



A proteome-wide atlas of lysine-reactive chemistry

Mikail E. Abbasov^{1,2}✉, Madeline E. Kavanagh¹ , Taka-Aki Ichu¹, Michael R. Lazear¹ , Yongfeng Tao¹, Vincent M. Crowley¹, Christopher W. am Ende³, Stephan M. Hacker^{1,4} , Jordan Ho⁵, Melissa M. Dix¹, Radu Suci¹, Matthew M. Hayward³, Laura L. Kiessling⁵ and Benjamin F. Cravatt^{1,5}✉

Recent advances in chemical proteomics have begun to characterize the reactivity and ligandability of lysines on a global scale. Yet, only a limited diversity of aminophilic electrophiles have been evaluated for interactions with the lysine proteome. Here, we report an in-depth profiling of >30 uncharted aminophilic chemotypes that greatly expands the content of ligandable lysines in human proteins. Aminophilic electrophiles showed disparate proteomic reactivities that range from selective interactions with a handful of lysines to, for a set of dicarboxaldehyde fragments, remarkably broad engagement of the covalent small-molecule-lysine interactions captured by the entire library. We used these latter ‘scout’ electrophiles to efficiently map ligandable lysines in primary human immune cells under stimulatory conditions. Finally, we show that aminophilic compounds perturb diverse biochemical functions through site-selective modification of lysines in proteins, including protein–RNA interactions implicated in innate immune responses. These findings support the broad potential of covalent chemistry for targeting functional lysines in the human proteome.

Small-molecule probes are critical to illuminate the biological functions of proteins and serve as leads for the discovery of therapeutics¹. At present, the vast majority of human proteins lack selective chemical probes and certain categories of proteins are considered potentially undruggable². Historical strategies for discovering chemical probes, such as high-throughput screening of large compound libraries³, have been more recently complemented by alternative approaches such as fragment-based drug discovery⁴ and covalent ligand development⁵, which have been applied proteome-wide by leveraging reactive chemical probes and quantitative mass spectrometry (MS) methods^{6–8}. By combining features of recognition and reactivity, electrophilic compounds can engage more shallow or dynamic binding pockets on proteins, thereby expanding the scope of proteins targeted by small molecules, and these irreversible interactions may also produce extended pharmacological effects that are maintained until protein targets physically turnover in the cells⁹.

Original covalent probes mainly targeted catalytic serine/threonine⁹ or cysteine¹⁰ residues located within the active sites of enzymes. More recently, covalent probes and drugs have been developed that target non-catalytic cysteine residues in functional sites of proteins, such as the ATP-binding pockets of kinases¹¹, the substrate recognition groove of the nuclear export receptor XPO1¹² and the oncogenic G12C variant in KRAS¹³. The propensity of the cysteine thiolate group to react with covalent probes and drugs is not surprising, given its greater relative nucleophilicity compared with other amino acid side chains under physiological conditions. Productive covalent binding to other amino acids, such as lysine, typically depends on the surrounding microenvironment, which may either perturb the pK_a of the lysine amino group or support high effective molarity of the reactive compound through tight reversible binding.

Given these constraints, examples of site-specific covalent binding to non-catalytic residues other than cysteine remain limited.

Expanding the scope of covalent probes also depends on understanding the reactivity and chemoselectivity of candidate electrophiles. Diverse electrophilic groups can form covalent adducts with proteinaceous lysines and include sulfonyl fluorides¹⁴, fluorosulfates¹⁵, dichlorotriazines¹⁶, activated esters¹⁷, activated sulfonamides¹⁸, vinyl sulfonamides¹⁹, imidoesters²⁰, isothiocyanates²¹, salicylaldehydes²², iminoboronates²³ and α,β -unsaturated carbonyls²⁴. These ‘aminophilic’ chemotypes show varying degrees of selectivity for lysine over other amino acids, but very few have been evaluated for reactivity on a proteome-wide scale, and, only in rare cases, have these chemotypes been leveraged to create chemical probes that can site-specifically target individual lysines to perturb protein function^{7,24,25}. Despite this, there remains great potential for aminophilic electrophiles to progress to more advanced chemical probes and drugs, as exemplified, for instance, by the recent development of voxelotor (GBT440). This salicylaldehyde drug is used to treat sickle cell disease and engages the N-terminal amine of haemoglobin in a reversible-covalent bond to increase haemoglobin affinity for oxygen²².

We previously described a chemical proteomic strategy to quantify the site-specific reactivity and small-molecule interactions of nucleophilic amino acid residues in native biological systems^{6,26,27}. In this activity-based protein profiling (ABPP) approach, libraries of electrophilic compounds are evaluated for their ability to ‘compete’ or block proteomic interactions of a chemical probe that displays broad, chemoselective reactivity with a specified amino acid. Such reactivity probes have been developed for cysteine²⁷ and lysine^{7,8}, and have more recently been extended to aspartate/glutamate^{28–30}, methionine³¹ and tyrosine³². Initial studies with the lysine-directed

¹Department of Chemistry, The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA, USA. ²Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY, USA. ³Medicine Design, Pfizer Worldwide Research and Development, Groton, CT, USA. ⁴Department of Chemistry, Technische Universität München, Garching, Germany. ⁵Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA. ✉e-mail: mikail.abbasov@cornell.edu; cravatt@scripps.edu

probe sulfotetrafluorophenyl pentynoate, however, only assessed a limited diversity of aminophilic electrophiles (activated esters and *N,N'*-diacylpyrazolecarboxamides) in the human proteome⁷. Consequently, our understanding of the types of electrophiles that can react site selectively with lysine residues to afford functional outcomes remains limited.

Motivated by these findings, we hypothesized that a more thorough understanding of the ligandability of lysines in the human proteome could be achieved by profiling a much broader array of aminophilic chemotypes in diverse human cell types and states. With this goal in mind, we report here the chemical proteomic analysis of lysine reactivity for ~180 compounds distributed across >30 aminophilic chemotypes, which include both covalent-reversible and -irreversible electrophiles, as well as terpene natural products. Across >14,000 total lysines mapped in human cancer cell line and primary human immune cell proteomes, we identified numerous sites that showed a preferential reactivity with distinct aminophilic chemotypes. These liganded lysines are found in structurally and functionally diverse proteins, and we show, in several cases, that site-specific engagement by aminophilic compounds affects protein function. We furthermore report the discovery of a remarkable set of dicarboxaldehydes that broadly landscape the ligandable lysine proteome, and thus constitute versatile 'scout' compounds for efficient mapping of small-molecule-lysine interactions in diverse biological systems. Finally, we show that cyanomethyl acyl sulfonamide compounds serve as the cell-active probes that perturb the RNA-binding interactions of the IFIT family of innate immune proteins by targeting a conserved lysine residue. These results underline how the proteome-wide analysis of lysine-reactive chemistries can uncover new chemical tools for heretofore unliganded proteins.

Results

Design of a lysine-reactive library of aminophilic compounds.

We synthesized a compound library of about ~180-members composed of 34 distinct aminophilic chemotypes tethered to structurally diversified molecular recognition (or binding) elements intended to promote interactions with distinct proteins and afford initial structure-activity relationships (SARs) both across and within chemotypes. (Fig. 1a,b; see Supplementary Data 1 for structures of the aminophilic compounds). The compounds, which can be subcategorized based on their predicted modes of reactivity (Fig. 1a), had an average molecular weight of 312 Da and were prepared using three or fewer synthetic steps. Most library members were low molecular weight fragments, but a subset had more elaborated structures, which represented structural modifications to natural products (**28n** and **33h**) and drugs (for example, **7a**, a derivative of glibenclamide, and **13d**, a derivative of celecoxib)^{33,34} (Fig. 1b and Supplementary Data 1). Attention was also paid in the library design to the installation of polarizable groups proximal to the aminophilic centre, which we hoped would promote reactivity with lysines at compound-protein interaction sites (for example, hydroxyl groups positioned adjacent to aminophilic centres with the potential to displace water molecules in the hydration shell of solvent-exposed lysines (**26a–26q**, **28a–28u** and **29a–29g**)). Most of the library was compliant with the Lipinski rule-of-five values for lead- and drug-like compounds³⁵ (Fig. 1c and Extended Data Fig. 1).

As an initial qualitative assessment of the proteome-wide reactivity and chemoselectivity of each compound, we performed gel-ABPP experiments in human cancer cell lysates with fluorescent, broad-spectrum probes that targeted individual nucleophilic amino acids—a lysine-directed probe Alexa-Fluor 488 (**P3**)⁷, a cysteine-directed probe iodoacetamide-rhodamine (**P7**)⁶ and a serine-directed probe fluorophosphonate-rhodamine (**P8**) (Supplementary Data 1)³⁶. Representative gel-ABPP results are shown in Fig. 1d–f for heterocyclic aldehydes **32a–32i**, which blocked several **P3**-protein interactions (Fig. 1d), but did not

show evidence of cross-reactivity with **P7**-labelled (Fig. 1e) or **P8**-labelled (Fig. 1f) proteins. Dicarboxaldehyde **32i** was notable in that it impaired **P3** reactivity with many proteins, but still preserved strong apparent chemoselectivity (negligible blockade of **P7**- or **P8**-reactive proteins) (Fig. 1d–f). The atypically broad lysine reactivity of dicarboxaldehydes is revisited below. Other aminophilic chemotypes blocked **P3**-labelled proteins with a selectivity that ranged from exclusive to preferential over **P7**- and **P8**-labelled proteins (Supplementary Data 2). These initial gel-ABPP experiments suggested that the aminophilic compound library engages diverse lysine residues in the human proteome but showed limited cross-reactivity with other amino acids.

Proteomic analysis of aminophilic compound-lysine interactions.

We next screened aminophilic compounds for blockade of probe **1** (**P1**; Supplementary Data 1) labelling of lysines by the mass spectrometry (MS)-based proteomic method isoTOP-ABPP (isotopic tandem orthogonal proteolysis-activity-based protein profiling)⁷. Experiments were performed using two human cancer cell line proteomes, a suspension haematological (Ramos, Burkitt's lymphoma) and an adherent epithelial (MDA-MB-231, breast) cancer cell line, which we previously found to display complementary protein content⁶. Cancer cell proteomes were pretreated with aminophilic compounds (10–100 μ M, 1 h, 23 °C) or dimethylsulfoxide (DMSO) followed by **P1** (100 μ M, 1 h, 23 °C), after which **P1**-labelled proteins in compound- and DMSO-treated proteomes were conjugated to isotopically differentiated azide-biotin tags (heavy and light, respectively) by copper-catalysed azide-alkyne cycloaddition (CuAAC)³⁷, combined, enriched by streptavidin, proteolytically digested on-bead by sequential exposure to trypsin and TEV protease, and the TEV-released **P1**-labelled peptides analysed by liquid chromatography-mass spectrometry (LC-MS) (Fig. 2a). Lysine residues were considered 'liganded' if they showed substantial reductions ($\geq 75\%$) in enrichment by **P1** in the presence of compounds compared with DMSO (MS1 chromatographic peak ratios (*R*) of ≥ 4 for DMSO/compound).

Most compounds were screened against both Ramos and MDA-MB-231 proteomes and against at least one of these proteomes in duplicate, which resulted in >460 total isoTOP-ABPP datasets (Supplementary Data 3). A median number of 2,593 lysines was quantified per dataset (Extended Data Fig. 2a), and we required that a lysine be quantified in at least 5 independent datasets for the assessment of small-molecule interactions, or 'ligandability'. From an aggregate tally of 13,785 quantified lysines on 3,552 unique proteins, we identified 818 lysines on 581 proteins that were liganded by one or more aminophilic compounds (Fig. 2b and Supplementary Data 3). About half of the liganded lysines (55%) were engaged by a single compound, and the remaining lysines showed distributed interactions profiles that ranged from 2 to many (>5) aminophilic compounds (Extended Data Fig. 2b, left panel). The majority of proteins contained zero or one liganded lysine, indicating that lysine ligandability may often be site specific within proteins^{7,8} (Extended Data Fig. 2b, right panel). The aminophilic chemotypes were found to engage discrete sets of liganded lysines and displayed marked differences in their overall lysine reactivity, which ranged from the least reactive benzoxazinones (**9**), heterocyclic sulfamates (**23**) and fluorosulfates (**18**) to the most reactive diacylphloroglucinols (**28**) and phthalaldehydes (**27**) (Fig. 2c).

We also compared the reactivity of representative members of each aminophilic chemotype with a model amine (*N*- α -acetyl-L-lysine-OMe) and found a generally good correlation with proteomic reactivity for the compounds (that is, aminophilic compounds with strong proteomic reactivity also tended to show strong model amine reactivity; Extended Data Fig. 3). We noticed that diacylphloroglucinols (**28j**, **28l** and **28n**) tended to show greater proteomic reactivity than model amine reactivity, which could reflect better

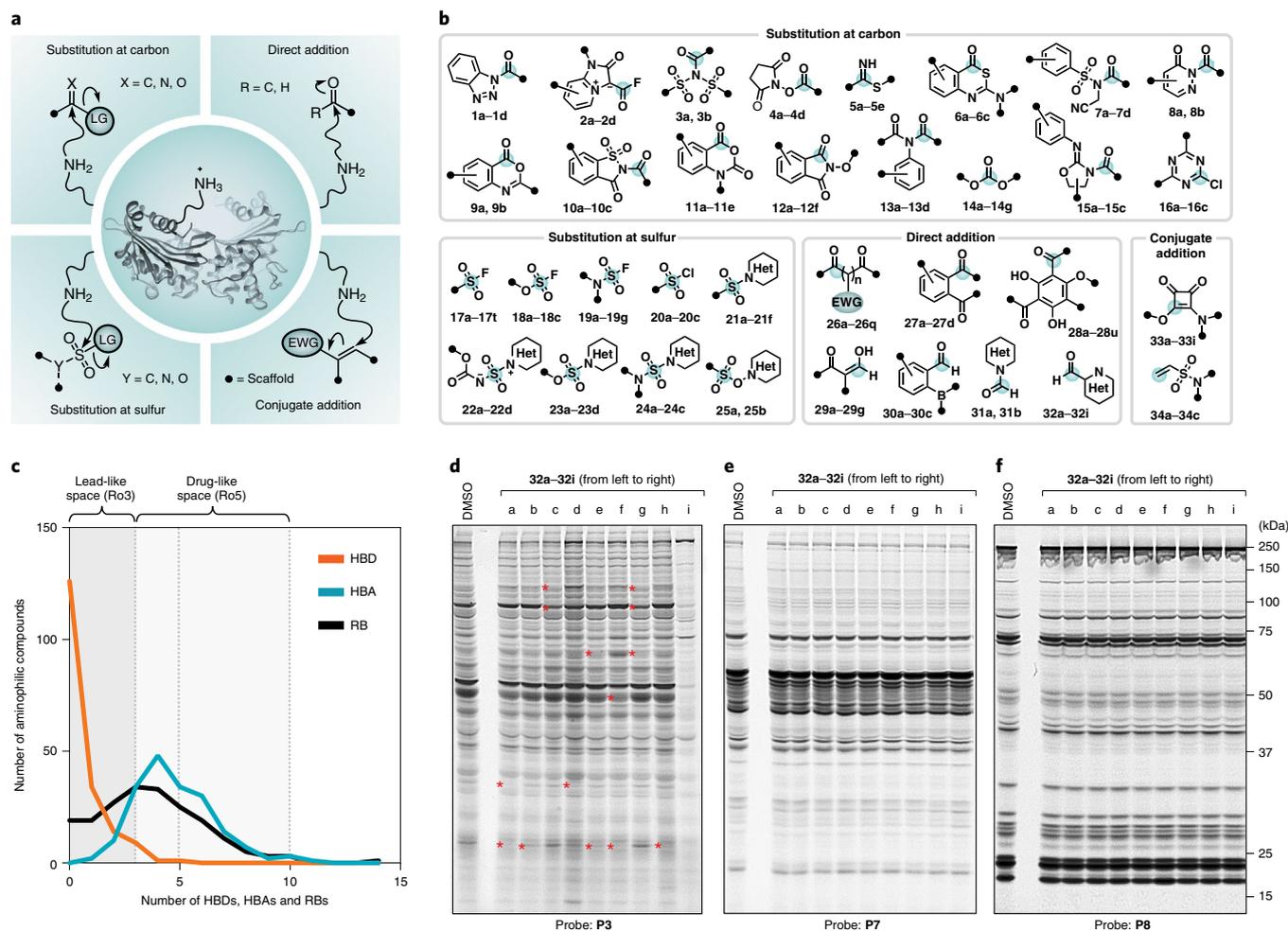


Fig. 1 | An aminophilic compound library for mapping small-molecule-lysine interactions in the proteome. **a**, Mechanism-based categorization of aminophilic chemotypes by their predicted modes of reactivity. **b**, Structural composition and diversity of representative aminophilic chemotypes clustered by reactivity modes (presumed electrophilic centres are highlighted by blue circles). See Supplementary Data 1 for a complete list of the compound structures for each chemotype. **c**, Distribution of aminophilic compounds by the number of HBDs, HBAs and RBs. The graph highlights compounds that follow Lipinski's 'rule-of-five' (Ro5) (drug-like space, light-grey box)³⁵ and Congreve's 'rule-of-three' (Ro3) (fragment-based lead-like space, dark-grey box)⁵⁹. **d-f**, Qualitative assessment of the apparent amino acid reactivity of representative compounds **32a-32i** from the heterocyclic aldehyde chemotype, as measured by competitive gel-ABPP with probes **P3'** (**d**), **P7'** (**e**) and **P8'** (**f**) in proteomic lysate of the MDA-MB-231 human breast cancer cell line. Competitive profiling experiments were generally performed as follows: the soluble proteome from MDA-MB-231 cells was treated with the indicated compounds (100 μ M, 1 h, 23 $^{\circ}$ C), followed by labelling with the indicated fluorogenic probe (2 μ M, 1 h, 23 $^{\circ}$ C) and analysis by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and in-gel fluorescence scanning. Red asterisks mark representative compound-competed proteins. This experiment was conducted twice ($n=2$) with similar results. HBDs, hydrogen-bond donors; HBAs, hydrogen-bond acceptors; RBs, rotatable bonds; EWG, electron-withdrawing group; Het, heterocyclic.

stabilization of the presumably reversible-covalent adducts with proteinaceous lysines. Furthermore, there were some exceptional compounds that showed strong model amine reactivity, but limited proteomic reactivity, such as the acyl pyridazinone **8a** and multiple acylcyclohexadione-containing polyketones (**26a-26d**). We are unsure of the basis for these differences, but, for **26a-26d**, a possible steric hindrance surrounding the most electrophilic ketone could result in more limited reactivity with proteinaceous lysines.

The extent of lysine engagement for chemotypes and individual compounds did not correlate with cLogP or molecular mass (Extended Data Fig. 4a). For instance, compounds **27d** and **28f**, which contained simple *ortho*-phthalaldehyde and elaborate 2-hydroxybenzaldehyde cores, respectively, engaged comparable numbers of lysines, despite varying considerably in their molecular weights (184 and 480 Da, respectively). Likewise, representative members of the carbonate chemotype, **14a** and **14f**, displayed a

similar overall lysine engagement, but substantially different cLogP values (1.18 and 4.96, respectively). We also found that the relative lysine ligandability values of aminophilic compounds aligned with the frequency of their representation across hydrogen-bond acceptor (HBA), hydrogen-bond donor (HBD) and rotatable bond (RB) categories (Extended Data Fig. 4b), indicating the potential for compounds of differing structures to engage lysines in the proteome.

Liganded lysines that showed broad cross-reactivity with diverse aminophilic chemotypes (Extended Data Fig. 4c) tended to correspond to residues that were also found to be liganded in our previous study that relied on activated esters as the competitor compounds⁷, suggesting that they might represent hot spots in the proteome for aminophilic compound reactivity. Even within this category, as well as more broadly across the entire set of liganded lysines, we found evidence for a substantial recognition component that directed small-molecule interactions, as reflected in individual liganded

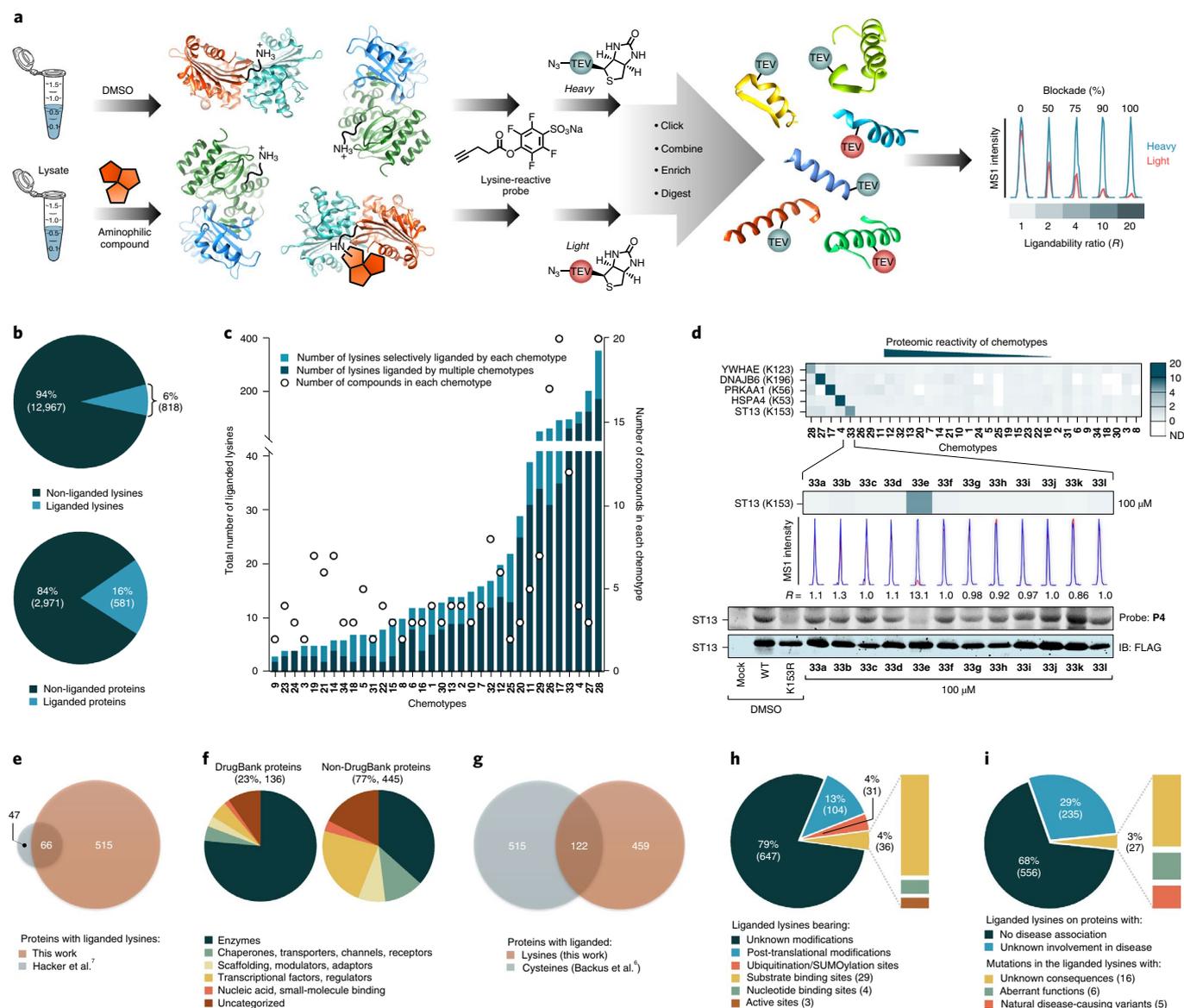


Fig. 2 | A global map of aminophilic compound-lysine interactions in the human proteome. a, General schematic for competitive isoTOP-ABPP experiments and experimental workflow to identify lysines liganded by aminophilic compounds. **b**, Fraction of total quantified lysines (top) and proteins (bottom) liganded by aminophilic compounds. **c**, Plot comparing the number of liganded lysines for each aminophilic chemotype, with blue and black designate lysines that were engaged by a single or multiple chemotypes, respectively. **d**, Top: heatmap showing R values of representative lysines preferentially liganded by a single chemotype (coloured in increasing grades of blue from values of 0–20). Middle: extracted MS1 chromatograms with the corresponding R values for K153 in ST13 showing a highly restricted SAR across individual members of the squarate chemotype (also see Supplementary Data 3). Bottom: corresponding recapitulation by gel-ABPP of the recombinant ST13 (also see Supplementary Data 4). **e**, Overlap of proteins with liganded lysines targeted by liganded lysines evaluated in this study and by activated esters evaluated in a previous study⁷. **f**, Functional class distribution of liganded DrugBank (left) and non-DrugBank (right) proteins. **g**, Overlap of proteins harbouring liganded lysines and liganded cysteines⁵ in Ramos and MDA-MB-231 proteomes. **h**, Distribution of liganded lysines based on the indicated functional categories. **i**, Distribution of liganded lysines on proteins that have human-disease relevance (as assessed by pathogenic mutations that lead to monogenic disorders defined in the OMIM database) and functional consequences of mutations of the liganded lysine residues themselves (in cases where these mutations are associated with disease). For **e–i**, lysines and proteins exclusively liganded by compounds **27c**, **28o** and **32i** were excluded from the analyses, as we revisit the preferred targets of these scout compounds in Fig. 5. TEV, tobacco etch virus; ND, not determined.

lysines displaying markedly distinct SARs (Fig. 2d), which, in some cases, opposed the overall reactivity profiles of the chemotypes. For instance, K153 in the Hsc70-interacting protein (ST13) was preferentially targeted by squarates (**33**) over other chemotypes that showed much greater proteome-wide reactivity (Fig. 2d, upper panel). Within-chemotype SAR was also apparent for this lysine, as it was engaged preferentially by **33e** over other squarates (Fig. 2d, lower

panel). Another clear example of a distinct SAR was observed for **17r**, which showed a broader and more selective engagement of kinase active-site lysines compared with those of other sulfonyl fluoride compounds (Extended Data Fig. 4d,e), reflecting a recognition element that preferentially binds to the ATP-binding pocket of kinases¹⁴.

The vast majority (~89%) of liganded lysines had not been previously identified to engage aminophilic small molecules (Fig. 2e)⁷,

which probably reflects the much broader array of chemotypes used in the current study. The proteins harbouring liganded lysines originated from diverse structural and functional classes (Fig. 2f), a modest fraction (~23%) of which, primarily enzymes, have established interactions with small molecules as reflected by their presence in the DrugBank database (Fig. 2f, left panel). The much larger fraction (~77%) of proteins with liganded lysines that were not represented in the DrugBank showed a broad functional class distribution that included transcription factors and/or regulators and scaffolding, modulator and/or adaptor proteins (Fig. 2f, right panel). Additionally, only a small fraction (~21%) of proteins with liganded lysines were found in other chemical proteomic studies to contain liganded cysteines (Fig. 2g)⁶.

Approximately one-quarter (21%) of the liganded lysines represented established 'functional' sites, which included residues that undergo post-translational modification (for example, acetylation ubiquitination and/or SUMOylation) or participate in substrate or cofactor binding (for example, active-site lysines in GLUD1/2 (K183) and UGP2 (K396) (Fig. 2h). A survey of the OMIM (Online Mendelian Inheritance in Man) database identified several human disease-relevant proteins with liganded lysines (Fig. 2i), and an interesting subset of these cases where disease-causing missense mutations occurred in the liganded lysine residues themselves, which included PMVK (K69→E) in individuals with porokeratosis 1³⁸, RPL10 (K78→E) in individuals with X-linked (MRXS35) microcephaly³⁹ and CPOX (K404→E) in patients with the harderoporphyria form of hereditary coproporphyrin⁴⁰ (Extended Data Fig. 4f–h). Such convergence of ligandability and human genetic data point to functional lysines with the potential to be targeted by chemical probes.

Characterization of aminophilic compound–lysine interactions.

We next aimed to verify and understand the functional consequences of representative aminophilic compound–lysine interactions. We noted that K404 of CPOX, which catalyses the aerobic oxidative decarboxylation of coproporphyrinogen-III to protoporphyrinogen-IX during haem biosynthesis⁴⁰, was liganded by only a single member of the aminophilic compound library—sulfonyl fluoride **17b** (*R* value = 12.6; Fig. 3a). This conserved lysine is enclosed within a cavity binding the haem precursor, coproporphyrinogen III (Fig. 3b, left panel), and other quantified lysines located elsewhere in CPOX (K347, K370 and K371) were unaffected by **17b** (Fig. 3a). Among a panel of commercially available aminophilic fluorescent probes used for a convenient gel-based analysis of ligandable lysines in recombinantly expressed proteins⁷ (Supplementary Data 1), we found that probe **P4** labelled wild-type (WT) CPOX, but not K404R or K404E mutants, in transfected HEK293T cell proteomes (Fig. 3b, right panel), and confirmed that **17b**, but not the structurally related sulfonyl fluoride **17c**, blocked **P4** labelling of WT-CPOX. Thioimido ester **5a** and formyl phloroglucinol **28p**, but not structurally analogues **5b** and **28j** (Fig. 3b, right panel), partially blocked **P4** labelling of K404 of CPOX, which matched the SAR profile acquired by chemical proteomics (*R* values of 3.9 and 3.3 for **5a** and **28p**, respectively).

A similarly strict SAR was observed by chemical proteomics for K53 of the glutathione *S*-transferase GSTT2B⁴¹. K53 is an active-site proximal residue (Fig. 3c, left panel) and was found by chemical proteomics to be preferentially engaged by ammonium-sulfonyl carbamate **22b**, *N*-hydroxyphthalimide **12a** and diketone **26h** compared with the analogues **22c**, **12c** and **26j** (Fig. 3c, right panel). Recombinant GSTT2B showed a similar SAR, as readout by site-selective profiling of K53 with probe **P1** (Fig. 3c, right panel). The most active compound **22b** engaged K53 of GSTT2B with an apparent half-maximum inhibitory concentration (*IC*₅₀) value of 3.7 μM (Fig. 3d).

Compelling SAR profiles were also observed for lysines in scaffolding proteins, as exemplified by the conserved and homologous lysines K155 and K73 in the transcriptional repressors SIN3A

and SIN3B, respectively. We previously found that K155, located in the first paired amphipathic helix (PAH1) domain of SIN3A, was liganded by an activated ester compound and this interaction blocked SIN3A interactions with binding partner TGIF1⁷. Here, we found that probe **P5** site-specifically modified the N-terminal PAH1 and PAH2 domains of SIN3A and SIN3B, but not their corresponding K155R and K73R mutants (Fig. 3e,f, respectively) and confirmed that diformyl phloroglucinol **28i**, *N*-succinimidyl ester **4a** and squarate **33e**, but not other compounds (**28h**, **4d** and **33f**), liganded both SIN3A and SIN3B, generally matching the SAR profiles for endogenous forms of these proteins. We also found that *N*-hydroxyphthalimide **12a**, which possesses the same 3,5-bis(trifluoromethyl)phenyl recognition element found in an activated ester that engages K155 of SIN3A⁷, liganded K155 and K73 with apparent *IC*₅₀ values of 0.54 and 0.92 μM, respectively (Fig. 3g). Taken together, these results demonstrate that diverse types of proteins possess lysines that can be liganded by aminophilic small molecules with interpretable SAR assignments that are preserved in the recombinantly expressed forms of the proteins.

We next prioritized proteins for functional analysis that lack chemical probes and/or were site-specifically liganded on lysines located at protein–protein interfaces. The liganded lysine K117 in the metabolic enzyme RIDA, which catalyses the hydrolytic deamination of toxic enamine and/or imine intermediates, lines the cleft of the putative substrate-binding site⁴² (Extended Data Fig. 5a). Recombinantly expressed WT-RIDA, but not a K117R mutant, reacted with probe **P2**, and this interaction was blocked by pretreatment with diketone **26l** and diformyl phloroglucinol **28h**, but not with structural analogues **26k** and **28g**, respectively (Extended Data Fig. 5b). This SAR matched the chemical proteomic data acquired for K117 of endogenous RIDA. We were particularly interested in **26l**, which showed a low micromolar activity (Extended Data Fig. 5c,d) and limited cross-reactivity with lysines across the proteome (Extended Data Fig. 5e). We found that both **26l** and **28h** blocked RIDA catalytic activity with similar *IC*₅₀ values to those measured by ABPP with probe **P2** (Extended Data Fig. 5f,g). Neither compound blocked the substrate hydrolysis mediated by a K117I mutant of RIDA, which retained near-WT levels of activity (Extended Data Fig. 5f,g) despite being unreactive with probe **P2** (Extended Data Fig. 5h). In contrast, **26l**, but not **28h**, retained inhibitory activity when tested against a K117R mutant (Extended Data Fig. 2g), which may indicate that the 1,3-dicarbonyl reactive group of **26l** can form covalent adducts with both lysine and arginine residues⁴³ (Extended Data Fig. 5i). Finally, we noted that the most common natural variant for K117 in the Exome Aggregation Consortium (ExAC) database is K117E, and the testing of this RIDA mutant revealed that it shows substantial reductions in catalytic activity and reactivity with probe **P2** (Extended Data Fig. 5h). These data thus demonstrate how chemical proteomics can identify residues for which engagement by small molecules or natural genetic mutation affect protein function.

Our chemical proteomic experiments furnished a rich map of quantified lysines in the DNA helicase XRCC6, one of which (K351) was liganded by diverse aminophilic compounds, including isatoic anhydride **11e** and squarate **33e** (Fig. 4a). XRCC6, also known as Ku70, together with XRCC5 (or Ku80), form the Ku70/Ku80 heterodimer that plays a pivotal role in non-homologous end-joining, an important pathway to repair DNA double-strand breaks in human cells⁴⁴. K351 is located at the heterodimer interface of the Ku70/Ku80 complex and engages in a salt bridge with D475 of Ku80 (Fig. 4b), and we found that pretreatment of recombinantly expressed Ku70 with **11e** or **33e** blocked co-immunoprecipitation with Ku80 in a concentration-dependent manner (Fig. 4c–e), with **11e** displaying a greater potency (*IC*₅₀ = 3.2 μM; Fig. 4e, right panel). Neither **11e** nor **33e** blocked the co-immunoprecipitation of a K351R mutant of Ku70 with Ku80 (Fig. 4d,e). However, preformation of the Ku70/Ku80 heterodimer prevented **11e** from disrupting the

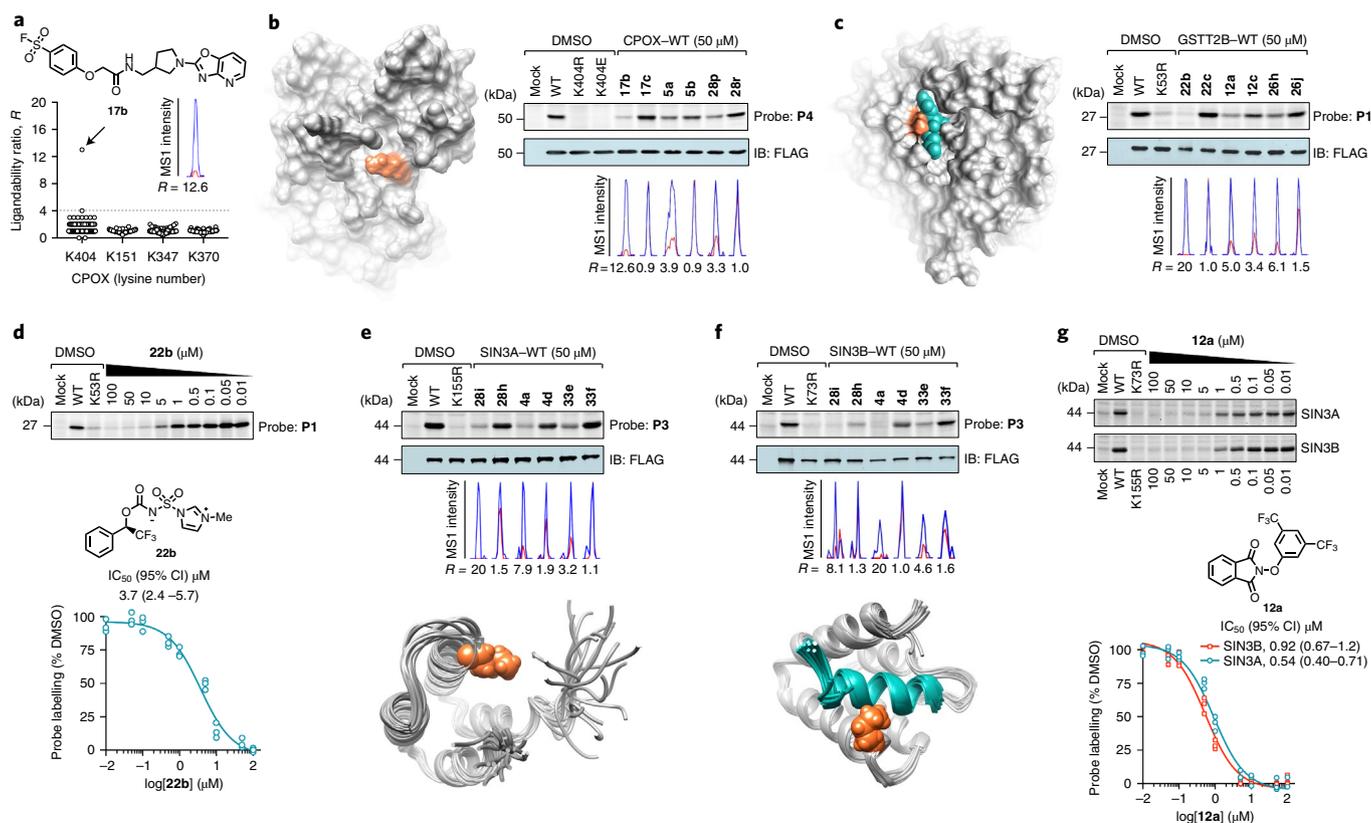


Fig. 3 | Confirmation and SAR analysis of aminophilic compound-lysine interactions with recombinantly expressed proteins. **a**, Structure of sulfonyl fluoride **17b** and the R values for quantified lysines in CPOX, identifying K404 as a liganded lysine in this protein. Each point represents a distinct aminophilic compound-lysine interaction quantified by isoTOP-ABPP. The dashed line marks the R value of 4 used to define a liganding event. The results shown are average ratios from three experiments ($n = 3$ biologically independent experiments). **b,c**, Site-specific aminophilic compound-lysine interactions are preserved in recombinant proteins. Right panels: lysates from HEK293T cells recombinantly expressing representative liganded proteins and their corresponding lysine-to-arginine mutants as FLAG epitope-tagged proteins were treated with the indicated aminophilic compounds (50 μM , 1 h) followed by treatment with the indicated lysine-reactive probes and analysis by gel-ABPP (top) and western blotting (middle). Bottom: extracted MS1 chromatograms depicting R values for the indicated aminophilic compound-lysine interactions mapped for endogenous proteins by competitive isoTOP-ABPP (also see Supplementary Data 3 and 4). Left panels: location of liganded lysines (orange) in the protein crystal structures (grey) of CPOX (**b**, PDB ID: 2AEX) and GSTT2B in complex with glutathione (blue) (**c**, PDB ID: 4MPG). **d**, Top: representative gel-ABPP data showing concentration-dependent blockade of probe **P1** labelling of recombinant GSTT2B in HEK293T cell lysates by ammoniumsulfonyl carbamate **22b**. Middle: structure of **22b**. Bottom: IC₅₀ curve for blockade of **P1** labelling by **22b** (IC₅₀ (95% CI) = 3.7 (2.4–5.7)). Data represent average values \pm s.d., $n = 3$ per group. **e,f**, SARs determined for aminophilic compound interactions with the conserved lysine in endogenous SIN3A (**e**) and SIN3B (**f**) by competitive isoTOP-ABPP recapitulated by gel-ABPP of recombinant proteins. Top two panels: HEK293T cells recombinantly expressing representative liganded proteins and their corresponding lysine-to-arginine mutants as FLAG epitope-tagged proteins were treated with the indicated aminophilic compounds (50 μM , 1 h) followed by treatment with the indicated lysine-reactive probes and analysis by gel-ABPP (top panel) and western blotting (bottom panel). Middle: extracted MS1 chromatograms depicting R values for the indicated aminophilic compound-lysine interactions identified by competitive isoTOP-ABPP (also see Supplementary Data 3 and 4). Bottom: liganded lysines (orange) mapped onto the protein crystal structures (grey) of SIN3A (**e**, PDB ID: 2RMR) and SIN3B (**f**, PDB ID: 2CZY) in complex with the neural repressor NRSF/REST (blue). **g**, Top: representative gel-ABPP data showing concentration-dependent blockade of probe **5** labelling by *N*-hydroxyphthalimide **12a** (middle) of recombinant SIN3A and SIN3B in HEK293T cell lysates. Bottom: corresponding fitted IC₅₀ curves (IC₅₀ (95% CI) = 0.54 (0.40–0.71) μM and IC₅₀ (95% CI) = 0.92 (0.67–1.2) μM , respectively). Data represent average values \pm s.d., $n = 3$ per group. For gel-ABPP data in **b–d** and **f**, the experiments were conducted three times ($n = 3$ biologically independent experiments) with similar results. PDB, protein database; CI, confidence interval; SARs, structure-activity relationships.

complex and its DNA-binding ability (Extended Data Fig. 6a,b). These results demonstrate that aminophilic compounds targeting K351 in Ku70 can block the formation of Ku70–Ku80 heterodimers, without disrupting pre-assembled ones.

Dicarboxaldehydes as scout fragments for mapping lysine ligandability. An overview of our chemical proteomic data identified three dicarboxaldehyde fragments, **27c**, **28o** and **32i** that liganded a remarkably high fraction (~58%) of the protein targets of the aminophilic compound library as a whole, as well as by activated ester

compounds previously profiled for lysine reactivity⁷ (Fig. 5a,b). These compounds represent fluorogenic reagents used to analyse amine metabolites and for traceless chemoselective bioconjugations with lysines^{45,46} (**27c** and **32i**), as well as a natural product with a reactive diformylphloroglucinol core (**28o**).

The dicarboxaldehyde fragments each liganded >200 lysine residues, which greatly exceeded the more-limited engagement profiles of other aminophilic compounds in the library (Fig. 5c), and showed overlapping but distinct lysine interaction profiles (Extended Data Fig. 7a). Previous chemical proteomic studies that

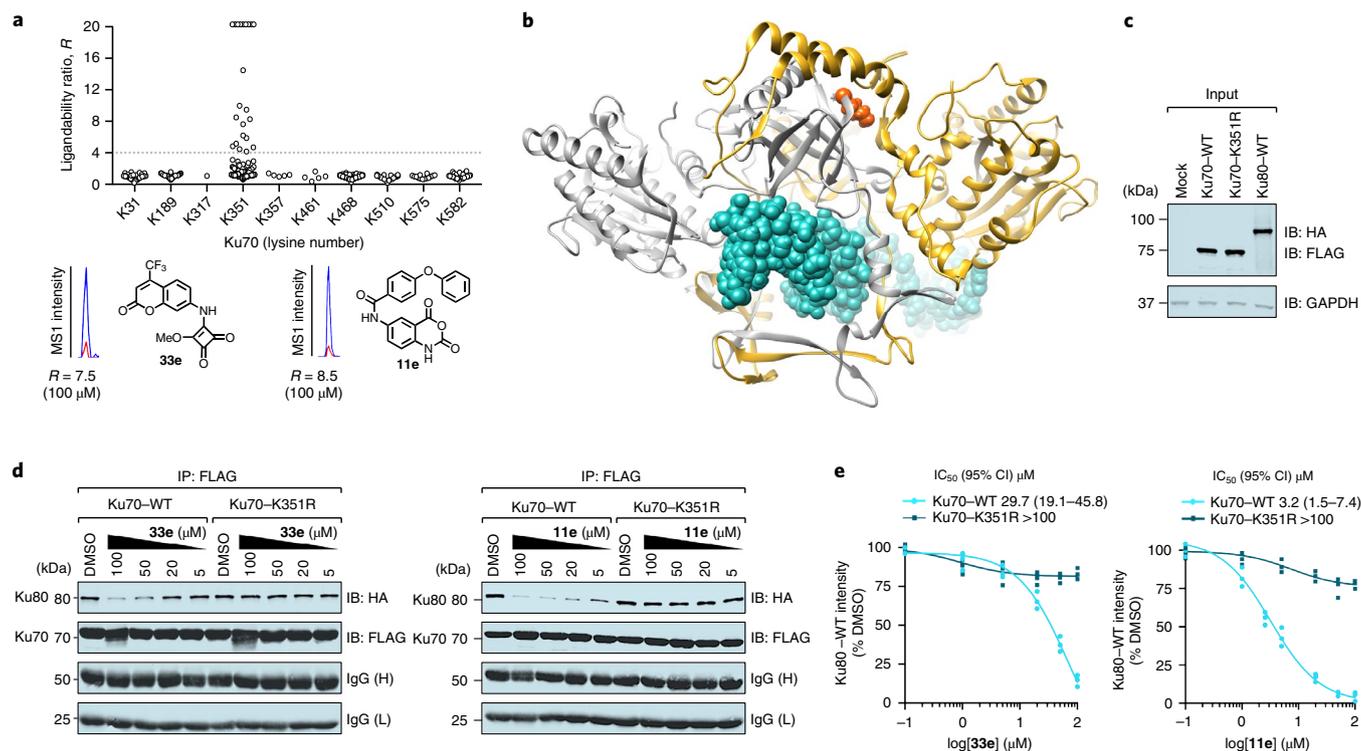


Fig. 4 | Functional impact of aminophilic compound-lysine interactions for representative proteins. **a**, Top: R values for quantified lysines in XRCC6 (or Ku70), identifying K351 as the only observed liganded lysine in this protein. Each point represents a distinct aminophilic compound-lysine interaction quantified by isoTOP-ABPP. The dashed line marks the R value of 4 used to define a liganded lysine event (also see Supplementary Data 3). The results shown are average ratios from three experiments ($n = 3$ biologically independent experiments). Bottom: structures of fragments **33e** and **11e** with extracted MS1 chromatograms depicting R values for their respective engagement of K351 in XRCC6 mapped by competitive isoTOP-ABPP. **b**, The location of liganded lysine K351 of Ku70 (orange) mapped onto the crystal structure of the Ku heterodimer (PDB ID: 1JEY) that consists of the Ku70 (grey) and Ku80 (yellow) bound to double-stranded DNA (teal). **c–e**, Aminophilic compounds engaging K351 of Ku70 block the formation of the Ku70–Ku80 heterodimer. **c**, Western blot analysis showing recombinantly expressed FLAG-tagged WT and K351R mutant forms of Ku70, as well as HA-tagged WT–Ku80 in HEK293T cells. **d**, Lysates of cells expressing Ku70 protein variants were treated with DMSO, **33e** (left) or **11a** (right) at the indicated concentrations (1 h, 23 °C) and then mixed with lysates expressing Ku80 protein (1 h, 23 °C), followed by co-immunoprecipitation (IP) with anti-FLAG antibody (1 h, 4 °C) and western blot analysis (also see Supplementary Data 4). **e**, Quantification of western blotting data for **33e** (left) and **11e** (right) from three biological replicates. Data represent average values \pm s.d., $n = 3$ per group from three biologically independent experiments. IgG, immunoglobulin G; H, heavy chain; L, light chain; HA, human influenza haemagglutinin.

evaluated cysteine-directed electrophilic fragments identified rare compounds that showed similarly broad patterns of reactivity⁶, and these fragments have since been used as ‘scouts’ to efficiently survey the cysteine ligandability of diverse biological systems⁴⁷, as well as to discover E3 ligases that support small-molecule-mediated protein degradation⁴⁸. We were therefore interested in understanding whether dicarboxaldehyde fragments could also be deployed as scouts for profiling lysine ligandability and functionality.

To explore the potential utility of **27c**, **28o** and **32i** as scout fragments for further expanding the fraction of lysines that can be targeted by aminophilic small molecules, we evaluated the reactivity of these compounds in primary human immune cell proteomes, specifically, the proteomes of human T cells activated by anti-CD3/CD28 antibodies and human peripheral blood mononuclear cells (PBMCs) with or without stimulation with bacterial lipopolysaccharides (LPS). From a total of 7,881 quantified lysines on 2,495 unique proteins across the immune cell proteomes treated with scout fragments, we identified 1,439 liganded lysines on 867 proteins (Fig. 5d and Supplementary Data 3). These liganded lysines were found in several immune-relevant proteins, defined as proteins with immune cell-enriched expression profiles and/or mutations that cause immune-related disorders in humans⁴⁹ (Fig. 5d). Gene ontology (GO) term analysis confirmed the enrichment of

diverse immune processes for proteins that harbour liganded lysines (Fig. 5e). A subset of liganded immune-relevant proteins also showed a heightened expression in LPS-stimulated PBMCs compared with quiescent PBMCs (Fig. 5f), underlining the importance of studying human immune cells in activated states to more broadly capture immune-relevant proteins.

We selected a liganded lysine (K221) in the immune-relevant protein LPCAT1, a lipid acyltransferase involved in phospholipid synthesis and remodelling⁵⁰, for further study due to its conservation among other LPCATs, as well as in more distantly related AGPAT (acylglycerol-3-phosphate O-acyltransferase) enzymes (Extended Data Fig. 7b,c). We found that the mutation of K221 to arginine blocked the LPCAT1 activity (Fig. 5g), suggesting that K221 may be involved in catalysis (three-dimensional structures of LPCAT1 and related LPCATs have not yet been determined). Consistent with this premise, each scout fragment inhibited LPCAT1 activity to a variable extent; **28o** showed the highest apparent potency (Fig. 5g) and matched the reduction in activity of HEK293T cells transfected with the inactive K221R-LPCAT1 mutant (Fig. 5g and Extended Data Fig. 7d). We next screened compounds from the parent chemotype (**28**, Extended Data Fig. 7e–g) and found that **28f** showed the strongest LPCAT1 inhibitory activity (Extended Data Fig. 7e) with an IC_{50} of 38 nM (Fig. 5h and Extended Data Fig. 7h). Previous

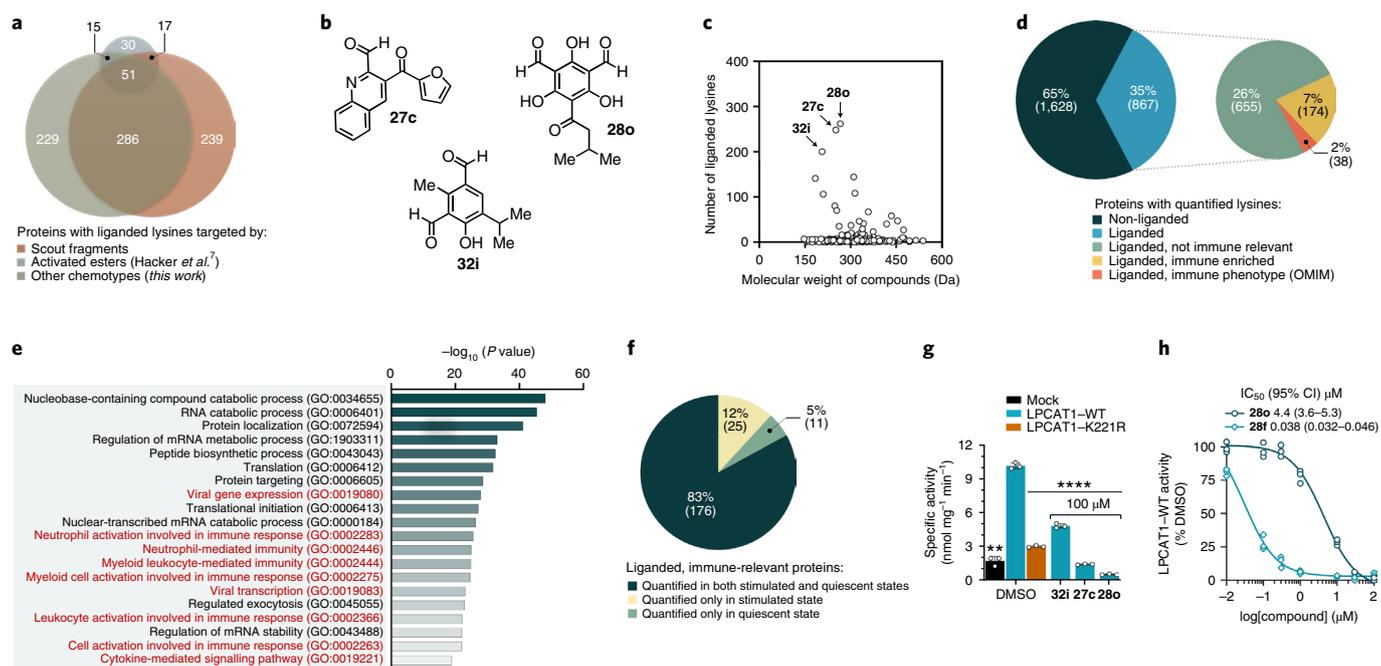


Fig. 5 | Dicarboxaldehyde scout fragments and their application for profiling lysine ligandability in human immune cells. a, Overlap of proteins with liganded lysines targeted by scout fragments (**27c**, **28o** and **32i**), other chemotypes from this study, and activated esters from a previous study⁷. **b**, Structures of scout fragments **27c**, **28o** and **32i**. **c**, Scout fragments **27c**, **28o** and **32i** engage a much larger number of lysines in human cancer cell proteomes compared to other aminophilic compounds. **d**, Fraction of proteins harbouring liganded lysines and a subset of these liganded proteins that are immune-relevant. **e**, Top-20 enriched clusters of biological processes from GO-term enrichment analysis of liganded proteins. Red font highlights immune-relevant biological processes. **f**, Fraction of liganded, immune-relevant proteins quantified in LPS-stimulated and quiescent PBMCs. **g**, Treatment with scout fragments inhibits the lyso-PC hydrolysis activity of membrane lysates of HEK293T cells recombinantly expressing WT-LPCAT1. Lysates expressing a K221R mutant of LPCAT1 also show impaired lyso-PC hydrolysis activity compared with lysates expressing WT-LPCAT1. The results shown are average ratios from three experiments ($n=3$ biologically independent experiments). **h**, Fitted IC_{50} (95% CI) curves for the concentration-dependent inhibition of the lyso-PC hydrolysis activity of recombinantly expressed WT-LPCAT1 in HEK293T cell lysates by **28o** and **28f**. Data represent average values \pm s.d., $n=3$ per group from three biologically independent experiments. P values were 0.0036 and 0.000027. Statistical significance was calculated for changes $>25\%$ in magnitude in comparison to DMSO-treated samples with unpaired two-tailed Student's t -tests: ** $P < 0.01$, **** $P < 0.0001$. GO, gene ontology; LPS, lipopolysaccharides; PBMCs, peripheral blood mononuclear cells; lyso-PC, lysophosphatidylcholine.

studies also indicated that K221 in mouse LPCAT1 was ubiquitinated³¹; however, we did not find evidence of ubiquitin modification of human LPCAT1 in a K221-dependent manner (Extended Data Fig. 7i–k).

We finally noted, in our scout fragment isoTOP-ABPP datasets, cases of differential ligandability of lysines in stimulated immune cells that occurred on proteins that did not show apparent alterations in expression. For example, K252 in the porphobilinogen synthase ALAD showed substantially weaker interactions with scout fragments in LPS-stimulated than in control PBMCs (for example, $R_{\text{control} \rightarrow \text{LPS}}$ of $9.3 \rightarrow 1.5$ for scout fragment **27c**) (Extended Data Fig. 8). LPS treatment had little effect on the reactivity of a different lysine in ALAD (K159) (Extended Data Fig. 8).

Cyanomethyl acyl sulfonamides inhibit IFIT RNA-binding proteins by engaging a conserved lysine. Proteins that possess scout fragment-sensitive lysines were found in 44 of the 47 immune cell-resolved functional modules (ME) established in a previous proteomic analysis of protein expression across diverse human immune cell types⁵² (Fig. 6a). Modules that lacked liganded proteins mainly correspond to rare immune cell types, such as plasmacytoid dendritic cells (ME28), plasma blasts (ME33) and basophils (ME35), which may not be sufficiently represented in PBMC preparations for evaluation by chemical proteomics. Modules strongly correlated with T-cell subtypes, such as ME4, harboured several proteins with liganded lysines (Fig. 6a), and GO annotation further revealed an

enriched network of RNA-related cellular functions (Figs. 5e and 6a inset), including a conserved lysine in the interferon-induced RNA-binding proteins IFIT1 (K151), IFIT3 (K148) and IFIT5 (K150) (Extended Data Fig. 9a,b), which suppress viral replication, in part, by binding to viral-specific RNA structures⁵³. IFIT1, IFIT2 and IFIT3 were identified in LPS-treated, but not control, PBMCs, consistent with their induced expression by immunostimulatory agents, whereas IFIT5, which is thought to have broader functions beyond antiviral immunity⁵⁴, was quantified in both LPS-treated and control PBMCs (Extended Data Fig. 9c). Previous literature demonstrated that the liganded lysines in IFIT1 and IFIT5 play an important role in binding viral RNA, based on both mutagenesis and structural studies⁵⁵, in which the lysine appears to directly interact with the 5'-triphosphate (5'-PPP) group of the RNA (Fig. 6b and Extended Data Fig. 9d). Considering further that, to our knowledge, chemical probes are lacking for IFITs, we pursued the further characterization of aminophilic compounds that engage the conserved lysine in IFITs.

We first confirmed that probe **P2** labelled recombinantly expressed WT-IFIT5, but not the K150R mutant of this protein (Fig. 6c), and that **P2** reactivity with recombinant WT-IFIT5 was blocked by aminophilic compounds with an SAR that generally matched our chemical proteomic data for the endogenous protein (Fig. 6c). Using an *in vitro* RNA pulldown assay, we established that recombinant WT-IFIT1 and IFIT5 were selectively pulled down by a biotinylated 5'-PPP-RNA probe, but not a 5'-hydroxyl-RNA (5'-OH-RNA)

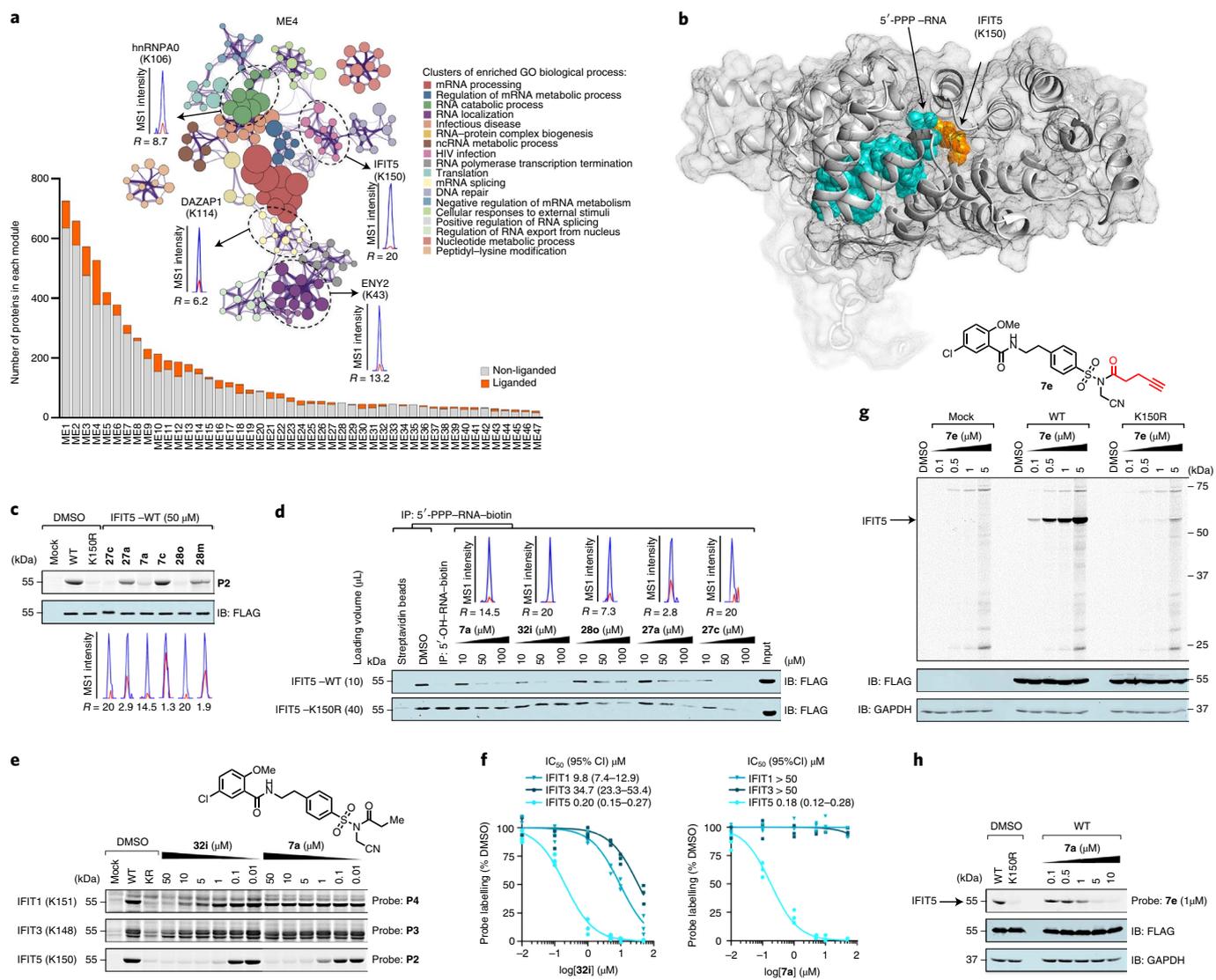


Fig. 6 | Identification of aminophilic compounds that inhibit the IFIT family of antiviral RNA-binding proteins. a, Number of proteins with lysines liganded by scout fragments within each functional module of the immune system (the modules are defined in a previous study⁵²). Inset: the enrichment network of module 4 shows protein clusters associated with RNA-related functions. Nodes are individual annotation terms, edges represent protein overlap between terms, node size represent annotation enrichment and coloured circles represent clusters. Extracted MS1 chromatograms are shown for representative liganded lysines in RNA-binding proteins involved in RNA metabolism, splicing, localization and response to viral infection. **b**, Location of ligandable lysine K150 (orange) mapped onto the crystal structures of IFIT5 (grey, PDB ID: 4HOT) bound to 5'-PPP-RNA (blue) in the nucleotide binding cleft. **c**, SAR for the aminophilic compound engagement of K150 of IFIT5, as determined by competitive isoTOP-ABPP, is recapitulated by gel-ABPP of recombinant IFIT5. Top: HEK293T cells recombinantly expressing WT-IFIT5 and the corresponding K150R mutant as FLAG epitope-tagged proteins were treated with the indicated aminophilic compounds (50 μ M, 1 h) followed by treatment with probe **3** and analysed by gel-ABPP (top panel) and western blotting (bottom panel). Bottom: extracted MS1 chromatograms depicting R values for the indicated aminophilic compound-IFIT5 (K150) interactions mapped by competitive isoTOP-ABPP (also see Supplementary Data 3 and 4). **d**, Extracted MS1 chromatograms with corresponding isoTOP-ABPP ratios (top) and western blot analysis (bottom) from biotinylated RNA pulldown experiments of WT-IFIT5 and the K150R-IFIT5 mutant from HEK293T cell lysates treated with the indicated concentrations of aminophilic compounds (also see Supplementary Data 4). **e, f**, Gel-ABPP data (**e**) and corresponding fitted IC₅₀ (95% CI) curves (**f**) for the concentration-dependent blockade of probe labelling of IFIT1, IFIT3 and IFIT5 by **32i** and **7a**. Data represent average values \pm s.d., $n = 3$ per group. **g**, Concentration-dependent in situ labelling of WT IFIT5, but not the K150R mutant of IFIT5, by an alkyne probe **7e** in transfected HEK293T cells (also see Supplementary Data 4). **h**, Representative gel-ABPP for the concentration-dependent blockade of the **7e**-WT-IFIT5 interaction by **7a** in transfected HEK293T cells. Data represent average values \pm s.d., $n = 2$ per group. For gel-ABPP and western blotting data in **c-e, g** and **h**, the experiments were conducted three times ($n = 3$ biologically independent experiments) with similar results. CI, confidence interval.

control probe (Fig. 6d and Extended Data Fig. 10a). The yield from the pulldown of the corresponding K151R and K150R mutants of IFIT1 and IFIT5, respectively, by the 5'-PPP-RNA probe was considerably lower, requiring a greater input load for detection, and was comparable in signal to the interactions of these mutant

proteins with the 5'-OH-RNA control probe (Fig. 6d and Extended Data Fig. 10a). Among the aminophilic ligands, we found that **7a** and **32i** showed a strong blockade of WT-IFIT5, but not K150R mutant, binding to the 5'-PPP-RNA probe, whereas other ligands blocked both WT and mutant protein interactions (Fig. 6d and

Extended Data Fig. 10a), possibly indicating that they engage additional lysines on IFIT5. Notably, we observed divergent SARs for blockade of IFIT1 and IFIT5 binding to 5'-PPP-RNA, which points to the potential to create subtype-selective IFIT chemical probes (Extended Data Fig. 10b). Consistent with this premise, using fluorescent probes that label each recombinantly expressed WT-IFIT, but not their corresponding lysine-to-arginine mutants (K151R for IFIT1, K148R for IFIT3 and K150R for IFIT5) (Fig. 6e,f), we found that **7a** blocked probe labelling of IFIT5 with an IC_{50} of $\sim 0.2 \mu\text{M}$, but did not inhibit probe labelling of IFIT1 and IFIT3 up to $50 \mu\text{M}$ (Fig. 6e,f, right panel). Compound **32i** also preferentially blocked fluorescent probe labelling of IFIT5, but cross-reacted with IFIT1 and IFIT3 at higher concentrations (Fig. 6e,f, left panel). The potency of the inhibition of probe **P3** labelling by **7a** was greater than that originally observed for blockade of IFIT5 interactions with 5'-PPP-RNA (Fig. 6d and Extended Data Fig. 10b); however, the latter assay contained non-ionic detergent, which we surmised might slow the rate of engagement of IFIT5 by **7a**. Consistent with this hypothesis, we found that the potency of **7a** blockade of 5'-PPP-RNA interactions with IFIT5 improved considerably when the pre-incubation time was extended from one to four hours before performing the 5'-PPP-RNA pulldown (Extended Data Fig. 10c). We next synthesized an alkyne analogue of **7a**, compound **7e** (Fig. 6g), for targeted labelling of IFIT5 using a CuAAC conjugation to azide reporter tags^{37,56}. We found that **7e** labelled WT-IFIT5 expressed in HEK293T cells both in vitro (Extended Data Fig. 10d–f) and in cellulo (Fig. 6g) at concentrations as low as $0.1 \mu\text{M}$ and showed limited cross-reactivity with other proteins in HEK293T cells below $1 \mu\text{M}$ (Fig. 6g and Extended Data Fig. 10e,f). Negligible labelling was observed for **7e** with the K150R mutant of IFIT5 (Fig. 6g and Extended Data Fig. 10e,f). We leveraged probe **7e** to measure a cellular (in situ) IC_{50} for **7a** of $1.3 \mu\text{M}$ (Fig. 6h and Extended Data Fig. 10g). We also found that **7a** exhibited a good selectivity in cells, where the compound ($1 \mu\text{M}$, 2 h) engaged few additional lysines beyond K150 of IFIT5 (Extended Data Fig. 10h,i). Taken together, these findings demonstrate that aminophilic compounds targeting a conserved lysine in human IFIT proteins with subtype selectivity can pharmacologically disrupt specific RNA–protein interactions implicated in viral replication and immune response.

Discussion

Several conclusions can be drawn from this large-scale study of the proteomic reactivity of aminophilic compounds that addresses both the opportunities and challenges facing the development of covalent ligands targeting lysines residues in proteins. First, we note that, despite identifying >800 liganded lysines, we still consider such events to be rare across the proteome, given that >14,000 lysines were quantified in our studies. It is, however, important to qualify that the total lysines quantified here represent a small fraction of all lysine residues in the human proteome, and it is therefore possible that our ligandability estimates may not reflect the broader potential for aminophilic compounds to engage lysines across the entirety of human proteins. Regardless, we are encouraged by the discovery of liganded lysines in structurally and functionally diverse proteins, including those that lack chemical probes, and underlines the potential of aminophilic compounds to expand the scope of the human proteome that can be targeted by small molecules. Indeed, our follow-up studies verified the ligandability and functionality of lysines not only at traditional druggable locations, such as enzyme active sites, but also at protein–protein (K351 in XRCC6) and protein–RNA (K150 in IFIT5) binding interfaces. In each case, we observed SARs that point to unique and substantial contributions of both the reactivity and recognition elements of aminophilic compounds. These findings highlight the potential for future optimization of potency and selectivity based on matching ligandable lysines with the preferred aminophilic chemotypes and

increasing the binding affinity through modifications to the recognition elements. We are particularly intrigued by the discovery of conserved, ligandable lysines involved in RNA binding, as targeting protein–RNA interactions with small molecules has, to date, proved challenging⁵⁷. Considering the high prevalence of lysines at protein–RNA interfaces, where these residues often bind to negatively charged RNA backbone phosphates, we speculate that aminophilic compounds may offer an advantaged type of chemical probe to perturb protein–RNA interactions.

Given the large number of lysines that preferentially or exclusively interacted with a single aminophilic chemotype, our data emphasize the value of the continued exploration of different types of aminophilic compounds to fully assess the ligandability of lysines in the human proteome. Across the chemotypes tested here, some stood out as potentially attractive starting points for a broader library construction and focused chemical-probe development. We call attention to both the squarates (**33e–33i**) and cyanomethyl acyl sulfonamides (**7a–7d**), which show atypical lysine reactivity profiles that furnished functional compounds targeting protein–protein and protein–RNA interfaces, respectively. A review of within-chemotype SAR further underlined certain features that may enhance lysine reactivity with specific compound classes. We note, for instance, that squarate **33e**, as well as **33b**, showed a broader lysine ligandability profile compared with other squarates, which could reflect the presence of a small, sterically unhindered methoxy leaving group that favours lysine modification by aza-Michael addition, along with a vicinal recognition scaffold bearing electron-withdrawing substituents that further activate the electrophilic β -carbon. Other aminophilic compounds served different purposes. The reversible-covalent dicarboxaldehydes showed a broad reactivity with ligandable lysines and were subsequently deployed as scout fragments to map covalent small-molecule–lysine interactions in primary human immune cells under different stimulation states. We anticipate that these dicarboxaldehyde scout fragments will offer versatile tools for future surveys of lysine ligandability in diverse biological systems. Finally, a recent and complementary study that explored the direct proteomic reactivity of diverse electrophilic groups also evaluated some of the same aminophilic compounds studied here, and provided additional evidence for preferential reactivity with lysine over other proteinaceous amino acids for several chemotypes (activate esters, cyanomethyl acyl sulfonamides and squarates), whereas for others showing a capacity to react with lysines and additional amino acids (sulfonyl fluorides)⁵⁸.

In considering the limitations of our studies, as well as future directions, we note that some aminophilic compound–lysine interactions may be overlooked by our approach of assessing these interactions in native proteomes, followed by confirmation with recombinant proteins (and lysine mutants of these proteins), if, for instance, the interactions require an intact cellular environment or involve proteins that are unstable in cell lysates or not straightforward to recombinantly express in heterologous systems. Future efforts to address these items could include using alternative lysis buffers, as well as establishing protocols for the in cellulo profiling of aminophilic compound–lysine interactions. Also, as the recognition element of aminophilic compounds is more extensively elaborated, we may encounter instances in which reversible rather than covalent binding blocks lysine reactivity in our chemical proteomic experiments. Indeed, this possibility should even be considered for **17r**, which is a sulfonyl fluoride that bears an ATP pocket-directed recognition element that we found to interact with many more active-site lysines in protein kinases than were engaged by other sulfonyl fluorides. Although we currently assume that the blockade of active-site lysine reactivity by **17r** reflects covalent modification, it is also possible that reversible binding by **17r** could disrupt interactions between probe **P1** and kinases. Of course, this outcome would point to another intriguing utility of lysine

reactivity profiling, namely, as a way to discover reversible small-molecule interactions that competitively disrupt probe **P1** labelling of lysines in druggable pockets in proteins. We further acknowledge that the aminophilic ligands discovered here require improvements in potency and selectivity to furnish advanced chemical probes, and this optimization would benefit from a deeper understanding of the SARs for aminophilic compound-lysine interactions, which include measurements of not only their concentration-dependency, but also their time dependency, as well as generating alkyne analogues of hit ligands, which allow for confirmation of direct and site-specific labelling of lysine residues on proteins (as we showed here for K150 in IFIT5) and provide tailored probes to assay such lysines in more diverse experimental settings. We are encouraged by the initial potency and selectivity observed for interactions such as compound **7a** with K150 of IFIT5, which may provide a path to the first chemical probes to study the contributions of this IFIT to antiviral immunity and other biological processes. Finally, the conservation of K150 across the broader IFIT family, combined with our initial evidence of divergent SARs for aminophilic compound interactions with K150 and K151 in IFIT5 and IFIT1, respectively, indicates the potential to create covalent probes with subtype selectivity for individual IFITs.

In summary, our in-depth chemical proteomic analysis of structurally diverse aminophilic chemotypes has uncovered many hundreds of ligandable lysines that include those residing at functional sites on proteins historically considered challenging to target with small molecules. We also show here how integrating these ligandability maps with human genetic information and cell-activation-state profiling can further refine our knowledge of lysines for which covalent modification by small molecules is likely to affect the activity of proteins. By defining the aminophilic chemotypes that prefer to react with such ligandable and functional lysines, our study provides attractive starting points for chemical probe development for a diverse array of proteins in the human proteome.

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/s41557-021-00765-4>.

Received: 17 September 2020; Accepted: 5 July 2021;

Published online: 9 September 2021

References

- Schreiber, S. L. et al. Advancing biological understanding and therapeutics discovery with small-molecule probes. *Cell* **161**, 1252–1265 (2015).
- Hopkins, A. L. & Groom, C. R. The druggable genome. *Nat. Rev. Drug Discov.* **1**, 727–730 (2002).
- Macarron, R. et al. Impact of high-throughput screening in biomedical research. *Nat. Rev. Drug Discov.* **10**, 188–195 (2011).
- Scott, D. E., Coyne, A. G., Hudson, S. A. & Abell, C. Fragment-based approaches in drug discovery and chemical biology. *Biochemistry* **51**, 4990–5003 (2012).
- Johnson, D. S., Weerapana, E. & Cravatt, B. F. Strategies for discovering and derisking covalent, irreversible enzyme inhibitors. *Future Med. Chem.* **2**, 949–964 (2010).
- Backus, K. M. et al. Proteome-wide covalent ligand discovery in native biological systems. *Nature* **534**, 570–574 (2016).
- Hacker, S. M. et al. Global profiling of lysine reactivity and ligandability in the human proteome. *Nat. Chem.* **9**, 1181–1190 (2017).
- Ward, C. C., Kleinman, J. I. & Nomura, D. K. NHS-esters as versatile reactivity-based probes for mapping proteome-wide ligandable hotspots. *ACS Chem. Biol.* **12**, 1478–1483 (2017).
- Bachovchin, D. A. & Cravatt, B. F. The pharmacological landscape and therapeutic potential of serine hydrolases. *Nat. Rev. Drug Discov.* **11**, 52–68 (2012).
- Kato, D. et al. Activity-based probes that target diverse cysteine protease families. *Nat. Chem. Biol.* **1**, 33–38 (2005).
- 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).
- Walker, C. J. et al. Preclinical and clinical efficacy of XPO1/CRM1 inhibition by the karyopherin inhibitor KPT-330 in Ph+ leukemias. *Blood* **122**, 3034–3044 (2013).
- Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A. & Shokat, K. M. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* **503**, 548–551 (2013).
- 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).
- 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).
- Shannon, D. A. et al. Investigating the proteome reactivity and selectivity of aryl halides. *J. Am. Chem. Soc.* **136**, 3330–3333 (2014).
- Choi, S., Connelly, S., Reixach, N., Wilson, I. A. & Kelly, J. W. Chemoselective small molecules that covalently modify one lysine in a non-enzyme protein in plasma. *Nat. Chem. Biol.* **6**, 133–139 (2010).
- 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).
- Suh, E. H. et al. Stilbene vinyl sulfonamides as fluorogenic sensors of and traceless covalent kinetic stabilizers of transthyretin that prevent amyloidogenesis. *J. Am. Chem. Soc.* **135**, 17869–17880 (2013).
- Hunter, M. J. & Ludwig, M. L. The reaction of imidoesters with proteins and related small molecules. *J. Am. Chem. Soc.* **84**, 3491–3504 (1962).
- Nakamura, T., Kawai, Y., Kitamoto, N., Osawa, T. & Kato, Y. Covalent modification of lysine residues by allyl isothiocyanate in physiological conditions: plausible transformation of isothiocyanate from thiol to amine. *Chem. Res. Toxicol.* **22**, 536–542 (2009).
- 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).
- Akçay, G. et al. Inhibition of Mcl-1 through covalent modification of a noncatalytic lysine side chain. *Nat. Chem. Biol.* **12**, 931–936 (2016).
- Pettinger, J. et al. An irreversible inhibitor of HSP72 that unexpectedly targets lysine-56. *Angew. Chem. Int. Ed.* **56**, 3536–3540 (2017).
- Cuesta, A. & Taunton, J. Lysine-targeted inhibitors and chemoproteomic probes. *Ann. Rev. Biochem.* **88**, 365–381 (2019).
- Wang, C., Weerapana, E., Blewett, M. M. & Cravatt, B. F. A chemoproteomic platform to quantitatively map targets of lipid-derived electrophiles. *Nat. Methods* **11**, 79–85 (2014).
- Weerapana, E. et al. Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790–795 (2010).
- Ma, N. et al. 2H-azirine-based reagents for chemoselective bioconjugation at carboxyl residues inside live cells. *J. Am. Chem. Soc.* **142**, 6051–6059 (2020).
- Bach, K., Beerkens, B. L. H., Zanon, P. R. A. & Hacker, S. M. Light-activatable, 2,5-disubstituted tetrazoles for the proteome-wide profiling of aspartates and glutamates in living bacteria. *ACS Cent. Sci.* **6**, 546–554 (2020).
- Cheng, K. et al. Tetrazole-based probes for integrated phenotypic screening, affinity-based proteome profiling, and sensitive detection of a cancer biomarker. *Angew. Chem. Int. Ed.* **56**, 15044–15048 (2017).
- Lin, S. et al. Redox-based reagents for chemoselective methionine bioconjugation. *Science* **355**, 597–602 (2017).
- Hahm, H. S. et al. Global targeting of functional tyrosines using sulfur-triazole exchange chemistry. *Nat. Chem. Biol.* **16**, 150–159 (2020).
- Balthaser, B. R., Maloney, M. C., Beeler, A. B., Porco, J. A. & Snyder, J. K. Remodelling of the natural product fumagillol employing a reaction discovery approach. *Nat. Chem.* **3**, 969–973 (2011).
- Lajkiewicz, N. J., Cognetta, A. B., Niphakis, M. J., Cravatt, B. F. & Porco, J. A. Remodeling natural products: chemistry and serine hydrolase activity of a rocaglate-derived β -lactone. *J. Am. Chem. Soc.* **136**, 2659–2664 (2014).
- Lipinski, C. A. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov. Today Technol.* **1**, 337–341 (2004).
- Patricelli, M. P., Giang, D. K., Stamp, L. M. & Burbaum, J. J. Direct visualization of serine hydrolase activities in complex proteomes using fluorescent active site-directed probes. *Proteomics* **1**, 1067–1071 (2001).
- Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective ‘ligation’ of azides and terminal alkynes. *Angew. Chem. Int. Ed.* **41**, 2596–2599 (2002).
- Zhang, Z. et al. Genomic variations of the mevalonate pathway in porokeratosis. *eLife* **4**, e06322 (2015).
- Brooks, S. S. et al. A novel ribosomopathy caused by dysfunction of RPL10 disrupts neurodevelopment and causes X-linked microcephaly in humans. *Genetics* **198**, 723–733 (2014).

40. Lee, D.-S. et al. Structural basis of hereditary coproporphyrin. *Proc. Natl Acad. Sci. USA* **102**, 14232–14237 (2005).
41. Hussey, A. J. & Hayes, J. D. Characterization of a human class-Theta glutathione S-transferase with activity towards 1-menaphthyl sulphate. *Biochem. J.* **286**, 929–935 (1992).
42. Schmiedeknecht, G. et al. Isolation and characterization of a 14.5-kDa trichloroacetic-acid-soluble translational inhibitor protein from human monocytes that is upregulated upon cellular differentiation. *Eur. J. Biochem.* **242**, 339–351 (1996).
43. Katritzky, A. R. & Yousaf, T. I. A C-13 nuclear magnetic resonance study of the pyrimidine synthesis by the reactions of 1,3-dicarbonyl compounds with amidines and ureas. *Can. J. Chem.* **64**, 2087–2093 (1986).
44. Kragelund, B. B., Weterings, E., Hartmann-Petersen, R. & Keijzers, G. The Ku70/80 ring in non-homologous end-joining: easy to slip on, hard to remove. *Front. Biosci.* **21**, 514–527 (2016).
45. Tung, C. L., Wong, C. T. T., Fung, E. Y. M. & Li, X. Traceless and chemoselective amine bioconjugation via phthalimidine formation in native protein modification. *Org. Lett.* **18**, 2600–2603 (2016).
46. Adhikari, S. et al. Colorimetric and fluorescence probe for the detection of nano-molar lysine in aqueous medium. *Org. Biomol. Chem.* **14**, 10688–10694 (2016).
47. Bar-Peled, L. et al. Chemical proteomics identifies druggable vulnerabilities in a genetically defined cancer. *Cell* **171**, 696–709.e623 (2017).
48. Zhang, X., Crowley, V. M., Wucherpennig, T. G., Dix, M. M. & Cravatt, B. F. Electrophilic PROTACS that degrade nuclear proteins by engaging DCAF16. *Nat. Chem. Biol.* **15**, 737–746 (2019).
49. Vinogradova, E. V. et al. An activity-guided map of electrophile–cysteine interactions in primary human T cells. *Cell* **182**, 1009–1026 (2020).
50. Shi, C., Qiao, S., Wang, S., Wu, T. & Ji, G. Recent progress of lysophosphatidylcholine acyltransferases in metabolic disease and cancer. *Int. J. Clin. Exp. Med.* **11**, 8941–8953 (2018).
51. Zou, C. et al. LPS impairs phospholipid synthesis by triggering β -transducin repeat-containing protein (β -TRCP)-mediated polyubiquitination and degradation of the surfactant enzyme acyl-coa:Lysophosphatidylcholine acyltransferase 1 (LPCAT1). *J. Biol. Chem.* **286**, 2719–2727 (2011).
52. Rieckmann, J. C. et al. Social network architecture of human immune cells unveiled by quantitative proteomics. *Nat. Immunol.* **18**, 583–593 (2017).
53. Fensterl, V. & Sen, G. C. Interferon-induced IFIT proteins: their role in viral pathogenesis. *J. Virol.* **89**, 2462–2468 (2015).
54. Lo, U.-G. et al. Interferon-induced IFIT5 promotes epithelial-to-mesenchymal transition leading to renal cancer invasion. *Am. J. Clin. Exp. Urol.* **7**, 31–45 (2019).
55. Abbas, Y. M., Pichlmair, A., Górna, M. W., Superti-Furga, G. & Nagar, B. Structural basis for viral 5'-PPP-RNA recognition by human IFIT proteins. *Nature* **494**, 60–64 (2013).
56. Speers, A. E., Adam, G. C. & Cravatt, B. F. Activity-based protein profiling in vivo using a copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition. *J. Am. Chem. Soc.* **125**, 4686–4687 (2003).
57. Krüger, D. M., Neubacher, S. & Grossmann, T. N. Protein–RNA interactions: structural characteristics and hotspot amino acids. *RNA* **24**, 1457–1465 (2018).
58. Zanon, P. R. A. et al. Profiling the proteome-wide selectivity of diverse electrophiles. Preprint at <https://doi.org/10.26434/chemrxiv.14186561.v1> (2021).
59. Congreve, M., Carr, R., Murray, C. & Jhoti, H. A 'rule of three' for fragment-based lead discovery? *Drug Discov. Today* **8**, 876–877 (2003).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2021, corrected publication 2021

Methods

Cell lines. All cell lines were purchased from ATCC, tested negative for mycoplasma contamination and were used without further authentication. HEK293T (CRL-3216) and MDA-MB-231 (HTB-26) cells were maintained at 37 °C with 5% CO₂ in DMEM (Corning, 15-013-CV) supplemented with 10% (v/v) fetal bovine serum (FBS, Omega Scientific, FB-11, Lot no. 441224), penicillin (100 U ml⁻¹), streptomycin (100 µg ml⁻¹) and L-glutamine (2 mM). Ramos (CRL-1596) cells were grown at 37 °C in a humidified 5% CO₂ atmosphere in RPMI-1640 medium (Corning, 15-040-CV) supplemented with 10% (v/v) FBS, penicillin (100 U ml⁻¹), streptomycin (100 µg ml⁻¹) and L-glutamine (2 mM). All the cell lines were maintained at a low passage number (≤10 passages).

Isolation of primary human T cells and peripheral blood mononuclear cells.

All the studies with primary human cells were performed with samples from human volunteers following protocols approved by The Scripps Research Institute Institutional Review Board. Blood from healthy donors (age 18 to 65 years) was obtained after informed donor consent. PBMCs were isolated over a Lymphoprep (STEMCELL Technologies, 07851) gradient using slightly modified manufacturer's instructions. Briefly, 25 ml of freshly isolated blood was carefully layered on top of 12.5 ml of Lymphoprep in a 50 ml Falcon tube, minimizing the mixing of blood with Lymphoprep. The tubes were centrifuged (931g, 20 min, 23 °C, with brakes off) and the plasma with Lymphoprep layers that contained PBMCs was transferred to new 50 ml Falcon tubes and diluted (2:1) with Dulbecco's phosphate-buffered saline (DPBS, VWR, 45000-434). The cells were pelleted (524g, 8 min, 4 °C) and washed with DPBS (20 ml). T cells were isolated by negative selection from freshly isolated PBMCs using an EasySep Human T Cell Isolation Kit (STEMCELL Technologies, 17951) according to the manufacturer's instructions.

Preparation of human cancer cell proteome for gel- and MS-based ABPP analysis.

Cells were grown to 95% confluence for MDA-MB-231 or until the cell density reached 2 × 10⁶ cells ml⁻¹ for Ramos. Cells were washed and scraped with cold DPBS, and cell pellets were isolated by centrifugation (1,400g, 3 min, 4 °C). Cell pellets were either directly processed or kept frozen at -80 °C until further use. Cell pellets were next lysed using a Branson Sonifier probe sonicator (14 pulses, 30% duty cycle, output setting = 4) and fractionated (100,000g, 45 min) to yield soluble (supernatant) and membrane (pellet) fractions, which were then adjusted to a final protein concentration of 1.8 mg ml⁻¹ for competitive isoTOP-ABPP experiments. Membrane pellets were resuspended in cold DPBS after separation by sonication. For gel-ABPP experiments, the protein concentration was adjusted to 1.0 mg ml⁻¹ for MDA-MB-231 and Ramos cell lysates, or HEK293T cell lysates that expressed the target proteins. The lysates were prepared fresh from frozen cell pellets directly before each experiment. Protein concentration was determined using the DC Protein Assay (Bio-Rad) and absorbance read using a Tecan Infinite F500 plate reader following manufacturer's instructions.

Activation of primary human T cells for MS-based ABPP analysis.

Non-tissue-culture-treated 6-well plates were precoated with αCD3 (5 µg ml⁻¹, BioXCell) and αCD28 antibodies (2 µg ml⁻¹, BioXCell) in DPBS (2 ml per well) and kept at 4 °C overnight. The plates were then transferred to an incubator (37 °C in a humidified 5% CO₂ atmosphere) for 1 h and washed with DPBS (2 × 5 ml per well). Freshly isolated T cells were resuspended in RPMI-1640 medium supplemented with 10% FBS, penicillin (100 U ml⁻¹), streptomycin (100 µg ml⁻¹) and L-glutamine (2 mM) at 1 × 10⁶ cells ml⁻¹, plated into pre-coated 6-well plates (8 ml per well) and kept at 37 °C in a humidified 5% CO₂ atmosphere for 3 days. Activated T cells were then combined into 50 ml Falcon tubes, pelleted (524g, 8 min, 4 °C), washed with DPBS (10 ml) and the cell pellets were flash-frozen and stored at -80 °C until in vitro treatments with lysine-reactive electrophiles.

Stimulation of human PBMCs for MS-based ABPP analysis. Freshly isolated PBMCs were resuspended in RPMI-1640 medium supplemented with 10% FBS, penicillin (100 U ml⁻¹), streptomycin (100 µg ml⁻¹) and L-glutamine (2 mM) to a cell density of 2 × 10⁶ cells ml⁻¹. PBMCs were then treated with bacterial LPS (100 ng ml⁻¹, Sigma-Aldrich, L2630, from *Escherichia coli* O111:B4) over a period of 18 h at 37 °C in a humidified 5% CO₂ atmosphere. Stimulated PBMCs were next combined into 50 ml Falcon tubes, pelleted (524g, 8 min, 4 °C), washed with DPBS (10 ml) and the cell pellets were flash-frozen and stored at -80 °C until in vitro treatments with lysine-reactive compounds.

In vitro treatment of cell lysates with lysine-reactive compounds. Lysine-reactive compounds were prepared as either 2, 5 or 10 mM stock solutions in DMSO (Sigma-Aldrich, D8418) and were used at a final concentration of 20, 50 or 100 µM, respectively. For each profiling sample, 500 µl of soluble or membrane proteomes (1.8 mg ml⁻¹) were treated with 5 µl of the 2, 5 or 10 mM fragment stock solutions or 5 µl of DMSO vehicle for 1 h at 23 °C. Samples were next labelled with 100 µM of lysine-reactive P1 (5 µl of a 10 mM stock solution in DMSO) for 1 h at 23 °C. Samples were then conjugated by CuAAC, as described below.

In situ treatment of live cells with lysine-reactive electrophiles. MDA-MB-231 cells were grown to 95% confluence and Ramos cells were grown to

2 × 10⁶ cells ml⁻¹ at the time of treatment. Cells were carefully washed with DPBS and replenished with fresh media that contained lysine-reactive compounds at the indicated concentrations or the DMSO vehicle, with the total DMSO content maintained below 0.3%. Cells were then harvested in cold DPBS by scraping, centrifuged (1,400g, 3 min, 4 °C) and the cell pellets were washed with cold DPBS (2 ×). Pellets were either directly processed or kept frozen at -80 °C until further use. Cell pellets were next resuspended in DPBS, lysed by sonication (14 pulses, 30% duty cycle, output setting = 4) and fractionated (100,000g, 45 min) to yield soluble and membrane fractions, which were then adjusted to a final protein concentration of 1.8 mg ml⁻¹. Fractions were treated with the lysine-reactive P1 at a final concentration of 100 µM and incubated for 1 h at 23 °C. Samples were then conjugated by CuAAC as described below.

CuAAC conjugation. Following the in vitro or in situ fragment treatment and subsequent probe labelling, samples (500 µl) were conjugated to either the light (fragment-treated) or heavy (DMSO-treated) isotopically labelled, TEV-cleavable biotin tags (TEV-tags) using a CuAAC reaction. CuAAC reagents were premixed prior to their addition to the proteome samples. TEV tags (light or heavy, 10 µl of 5 mM stock in DMSO to a final concentration of 100 µM), tris(benzyltriazolylmethyl)amine ligand (30 µl of 1.7 mM stock in DMSO/BuOH 1:4 to a final concentration of 100 µM), tris(2-carboxyethyl) phosphine hydrochloride (10 µl of freshly prepared 50 mM stock in H₂O to a final concentration of 1 mM) and CuSO₄ (10 µl of 50 mM stock in H₂O at a final concentration of 1 mM) were combined in an Eppendorf tube, vortexed and added to the proteomic samples (55 µl per 500 µl sample). The CuAAC reaction mixture that contained the heavy TEV tag was added to DMSO-treated samples and the CuAAC reaction mixture that contained the light TEV tag was added to fragment-treated samples. The reaction was allowed to proceed at 23 °C for 1 h, heavy and light samples were combined pairwise in 15 ml conical Falcon tubes on ice that contained 4 ml of MeOH (precooled to -80 °C), 1 ml of CHCl₃ (precooled to 0 °C) and 1 ml of H₂O (precooled to 4 °C). Eppendorf tubes from the reaction mixtures were washed with additional cold H₂O (1 ml each) and washes were added to the same Falcon tube to a final ratio of 4:4:1 (H₂O/MeOH/CHCl₃). After centrifugation (5,000g, 10 min, 4 °C), a protein disk formed at the interface of CHCl₃ and MeOH/H₂O layers. The top MeOH/H₂O layer was carefully aspirated without perturbing the disk, and additional MeOH (2 ml, precooled to -80 °C) was added and the suspension mixed by vortexing. The proteins were pelleted (5,000g, 10 min, 4 °C) and the resulting pellets were solubilized in 1.2% SDS in DPBS (1 ml) with sonication (Branson Sonifier probe sonicator, 10 pulses, 40% duty cycle, output setting = 4) and heating (95 °C, 5 min). The insoluble materials were further removed by an additional centrifugation step (5,000g, 10 min, 23 °C).

Streptavidin enrichment. The SDS-solubilized protein mixture (1 ml) was diluted with DPBS (4.5 ml) to a final SDS concentration of 0.2%. The streptavidin-agarose beads (Pierce, 20349; 100 µl slurry per sample) were washed with 10 ml of DPBS (3 ×) and resuspended in DPBS (0.5 ml per sample) prior to addition. The final mixture was rotated for 3 h at 23 °C. After this enrichment step, the beads were pelleted by centrifugation (2,000g, 2 min) and washed to remove non-specifically bound proteins (2 × 10 ml of 0.2% SDS in DPBS, 2 × 10 ml of DPBS and 2 × 10 ml of H₂O).

Trypsin and TEV digestion. The beads were transferred to Eppendorf tubes (2 × 500 µl of H₂O), pelleted (2,000g, 2 min) and resuspended in 6 M urea in DPBS (500 µl). To this slurry was added dithiothreitol (25 µl of a freshly prepared 200 mM stock in H₂O to a final concentration of 10 mM) and samples were incubated at 65 °C for 15 min. Then, iodoacetamide (25 µl of a freshly prepared 400 mM stock in H₂O to a final concentration of 20 mM) was added and samples were incubated at 37 °C with shaking for 30 min. The bead mixtures were next diluted with 800 µl of DPBS, pelleted by centrifugation (2,000g, 2 min) and washed with 2 M urea in DPBS (1 mL). The samples were resuspended in 2 M urea in DPBS (200 µl) and to this slurry was added sequencing grade trypsin (Promega, 2 µg in 4 µl of a trypsin resuspension buffer that contained 1 mM CaCl₂). The samples were allowed to digest overnight at 37 °C with shaking. The beads were pelleted (2,000g, 2 min) and the tryptic digest aspirated. The beads were then washed with DPBS (3 × 1 ml), H₂O (3 × 1 ml) and TEV buffer (500 µl, 50 mM Tris, pH 8, 0.5 mM EDTA, 1 mM dithiothreitol), and resuspended in TEV buffer (140 µl). TEV protease (4 µl per sample, 80 µM) was then added and the beads were incubated at 30 °C overnight with rotation. After the TEV digestion, the beads were pelleted by centrifugation (2,000g, 2 min) and the TEV digest was separated from the beads using Micro Bio-Spin columns (Bio-Rad) with centrifugation (800g, 30 s). The beads were washed with H₂O (100 µl, and centrifuged at 16,000g for 1 min) and the eluents (300 µl) were acidified by the addition of formic acid (0.1%, 15 µl per sample to a final concentration of 5% v/v) and stored at -80 °C prior to analysis.

Liquid chromatography-mass spectrometry analysis. TEV-digested samples were pressure loaded onto a 250 µm (inner diameter) fused silica capillary column packed with C18 resin (Aqua 5 µm, Phenomenex) and analysed by multidimensional liquid chromatography tandem (MudPIT) MS using an LTQ-Velos Orbitrap mass spectrometer (Thermo Scientific) coupled to an Agilent 1200-series quaternary pump. The peptides were eluted onto a biphasic

Table 1 | Amplification primers

Protein	Fwd/Rev	Sequence	Vector	Restriction sites	Affinity tag
CPOX	Fwd	CTGCAGGTCGACGACTTCGCTGGGGAGGCCGGAGG	pRK5	Sall/NotI	N-terminal FLAG
CPOX	Rev	GTTAGCGCGGCCGCTCAACGCACCCAGTCCCTTGGATGGC	pRK5	Sall/NotI	N-terminal FLAG
SIN3A	Fwd	TGGTGCCTCGACGATGAAGCGGCGTTTGGATGA	pRK5	Sall/NotI	N-terminal FLAG
SIN3A	Rev	TACTTAGCGGCCGCTTACACAGAATCAACCTTCTCAGCAGTTG	pRK5	Sall/NotI	N-terminal FLAG
SIN3B	Fwd	CTGCAGGTCGACGATGGCGCACGCTGGCGGTGG	pRK5	Sall/NotI	N-terminal FLAG
SIN3B	Rev	GTTAGCGCGGCCGCTCAGCTGGATCTATGCGCTTGCAGGATGC	pRK5	Sall/NotI	N-terminal FLAG
GSTT2B	Fwd	CTGCAGGTCGACGATGGGCCCTAGAGCTGTTC	pRK5	Sall/NotI	N-terminal FLAG
GSTT2B	Rev	GTTAGCGCGGCCGCTCAGGGGATCCTGGCGATTCT	pRK5	Sall/NotI	N-terminal FLAG
RIDA	Fwd	CTGCAGGTCGACGATGTCGTCCTTGATCAGAAG	pRK5	Sall/NotI	N-terminal FLAG
RIDA	Rev	GTTAGCGCGGCCGCTTATAGTGATGCCGTTGTC	pRK5	Sall/NotI	N-terminal FLAG
Ku70	Fwd	CTGCAGGTCGACGATGTCAGGGTGGGAGTCATAT	pRK5	Sall/NotI	N-terminal FLAG
Ku70	Rev	GTTAGCGCGGCCGCTCAGTCTGGAAGTGCTTGG	pRK5	Sall/NotI	N-terminal FLAG
Ku80	Fwd	CTGCAGGTCGACGATGGTGCGGTGGGGAATAAG	pRK5	Sall/NotI	N-terminal HA
Ku80	Rev	GTTAGCGCGGCCGCTTATATCATGTCCAATAAATCGTCCACATCAC	pRK5	Sall/NotI	N-terminal HA
IFIT1	Fwd	CTGCAGGTCGACGATGAGTACAAATGGTGATG	pRK5	Sall/NotI	N-terminal FLAG
IFIT1	Rev	GTTAGCGCGGCCGCTAAGGACCTTGTCTCACAG	pRK5	Sall/NotI	N-terminal FLAG
IFIT5	Fwd	CTGCAGGTCGACGATGAGTAAATTCGTAAGGAC	pRK5	Sall/NotI	N-terminal FLAG
IFIT5	Rev	GTTAGCGCGGCCGCTTAAATGAAAGTCGGAGCTC	pRK5	Sall/NotI	N-terminal FLAG

Fwd, forward; Rev, reverse.

column with a 5 μm tip (100 μm fused silica, packed with 10 cm of C18 resin and 4 cm of bulk strong cation exchange resin (SCX, Phenomenex) in a five-step MudPIT experiment using 0, 30, 60, 90 and 100% salt 'bumps' of 500 mM aqueous ammonium acetate and 5 \rightarrow 100% gradient of buffer B in buffer A (buffer A, 95% water, 5% acetonitrile, 0.1% formic acid; buffer B, 5% water, 95% acetonitrile, 0.1% formic acid) as previously described²⁷. The acquired data were collected in a data-dependent acquisition mode with dynamic exclusion enabled (20 s, repeat count of 2). One full MS (MS1) scan (400–1,800 m/z) was followed by 30 MS2 scans (ion trap mass spectrometry) of the n th most abundant ions.

Peptide identification and quantification. From each of the five raw files (one for each salt bum) generated by the instrument (Xcalibur software), the MS2 spectra for all fragmented parent ions were extracted from the raw file using RAW Converter (version 1.1.0.22, available at <http://fields.scripps.edu/rawconv/>). The generated MS2 spectral files (.ms2 files) were uploaded and searched using the ProLuCID algorithm (available at <http://fields.scripps.edu/downloads.php>) using a reverse concatenated, non-redundant (gene-centric) variant of the Human UniProt database (release-2012_11). Cysteine residues were searched with a static modification for S-carboxyamidomethylation (+57.02146). For all the competitive and reactivity profiling experiments, lysine residues were searched with up to one differential modification for either the light or heavy TEV tags (+464.24957 or +470.26338, respectively). Peptides were required to have at least one tryptic terminus and to contain the TEV modification. ProLuCID data were filtered through DTASelect (version 2.0) to achieve a peptide false-positive rate below 1%.

R-value calculation and data processing. The ratios of heavy (DMSO)/light (fragment treated) MS1 peaks (R values) for each unique peptide were quantified with in-house CIMAGE software⁷ using default parameters (3 MS1 acquisitions per peak and the signal-to-noise threshold set to 2.5). The site-specific engagement of lysine residues was assessed by the blockade of pentynoic acid sulfotetrafluorophenyl ester **PI** (Lumiprobe) labelling. A maximal ratio of 20 was assigned for peptides that showed a $\geq 95\%$ reduction in MS1 peak area in the fragment-treated proteome (light TEV tag) compared with that in the DMSO-treated (control) proteome (heavy TEV tag). Ratios for unique peptide sequences were calculated for each experiment; overlapping peptides with the same modified lysine (for example, different charge states, chromatographic elution times or tryptic termini) were grouped together and the median ratio was reported as the final ratio (R). Additionally, ratios for peptide sequences that contained multiple lysines were grouped together. When aggregating data across experimental replicates, the mean of each experimental median R was reported. The peptide ratios reported by CIMAGE were further filtered to ensure the removal or correction of low-quality ratios in each individual dataset. The quality filters applied were: (1) removal of peptides with co-elution correlation score R^2 values ≤ 0.8 , (2) removal of reverse peptide sequences, (3) removal of half-tryptic peptides, (4) removal of peptide sequences with tryptic-site modified lysines (for

example, K.K*, R.K*, K*.K, and K*.R), (5) removal of peptides with $R=20$ and only a single MS2 event triggered during the elution of the parent ion and (6) removal of peptides with $R=20$ and a coefficient of variation ≥ 0.6 . For peptide ratios with standard deviations $\geq 90\%$ from the median, the lowest ratio was taken instead of the median. For each biological replicate, the reported ratio of a given peptide is the median ratio. Across biological replicates for a single fragment: (1) peptides with $R=20$ are only reported if they were quantified and liganded ($R \geq 4 < 20$) in at least one other dataset across all datasets and (2) peptides with $R \geq 4 < 20$ are reported if peptides were quantified (but not necessarily liganded) in at least one other dataset across all datasets. The remaining peptides with $R=20$ were manually annotated. Where fragments are aggregated, the reported ratio for a given peptide is the median ratio across the biological replicates. Where chemotypes are aggregated, the reported ratio is the maximum ratio of the constituent fragments.

Recombinant expression of proteins by transient transfection. HEK293T cells were grown to 60% confluency under standard growth conditions in 10 cm tissue-culture dishes. To 5 μg of DNA diluted in 250 μl of serum-free DMEM was added 15 μl of aqueous polyethyleneimine 'MAX' (1 mg ml^{-1} , molecular mass 40,000, polyethylenimine; Polysciences, Inc.). 'Mock' transfected HEK293T cells were transfected with an empty pRK5 vector. The mixture was incubated at room temperature for 20 min and added dropwise to the cells. Cells were grown for 48 h at 37 $^{\circ}\text{C}$ in a humidified 5% CO_2 atmosphere. Cells were then harvested in cold DPBS by scraping, centrifuged (1,400g, 3 min, 4 $^{\circ}\text{C}$) and cell pellets were washed with cold DPBS (2 \times). Pellets were either directly processed or kept frozen at -80°C until further use. Cell pellets were next lysed by sonication (6 pulses, 30% duty cycle, output setting = 4) and fractionated (100,000g, 45 min) to yield soluble and membrane fractions, which were then adjusted to a final protein concentration of 1.0 mg ml^{-1} .

Subcloning and site-directed mutagenesis. Full-length genes that encoded the proteins of interest were PCR-amplified from a complementary (cDNA) library derived from low-passage HEK293T cells using the Ribozol RNA extraction reagent (Amresco) and the iScript Reverse Transcription Supermix kit (Bio-Rad). For the following proteins, cDNA clones were used for PCR-amplification: CPOX (OHu18833, GenScript), SIN3B (OHu28835, GenScript), IFIT3 (OHu10416, GenScript), and RIDA (OHu25061, GenScript). Gene products were subcloned into the pRK5 vector with a C-terminal FLAG tag using Sall (N-terminal) and NotI (C-terminal) restriction sites. DNA was amplified with custom forward and reverse primers (Table 1) using Phusion Polymerase (NEB, M0530S), following the manufacturers' instructions, digested with the indicated restriction enzyme and ligated into the pRK5 vector with the appropriate affinity tag. Lysine mutants were generated using QuikChange site-directed mutagenesis with Phusion High-Fidelity DNA Polymerase and custom primers that contained the desired mutations and their respective complements (Table 2). All clone sequences were verified (Eton Bioscience).

Table 2 | Mutagenesis primers

Protein	Mutation	Fwd/Rev	Sequence
CPOX	K404R	Fwd	GTATGATCGGGGCACAAGGTTTGGCCTCTC
CPOX	K404R	Rev	CCTTGTGCCCGATCATACAGCAGATTAATTC
CPOX	K404E	Fwd	GTATGATCGGGGCACAGAGTTTGGCCTCTC
CPOX	K404E	Rev	CTCTGTGCCCGATCATACAGCAGATTAATTC
SIN3A	K155R	Fwd	CCTTGACATCATGAAGGAATTCGTCTCAGAGCATCGACACCCC
SIN3A	K155R	Rev	GGGGTGTGATGCTCTGAGAACGAAATTCCTTCATGATGCAAGG
SIN3B	K73R	Fwd	AGATCATGAAGGAGTTCAGGAGCCAGAGCA
SIN3B	K73R	Rev	CCTGAACCTCTCATGATCTCCAGGAAGCC
GSTT2B	K75R	Fwd	CAACAGCCTGGGGAGGCTGCCGACGCTCAAG
GSTT2B	K75R	Rev	CCTCCCCAGGCTGTTGATCTGCAAGAATCC
RIDA	K117R	Fwd	CAAGTTGCTGCTTTACCCAGGGGACGCCGAATTG
RIDA	K117R	Rev	CCTGGGTTAAAGCAGCAACTTGGTAAAGCAGCTCTA
RIDA	K117E	Fwd	GTTGCTGCTTTACCCAGGGGACGCCGAATTG
RIDA	K117E	Rev	CTCGGGTTAAAGCAGCAACTTGGTAAAGCAGC
RIDA	K117Q	Fwd	GTTGCTGCTTTACCCAGGGGACGCCGAATTG
RIDA	K117Q	Rev	CTGGGGTTAAAGCAGCAACTTGGTAAAGCAGC
RIDA	K117I	Fwd	GTTGCTGCTTTACCCATCGGCAGCCGAATTG
RIDA	K117I	Rev	GATGGGTTAAAGCAGCAACTTGGTAAAGCAGC
Ku70	K351R	Fwd	GATGCTCATGGGTTTCAGGCCGTTGTAAGTCTG
Ku70	K351R	Rev	CCTGAAACCCATGAGCATCAAACCTGGATCATC
IFIT1	K151R	Fwd	GGAAGGATGGGCCTTGCTGAGGTGTGGAGGAAAG
IFIT1	K151R	Rev	CCTCAGCAAGGCCCATCTTCTCACAGTCTATTTTC
IFIT5	K150R	Fwd	GAAAGGCTGGGCACTCTGAGGTTTGGAGGAAAG
IFIT5	K150R	Rev	CCTCAAGAGTGCCAGCCTTCTCACAGTCAGTC

Western blot analysis. Cells were collected and lysed in a 1% NP-40 lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 10% glycerol, 1% Nonidet P-40) with a complete protease inhibitor cocktail (Roche). Cells were vortexed and sonicated (6 pulses, 30% duty cycle, output setting = 4), and the supernatant was collected after centrifugation (16,000g, 10 min, 4°C). Protein concentration was determined by a detergent-compatible assay (5000112, Bio-Rad). Protein lysate was heated at 95°C for 5 min in Laemmli sample buffer (1×). Proteins were resolved by 12 or 14% Novex Tris-glycine mini gels (Invitrogen) and transferred to 0.45 μm nitrocellulose membrane (GE Healthcare). The membrane was blocked with 5% milk in Tris-buffered saline (20 mM Tris-HCl, pH 7.6, 150 mM NaCl) with Tween (TBST) buffer (0.1% Tween 20, 20 mM Tris-HCl, pH 7.6, 150 mM NaCl) for 1 h at 23°C with gentle rocking. The primary antibody (anti-FLAG) was diluted (1:5000) with 5% milk in TBST buffer and incubated with the membrane for 1 h at 23°C or overnight at 4°C with gentle rocking. The membrane was washed with TBST buffer (3×, 5 min) and incubated with the secondary antibody (1:5000 dilution in 5% milk in TBST buffer) for 1 h at 23°C with gentle rocking. The membrane was washed with TBST buffer (3×, 5 min) and western blots were visualized on a LICOR Odyssey scanner. Relative band intensities were quantified using ImageJ software (<https://imagej.nih.gov/ij/>).

RIDA deiminase activity assay. Soluble proteome (100 μl, 1.0 mg ml⁻¹) from HEK293T cells that express human RIDA (WT or K117R, K117Q, K117E or K117I mutants) or mock transfected cells (empty vector, negative control) were prepared in a 50 mM potassium pyrophosphate (pH 8.5) assay buffer and added into a clear-bottom 96-well plate. For compound treatments, 1.0 μl of the lysine-reactive compound (in DMSO) or 1.0 μl of DMSO (positive control) were added and the reactions were incubated for 1 h at 23°C. A mixture that contained 10 μl of semicarbazide-HCl (100 mM in assay buffer, Sigma-Aldrich, S2201), 10 μl of catalase from bovine liver (10 μg in an assay buffer, Sigma-Aldrich, C9322) and 10 μl of L-amino acid oxidase from *Crotalus adamanteus* (10 μg in assay buffer, Sigma-Aldrich, A9253) was added to each well and the reaction was started by the addition of 10 μl of L-methionine (2 mM in assay buffer). The absorbance of the semicarbazone formation was measured at 248 nm every minute for 20 min at 23°C.

Ku70 and Ku80 heterodimerization assay. HEK293T cell lysates that expressed human FLAG-tagged Ku70 (WT or K351R mutant) were lysed by sonication (5 pulses, 40% duty cycle, output setting = 4) in 1% NP-40 lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 10% glycerol, 1% Nonidet P-40) that contained a complete protease inhibitor cocktail (Roche). Samples were rotated for 30 min at 4°C to complete lysis, clarified by centrifugation (16,000g, 10 min, 4°C), and

protein concentration was measured using the DC Protein Assay (Bio-Rad) and normalized to 1.0 mg ml⁻¹. Normalized lysates of cells that expressed human HA-tagged Ku70 (WT or K351R mutant) were treated with lysine-reactive compounds or DMSO (control) at the indicated concentrations (1 h, 23°C) and then mixed with lysates that expressed the WT Ku80 protein (1.0 mg ml⁻¹ in 1% NP-40 buffer) for 1 h at 23°C. Samples were then co-immunoprecipitated with ANTI-FLAG M2 affinity gel (20 μl slurry per sample; Sigma-Aldrich, A2220) by rotation (1 h, 4°C), washed with 1.0 ml of 0.2% NP-40 washing buffer (4×, 25 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.2% Nonidet P-40), and heated at 95°C for 10 min in Laemmli sample buffer (2×), followed by western blot analysis with anti-HA immunoblotting.

RNA probe synthesis for IFIT pulldown assay. 3'-Biotinylated 5'-PPP and 5'-hydroxy (OH) RNA probes were synthesized by in vitro transcription using a MEGAScript T7 Transcription Kit (Invitrogen, AM1334), according to the product guidelines. A 300 nt double-stranded DNA oligo that contained a T7 promoter sequence was used as a template, and biotin-16-UTP (Roche, 11388908910) was incorporated into the reaction mixture at a 1:5 biotin-UTP:UTP ratio. Residual DNA was digested with Turbo DNase and biotinylated probes were subsequently purified using an RNeasy Mini Kit (Qiagen, no. 74104). 5'-OH-RNA probes were prepared by dephosphorylation of 5'-PPP RNA probes using calf intestinal alkaline phosphatase (NEB, M0290) and then purified using an RNeasy Mini Kit). Mock dephosphorylated 5'-PPP-RNA probes were prepared alongside 5'-OH-RNA, with the calf intestinal alkaline phosphatase replaced with water, and were shown to bind comparably to IFITs as untreated 5'-PPP-RNA probes.

IFIT pulldown assay. Affinity enrichment, resins were prepared by coupling biotinylated RNA probes to streptavidin resin (1 μg of RNA per 50 μl of agarose slurry). Streptavidin agarose was initially washed with TAP buffer (3×, 50 mM Tris-HCl pH 7.5, 100 mM NaCl, 5% glycerol, 0.2% Nonidet-P40, 1.5 mM MgCl₂) and then incubated with biotinylated probes for 1 h at 4°C. Unbound probe was removed by centrifugation, and the coupled resin was washed with TAP buffer (1×), prior to dilution for the addition to cell lysates. HEK293T cells that expressed FLAG-tagged IFIT1/5 (WT or K151R and/or K150R mutants) were resuspended in TAP buffer, which contained complete EDTA-free protease inhibitor tablets (Roche, 04693159001), and allowed to lyse on ice for 15 min. Lysates were briefly sonicated (5 pulses, 40% duty cycle, output setting = 4), and then cleared by centrifugation (16,000g, 5 min, 4°C). The soluble proteome from IFIT5 (WT or K150R mutant) was normalized to 0.25 mg ml⁻¹ or 1.0 mg ml for IFIT1 (WT or K151R mutant) and treated with lysine-reactive compounds (5 μl) or DMSO (5 μl, control) at the indicated concentrations for 1 h at 23°C. Pulldown assays

were carried out by rotating samples with 75 μl of IFIT1 (or 50 μl of IFIT5) of RNA-coupled streptavidin for 2 h at 4 °C. Resins were washed with TAP buffer (4 \times , 1.0 ml), and bound proteins were eluted with 2 \times SDS-PAGE sample buffer. Samples were heated (10 min, 95 °C) and resolved by gel electrophoresis on Novex 10% Tris-glycine precast gels (Invitrogen), followed by western blot analysis.

LPCAT1 acyltransferase assay. HEK293T cells that expressed human LPCAT1 (WT or K221R mutant) were resuspended in assay buffer (10 mM Tris-HCl pH 7.4, 1 mM EDTA, 150 mM NaCl) and lysed by sonication using a probe sonicator (15 pulses, 30% duty cycle, output setting = 3). The lysate was centrifuged (16,000g, 45 min, 4 °C) to collect the membrane fraction. The membrane pellet was then resuspended in assay buffer by sonication (5 pulses, 30% duty cycle, output setting = 3) and diluted to 0.05 mg ml⁻¹. For the acyltransferase assay, 100 μl of 0.05 mg ml⁻¹ lysate was treated with lysine-reactive compounds at the indicated concentrations (1 h, 23 °C). After the incubation, 10 μl of an 11 \times substrate cocktail (550 μM 15:0 lyso-PC and 550 μM 10:0 CoA in assay buffer; Avanti Polar Lipids) was added to the sample and incubated for 10 min at 23 °C. The reaction was quenched by adding 300 μl of CHCl₃/MeOH (2:1, v/v) that contained 1 nmol of phosphatidylcholine (PC) (12:0/12:0; Avanti Polar Lipids) as an internal standard. The suspension was vortexed vigorously and centrifuged (2,000g, 5 min, 4 °C). The bottom layer (150 μl) was collected and mixed with 75 μl of MeOH, and 2.5 μl of the extract was used for MS analysis to measure the production of PC (15:0/10:0). The amount of PC (15:0/10:0) was quantified using an LC-MS-based multiple reaction monitoring method in positive mode (Agilent Technologies 6460 Triple Quad). MS analysis was performed using electrospray ionization with the following parameters: drying gas temperature, 350 °C; drying gas flow, 9 l min⁻¹; nebulizer pressure, 45 p.s.i. (310 kPa); sheath gas temperature, 375 °C; sheath gas flow, 10 l min⁻¹; fragmentor voltage, 100 V; capillary voltage, 3.5 kV. Ammonium acetate (20 mM in H₂O) and ammonium acetate (20 mM in MeOH) were used as buffer A and B, respectively. After injection, the LC gradient was: start from 90% B at 0.8 ml min⁻¹, increase to 99% B at 0.8 ml min⁻¹ for 5 min, stay at 99% B at 0.8 ml min⁻¹ for 1 min, return to 90% B at 0.8 ml min⁻¹, and then equilibrate for 1.5 min. The multiple reaction monitoring transitions for PC (15:0/10:0) and PC (12:0/12:0) were 636.5 \rightarrow 184.1 and 622.4 \rightarrow 184.1, respectively. The amount of PC (15:0/10:0) was quantified by measuring areas under the curve in comparison with those for the corresponding PC (12:0/12:0) curve. The hydrolysis activity of LPCAT1 (WT or K221R mutant) was calculated by normalizing to the amount of PC (15:0/10:0) produced against the proteome amount and the incubation time.

LPCAT1 ubiquitination assay. HA-tagged ubiquitin (2 μg) and FLAG-tagged LPCAT1 (2 μg , WT or K221R mutant) or an empty FLAG-tagged pRK5 vector (2 μg , control) or FLAG-tagged green fluorescent protein (2 μg , control) were co-expressed in HEK293T cells prior to treatment with or without proteasome inhibitor MG132 (10 μM , SelleckChem) for 2 or 14 h (37 °C, 5% CO₂). Cells were then collected and lysed by sonication (5 pulses, 40% duty cycle, output setting = 4) in 1% NP-40 lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 10% glycerol, 1% Nonidet P-40) that contained complete protease inhibitor cocktail (Roche). Samples were rotated for 30 min at 4 °C to complete lysis, clarified by centrifugation (16,000g, 10 min, 4 °C) and the protein concentration was measured using the DC Protein Assay (Bio-Rad) and normalized to 1.0 mg ml⁻¹. Samples were then co-immunoprecipitated with ANTI-FLAG M2 affinity gel (20 μl of slurry per sample; Sigma-Aldrich, A2220) by rotation (1 h, 4 °C), washed with 1.0 ml of 0.2% NP-40 washing buffer (4 \times , 25 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.2% Nonidet P-40) and heated at 95 °C for 10 min in Laemmli sample buffer (2 \times), followed by western blot analysis with anti-HA immunoblotting. For the endogenous ubiquitination of LPCAT1, FLAG-tagged LPCAT1 (5 μg , WT or K221R mutant) were overexpressed in HEK293T cells prior to treatment with MG132 (10 μM , 2 h). Cell lysates were subjected to anti-FLAG immunoprecipitation as described above, and the affinity-enriched precipitates were analysed by anti-Ubiquitin immunoblotting.

Calculation of relative activity or percent inhibition. For RIDA, the slope of the linear regression of the linear portion of the absorbance over time was used as the measure of activity. Apparent activity was calculated relative to the WT. Percentage inhibition was calculated relative to the positive and negative control and used to calculate IC₅₀ values by non-linear regression analysis from a dose-response curve generated using GraphPad Prism 7. For quantification of the inhibition and

apparent IC₅₀ determination in competitive gel-ABPP experiments, the percentage of labelling was determined by quantifying the integrated optical intensity of the bands using ImageLab 5.2.1 software (Bio-Rad).

Statistical analysis. Unless otherwise stated, quantitative data are expressed in bar and line graphs with mean \pm s.d. (error bars) shown. Differences between two groups were examined using an unpaired two-tailed Student's *t*-test with equal or unequal variance as noted. Significant *P* values are indicated (**P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001).

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

Data availability

All data associated with this study are available in the published article and its Supplementary Information. All raw proteomics data have been uploaded to the PRIDE repository with PRIDE ID PXD025829. Source data are provided with this paper.

Code availability

Custom code used for proteomic data processing are available at <https://github.com/cravattlab/abbasov>.

References

- Sander, T., Freyss, J., von Korff, M. & Rufener, C. DataWarrior: an open-source program for chemistry aware data visualization and analysis. *J. Chem. Inf. Model.* **55**, 460–473 (2015).
- Herdendorf, T. J. & Mizioro, H. M. Functional evaluation of conserved basic residues in human phosphomevalonate kinase. *Biochemistry* **46**, 11780–11788 (2007).

Acknowledgements

This work was supported by the NIH (CA231991, AI-126592), a Hewitt Foundation for Medical Research Fellowship (M.E.A.), a Sir Henry Wellcome Postdoctoral Fellowship, Wellcome Trust (M.E.K.), Pfizer and Vividion Therapeutics.

Author contributions

M.E.A. and B.F.C. conceived the research, designed the experiments and analysed the data. M.E.A. performed mass spectrometry experiments and data analysis. M.E.A. designed and synthesized compounds, cloned, expressed and purified the proteins, and conducted RIDA and Ku70-Ku80 studies. M.E.A., M.R.L., R.S. and M.M.D. compiled and analysed mass spectrometry data. M.E.K. and M.E.A. conducted IFIT1 and IFIT5 studies and data analysis. T.-A.I. and M.E.A. conducted LPCAT1 studies and data analysis. M.E.A., Y.T. and V.M.C. characterized synthetic compounds. V.M.C. conducted reactivity studies with a model amine nucleophile. C.Wa.E. and M.M.H. designed and synthesized compounds 17a–17i and 17r, 18a–18c and 19a–19g. S.M.H. designed and synthesized compounds 17m–17q, 17s and 17t. J.H. and L.L.K. designed and synthesized compounds 33a and 33b. M.E.A. and B.F.C. wrote the manuscript.

Competing interests

B.F.C. is a founder and scientific advisor to Vividion Therapeutics, a biotechnology company interested in developing small-molecule therapeutics.

Additional information

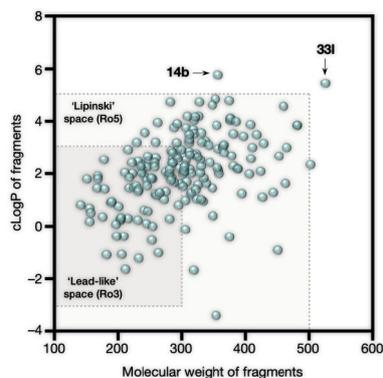
Extended data is available for this paper at <https://doi.org/10.1038/s41557-021-00765-4>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41557-021-00765-4>.

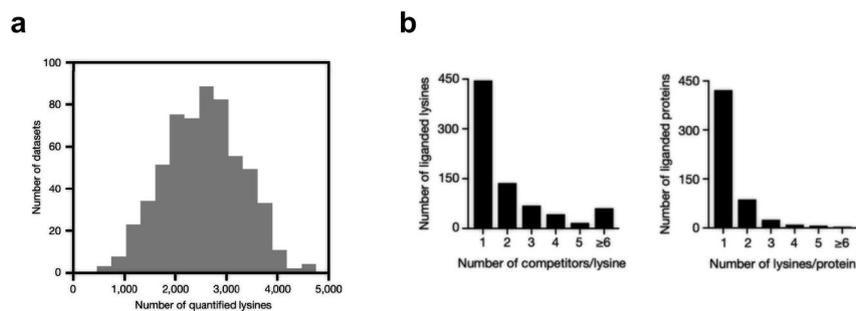
Correspondence and requests for materials should be addressed to Mikail E. Abbasov or Benjamin F. Cravatt.

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

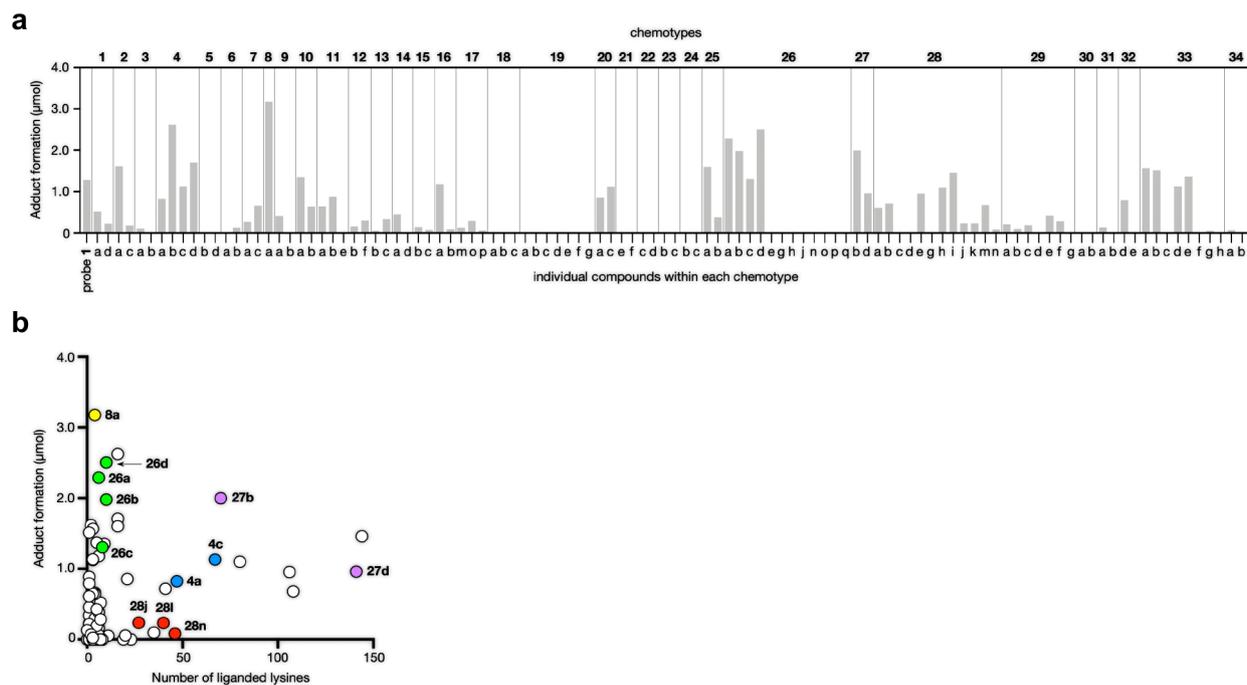
Reprints and permissions information is available at www.nature.com/reprints.



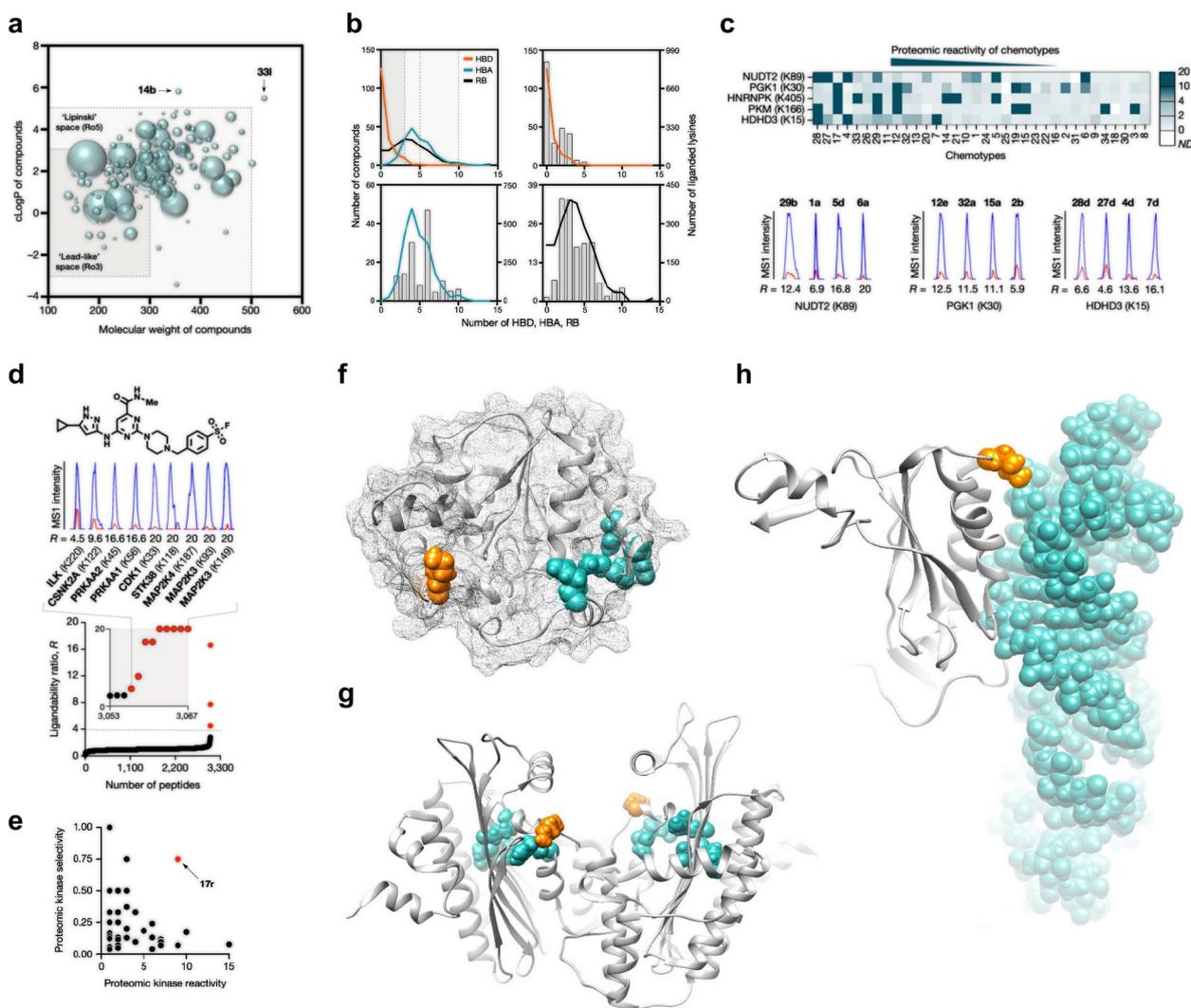
Extended Data Fig. 1 | Properties of aminophilic compound library for mapping small molecule-lysine interactions in the proteome. cLogP versus molecular weight plot showing aminophilic compounds that follow Lipinski's drug-likeness 'rule-of-five' (Lipinski space Ro5, light gray box)³⁵ and Congreve's fragment-based lead-likeness 'rule-of-three' (Lead-like space Ro3)^{59,60}.



Extended Data Fig. 2 | Features of aminophilic compound-lysine interaction map in human cancer cell proteomes. a, Histogram showing number of quantified lysines across all isoTOP-ABPP datasets. **b**, Number of aminophilic compound hits per liganded lysine (left) and the number liganded lysines per protein (right). The results shown are average ratios from three experiments ($n = 3$ biologically independent experiments).

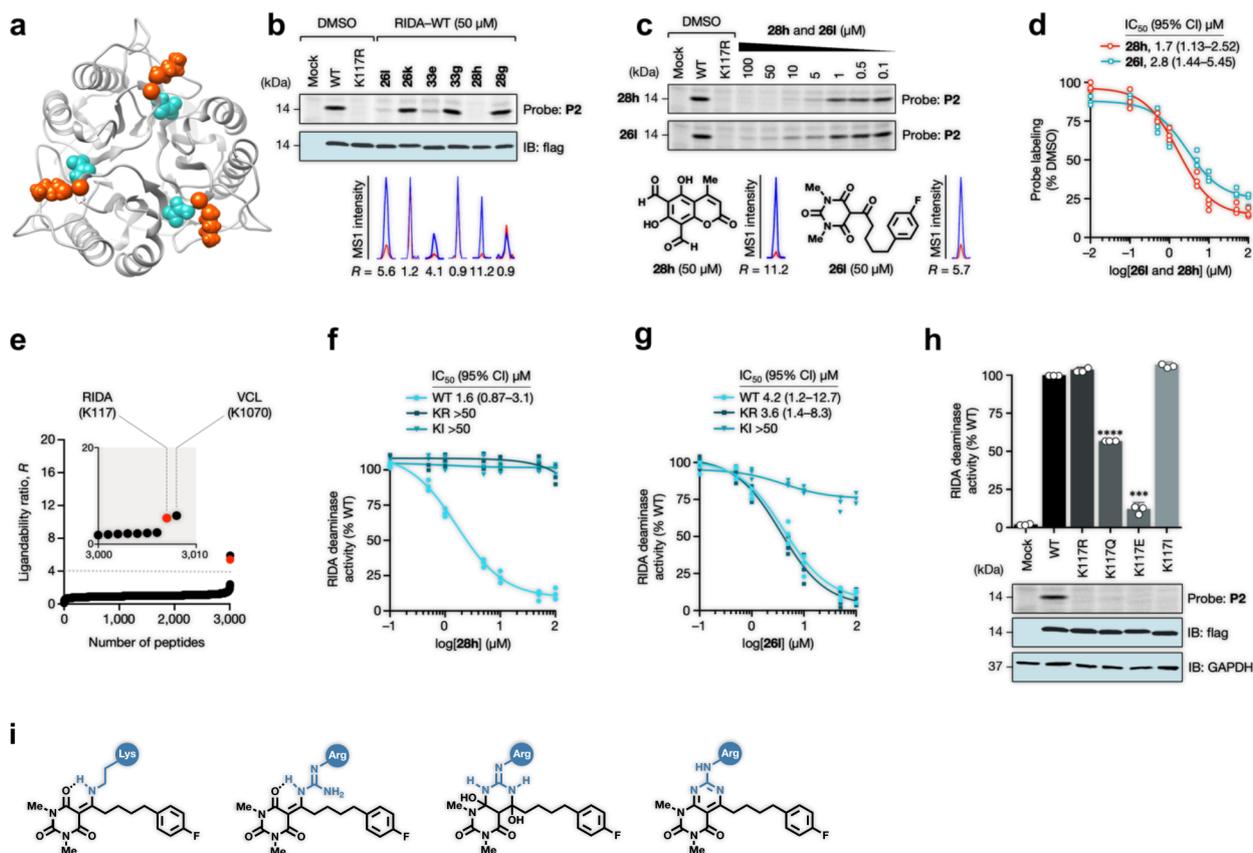


Extended Data Fig. 3 | Reactivity profiles of representative aminophilic compounds with a model amine nucleophile. a, Aminophilic compounds (125 μM) were incubated at room temperature with the amine nucleophile $\text{N}\alpha$ -acetyl-L-lysine-OMe (2 M, 1 h, at pH 10 (0.05 M NaHCO_3)). All samples contained 5 μM $\text{N}\alpha$ -acetyl-L-methionine-OH as an internal standard. Samples were neutralized with formic acid and 20 μL of the resulting solution was inject on to an Agilent 6100 series single quadrupole LC/MS system. Samples were run with the following gradient of Buffer A (95/5 Water/MeCN with 0.1% formic acid) and Buffer B (5/95 Water/MeCN with 0.1% formic acid): 100% A from 0–1 min, 100% A \rightarrow 100% B from 1–11 min, 100% B from 11–13 min, and 100% A from 13–15 min. Peaks corresponding to the amine nucleophile adducts were quantified using Agilent Open Lab software. **b**, Correlation plot comparing amine nucleophile adduct formation to liganded lysines for each compound (also see Supplementary Table 1). Representative aminophilic compound chemotypes are color-coded. For **a** and **b**, data represent average values \pm SD; $n=2$ per group ($n=2$ independent experiments).

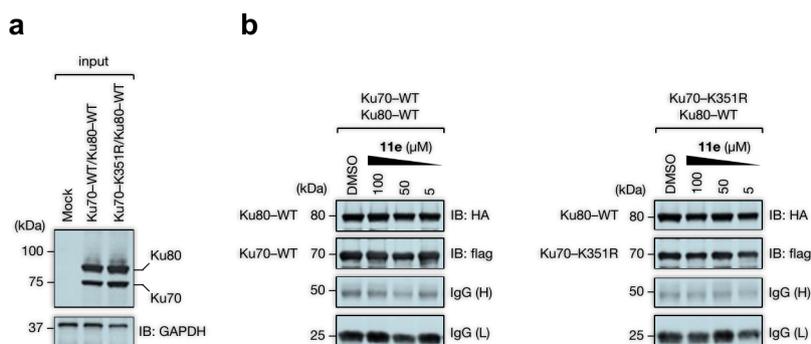


Extended Data Fig. 4 | Relating aminophilic compound-lysine interaction map to compound properties and representative features of liganded lysines.

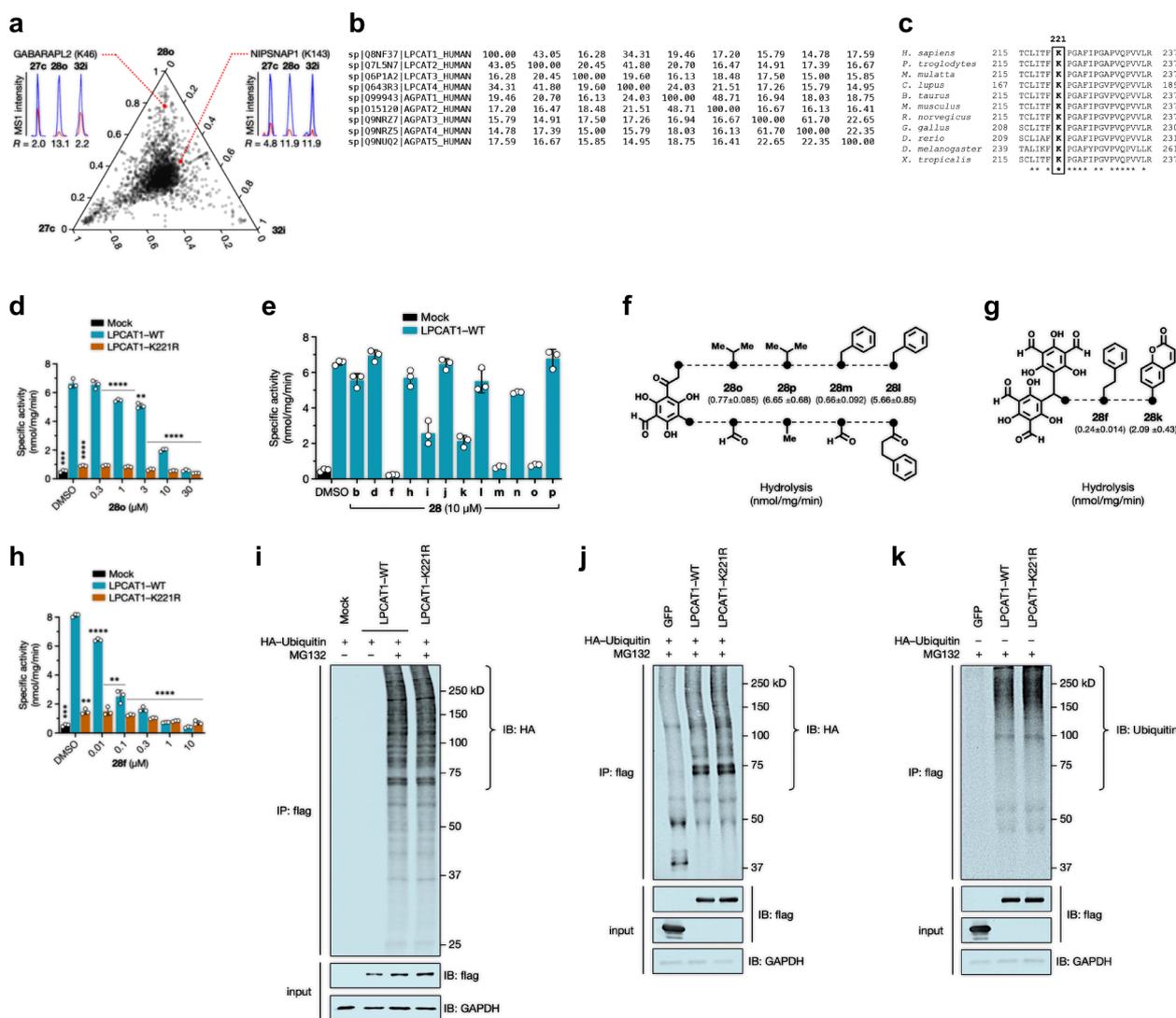
a, cLogP versus molecular weight plot showing aminophilic compounds that follow Lipinski's 'rule of five' (Lipinski space Ro5) and lead-likeness 'rule of three' (Lead-like space Ro3). The size of each bubble represents the number of liganded lysines per compound. **b**, Distribution of compounds (top, left) by the number of hydrogen-bond donors (HBDs, orange line), hydrogen-bond acceptors (HBAs, blue line) and rotatable bonds (RBs, black line)⁶⁰. Correlation between the compound distribution and the number of liganded lysine interactions (gray bars, right y-axis) as relates to the number of HBDs (top, right), HBAs (bottom, left) and RBs (bottom, right). **c**, Heatmap (top) and extracted MS1 chromatograms (bottom) of representative liganded lysines that show broad reactivity with aminophilic compounds (also see Supplementary Data 3). **d**, isoTOP-ABPP ratio plot for the sulfonyl fluoride **17r** containing a kinase-directed recognition element. Red points represent liganded active-site lysines in kinases and their corresponding extracted MS1 chromatograms. The dashed line marks the R value of 4 used to define a lysine liganding event (also see Supplementary Data 3). **e**, Comparison of reactivity of aminophilic compounds toward kinase lysines as a function of selectivity toward kinase lysines across the proteome (right panel). The kinase reactivity of individual compounds was defined by the total number of liganded kinase lysines. The selectivity of individual compounds toward kinase lysines was defined by the fraction of liganded kinase to non-kinase lysines. **f-h**, Location of liganded lysines that are also missense mutated in human disease (orange) in protein crystal structures (gray) of PMVK (K69) (**f**, PDB ID: 3CH4), CPOX (K404) (**g**, PDB ID: 2AEX), and RPL10 (K78) (**h**, PDB ID: 6OLE). Also shown highlighted in blue are active site residues or protein-RNA interaction regions of the proteins where the indicated lysines reside. Note the proximity of K404 in CPOX and K78 in RPL10 to the active site and RNA-interaction region of these proteins, respectively. K69 of PMVK is distant from the active site of the enzyme, but the missense mutation of this lysine causes substantial catalytic defects^{38,61}, pointing to an allosteric regulatory function.



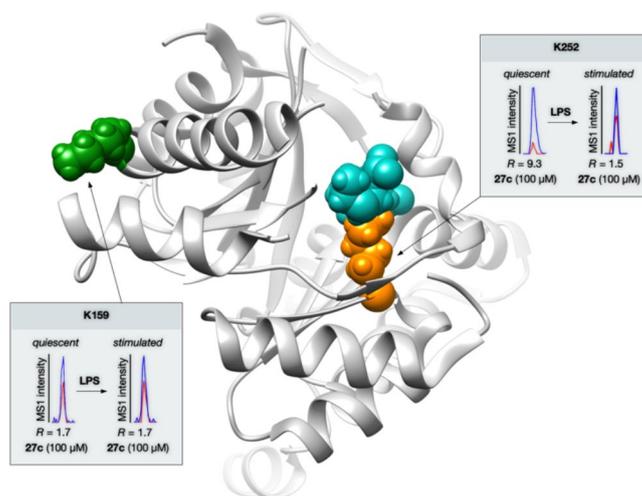
Extended Data Fig. 5 | Functional impact of aminophilic compound-lysine interactions for representative proteins. **a**, The location of liganded lysine K117 (orange) in the RIDA crystal structure (gray, PDB ID: 1ONI). Also shown is bound pyruvate (teal) in each of the three active sites at the interfaces of adjacent monomers. **b**, SAR for aminophilic compound engagement of K117 in RIDA, as determined by competitive isoTOP-ABPP is recapitulated by gel-ABPP of recombinant protein (also see Supplementary Data 3 and 4). Top, HEK293T cells recombinantly expressing WT-RIDA and the corresponding K117R mutant as Flag epitope-tagged proteins were treated with the indicated aminophilic compounds (50 μ M, 1h) followed by treatment with probe **P2** and analyzed by gel-ABPP (top panel) and western blotting (bottom panel). Bottom, Extracted MS1 chromatograms depicting *R* values for the indicated aminophilic compound-RIDA-K117 interactions mapped by competitive isoTOP-ABPP (also see Supplementary Data 3). **c**, Top, gel-ABPP data showing concentration-dependent blockade of **P2** labelling of recombinantly expressed WT-RIDA by **28h** and **26l** in HEK293T cell lysates. Bottom, structures of **28h** and **26l** with extracted MS1 chromatograms depicting *R* values for their respective engagement of K117 or RIDA determined by competitive isoTOP-ABPP (also see Supplementary Data 3). **d**, Corresponding fitted IC_{50} curves for blockade of probe **2** labelling of WT-RIDA. Data represent average values \pm SD; $n=3$ per group. CI, confidence interval. **e**, Representative isoTOP-ABPP ratio plot showing proteome-wide lysine reactivity profile for **26l** (50 μ M). Among ~3,000 quantified lysines, only two - K117 of RIDA and K1070 of VCL - were liganded. The dashed line marks the *R* value of 4 used to define a liganded lysine event (also see Supplementary Data 3). **f**, **g**, Fitted IC_{50} curves for the concentration-dependent inhibition of the deaminase activity of recombinantly expressed WT- and K117R and K117I mutants of RIDA in HEK293T cell lysates by **28h** (**f**) and **26l** (**g**). Data represent average values \pm SD; $n=3$ per group. CI, confidence interval. **h**, Catalytic activity (upper panel) and gel-ABPP analysis of **P2** labelling (lower panel) of WT- and indicated K117 mutants. **i**, Presumed reversible-covalent and irreversible adducts formed between **26l** with K117 and R11742. Data represent average values \pm SD; $n=3$ per group. P values were 0.00081 and 0.000066. For western blot and gel-ABPP data in **b**, **c**, and **h**, experiments were conducted three times ($n=3$ biologically independent experiments) with similar results. Statistical significance was calculated for changes >25% in magnitude in comparison to DMSO-treated samples with unpaired two-tailed Student's *t*-tests: *** $P < 0.001$, **** $P < 0.0001$.



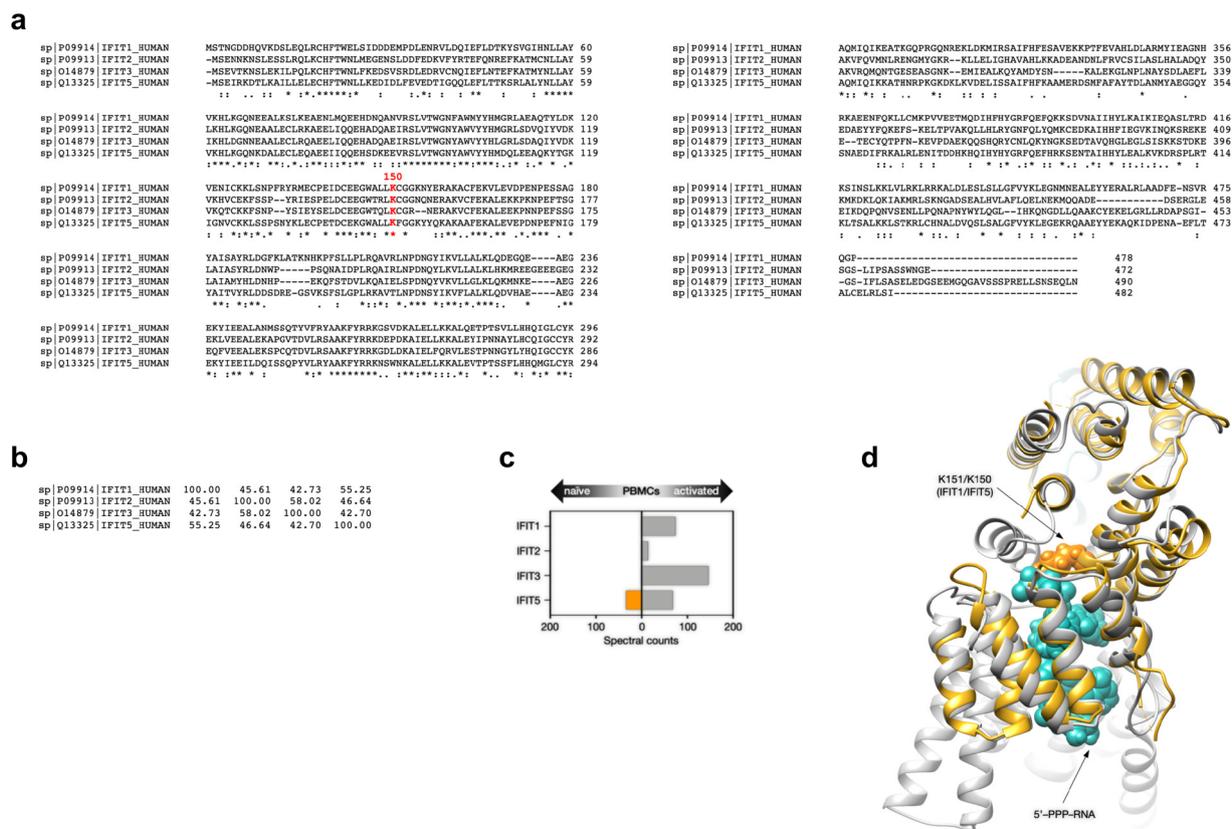
Extended Data Fig. 6 | Compound 11e does not block preformed Ku70-Ku80 complex. **a**, Western blot showing recombinantly co-expressed HA-tagged WT Ku80 with Flag-tagged WT and K351R mutant forms of Ku70 in HEK293T cells. **b**, Lysates of HEK293T cells co-expressing WT Ku80 with WT (left panel) and K351R (right panel) mutant forms of Ku70 were co-immunoprecipitated with anti-Flag antibody (1 h, 4 °C), treated with DMSO or **11e** at the indicated concentrations (1 h, 23 °C), washed, and analyzed by Western blotting. Western blots in **a** and **b** are representative of four independent experiments.



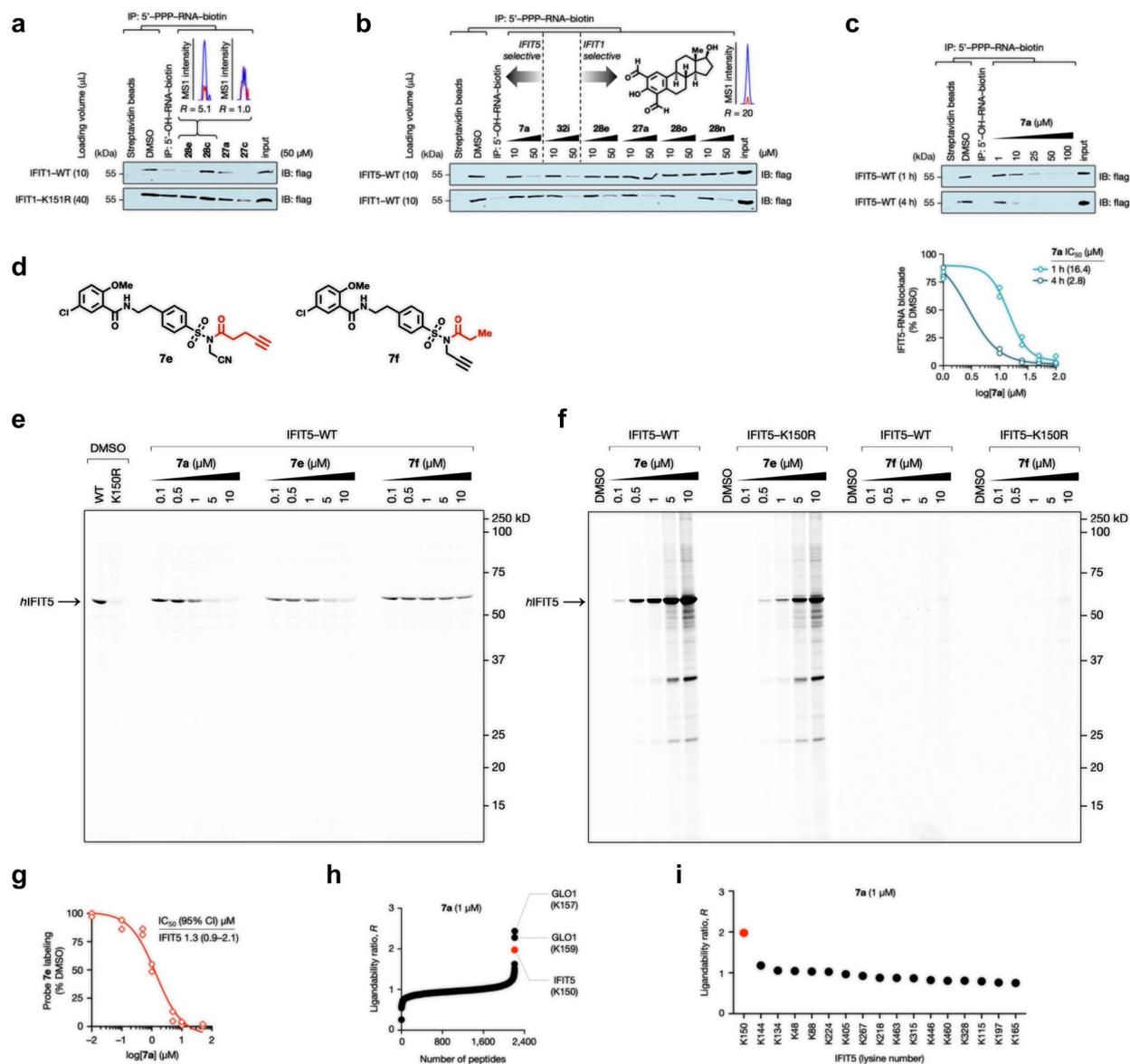
Extended Data Fig. 7 | Dicarboxaldehyde scout fragments and their functional effects on LPCAT1. **a**, Ternary plot showing the proportional lysine reactivity of **27c**, **28o** and **32i** for each lysine. Each point represents a different composition of the three scout fragments based on their individual lysine reactivity ratio (R) values, with the maximum proportion (100%) of each fragment in each corner of the triangle and the minimum proportion (0%) at the opposite line. Extracted MS1 chromatograms of representative competed lysines targeted by scout fragments with differential R values (also see Supplementary Data 3). **b**, Percent identity matrix of human LPCAT1-4 and AGPAT1-4 (<https://www.ebi.ac.uk/Tools/msa/clustalo>). **c**, Conservation of K221 of LPCAT1 across species (<https://www.ncbi.nlm.nih.gov/homologene>). **d**, **28o** produces concentration-dependent blockades of WT-LPCAT1 activity. Data represent average values \pm SD; $n=3$ per group from three biologically independent experiments. P values were 0.00074, 0.000028, 0.000065, 0.0050, and 0.0000062. Statistical significance was calculated for changes $>25\%$ in magnitude in comparison to DMSO-treated samples with unpaired two-tailed Student's t -tests: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. **e**, SAR for aminophilic compound blockade of lyso-PC hydrolysis activity of recombinantly expressed LPCAT1 in HEK293T cell lysates. Data represent average values \pm SD; $n=3$ per group. **f**, Compounds **28o** and **28m** produced greater blockade of LPCAT1 enzymatic activity compared with structural analogs **28p** or **28l**. **g**, Compound **28f** produced greater blockade of LPCAT1 enzymatic activity compared with structural analog **28k**. **h**, **28f** produced concentration-dependent blockades of WT-LPCAT1 activity. Data represent average values \pm SD; $n=3$ per group from three biologically independent experiments. Statistical significance was calculated for changes $>25\%$ in magnitude in comparison to DMSO-treated samples with unpaired two-tailed Student's t -tests: $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. **i, j**, HA-tagged ubiquitin and FLAG-tagged WT or K221R LPCAT1 were co-expressed in HEK293T cells in the presence of proteasome inhibitor MG132 (10 μ M) for 14 h (**i**) or 2 h (**j**), after which cell lysates were subjected to *anti*-FLAG immunoprecipitation, and the affinity-enriched precipitates analyzed by *anti*-HA immunoblotting. For **i**, mock-transfected cells and LPCAT1-WT cells not treated with MG132 were used as controls. For **j**, FLAG-tagged GFP-transfected cells were used as a control. **k**, FLAG-tagged WT or K221R LPCAT1, or GFP-transfected HEK293T cells were treated with MG132 (10 μ M, 2 h), followed by *anti*-FLAG immunoprecipitation and the affinity-enriched precipitates were analyzed by *anti*-ubiquitin immunoblotting. Western blots in **i-k** are representative of four independent biological experiments.



Extended Data Fig. 8 | Example of differential lysine ligandability event in stimulated immune cells. The location of liganded lysine K252 (orange) mapped onto the crystal structure of ALAD (gray, PDB ID: 1PV8). K252 showed substantially weaker interactions with **27c** (100 μ M, 23 $^{\circ}$ C, 1 h) in LPS-stimulated ($R = 1.5$) vs quiescent PBMCs ($R = 9.3$), whereas the reactivity of K159 (green) remained largely unchanged by LPS treatment ($R = 1.7$). K252 is an active-site residue responsible for reversible Schiff-base formation with substrate (blue).



Extended Data Fig. 9 | Characterization of aminophilic compounds that selectively inhibit IFIT family of RNA-binding proteins. a–b, Multiple sequence alignment (**a**) and percent identity matrix (**b**) of human IFIT paralogs (<https://www.ebi.ac.uk/Tools/msa/clustalo>). The red highlight marks a conserved and liganded lysine. **c**, Aggregate spectral counts for quantified lysine-containing peptides for IFIT proteins in human PBMCs \pm LPS treatment. Data represent average values \pm SD; $n=3$ per group from three biologically independent experiments. **d**, Location of liganded lysines (orange) mapped onto the aligned crystal structures of N-terminal domains in IFIT5 (gray, PDB ID: 4HOT) and IFIT1 (yellow, PDB ID: 4HOU) displaying 5'-PPP-RNA (blue) in the nucleotide binding cleft.



Extended Data Fig. 10 | Characterization of aminophilic compounds that inhibit the IFIT family of antiviral RNA-binding proteins. **a**, Extracted MS1 chromatograms with corresponding isoTOP-ABPP ratios (top) and Western blot analysis (bottom) from biotinylated RNA pull-down experiments of WT-IFIT1 and the K151R-IFIT1 mutant from HEK293T cell lysates treated with the indicated concentrations of aminophilic compounds. Western blot is representative of three independent experiments). Also see Supplementary Data 4. **b**, Western blot analysis from biotinylated RNA pull-down experiments of WT-IFIT1 and IFIT5 from HEK293T cell lysates showing concentration-dependent blockade of RNA binding by indicated aminophilic compounds. Western blot is representative of three independent experiments). Also see Supplementary Data 4. **c**, Concentration-dependent blockade (upper panel) and fitted IC_{50} curve (lower panel) of RNA binding of WT-IFIT5 by **7a** after 1 versus 4 h of pre-incubation ($n=2$ biologically independent experiments). **d**, Structures of **7e** containing an alkyne moiety on 'staying group' and **7f** with an alkyne moiety on 'leaving group'. Highlighted in red are 'staying groups' in both compounds. **e**, Representative competition gel showing concentration-dependent blockade of probe **3** labelling by **7a**, **7e**, and **7f** of recombinant WT-IFIT5 in HEK293T cell lysates. **f**, Concentration-dependent labelling of recombinantly expressed WT-IFIT5 and the K150R mutant in HEK293T cell lysates by the clickable probes **7e** and **7f**. gel-ABPP data in **e** and **f** are representative of three independent experiments. **g**, Fitted *in situ* IC_{50} curve for the concentration-dependent blockade of the **7e**-WT-IFIT5 interaction by **7a** in transfected HEK293T cells ($n=4$ biologically independent experiments). **h**, Average ratio values for lysines quantified by isoTOP-ABPP in IFIT5-transfected HEK293T cells treated *in situ* with **7a** (1 μM, 2 h) ($n=2$ independent experiments; also see Supplementary Data 3). **i**, R values for quantified lysines in IFIT5 of experiment described in part **h**.

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	Xcalibur v2.2 was used for mass spectrometry data collection. Image Lab v5.2.1 was used for fluorescence gel scanning. CLARIOstar v5.4 was used for absorbance (Pierce BCA Protein Assay) data collection. Li-COR Odyssey v3.0.30 was used for IRDye secondary antibodies imaging.
Data analysis	IP2 v6.0.02 and CIMAGE were used to analyze all mass spectrometry data. Image Lab v5.2.1 was used for relative quantization of fluorescence gel scanning. Image Studio Lite v5.2.5 was used to process western blot images. GraphPad Prism v7.03 was used to perform statistical tests and to generate all graphs and plots.

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

All data associated with this study are available in the published article and its supplementary information. All raw proteomics data will be uploaded to the PRIDE repository upon publication and corresponding accession information will be provided in the published article.

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	Human blood was acquired from healthy patient donors on an as needed basis therefore no sample-size calculation was performed.
Data exclusions	No data was excluded
Replication	All attempts at replication were successful
Randomization	Cell lines used in experiments were grown under identical conditions, and therefore randomization was not used. Human blood was acquired from de-identified healthy patient donors. No experimental human groups were designated in this study, therefore randomization was not used.
Blinding	Cell lines used in experiments were grown under identical conditions, and therefore blinding was not used. Human blood was acquired from de-identified healthy patient donors (male and female, aged 30-65)

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	Included 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 type="checkbox"/>	<input checked="" 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	Included 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	Mouse anti-FLAG M2 (Sigma Aldrich, F1804); IRDye 800CW antimouse secondary antibody (LICOR, 925-32210)
Validation	Mouse anti-FLAG M2 (Sigma Aldrich, F1804): validated for western blotting by manufacturer using extracts from CHO cells with spiked FLAG-BAP protein.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	MDA-MB-231 (ATCC), Ramos (ATCC), HEK293T (ATCC)
Authentication	All cell lines used in this study were authenticated by the vendors
Mycoplasma contamination	All cell lines used in this study were tested negative for mycoplasma
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in this study

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics

Blood from healthy donors (female and male aged 30 to 65) were obtained after informed consent.

Recruitment

All studies with samples from healthy human volunteers following protocols approved by the Scripps Research Institute institutional review board.

Ethics oversight

The Scripps Research Institute Institutional Review Board.

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