shown to engage the kinase catalytic lysine in animals. We hypothesized that reversible, lysine-targeted inhibitors could provide sustained kinase engagement in vivo, with selectivity driven in part by differences in residence time. By strategically linking benzaldehydes to a promiscuous kinase binding scaffold, we developed chemoproteomic probes that reversibly and covalently engage >200 protein kinases in cells and mice. Probe–kinase residence time was dramatically enhanced by a hydroxyl group *ortho* to the aldehyde. Remarkably, only a few kinases, including Aurora A, showed sustained, quasi-irreversible occupancy in vivo, the structural basis for which was revealed by X-ray crystallography. We anticipate broad application of salicylaldehyde-based probes to proteins that lack a druggable cysteine.

Targeted covalent inhibitors are an emerging class of drugs, evidenced by the recent approval of several cysteine-targeted inhibitors, including ibrutinib[1,2], osimertinib[3,4] and sotorasib[5,6]. A covalent inhibition strategy has two principal advantages: sustained target engagement in vivo, and in many cases, increased selectivity[7]. With the exception of protease and lipid hydrolase inhibitors, most targeted covalent drugs have been designed to react with a solvent-exposed, noncatalytic cysteine[8,9]. However, the low prevalence of cysteine near ligand binding sites restricts the scope of therapeutic targets that could potentially benefit from this approach[10].

Lysine, which is frequently located near functional and ligand binding sites, presents an alternative nucleophile for covalent targeting. Despite the reduced chemical reactivity of lysine, recent efforts have led to the development of irreversible, lysine-targeted inhibitors of phosphoinositide 3-kinase δ (PI3Kδ)[11], X-linked inhibitor of apoptosis protein (XIAP)[12], heat-shock protein 90 (Hsp90)[13,14] and eukaryotic translation initiation factor 4E (eIF4E)[15], among other diverse targets[16,17]. An unresolved question concerns whether the electrophiles used in these irreversible inhibitors are suitable for use in animals. More importantly, reversible covalent, lysine-targeted inhibitors could potentially provide advantages over their irreversible counterparts, such as increased proteome-wide selectivity driven by distinct residence times[18,19]. To the best of our knowledge, the concept of residence time-based selectivity has not been explored with reversible, lysine-targeted inhibitors.

Aldehydes are versatile electrophiles that can react rapidly and reversibly with biological amines, including the ε-amine of lysine, to form imines under physiological conditions. The enzyme cofactor, pyridoxal phosphate (PLP), forms an imine with an active-site lysine in PLP-dependent enzymes and mediates diverse chemical transformations of biological amines[20]. Voxelotor, a covalent drug that was recently approved for sickle cell disease, leverages a salicylaldehyde to reversibly target the amino-terminal amine of hemoglobin[21]. These examples suggest that aryl aldehydes may have the requisite balance of stability, reactivity and reversibility for in vivo applications[22].

Protein kinases represent a large class of validated therapeutic targets. Several kinases have been targeted by covalent drugs and chemical probes, including irreversible[11,23–25] and reversible[26,27] inhibitors that covalently modify the catalytic lysine[10,16]. However, to the best of our knowledge, no covalent protein kinase inhibitors have been shown to reversibly engage the catalytic lysine in humans or animal models. We envisioned that aldehyde-based kinase inhibitors could fill this gap. More importantly, we hypothesized that reversible imine formation—despite targeting the conserved catalytic lysine—could form the basis of residence time-based selectivity, in which distinct imine hydrolysis (and reformation) kinetics could lead to differential kinase engagement upon removal of excess unbound inhibitor (for example, by metabolic clearance).

Here, we describe benzaldehyde-based chemical probes that efficiently and reversibly engage the conserved catalytic lysine of protein kinases in animals. Using quantitative chemoproteomic methods, we estimated the residence times of clickable salicylaldehyde probes for endogenous kinases in cultured human cells, as well as in treated mice. These target engagement studies reveal a surprising degree of selectivity, driven by distinct probe–kinase residence times. Finally, we used a salicylaldehyde probe as a competitive occupancy probe in mice to study in vivo kinase engagement by PF-06873600, a recently developed cyclin-dependent kinase inhibitor in clinical trials for metastatic cancer[28,29].

Results

Design and characterization of aldehyde-based kinase probes. To test the feasibility of covalently targeting kinases with a benzaldehyde, we designed probe **1** based on the clickable sulfonyl fluoride XO44, a broad-spectrum kinase probe developed previously in our lab (Fig. 1a)[25]. Like XO44, probe **1** utilizes an aminopyrazole to bind

[1]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, CA, USA. [2]Department of Pharmaceutical Chemistry, University of California, San Francisco, CA, USA. [3]Pfizer Global Research and Development La Jolla, San Diego, CA, USA. ✉e-mail: jack.taunton@ucsf.edu

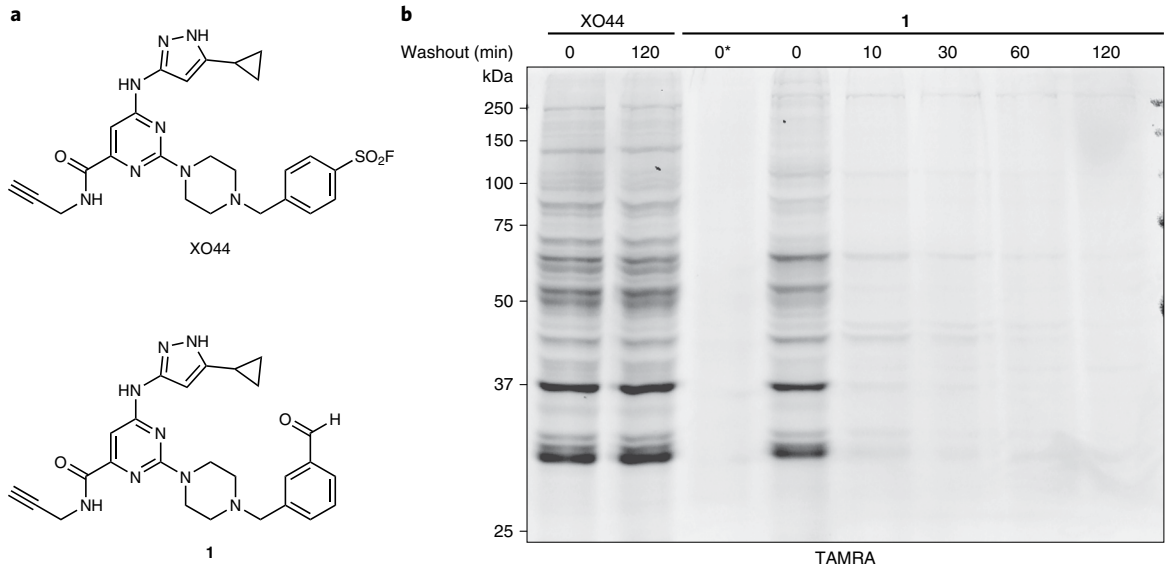


Fig. 1 | Benzaldehyde probe 1 reversibly labels cellular proteins. a, Chemical structures of XO44 and probe **1**. **b**, Jurkat cells were treated with XO44 or **1** (2 μM, 30 min), followed by compound washout for the indicated times. Cells treated with **1** were lysed in the presence of 25 mM sodium cyanoborohydride, except as indicated (*). After copper-promoted click conjugation with TAMRA azide, samples were analyzed by in-gel fluorescence and Coomassie blue staining (Supplementary Fig. 1). Data are representative of two independent experiments.

the ATP-site hinge region, as well as a piperazine linker to orient the benzaldehyde toward the catalytic lysine. We first compared probe **1** with XO44, which was previously shown to covalently modify >100 cellular kinases. Jurkat cells were treated briefly with probe **1** (2 μM, 30 min) and then lysed in buffer containing sodium cyanoborohydride to irreversibly trap imine adducts with cellular proteins. Following copper-catalyzed click conjugation of the fluorophore tetramethylrhodamine (TAMRA) azide, probe-modified proteins were resolved by SDS–polyacrylamide gel electrophoresis (SDS–PAGE) and visualized by in-gel fluorescence scanning. This analysis revealed multiple fluorescent protein bands with a profile similar to that of XO44-treated cells (Fig. 1b). Detection of TAMRA-labeled proteins was abolished when sodium cyanoborohydride was omitted from the lysis buffer, consistent with formation of fully reversible imine adducts by probe **1**. Reversibility was further confirmed by a washout experiment, in which fluorescent labeling intensity decreased rapidly in cells treated with probe **1** (dissociation half-time, $t_{1/2} \ll 10$ min), but not in cells treated with the irreversible probe XO44. The *para*-substituted benzaldehyde **2** (regioisomer of **1**) behaved similarly, albeit with decreased overall labeling intensity (Extended Data Fig. 1), suggesting that the orientation of the aldehyde or imine affects the rate of imine bond formation and/or hydrolysis.

Salicylaldehydes can form imines with enhanced hydrolytic stability conferred by an intramolecular hydrogen bond[30,31]. Given the rapid loss of probe **1** labeling upon washout, we hypothesized that salicylaldehyde-based probes **3** (also referred to as YTP-2137) and **4** (Fig. 2a) would exhibit longer residence times. Like probe **1**, salicylaldehydes **3** and **4** labeled several cellular proteins in Jurkat cells, as revealed by TAMRA azide click conjugation in lysates treated with sodium borohydride (Fig. 2b,c; sodium borohydride was superior to cyanoborohydride for trapping salicylaldehyde–protein adducts). In contrast to the rapid decrease in labeling observed after the washout of probe **1**, covalent protein modification by probes **3** and **4** persisted for at least 30 min after washout. At longer time points, reduced labeling was observed. However, certain fluorescent protein bands were detected even after a 2 h washout, suggesting that salicylaldehyde probes **3** and **4** form relatively stable imines with a subset of cellular proteins (Fig. 2b,c).

Cellular target engagement and residence times. To identify proteins targeted by probes **1**, **3** and **4**—and more ambitiously, to estimate probe–target residence times in cells—we turned to affinity enrichment and mass spectrometry. Jurkat cells were treated with dimethylsulfoxide (DMSO) or the aldehyde probes (2 μM, 30 min) in biological triplicate, washed with PBS and then incubated with compound-free media. Cells treated with probes **3** and **4** were collected at 0 h, 1 h, 3 h and 6 h after washout and processed as described above, except that a biotin picolyl azide reagent (Supplementary Note) was used instead of TAMRA azide, to provide a total of 33 affinity-enriched samples for chemoproteomic analysis (DMSO-treated cells were collected only at $t = 0$, and probe **1**-treated cells were collected only at $t = 0$ and 1 h after washout). After click conjugation to biotin picolyl azide, probe-modified proteins were enriched with NeutrAvidin agarose beads, followed by on-bead trypsinization and analysis of the tryptic peptide fraction by high-resolution liquid chromatography–tandem mass spectrometry (LC–MS/MS). Because the chemoproteomic workflow involves protein denaturing conditions at multiple steps after cell lysis, we assumed that specifically enriched kinases were irreversibly linked to the probes as a result of borohydride-mediated imine reduction in nondenaturing lysis buffer. Across the three aldehyde probes, a total of 159 kinases were enriched versus DMSO controls at the first time point ($t = 0$ after washout), with 95 kinases shared among the three compounds (Fig. 2d and Supplementary Datasets 1–3; 'enriched' kinase criteria: identified in three out of three biological replicates with intensity values ≥10-fold higher than those of DMSO controls). Based on label-free quantitation (LFQ), the median kinase intensity (summed over all unique peptides for each kinase) for a given probe was 4–8-fold greater than the median nonkinase intensity and 8–12-fold greater than the median kinase intensity in the DMSO control sample (Fig. 2e and Supplementary Dataset 4). Together, these data demonstrate that aldehyde probes **1**, **3** and **4** can preferentially enrich protein kinase targets from living cells. We additionally enriched probe **3** targets in a second cell line (K562) and found 179 kinases, of which 69 were uniquely identified in K562 cells (Supplementary Dataset 5). Based on the combined datasets from Jurkat and K562 cells, probe **3** reproducibly engages 204 kinases.

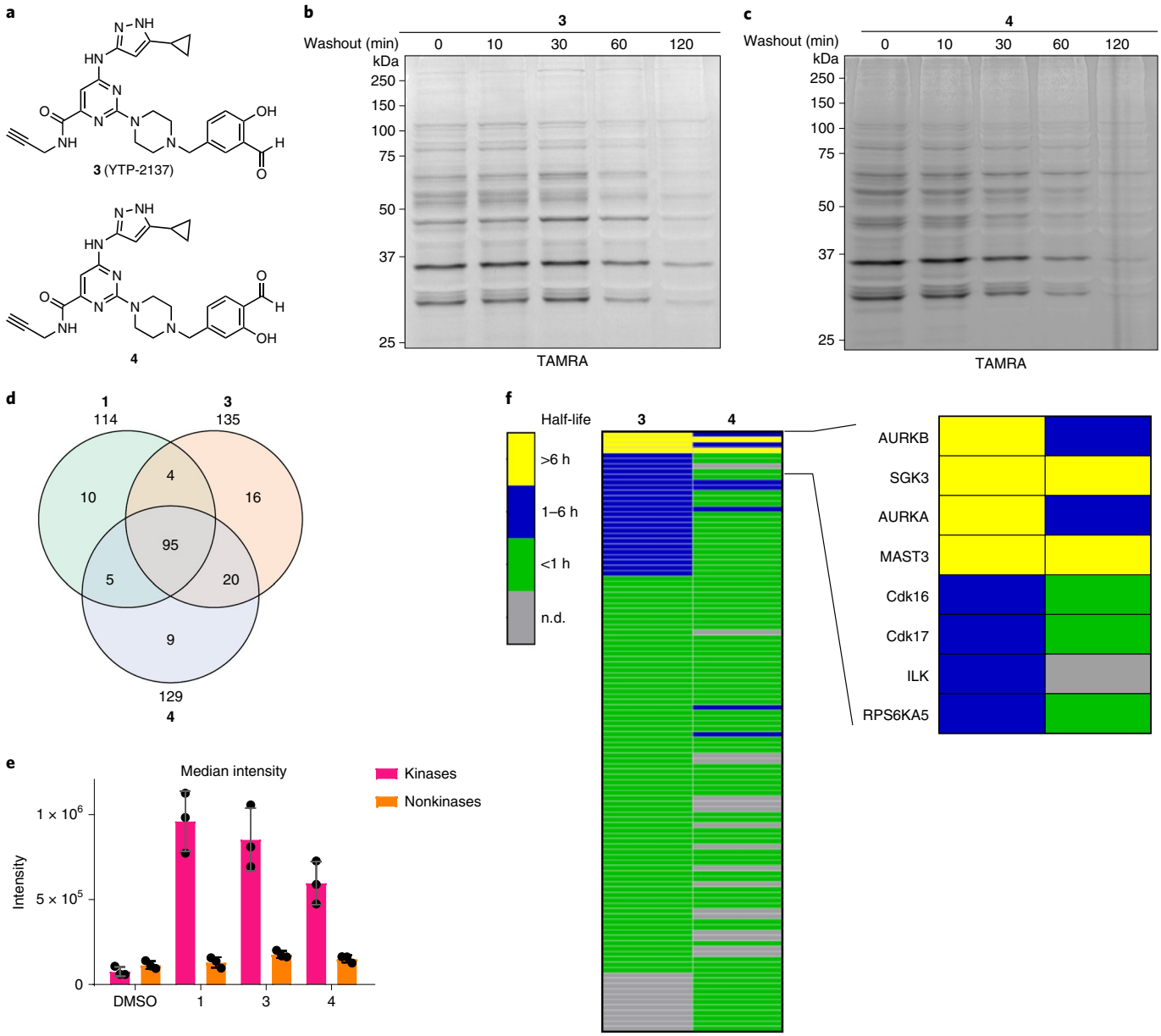


Fig. 2 | Salicylaldehyde probes exhibit prolonged kinase residence time in cells. a, Salicylaldehyde probes **3** and **4**. **b,c**, Jurkat cells were treated with **3** (**b**) or **4** (**c**) (2 μM, 30 min), followed by compound washout for the indicated times and lysis in the presence of sodium borohydride. After click conjugation with TAMRA azide, samples were analyzed by in-gel fluorescence and Coomassie blue staining (Supplementary Fig. 2). Data are representative of two independent experiments. **d**–**f**, Jurkat cells were treated with DMSO or probes **1**, **3** and **4** (2 μM, 30 min), followed by washout as described in the main text. After click conjugation with biotin azide, modified proteins were enriched with NeutrAvidin agarose and digested on-bead with trypsin. Enriched peptides were analyzed by LC–MS/MS and LFQ. **d**, Shared and unique kinases enriched by each probe. **e**, Median intensity values (LFQ) for identified kinases and nonkinases (mean ± s.d., $n = 3$). **f**, Heat map showing estimated $t_{1/2}$ values of kinases enriched from cells treated with probes **3** and **4**, followed by washout for 0 h, 1 h, 3 h and 6 h. Inset depicts kinases with prolonged engagement after washout. n.d., kinase not detected or did not meet criteria for $t_{1/2}$ determination (see main text).

We estimated cellular residence times for each kinase in Jurkat cells by comparing probe-enriched kinase intensities at each time point after washing cells into probe-free media. To increase the robustness of our baseline intensity measurements at $t = 0$, we included only those kinases identified by three or more unique peptides in all three biological replicates. Based on these criteria, 98, 101 and 96 kinases were included in our residence time analysis for probes **1**, **3** and **4**, respectively (Supplementary Datasets 6–8). All of the kinases enriched by probe **1** at $t = 0$ showed >90% reduced intensity after a 1 h washout (Supplementary Dataset 6), consistent with the results of in-gel fluorescence (Fig. 1). By contrast, the salicylaldehyde probes **3** and **4** engaged 27 and 9 kinases, respectively, with estimated $t_{1/2}$ values of >1 h (Fig. 2f, blue and yellow bars, and Supplementary Dataset 9).

Strikingly, probes **3** and **4** engaged four kinases (Aurora kinase A (AURKA), Aurora kinase B (AURKB), microtubule-associated serine/threonine kinase 3 (MAST3) and serum/glucocorticoid regulated kinase 3 (SGK3)) and two kinases (MAST3 and SGK3), respectively, in a sustained manner, with $t_{1/2} \gg 6$ h (Fig. 2f, yellow bars). Hence, although the salicylaldehyde probes robustly engage

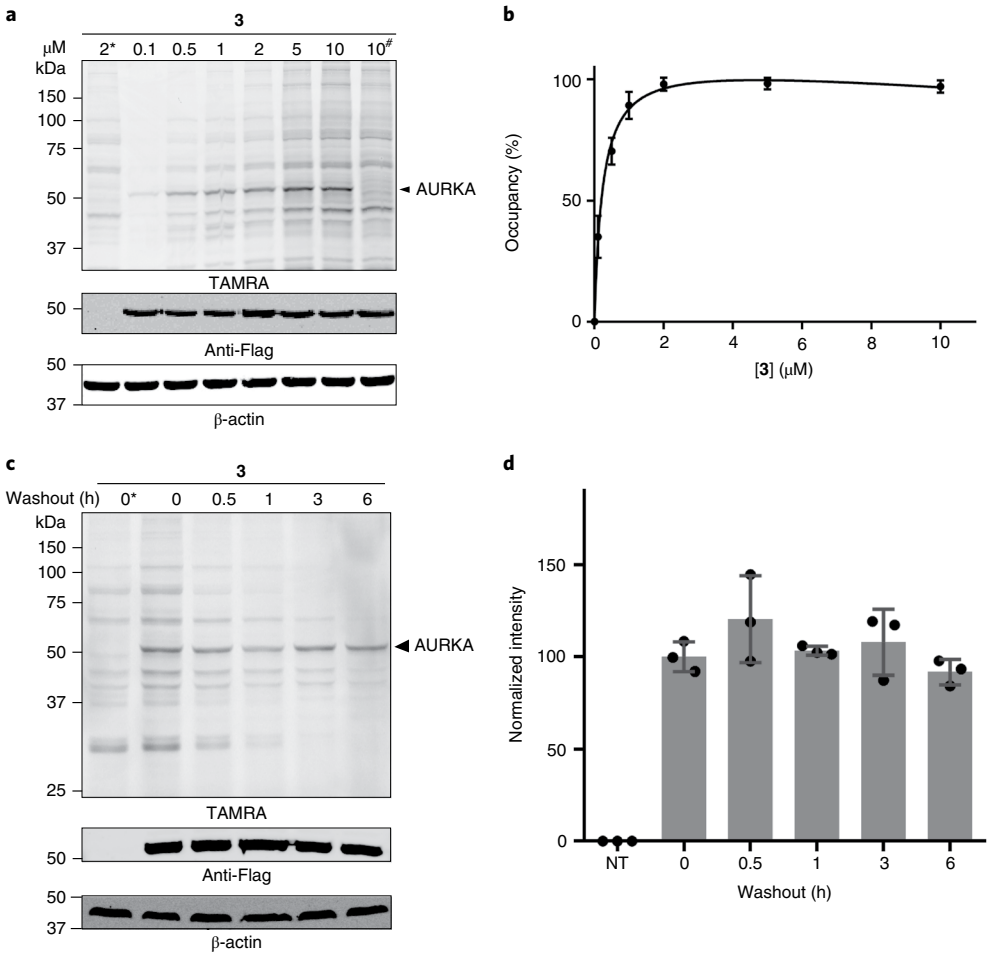


Fig. 3 | Quasi-irreversible engagement of AURKA in cells. a, COS-7 cells were transfected with WT or K162Q (#) Flag-tagged AURKA and treated with probe **3** (30 min). After lysis in the presence of sodium borohydride and TAMRA azide conjugation, samples were analyzed by in-gel fluorescence and western blotting. **b**, Concentration-dependent labeling of Flag-tagged AURKA (~50 kDa TAMRA band, mean ± s.d., $n = 3$). **c**,**d**, COS-7 cells were transfected with Flag-tagged AURKA and treated with probe **3** (2 μM, 30 min), followed by washout for the indicated times. Samples were analyzed as described above (mean ± s.d., $n = 3$). * and NT indicate nontransfected cells.

~100 kinases in cells (Supplementary Dataset 9), broad target occupancy decreases upon washout such that only a small subset of kinases remain substantially engaged after 6 h (Fig. 2f). Because the structurally related benzaldehyde **1** dissociated rapidly from all kinases, we hypothesized that prolonged kinase engagement by salicylaldehydes **3** and **4** is due to the formation of a kinetically stable imine with a subset of kinase catalytic lysines.

To validate the chemoproteomic data, we focused on two kinases, AURKA and Src, which were found to engage probe **3** with long ($t_{1/2} \gg 6$ h) and short ($t_{1/2} < 1$ h) residence times, respectively. First, we tested the ability of probe **3** to covalently modify overexpressed AURKA. COS-7 cells were transfected with Flag-tagged AURKA, followed by treatment with probe **3** for 30 min. After cell lysis (with $NaBH_4$ reduction) and TAMRA azide click conjugation, in-gel fluorescence revealed a major 50-kDa fluorescent band in AURKA-transfected cells, which appeared to reach saturation at ~2 μM probe **3** (half-maximum effective concentration (EC_{50}) ~300 nM; Fig. 3a,b). This band was not observed in nontransfected cells or in cells transfected with K162Q AURKA, in which the catalytic lysine is mutated to glutamine (Fig. 3a). Consistent with the chemoproteomic data, AURKA engagement by probe **3** was quasi-irreversible in cells, as shown by a washout experiment (Fig. 3c,d). By contrast, probe **3** rapidly dissociated from overexpressed Src with a $t_{1/2}$ of ~30 min, despite saturably modifying the Src catalytic lysine with an EC_{50} of ~600 nM (Extended Data Figs. 2 and 3). This striking difference in residence time was recapitulated in biochemical off-rate measurements using purified recombinant AURKA and Src kinase domains (Extended Data Fig. 4).

Cocrystal structure of AURKA bound to probe 3. To gain molecular insight into quasi-irreversible binding of probe **3** to AURKA, we cocrystallized the reversible covalent complex without borohydride reduction and determined the structure at a resolution of 2.3 Å (Fig. 4). Similar to previous kinase structures bound to the same scaffold[25], the aminopyrazole forms three hydrogen bonds with the AURKA hinge region. Imine formation between the ε-amine of K162 and the aldehyde of probe **3** was evident from the electron density map (Fig. 4a). Strikingly, the orientation of the lysine side chain relative to the salicylaldehyde allows the formation of an intramolecular hydrogen bond between the imine and the *ortho*-phenol. This hydrogen bonding interaction has previously been postulated to endow salicylaldehyde-derived imines with enhanced hydrolytic stability[30]. Furthermore, the imine carbon–nitrogen double bond, which must be attacked by a water molecule to reverse adduct formation, is shielded from bulk solvent by residues in the glycine-rich loop, including F144, as well as the N-terminal

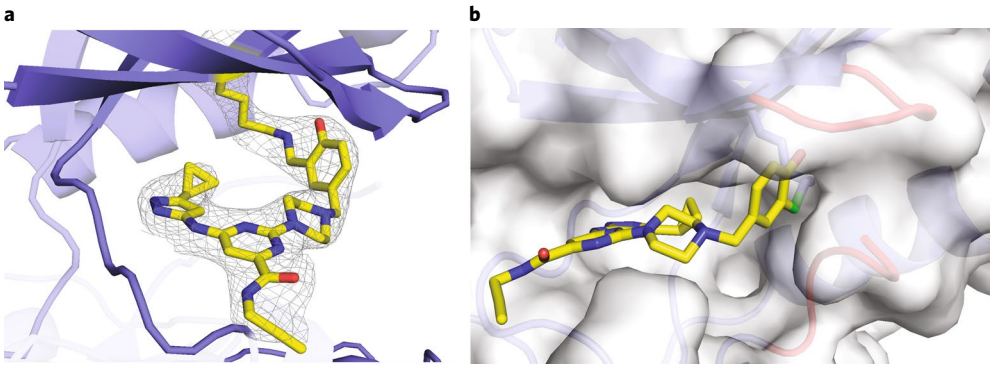


Fig. 4 | Cocrystal structure of AURKA bound to salicylaldehyde 3. a, Salicylaldehyde **3** forms an imine bond with the catalytic lysine of AURKA. Electron density map ($2F_o$-F_c) is shown at a contour level of 1σ. **b**, Surface rendering of AURKA bound to probe **3**, highlighting the Gly-rich loop (red, top right) and the activation segment (red, bottom right), both of which shield the salicylaldimine from bulk water (imine carbon in green).

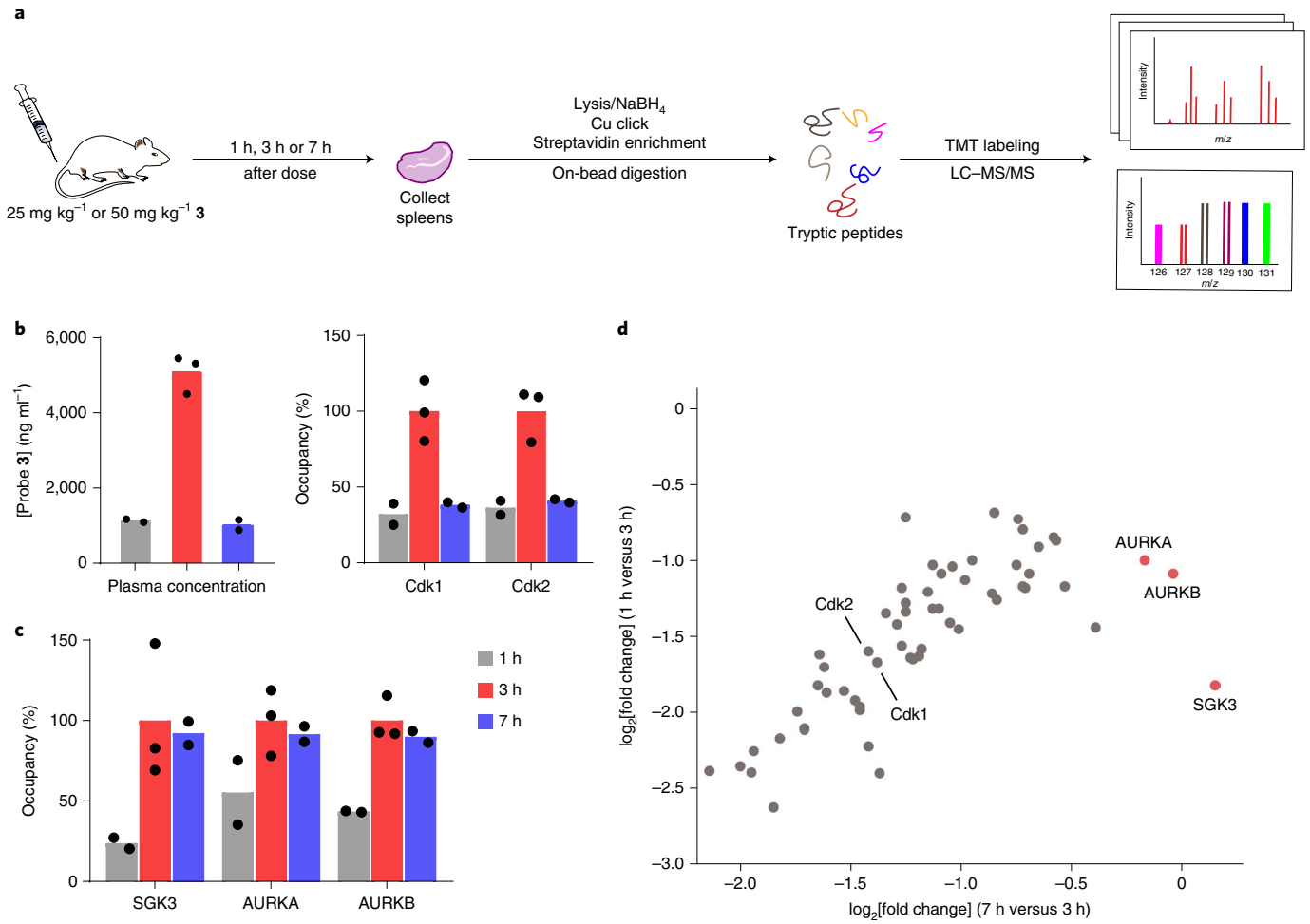


Fig. 5 | Quantifying kinase occupancy by salicylaldehyde 3 in mice. a, Mice were dosed with vehicle or probe **3** (25 mg kg^{-1} or 50 mg kg^{-1}) by subcutaneous injection. At 1 h, 3 h and 7 h after dose, plasma and spleens were collected for quantification of probe **3** concentrations and kinase occupancy, respectively. Spleens from each time point were selected for chemoproteomic analysis (biotin azide conjugation, streptavidin pulldown, 10-plex TMT quantitation) based on the probe **3** plasma levels quantified from the same mice (see main text and Supplementary Table 2). Spleens from vehicle-treated mice were collected at $t = 7$ h after injection. **b**,**c**, Plasma concentrations of probe **3** and relative occupancy of the indicated kinases (normalized to $t = 3$ h) are plotted for each time point ($t = 1$ and 7 h, $n = 2$; $t = 3$ h, $n = 3$). **d**, Scatter plot comparing relative kinase occupancy ($\log_2$[fold change] versus kinase occupancy at $t = 3$ h) at $t = 7$ h versus $t = 1$ h. Each circle represents one kinase, with AURKA, AURKB and SGK3 showing the highest occupancy at $t = 7$ h, despite decreased plasma levels of probe **3**.

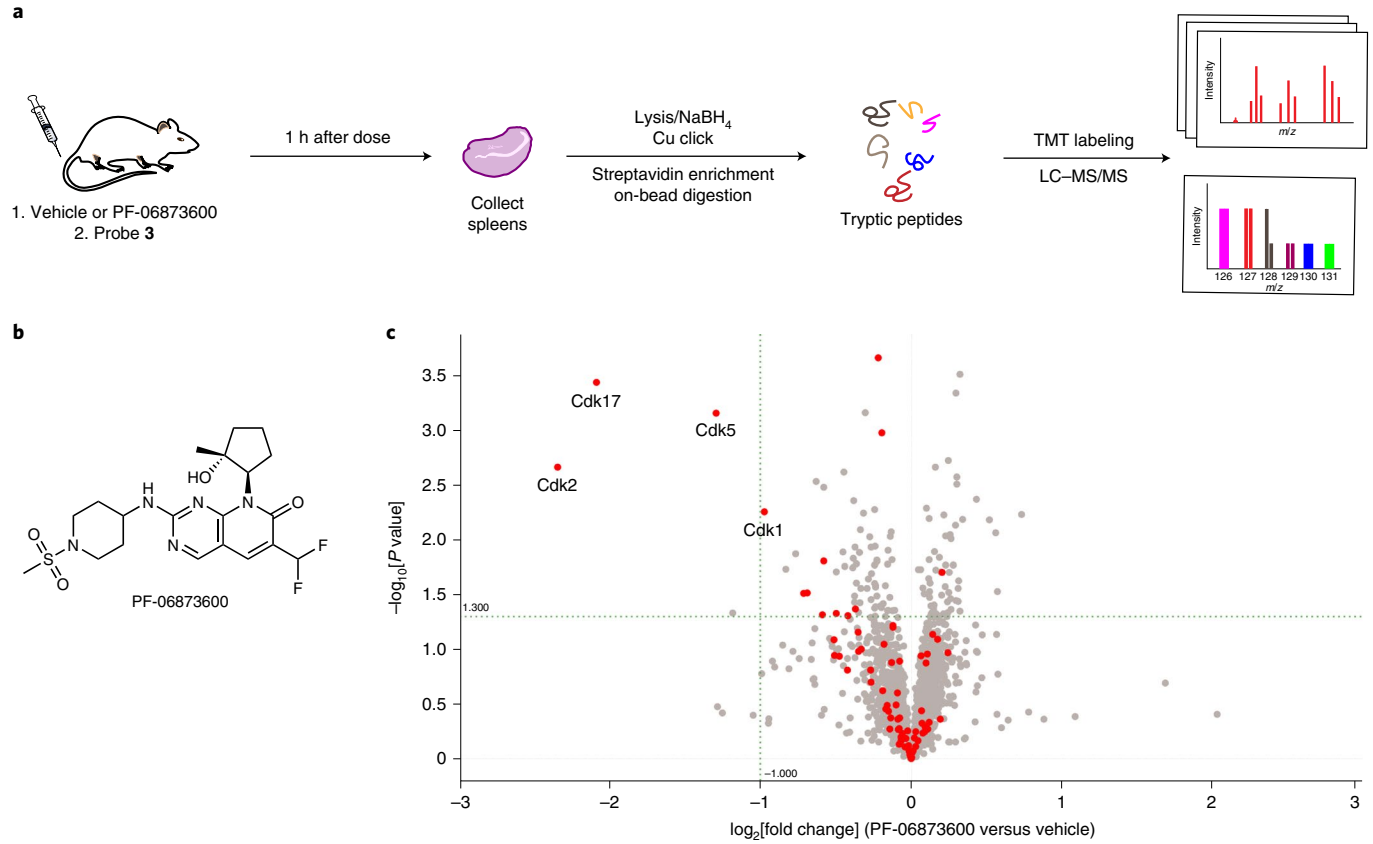


Fig. 6 | Salicylaldehyde probe 3 reveals PF-06873600 target engagement in vivo. a, Mice were dosed orally with vehicle ($n=5$) or PF-06873600 (50 mg kg^{-1}, $n=5$), and all mice were co-dosed with probe **3** (10 mg kg^{-1}) by subcutaneous injection. After 1 h, spleens were collected for chemoproteomic analysis. **b**, Chemical structure of PF-06873600. **c**, Volcano plot showing $\log_2$[fold change] and significance ($-\log_{10}$[*P* value]) between proteins enriched by probe **3** from mice treated with PF-06873600 versus vehicle. Red, kinases; gray, nonkinases.

portion of the activation segment (D274–S278), which is partially ordered in the AURKA–**3** crystal structure (Fig. 4b). In addition to affecting the imine hydrolysis rate, these interactions may also contribute to a faster on-rate for AURKA relative to other kinases. We speculate that the simultaneous formation of cooperatively stabilizing interactions is only possible with a subset of protein kinases, including the four kinases found to engage probe **3** in a sustained, washout-resistant manner. Nevertheless, we have not been able to identify conserved structure or sequence features that can account for quasi-irreversible binding to all four kinases.

Residence time-based selectivity in vivo. Given the encouraging results with probe **3** in cultured cells, we were motivated to test its suitability as a kinase probe in vivo. We dosed mice with vehicle or probe **3** (25 mg kg^{-1} or 50 mg kg^{-1}, subcutaneous injection) and collected plasma and spleen tissue at 1 h, 3 h and 7 h after dose (Fig. 5a). Depending on the dose, mean plasma concentrations of probe **3** were ~1 μg ml^{-1} at the earliest time point ($t=1$ h), increased to ~5 μg ml^{-1} at 3 h and finally returned to ~1 μg ml^{-1} at 7 h (Supplementary Table 2). To assess target engagement in vivo, we used our chemoproteomic workflow with 10-plex tandem mass tags (TMTs) for relative quantification. Our objective was to quantify kinase engagement in one TMT10 experiment across all three time points (probe **3** for 1 h, 3 h or 7 h) with a vehicle control (ten samples total, with two or three mice for each of the four conditions). The most important samples for comparing kinase residence times were from the 3 h and 7 h time points, corresponding to the maximum and subsequent minimum plasma concentrations of probe **3**. We chose the 50 mg kg^{-1} dose group for these samples, because the difference in probe **3** plasma concentrations between $t=3$ h and 7 h was greatest for these mice. Spleens from the lower dose group (25 mg kg^{-1}) were used for the $t=1$ h time point, because the mean plasma concentrations of **3** closely matched that of the 50 mg kg^{-1} dose group at $t=7$ h (Supplementary Table 2). Direct comparison of these samples would thus allow assessment of whether kinase engagement correlated with plasma concentrations of probe **3** across time points ($t=1$ h, 3 h or 7 h after dose). After homogenization, spleen lysates were immediately treated with 5 mM sodium borohydride prior to click conjugation, affinity enrichment and on-bead trypsinization. Tryptic peptides from each sample were TMT labeled, then combined and analyzed by mass spectrometry using synchronous precursor selection (SPS)-MS3 quantification with real-time search[32]. MS analysis revealed specific enrichment of 59 kinases at 3 h after dose ($\log_2$[fold change] ≥ 1 versus vehicle control, $P<0.05$; Supplementary Dataset 10), including six kinases that were not identified in Jurkat or K562 cells (Extended Data Fig. 5 and Supplementary Dataset 10). Most kinases behaved like cyclin-dependent kinases 1 and 2 (Cdk1 and Cdk2) (Fig. 5b,d), in that kinase occupancy versus time correlated with the plasma concentrations of probe **3**. This pharmacodynamic behavior is consistent with relatively rapid dissociation of probe **3** from the majority of kinases between 3 h and 7 h after dosing, concomitant with clearance of probe **3** from the circulation. Strikingly, three kinases—AURKA, AURKB and SGK3—exhibited a distinct pharmacodynamic pattern, maintaining similar levels of probe **3** occupancy at 3 h and 7 h, despite the ~5-fold reduction in probe **3** plasma concentrations (Fig. 5c,d and Supplementary Dataset 11). That probe **3** was found to exhibit sustained occupancy of AURKA, AURKB and SGK3 in mice is remarkable and provides

independent corroboration of our chemoproteomic findings in Jurkat cells.

Use of probe 3 to quantify drug–kinase engagement in mice. As a further demonstration of the utility of the clickable salicylaldehyde probe **3**, we performed a competitive labeling experiment in mice to quantify kinase engagement by PF-06873600 (Fig. 6a,b). PF-06873600, a Cdk2, Cdk4 and Cdk6 inhibitor recently disclosed in phase II clinical trials for metastatic breast and ovarian cancer (NCT03519178), exhibits greater biochemical and cellular potency versus Cdk2 than versus Cdk1, a closely related off-target kinase[28]. To assess target engagement in vivo, mice were dosed orally with PF-06873600 (50 mg kg^{-1}, $n=5$) or vehicle ($n=5$), followed by subcutaneous injection with probe **3** (10 mg kg^{-1}), which we used at a lower dose than in the study shown in Fig. 5. After 1 h, spleens were collected, processed and analyzed as described above using 10-plex TMT to quantify relative kinase enrichment by probe **3** in the presence and absence of PF-06873600 (Fig. 6c and Supplementary Dataset 12). Chemoproteomic analysis revealed significantly decreased enrichment of Cdk2, as well as the poorly characterized kinase Cdk17 ($\log_2$[fold change] < −2, $P<0.05$), from mice treated with PF-06873600. To a lesser extent, PF-06873600 also reduced probe **3** engagement of the off-target kinases, Cdk5 and Cdk1 ($\log_2$[fold change] ~ −1, $P<0.05$). Hence, PF-06873600, which is well tolerated and efficacious in Cdk2-driven tumor models[28], significantly engages Cdk2 in spleen tissue, with an estimated occupancy of >80% based on competition with probe **3**. Our chemoproteomic results demonstrate the feasibility of conducting competitive kinase engagement experiments in animals using a reversible, lysine-targeted occupancy probe.

Discussion

In this study, we have developed and characterized salicylaldehyde-based reversible covalent kinase probes suitable for in vivo experiments. Using our clickable aldehyde probes and chemoproteomic pipeline, we identified >200 kinases across experimental conditions, including 59 kinases enriched from mice dosed with YTP-2137 (probe **3**). Despite broad kinase engagement—as expected for a promiscuous scaffold that covalently modifies the conserved catalytic lysine—we uncovered striking kinetic discrimination based on distinct salicylaldehyde–kinase residence times. This residence time-based discrimination represents a previously unexplored concept for achieving selectivity with covalent kinase inhibitors that target the catalytic lysine. Kinase selectivity, driven by differences in both on-rates and off-rates, could be improved further by linking the salicylaldehyde to a more selective noncovalent-recognition scaffold. The precise geometry of the salicylaldehyde, and hence the resulting imine, likely contributes to residence time (as well as the on-rate), as evidenced by comparing the salicylaldehyde regioisomers YTP-2137 and **4** (Fig. 2f). Finally, our results with PF-06873600 demonstrate the feasibility of using clickable salicylaldehyde YTP-2137 as a competitive occupancy probe in mice.

The kinases that bind most tightly to probes YTP-2137 and **4** play disparate roles in cell biology and disease. AURKA and AURKB have distinct functions in cell cycle progression and cancer[33]. AURKA has been shown to drive neuroblastoma progression through its interactions with MYCN[34–36]. Moreover, recent work suggests that inhibition of AURKA can overcome acquired or intrinsic resistance to PI3K inhibitors in preclinical breast cancer models[37]. MAST3 regulates the response to changes in cyclic AMP in neurons[38], and recurrent, potential gain-of-function mutations in the MAST3 kinase domain have been implicated in epilepsy[39]. By contrast, SGK3 acts downstream of VPS34 and class I PI3Ks and can substitute for AKT signaling in various cancers[40–42]. Few selective inhibitors have been reported for either MAST3 or SGK3. Our results suggest the possibility of developing salicylaldehyde-based inhibitors with prolonged residence times for these and many other therapeutically relevant kinases, which lack a druggable cysteine.

Online content

Any methods, additional references, Nature Research 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/s41589-022-01019-1.

Received: 29 August 2021; Accepted: 23 March 2022;

References

1. Honigberg, L. A. et al. The Bruton tyrosine kinase inhibitor PCI-32765 blocks B-cell activation and is efficacious in models of autoimmune disease and B-cell malignancy. *Proc. Natl Acad. Sci. USA* **107**, 13075–13080 (2010).
2. Byrd, J. C. et al. Targeting BTK with ibrutinib in relapsed chronic lymphocytic leukemia. *N. Engl. J. Med.* **369**, 32–42 (2013).
3. Finlay, M. R. V. et al. Discovery of a potent and selective EGFR inhibitor (AZD9291) of both sensitizing and T790M resistance mutations that spares the wild type form of the receptor. *J. Med. Chem.* **57**, 8249–8267 (2014).
4. Jänne, P. A. et al. AZD9291 in EGFR inhibitor–resistant non–small-cell lung cancer. *N. Engl. J. Med.* **372**, 1689–1699 (2015).
5. Lanman, B. A. et al. Discovery of a covalent inhibitor of KRAS^{G12C} (AMG 510) for the treatment of solid tumors. *J. Med. Chem.* **63**, 52–65 (2020).
6. Hong, D. S. et al. KRAS^{G12C} inhibition with sotorasib in advanced solid tumors. *N. Engl. J. Med.* **383**, 1207–1217 (2020).
7. Singh, J., Petter, R. C., Baillie, T. A. & Whitty, A. The resurgence of covalent drugs. *Nat. Rev. Drug Discov.* **10**, 307–317 (2011).
8. Liu, Q. et al. Developing irreversible inhibitors of the protein kinase cysteinome. *Chem. Biol.* **20**, 146–159 (2013).
9. Chaikuad, A., Koch, P., Laufer, S. A. & Knapp, S. The cysteinome of protein kinases as a target in drug development. *Angew. Chem. Int. Ed.* **57**, 4372–4385 (2018).
10. Cuesta, A. & Taunton, J. Lysine-targeted inhibitors and chemoproteomic probes. *Annu. Rev. Biochem.* **88**, 365–381 (2019).
11. Dalton, S. E. et al. Selectively targeting the kinome-conserved lysine of PI3Kδ as a general approach to covalent kinase inhibition. *J. Am. Chem. Soc.* **140**, 932–939 (2018).
12. Gambini, L. et al. Covalent inhibitors of protein–protein interactions targeting lysine, tyrosine, or histidine residues. *J. Med. Chem.* **62**, 5616–5627 (2019).
13. Tamura, T. et al. Rapid labelling and covalent inhibition of intracellular native proteins using ligand-directed *N*-acyl-*N*-alkyl sulfonamide. *Nat. Commun.* **9**, 1870 (2018).
14. 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).
15. Wan, X. et al. Discovery of lysine-targeted eIF4E inhibitors through covalent docking. *J. Am. Chem. Soc.* **142**, 4960–4964 (2020).
16. Pettinger, J., Jones, K. & Cheeseman, M. D. Lysine-targeting covalent inhibitors. *Angew. Chem. Int. Ed.* **56**, 15200–15209 (2017).
17. Hacker, S. M. et al. Global profiling of lysine reactivity and ligandability in the human proteome. *Nat. Chem.* **9**, 1181–1190 (2017).
18. Copeland, R. A. The drug–target residence time model: a 10-year retrospective. *Nat. Rev. Drug Discov.* **15**, 87–95 (2016).
19. Bradshaw, J. M. et al. Prolonged and tunable residence time using reversible covalent kinase inhibitors. *Nat. Chem. Biol.* **11**, 525–531 (2015).
20. Eliot, A. C. & Kirsch, J. F. Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annu. Rev. Biochem.* **73**, 383–415 (2004).
21. Oksenberg, D. et al. GBT440 increases haemoglobin oxygen affinity, reduces sickling and prolongs RBC half-life in a murine model of sickle cell disease. *Br. J. Haematol.* **175**, 141–153 (2016).
22. Gampe, C. & Verma, V. A. Curse or cure? A perspective on the developability of aldehydes as active pharmaceutical ingredients. *J. Med. Chem.* **63**, 14357–14381 (2020).
23. Gushwa, N. N., Kang, S., Chen, J. & Taunton, J. Selective targeting of distinct active site nucleophiles by irreversible Src-family kinase inhibitors. *J. Am. Chem. Soc.* **134**, 20214–20217 (2012).
24. Patricelli, M. P. et al. Functional interrogation of the kinome using nucleotide acyl phosphates. *Biochemistry* **46**, 350–358 (2007).
25. Zhao, Q. et al. Broad-spectrum kinase profiling in live cells with lysine-targeted sulfonyl fluoride probes. *J. Am. Chem. Soc.* **139**, 680–685 (2017).

26. Bell, I. M. et al. Biochemical and structural characterization of a novel class of inhibitors of the type 1 insulin-like growth factor and insulin receptor kinases. *Biochemistry* **44**, 9430–9440 (2005).
27. Quach, D. et al. Strategic design of catalytic lysine-targeting reversible covalent BCR-ABL inhibitors. *Angew. Chem. Int. Ed.* **60**, 17131–17137 (2021).
28. Freeman-Cook, K. D. et al. Discovery of PF-06873600, a CDK2/4/6 inhibitor for the treatment of cancer. *J. Med. Chem.* **64**, 9056–9077 (2021).
29. Freeman-Cook, K. et al. Expanding control of the tumor cell cycle with a CDK2/4/6 inhibitor. *Cancer Cell* **39**, 1404–1421.e11 (2021).
30. Bruyneel, W., Charette, J. J. & De Hoffmann, E. Kinetics of hydrolysis of hydroxy and methoxy derivatives of *N*-benzylidene-2-aminopropane. *J. Am. Chem. Soc.* **88**, 3808–3813 (1966).
31. Herscovitch, R., Charette, J. J. & De Hoffmann, E. Physicochemical properties of Schiff bases. III. Substituent effects on the kinetics of hydrolysis of *N*-salicylidene-2-aminopropane derivatives. *J. Am. Chem. Soc.* **96**, 4954–4958 (1974).
32. 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).
33. Willems, E. et al. The functional diversity of Aurora kinases: a comprehensive review. *Cell Div.* **13**, 7 (2018).
34. Otto, T. et al. Stabilization of N-Myc is a critical function of Aurora A in human neuroblastoma. *Cancer Cell* **15**, 67–78 (2009).
35. Gustafson, W. C. et al. Drugging MYCN through an allosteric transition in Aurora kinase A. *Cancer Cell* **26**, 414–427 (2014).
36. Richards, M. W. et al. Structural basis of N-Myc binding by Aurora-A and its destabilization by kinase inhibitors. *Proc. Natl Acad. Sci. USA* **113**, 13726–13731 (2016).
37. Donnella, H. J. et al. Kinome rewiring reveals AURKA limits PI3K-pathway inhibitor efficacy in breast cancer. *Nat. Chem. Biol.* **14**, 768–777 (2018).
38. Musante, V. et al. Reciprocal regulation of ARPP-16 by PKA and MAST3 kinases provides a cAMP-regulated switch in protein phosphatase 2A inhibition. *eLife* **6**, e24998 (2017).
39. Spinelli, E. et al. Pathogenic *MAST3* variants in the STK domain are associated with epilepsy. *Ann. Neurol.* **90**, 274–284 (2021).
40. Gasser, J. A. et al. SGK3 mediates INPP4B-dependent PI3K signaling in breast cancer. *Mol. Cell* **56**, 595–607 (2014).
41. Bago, R. et al. The hVps34-SGK3 pathway alleviates sustained PI3K/Akt inhibition by stimulating mTORC1 and tumour growth. *EMBO J.* **35**, 2263–2263 (2016).
42. Vasudevan, K. M. et al. AKT-independent signaling downstream of oncogenic *PIK3CA* mutations in human cancer. *Cancer Cell* **16**, 21–32 (2009).

Methods

Cell culture. All cell lines were obtained from the American Type Culture Collection (ATCC), tested negative for mycoplasma contamination and were used without further authentication. Jurkat (ATCC, CRL-2570) and K562 (ATCC, CRL-3344) cells were cultured in RPMI 1640, GlutaMAX (Thermo Fisher Scientific, 61870036), 10% fetal bovine serum (Axenia BioLogix, F001) and 1% penicillin-streptomycin (Gibco, 15140122). COS-7 cells (ATCC, CRL-1651) were cultured in DMEM, GlutaMAX (Thermo Fisher Scientific, 10569069), 10% fetal bovine serum and 1% penicillin-streptomycin. Cells were maintained in a humidified 37 °C incubator with 5% CO_2.

Probe synthesis. Synthesis methods can be found in the Supplementary Note.

Compound treatment and preparation of cell lysates. Jurkat or K562 cells (1×10^6 per ml) were treated with probes **1**–**4** or XO44 (2 μM) at 37 °C for 30 min. Cells were pelleted by centrifugation at 500*g* for 5 min and washed with warm PBS. Then, they were resuspended in warm compound-free media and incubated at 37 °C for the duration of the washout period. Finally, cell pellets were collected by centrifugation at 500*g* at 4 °C and lysed in 100 mM HEPES pH 7.5, 150 mM NaCl, 0.1% NP-40, 1 mM phenylmethyl sulfonyl fluoride (PMSF), 1X complete EDTA-free protease inhibitor cocktail (Sigma-Aldrich, 11873580001); 25 mM $NaCNBH_3$ (probes **1** and **2**) or 5 mM $NaBH_4$ (probes **3** and **4**) were added immediately before lysing the cells. Lysates were cleared by centrifugation (16,000*g*, 30 min, 4 °C). Protein concentration was determined by the BCA assay (Thermo Fisher, 23225). Cell lysates were normalized to 5 mg ml^{-1} protein with lysis buffer for the subsequent Cu(I)-catalyzed click reactions.

Click chemistry with TAMRA azide and in-gel fluorescence (Jurkat cell lysates). Click chemistry was performed by combining 42 μl of cell lysate with 8 μl of click chemistry cocktail, resulting in a final concentration of 1% SDS, 100 μM TAMRA azide (Click Chemistry Tools, 1245-5), 1 mM Tris(2-carboxyethyl)phosphine (TCEP), 100 μM Tris((1-benzyl-4-triazolyl)methyl)amine (TBTA) (from a 2 mM stock prepared in 1:4 DMSO:*t*-butyl alcohol) and 1 mM $CuSO_4$. After incubation at room temperature (~22 °C) for 90 min, the click reaction was quenched by adding 10 μl of 6X Laemmli sample buffer. Proteins were resolved by SDS–PAGE. Gels were washed with deionized water and scanned for TAMRA fluorescence (Typhoon imaging system, Molecular Dynamics). Gels were then processed for western blotting or stained with Coomassie blue. All in-gel fluorescence images were processed by ImageJ bundled with 64-bit Java 1.8.0_172 software (National Institutes of Health (NIH)), and contrast was adjusted appropriately. Western blots were analyzed with Image Studio Lite software (LI-COR Biosciences).

Pulldown of probe-modified proteins and on-bead digestion (Jurkat and K562 cells). Lysates from probe-treated Jurkat or K562 cells (1.2 ml, 5 mg ml^{-1}) were incubated with 40 μl of settled streptavidin agarose beads (Thermo Fisher, 20353) at 4 °C overnight to remove endogenous biotinylated proteins. Beads were removed by filtration (Pall, 4650). The filtrate (1 ml) was combined with 191 μl of click chemistry cocktail, resulting in a final concentration of 1% SDS, 5 μM biotin dimethylthiophosphate (DMTP) picolyl azide (synthesis described below), 1 mM TCEP, 100 μM TBTA (from a 2 mM stock prepared in 1:4 DMSO:*t*-butyl alcohol) and 1 mM $CuSO_4$. After incubation at room temperature for 90 min, 40 μl of settled high-capacity NeutrAvidin agarose beads (Thermo Fisher, 29204) were added to each sample and incubated at 4 °C overnight. The beads were then washed with 1% NP-40, 0.1% SDS in PBS (3 × 10 min, room temperature), freshly prepared 6 M urea in PBS (3 × 30 min, 4 °C), and PBS (3 × 10 min, room temperature), before a brief wash in digestion buffer (20 mM Tris, 2 mM $CaCl_2$, pH 8). Disulfide reduction was performed with 5 mM dithiothreitol (DTT) at 56 °C for 30 min, followed by alkylation with 20 mM iodoacetamide at room temperature for 30 min in the dark. On-bead digestion was performed by adding 1 μg sequencing grade trypsin (Promega, V5113) in 100 μl digestion buffer to each sample, and incubating overnight at 37 °C. Digestion was quenched by adding 1% formic acid. The resulting peptides were desalted with Omix C18 Tips (Agilent, A57003100) and eluted with 50% acetonitrile (MeCN), 0.1% formic acid in water. The samples were dried down by SpeedVac.

LC–MS/MS analysis of tryptic peptides prepared by on-bead digestion (Jurkat and K562 cells). Tryptic peptides were reconstituted in 0.1% formic acid and analyzed on a Fusion Lumos Tribrid (Thermo Fisher) connected to an ACQUITY M-Class ultra-performance liquid chromatography (UPLC) system (Waters) (Jurkat cells), or an Orbitrap Eclipse Tribrid Mass Spectrometer (Thermo Fisher) connected to an UltiMate 3000 RSLCnano system (K562 cells), and an EASY-Spray 75 μm × 15 cm C18 column with 3 μm particle size (Thermo Fisher, ES800). Peptides were loaded onto a column warmed to 45 °C, and equilibrated with 5% solvent B (0.1% formic acid in MeCN) for 3 min at 600 nl min^{-1}, and then eluted with a flow rate of 300 nl min^{-1} using 5% solvent B for 3 min, followed by 5–30% B for 72 min, 50% B for 2 min and 5% B for 6 min. Precursor ion mass to charge ratios (MS1) were measured in the Orbitrap at a resolution of 120 K, a scan range of 375–1500 *m*/*z*, an automatic gain control (AGC) target of 40,000 and a maximum injection time of 50 ms with internal calibration enabled. Ions with a peptide-like isotopic distribution (monoisotopic precursor selection (MIPS) set to 'peptide') that exceeded an intensity threshold of 20,000 and contained a charge between 2 and 7 were selected for higher-energy C-trap dissociation (HCD) fragmentation. MS2 spectra of HCD-fragmented peptides were collected using a 1.6 *m*/*z* isolation window and an HCD collision energy of 30%. Fragment ions were measured in the Orbitrap at a resolution of 30 K and were collected with an AGC target of 50,000 and a maximum injection time of 100 ms. Peptides selected for fragmentation were dynamically excluded for the following 30 s using a 10 ppm window. The maximum duty cycle was set to 3 s.

Database matching and peptide quantification (Jurkat and K562 cells). Raw files were converted into peaklists using in-house software. A database search was performed using Batch-Tag within Protein Prospector 5.23.0 against a human Swiss-Prot database concatenated with a corresponding randomized decoy database. Only tryptic peptides with one or fewer missed cleavages were considered. Carbamidomethylation of cysteine was included as a constant modification; methionine oxidation, acetylation of the N-terminus, N-terminal glutamine to pyroglutamate conversion, and loss of the N-terminal methionine were included as variable modifications. Precursor tolerance was set to 20 ppm, and fragment tolerance was set to 30 ppm. Following database matching, quantification of peptide intensity was performed using Search Compare within Protein Prospector. The maximum peptide false discovery rate was set to 1%. Extracted ion chromatograms (XICs) were prepared for each peptide and were used to calculate intensities with a retention time window of ±20 s.

Label-free protein quantification and estimation of probe–kinase dissociation rates. Protein intensities were calculated from the sum of the individual unique peptide XICs using an in-house script. Relevant to Fig. 2d, a kinase was defined as specifically enriched by a probe if it was identified with one or more unique peptides in all three biological replicates and had a protein intensity value ≥10-fold higher than that of the DMSO control. Enriched kinases identified with three or more unique peptides in all three biological replicates were selected for estimation of probe–kinase dissociation half-life (Fig. 2f). The $t_{1/2}$ of a given probe for each kinase was estimated by fitting kinase intensity values at each time point (after washout) to a single-exponential decay function using the lmfit Python curve-fitting module.

Plasmid DNA transfection and compound treatment. COS-7 cells were seeded in six-well plates and grown until they reached ~80% confluency. To prepare the transfection complex, 2.5 μg of plasmid DNA (pCMV7.1 containing human AURKA or human Src with an N-terminal 3xFlag tag) in 150 μl of Opti-MEM I Reduced-Serum Medium (Gibco, 31985062) were mixed with 7.5 μl of Lipofectamine 2000 transfection reagent (Thermo Fisher, 11668019) in 150 μl of OPTI-MEM, and incubated at room temperature for 30 min. Then, 300 μl of DNA–Lipofectamine complexes were added to each well, and cells were incubated at 37 °C for 48 h. Media was exchanged with fresh media 3 h before compound treatment. Cells were treated with probe **3** for 30 min at 37 °C, washed with warm PBS and incubated with compound-free media for the duration of the washout period. Cells were then rinsed three times with cold PBS and frozen at −80 °C. Lysates were prepared by scraping the cells in PBS containing 2X complete EDTA-free protease inhibitor cocktail (Sigma-Aldrich, 11873580001) and 5 mM $NaBH_4$, followed by a 30 min incubation on ice. Lysates were cleared by centrifugation at 16,000*g* for 30 min at 4 °C. Protein concentration was determined by the BCA assay. Cell lysates were normalized to 5 mg ml^{-1} with lysis buffer for the subsequent Cu(I)-catalyzed click reactions.

Western blotting. After in-gel fluorescence scanning, proteins were transferred from the gels to 0.45 μm nitrocellulose membranes (Bio-Rad, 1620115), followed by blocking with 2% BSA in TBST buffer (0.1% Tween-20 in Tris-buffered saline (TBS) buffer) at room temperature for 1 h. Membranes were then incubated with primary antibody (Flag: Sigma-Aldrich, F3165; β-actin: Cell Signaling Technology, 3700; 1:1,000 dilution) at 4 °C overnight, followed by TBST wash for 3 × 10 min at room temperature. Secondary antibody (LI-COR Biosciences, IRDye 800CW goat anti-mouse IgG, 926-32210, 1:10,000 dilution; IRDye 680CW goat anti-mouse IgG, 926-68070, 1:10,000 dilution) incubation was performed at room temperature for 1 h. Membranes were washed with TBST for 3 × 10 min at room temperature, scanned on an Odyssey infrared imager (LI-COR Biosciences) and analyzed with Image Studio Lite software.

Protein expression and purification. His-tagged AURKA (residues 122–403) bearing C290A/C393A mutations in pET28a with a PreScission protease cleavage site was co-transformed with lambda phosphatase in a pCDFDuet vector (Promega) into BL-21(DE3) *Escherichia coli* and purified as previously described[35,43]. The final protein was stored in size exclusion buffer (20 mM Tris pH 7.0, 200 mM NaCl, 5 mM $MgCl_2$, 10% glycerol) and used directly for crystallization. His-tagged recombinant chicken c-Src kinase domain (residues 251–533) was purified as previously described[25].

Determination of the dissociation rates of probe 3 on AURKA and Src by LC–MS. Recombinant AURKA or Src (5 μM) was treated with probe **3** (5.1 μM) in 50 mM HEPES, pH 8.0 at room temperature for 5 min, followed by 20-fold

dilution into HEPES buffer containing 10 µM XO44. Aliquots were removed at the indicated time points, quenched with 2.5% formic acid and analyzed by intact-protein LC–MS (Waters Xevo G2-XS QTof). The percentage of XO44-unmodified kinase was plotted versus time and fitted to a one-phase decay in Prism 8 (GraphPad).

Protein crystallization. AURKA (5 µM) was incubated with probe **3** (15 µM) in 20 mM Tris pH 7.0, 200 mM NaCl, 5 mM $MgCl_2$, 10% glycerol at room temperature for 1 h. The AURKA complex with probe **3** was concentrated to 18 mg ml^{-1} for crystallization. Hanging drops were prepared as a 1:1 mixture of protein–ligand complex and precipitation solution (0.2 M lithium sulfate, 0.1 M BIS-TRIS pH 5.5, 25% w/v polyethylene glycol 3350). Crystals were grown at room temperature for 3–5 days.

Crystal structure determination. Crystals were cryoprotected in the precipitation solution supplemented with 20% glycerol and stored in liquid nitrogen. Diffraction data were collected at beamline 8.3.1 of the Advanced Light Source (Lawrence Berkeley National Laboratory). An automated script for the XDS package[44] was used to index, integrate and scale the reflections. Data collection and refinement statistics are summarized in Supplementary Table 1. Molecular replacement was conducted using Phaser[45] against 4CEG (PDB code) and the model was iteratively refined using phenix.refine[46] with manual modeling and adjustments carried out in Coot[47]. The description of the ligand and the covalent link was generated with JLigand[48] and refined with Refmac[49].

Mouse studies. All in vivo studies were conducted in accordance with the current guidelines for animal welfare (National Research Council Guide for the Care and Use of Laboratory Animals, 2011). The procedures used were reviewed and approved by the Institutional Animal Care and Use Committee. Male CD-1 (Crl:CD1(ICR)) mice approximately 7–9 weeks old weighing between 20 g and 40 g were housed (up to three per cage) at standard temperature, humidity and lighting conditions with purified water and rodent diet provided ad libitum. Mice were acclimated to the laboratory environment for a minimum of 5 days prior to the initiation of dosing, then randomized to groups based on body weights. Probe **3** was formulated in a vehicle consisting of 10% ethanol, 50% polyethylene glycol 400, 10% Kolliphor EL and 30% sterile water. PF-06873600 was formulated in a vehicle consisting of 0.5% methylcellulose A4M and 0.1% Tween 80 in sterile water. In the experiment presented in Fig. 5, groups of mice were administered a single dose of vehicle, 25 mg kg^{-1} or 50 mg kg^{-1} of probe **3** by the subcutaneous route then observed after dosing for clinical signs. After euthanasia by exposure to isoflurane gas and exsanguination, plasma was collected (1 h, 3 h or 7 h after dosing) using lithium heparin separation tubes, then flash frozen in liquid nitrogen. Tissues were collected, then flash frozen in liquid nitrogen. In the experiment presented in Fig. 6, a single dose of vehicle or 50 mg kg^{-1} of PF-06873600 was administered by oral gavage, followed by a single 10 mg kg^{-1} dose of probe **3** administered by the subcutaneous route 30 min later. Plasma and tissue samples were collected after 1 h, and flash frozen as previously described, corresponding to 90 min after administration of PF-06873600. Plasma samples or standard curves made with spiked analyte were subjected to protein precipitation, then centrifuged to obtain supernatants for analysis of analyte concentrations on a Sciex QTRAP 5500-02 LC–MS/MS instrument.

Mouse spleen sample preparation for LC–MS/MS analysis. Frozen spleens from vehicle-treated and probe **3**-treated mice were homogenized on ice in 100 mM HEPES pH 7.5, 150 mM NaCl, 0.1% NP-40 using a Tissue-Tearor (BioSpec model 985370-395). Proteins were clarified by centrifugation (16,000*g*, 4 °C, 30 min). Imine reduction was performed by adding 5 mM $NaBH_4$ for 30 min on ice. Protein concentrations were quantified by BCA and normalized to 5 mg ml^{-1}. Spleen lysates (1.2 ml) were incubated with 40 µl of settled streptavidin agarose beads at 4 °C overnight to remove endogenous biotinylated proteins. Beads were removed by filtration. The filtrate (1 ml) was reacted with 191 µl of click chemistry cocktail, resulting in a final concentration of 1% SDS, 100 µM biotin DMTP picolyl azide, 1 mM TCEP, 100 µM TBTA (from a 2 mM stock prepared in 1:4 DMSO:*t*-butyl alcohol) and 1 mM $CuSO_4$. After incubation at room temperature for 90 min, the proteins were precipitated by adding 10 ml of prechilled acetone and incubating overnight at −20 °C. The precipitated proteins were pelleted by centrifugation (3500*g*, 4 °C, 30 min), resuspended in cold methanol and re-pelleted. The pellet was solubilized in 1% SDS in PBS, diluted to a final detergent concentration of 0.4% SDS, 0.6% NP-40 in PBS, and desalted on a NAP-10 column (Cytiva, 17-0854-02) using 0.1% SDS, 1% NP-40 in PBS. The column eluate was incubated with 40 µl of settled high-capacity streptavidin magnetic beads (Thermo Fisher, 88817) overnight at 4 °C. The beads were then washed with 1% NP-40, 0.1% SDS in PBS (3 × 10 min, room temperature), freshly prepared 6 M urea in PBS (3 × 30 min, 4 °C) and PBS (3 × 10 min, room temperature), before a brief wash in digestion buffer (20 mM Tris, 2 mM $CaCl_2$, pH 8). A disulfide reduction was performed with 5 mM DTT at 56 °C for 30 min, followed by alkylation with 20 mM iodoacetamide at room temperature for 30 min in the dark. On-bead digestion was performed by adding 1 µg sequencing grade trypsin in 100 µl digestion buffer to each sample, and incubating overnight at 37 °C. Digestion was quenched by adding 1% formic acid. The resulting peptides were desalted with Omix C18 Tips and eluted with 50% MeCN, 0.1% formic acid. The samples were dried down by SpeedVac.

TMT labeling of tryptic peptides from mouse spleen samples. TMT labeling was performed with the TMT10plex kit (Thermo Fisher, SK257743) according to the manufacturer's instructions with minor modifications. In brief, peptides were reconstituted in 10 µl of 30% MeCN in 200 mM HEPES buffer pH 8.5. TMT reagents were reconstituted in 20 µl of MeCN per vial, and 1.3 µl of this stock solution was added to each sample for 1 h at room temperature. Reactions were quenched by adding 1.2 µl of 5% hydroxylamine and incubated at room temperature for 15 min, followed by adding 1.3 µl of 5% trifluoroacetic acid (TFA) to acidify the solution. TMT-labeled samples were pooled and concentrated by SpeedVac to remove MeCN, and desalted using Omix C18 Tips. Peptides were eluted with 50% MeCN, 0.1% TFA, and dried by SpeedVac.

LC–MS/MS analysis of mouse spleen samples. TMT-labeled tryptic peptides were reconstituted in 5% MeCN, 0.1% TFA in water, and analyzed on a Orbitrap Eclipse Tribrid Mass Spectrometer connected to an UltiMate 3000 RSLCnano system with 0.1% formic acid as buffer A, and 95% MeCN and 0.1% formic acid as buffer B. Peptides were separated on an EASY-Spray 3 µm, 75 µm × 15 cm C18 column with the following LC settings: flow rate at 0.3 µl min^{-1}, loading samples at 4% B for 20 min, then 4–11% B over 2 min, 11–37% B over 73 min, 37–84% B over 2 min and finally 84% B for 3 min. Data were acquired in a data-dependent mode. MS1 scans were acquired at a resolution of 120 K with an AGC target of 4×10^5, m/z scan range of 400–1,600, a maximum ion injection time of 50 ms, charge states of 2–8 and a 60 s dynamic exclusion time. MS2 spectra were acquired through collision-induced dissociation (CID) at a collision energy of 35%, in the ion trap with an AGC target of 1×10^4, isolation width of 0.7 m/z and maximum ion injection time set to 'auto'. For real time search, MS2 spectra were searched against the *Mus musculus* reviewed Swiss-Prot FASTA database (downloaded on 12 July 2020) with the digestion enzyme set to trypsin. Methionine oxidation was set as a variable modification; carbamidomethylation of cysteine and TMT modification were set as constant modifications. For MS3 acquisition, an SPS of 10 fragments was acquired in the Orbitrap for a maximum ion injection time of 105 ms with an AGC target of 1.5×10^5. MS3 spectra were collected at a resolution of 50 K with HCD collision energy of 65%.

Peptide identification and TMT quantification (mouse spleen samples). Raw files were analyzed with Proteome Discoverer 2.4 software (Thermo Fisher) against the *Mus musculus* reviewed Swiss-Prot FASTA database. Trypsin was selected as the digestion enzyme with a maximum of two missed cleavages and a minimum peptide length of six. Carbamidomethylation of cysteine and TMT modification of lysine/peptide N-terminus were set as fixed modifications; methionine oxidation and protein N-terminus acetylation were set as variable modifications. Precursor tolerance was set to 10 ppm, and fragment tolerance was set to 0.6 Da. Peptide-spectrum match (PSM) and protein false discovery rate (FDR) were set to 1% and 5%, respectively. Reporter ion intensities were adjusted to correct for impurities during synthesis of different TMT reagents according to the manufacturer's specifications. For quantification, PSMs with an average reporter signal-to-noise <9 and SPS mass match <75% were removed from the final dataset. Relative protein levels were determined by summing TMT reporter ion intensities across all corresponding PSMs (only PSMs corresponding to a unique protein match in the Swiss-Prot FASTA database were considered). For the competition experiment in Fig. 6, all mice were dosed with probe **3** (10 mg kg^{-1}) along with either vehicle ($n = 5$) or PF-06873600 ($n = 5$). To account for slight differences in total protein recovery across parallel affinity enrichment samples ($n = 10$ mouse spleens, each corresponding to a separate TMT 10-plex reporter), the median protein intensity for each TMT channel was used to normalize protein intensities across all TMT channels ($n = 10$, vehicle and PF-06873600-treated mice). Mean protein intensities from each treatment arm (vehicle, $n = 5$; PF-06873600, $n = 5$) were $\log_2$-transformed and used to calculate the $\log_2$[fold change] between each treatment condition (Fig. 6c).

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

Data availability

All mass spectrometry raw files have been deposited into the MassIVE database (http://massive.ucsd.edu) and can be downloaded by the identifier MSV000088924, as well as in ProteomeXchange (http://www.proteomexchange.org) with accession number PXD031899. Source data are provided in a source data file or in the Supplementary Datasets. Coordinates and structure factors have been deposited in the PDB under accession code 7FIC. Source data are provided with this paper.

Code availability

The script used for LFQ quantification is available upon request.

References

43. Burgess, S. G. & Bayliss, R. The structure of C290A:C393A Aurora A provides structural insights into kinase regulation. *Acta Crystallogr. Sect. F* **71**, 315–319 (2015).
44. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26**, 795–800 (1993).
45. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
46. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D* **68**, 352–367 (2012).
47. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D* **66**, 486–501 (2010).
48. Lebedev, A. A. et al. JLigand: a graphical tool for the CCP4 template-restraint library. *Acta Crystallogr. Sect. D* **68**, 431–440 (2012).
49. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D* **53**, 240–255 (1997).

Acknowledgements

Funding for this study was provided by the National Cancer Institute (NCI) (NIH NCI F31CA214028, A.C.), Ono Pharma Foundation (J.T.) and Pfizer. Mass spectrometry was supported in part by the University of California, San Francisco (UCSF) Program for Breakthrough Biomedical Research and the Adelson Medical Research Foundation (A.L.B.).

Author contributions

T.Y. and J.T. conceived the project, designed the experiments and analyzed the data. T.Y. synthesized the aldehyde probes, evaluated the probes in biochemical and chemoproteomics experiments, and acquired and analyzed LFQ and TMT proteomics data. A.C. acquired and analyzed LFQ proteomics data. X.W. crystallized and determined the structure of AURKA–probe **3**. G.B.C. refined the X-ray structure. B.H., P.K. and J.R.M. performed mouse pharmacology experiments. J.D.C. analyzed TMT proteomics data. T.Y. and J.T. wrote the manuscript with input from all of the authors. J.C.K., J.D.L., S.N. and A.L.B. edited the manuscript.

Competing interests

T.Y., B.H., P.K., J.R.M., J.C.K., J.D.L., S.N. and J.D.C. are current or former employees of Pfizer. J.T. is a founder of Global Blood Therapeutics, Principia Biopharma, Kezar Life Sciences, Cedilla Therapeutics and Terremoto Biosciences, and is a scientific advisor to Entos.

Additional information

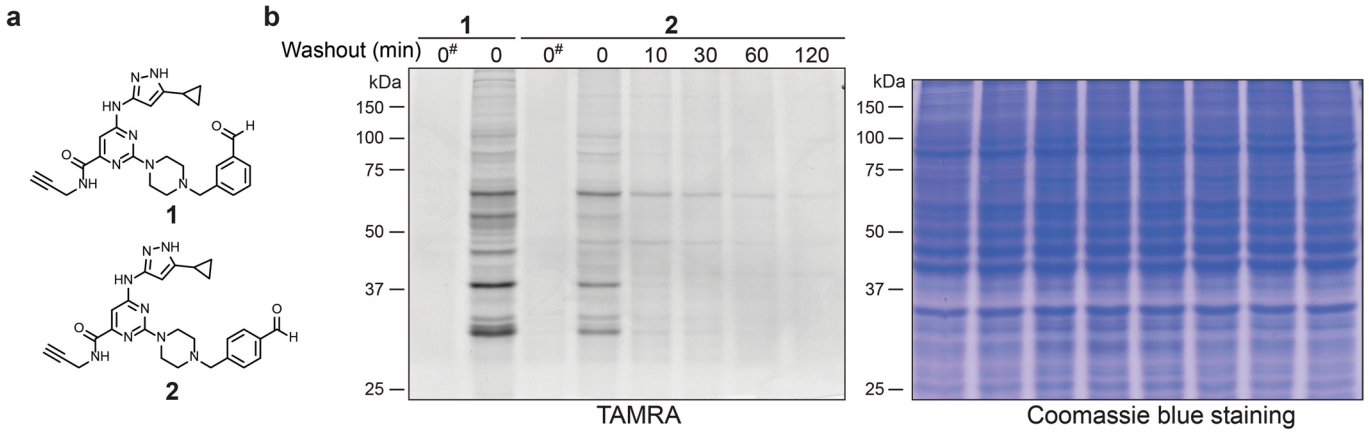
Extended data is available for this paper at https://doi.org/10.1038/s41589-022-01019-1.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41589-022-01019-1.

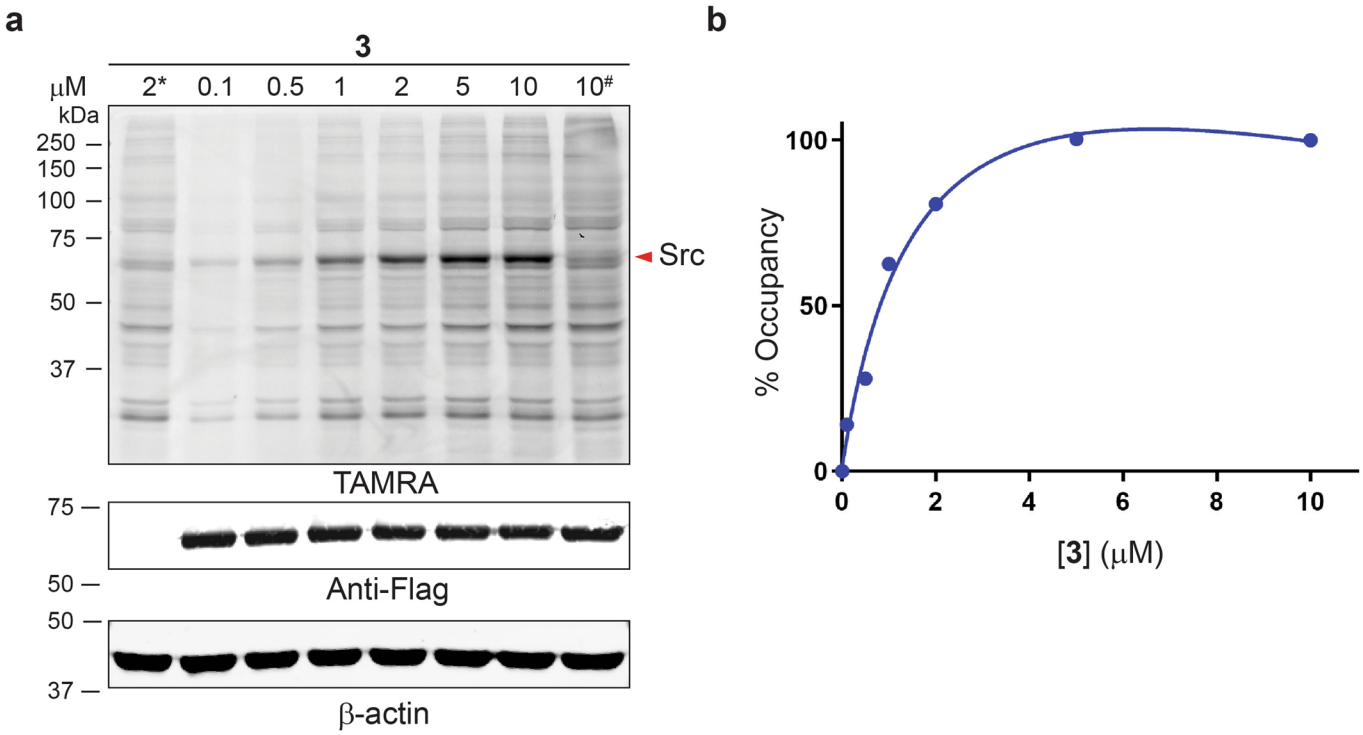
Correspondence and requests for materials should be addressed to Jack Taunton.

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

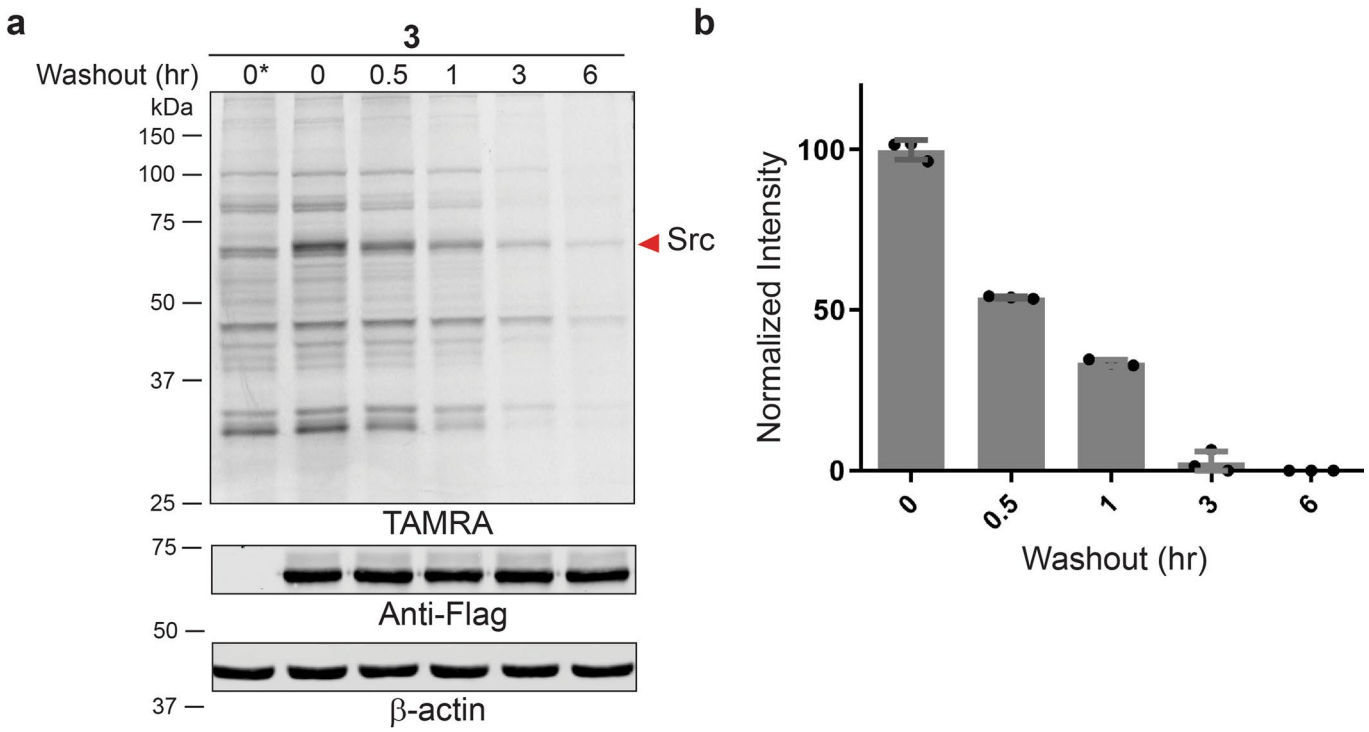
Reprints and permissions information is available at www.nature.com/reprints.



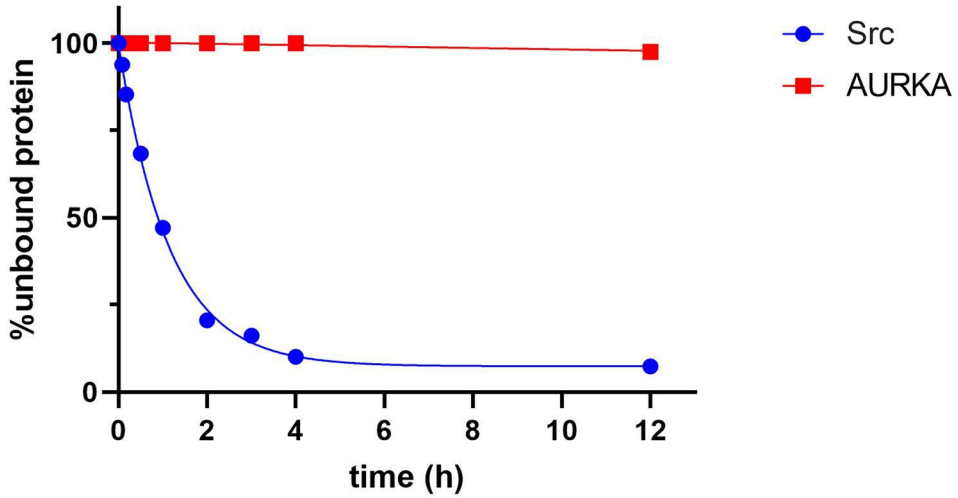
Extended Data Fig. 1 | Reversible protein labeling by benzaldehyde probe 2. (**a**) Chemical structures of **1** and **2**. (**b**) Jurkat cells were treated with **1** or **2** (2 μM, 30 min), followed by compound washout for the indicated times. Cells were lysed in the presence of 25 mM sodium cyanoborohydride, except as indicated (#). After copper-promoted click conjugation with TAMRA-azide, samples were analyzed by in-gel fluorescence and Coomassie blue staining. Data are representative of two independent experiments.



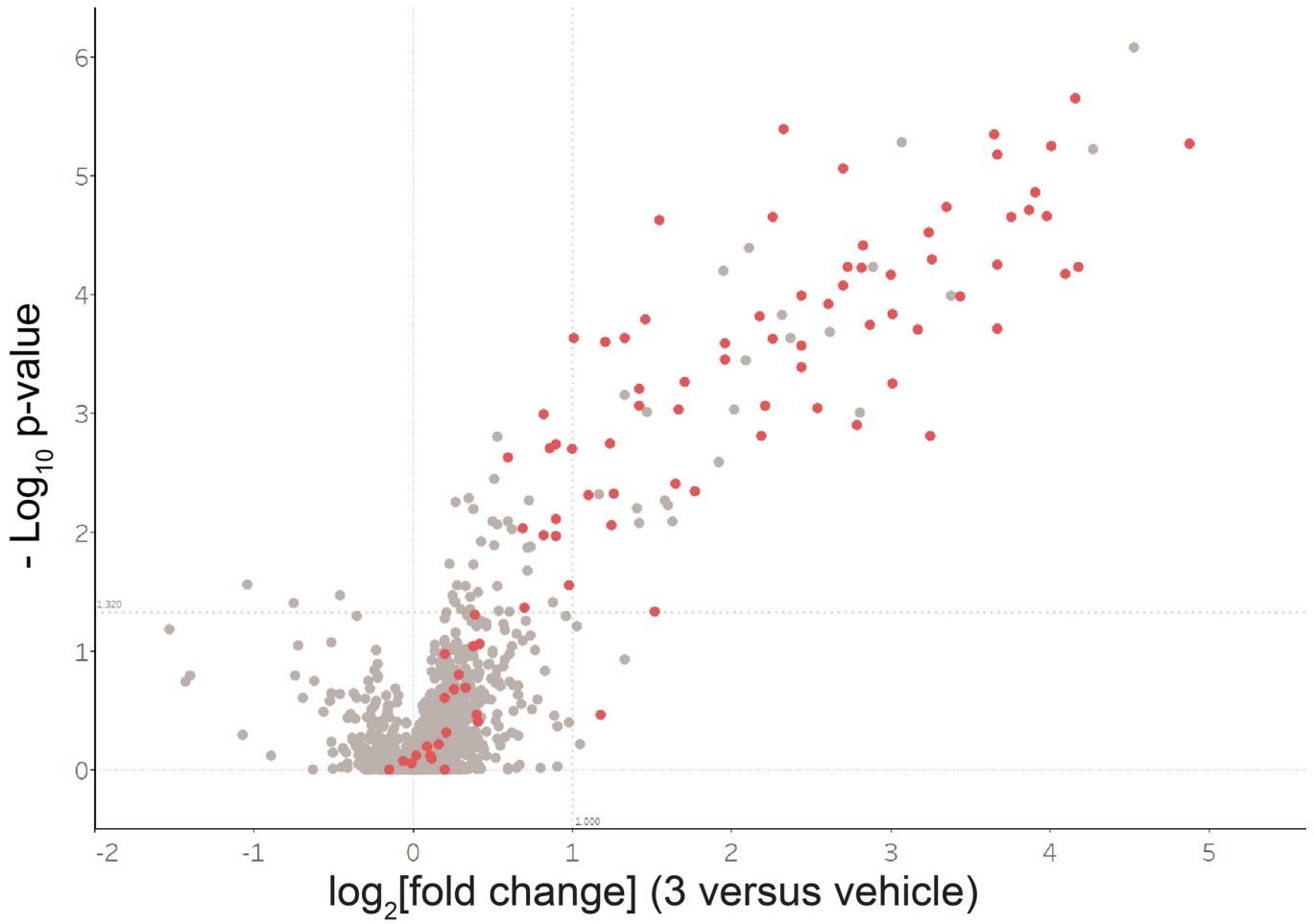
Extended Data Fig. 2 | Probe 3 modification of overexpressed Src. (**a**) COS-7 cells were transfected with WT or K295Q (#) Flag-Src, or not transfected (*), and then treated with the indicated concentrations of probe **3** (30 min). After lysis in the presence of sodium borohydride and TAMRA-azide conjugation, samples were analyzed by in-gel fluorescence and western blotting. Data are representative of two independent experiments. (**b**) Concentration-dependent labeling of Flag-Src (n = 2, mean values from two independently performed experiments were plotted).



Extended Data Fig. 3 | Rapid dissociation of probe 3 from overexpressed Src. (**a**) COS-7 cells were transfected with Flag-Src, or not transfected (*), and treated with probe **3** (2 μM, 30 min), followed by washout for the indicated times. After lysis in the presence of sodium borohydride and TAMRA-azide conjugation, samples were analyzed by in-gel fluorescence and western blotting. (**b**) Normalized fluorescence intensity of ~65 kDa band corresponding to Flag-Src (mean ± SD, n = 3).



Extended Data Fig. 4 | Dissociation of probe 3 from recombinant AURKA and Src. AURKA or Src (5 μM) was treated with probe **3** (5.1 μM) in 50 mM HEPES, pH 8.0 at RT for 5 min, followed by 20-fold dilution into buffer containing 10 μM XO44. The percentage of XO44-modified and unmodified kinase was quantified by LC-MS at the indicated time points, and % unmodified kinase (corresponding to probe **3**-bound kinase) was plotted vs. time (n = 2, mean values from two independently performed experiments were plotted).



Extended Data Fig. 5 | Related to Fig. 5. Volcano plot showing log_2 fold-change and significance ($-log_{10}$ p-value; two tailed t-test assuming unequal variance) of proteins enriched by probe **3** (t = 3 h post-dose) vs. vehicle. Red, kinases; gray, non-kinases.

nature research

Corresponding author(s): Jack Taunton

Last updated by author(s): 2021/09/03

Reporting Summary

Nature Research 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 Research 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 RSLCnano system, or a Fusion Lumos Tribrid (Thermo) connected to an ACQUITY M-Class UPLC system (Waters). X-ray diffraction data were collected at beamline 8.3.1 of the Advanced Light Source (Lawrence Berkeley National Laboratory). NMR data were collected on a 400 MHz Varian spectrometer or a 500 MHz Bruker spectrometer.
Data analysis	All software used in current study were commercially or publicly available and described in the Methods section. All in-gel fluorescence images (Figures 1b, 2b, 2c, 3a, 3c, S2, S4 and S5) were processed by Windows ImageJ bundled with 64-bit Java 1.8.0_172 software (National Institutes of Health), and contrast was adjusted appropriately. Western blots were analyzed with Image Studio Lite software (LI-COR Biosciences). Data displayed in graphs (Figures 2e, 2f, 3b, 3d, S4b and S5b) were generated and analyzed using Prism 8 (GraphPad). IC50s were fit by Prism 8 (GraphPad). LC-MS/MS data were analyzed with Proteome Discoverer (Thermo) or Protein Prospector 5.23.0 (UCSF).

Crystal structures were determined using CCP4, WinCoot, PHASER and PHENIX software (details of structure determination and refinement have been provided in Methods section).

Data displayed in graphs (Figure 4) were generated using Pymol (Schrödinger).

Data displayed in graphs (Figures 5d, 6c and S6) were generated using Tableau software.

NMR data were analyzed using MestReNova.

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 Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

The data supporting the findings of this work are available within the article and its Supplementary Information files. Source data of Figures 2d, 2e, 2f, 4, 5b, 5c, 5d, 6c, supplementary figure S6 are provided in the supplementary table. All data are available from the authors upon reasonable request. Source data are provided with this paper.

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

Life sciences study design

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

Sample size	Majority of experiments were performed at least twice using two biological replicates in each experiment.
Data exclusions	No data were excluded.
Replication	Cell based experiments were replicated at least twice in each experiment, as indicated in the figure legends. Western blots were run for two or more independent samples and representative blots are shown in the manuscript. Replications of cell based studies and western blots were successful in all cases.
Randomization	Mice from the study associated with Figures 5, 6 and S6 were randomized to groups based on body weights.
Blinding	Blinding is not relevant to this study since no group allocation was required.

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
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Antibodies used in this study were anti-FLAG antibody (F3165, Sigma-Aldrich) and β-Actin antibody (3700, Cell Signaling)
Validation	The applications of the antibodies used in this study have been validated by the manufactures based on their on-line statements. Specifically: For anti-FLAG antibody, see https://www.sigmaaldrich.com/US/en/product/sigma/f3165?gclid=Cj0KCQjw1dGJBhD4ARIsANb6OdnUSOeBHk0X9PO-j-tGpkfvxJ3rJM9Dq3qw-S4JaPeU4aA8p9wlEzkaAuDrEALw_wcB For β-Actin antibody, see https://www.cellsignal.com/products/primary-antibodies/b-actin-8h10d10-mouse-mab/3700

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Jurkat and COS-7 cells were purchased from American Type Culture Collection (ATCC).
Authentication	Cell lines were acquired from ATCC and no additional authentication was performed.
Mycoplasma contamination	All the cell lines were routinely tested for mycoplasma, showing mycoplasma-negative results.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Male CD-1 (Crl:CD1[ICR]) mice approximately 7 – 9 weeks old weighing between 12 20 – 40 grams
Wild animals	This study didn't involve wild animals.
Field-collected samples	This study didn't involve samples collected from the field.
Ethics oversight	All in vivo studies were conducted in accordance with the current guidelines for animal welfare (National Research Council Guide for the Care and Use of Laboratory Animals, 2011). The procedures used were reviewed and approved by the Institutional Animal Care and Use Committee.

Note that full information on the approval of the study protocol must also be provided in the manuscript.