Expedited Mapping of the Ligandable Proteome Using Fully Functionalized Enantiomeric Probe Pairs

Benjamin Cravatt, Yujia Wang, melissa dix, jarrett remsberg, hsin-yu lee, marian kalocsay, steven gygi, gregory vite, michael lawrence, Christopher Parker

Submitted date: 25/02/2019 • Posted date: 27/02/2019
Licence: CC BY-NC-ND 4.0

Citation information: Cravatt, Benjamin; Wang, Yujia; dix, melissa; remsberg, jarrett; lee, hsin-yu; kalocsay, marian; et al. (2019): Expedited Mapping of the Ligandable Proteome Using Fully Functionalized Enantiomeric Probe Pairs. ChemRxiv. Preprint.

A fundamental challenge in chemical biology and medicine is to understand and expand the fraction of the human proteome that can be targeted by small molecules. We recently described a strategy that integrates fragment-based ligand discovery with chemical proteomics to furnish global portraits of reversible small molecule-protein interactions in human cells. Excavating clear structure-activity relationships from these “ligandability” maps, however, was confounded by the distinct physicochemical properties and corresponding overall protein-binding potential of individual fragments. Here, we describe a compelling solution to this problem by introducing a next-generation set of fully functionalized fragments (FFFs) differing only in absolute stereochemistry. Using these enantiomeric probe pairs, or “enantioprobes”, we identify numerous stereoselective protein-fragment interactions in cells and show that these interactions occur at functional sites on proteins from diverse classes. Our findings thus indicate that incorporating chirality into FFF libraries provides a robust and streamlined method to discover ligandable proteins in cells.

File list (5)

| Name                                                                 | Size          | Enlarge | Download |
|---------------------------------------------------------------------|---------------|---------|----------|
| Wang_etal_2019_Manuscript_Final.pdf                                 | 3.78 MiB      | view on ChemRxiv | download file |
| Wang_etal_2019_SupplementaryInformation_Final.pdf                   | 11.39 MiB     | view on ChemRxiv | download file |
| Wang_etal_2019_SupplementaryTable3.xlsx                             | 10.18 MiB     | view on ChemRxiv | download file |
| Wang_etal_2019_SupplementaryTable2.xlsx                             | 30.62 MiB     | view on ChemRxiv | download file |
| Wang_etal_2019_SupplementaryTable1.xlsx                             | 329.95 KiB    | view on ChemRxiv | download file |
Expedited mapping of the ligandable proteome using fully functionalized enantiomeric probe pairs

Yujia Wang¹, Melissa M. Dix¹, Jarrett R. Remsberg¹, Hsin-Yu Lee¹, Marian Kalocsay², Steven P. Gygi², Gregory Vite³, R. Michael Lawrence³, Christopher G. Parker¹,4*, Benjamin F. Cravatt¹*

¹ Department of Chemistry, The Scripps Research Institute, La Jolla, CA, USA
² Department of Cell Biology, Harvard Medical School, Boston, MA, USA
³ Research and Development, Bristol-Myers Squibb Company, Princeton, NJ, USA
⁴ Department of Chemistry, The Scripps Research Institute, Jupiter, FL, USA

To whom correspondence should be addressed: cparker@scripps.edu; cravatt@scripps.edu
A fundamental challenge in chemical biology and medicine is to understand and expand the fraction of the human proteome that can be targeted by small molecules. We recently described a strategy that integrates fragment-based ligand discovery with chemical proteomics to furnish global portraits of reversible small molecule-protein interactions in human cells. Excavating clear structure-activity relationships from these “ligandability” maps, however, was confounded by the distinct physicochemical properties and corresponding overall protein-binding potential of individual fragments. Here, we describe a compelling solution to this problem by introducing a next-generation set of fully functionalized fragments (FFFs) differing only in absolute stereochemistry. Using these enantiomeric probe pairs, or “enantioprobes”, we identify numerous stereoselective protein-fragment interactions in cells and show that these interactions occur at functional sites on proteins from diverse classes. Our findings thus indicate that incorporating chirality into FFF libraries provides a robust and streamlined method to discover ligandable proteins in cells.
Chemical probes are versatile tools to interrogate the functions of proteins in biological systems and complement genetic approaches\(^1\) by producing reversible and graded gains or losses of protein activity, as well as, in certain instances, neo-functional outcomes\(^2\)\(^{\rightarrow}5\). Small molecules also represent a principal category of clinically approved drugs, and quality chemical probes are needed to pharmacologically validate novel targets on the path to developing therapeutic agents.

Despite their basic and translational value, chemical probes are lacking for the vast majority of human proteins\(^6\). Methods for the discovery of new chemical probes often rely on high-throughput screening (HTS) of large libraries (~10\(^6\)) of relatively high molecular weight and structurally diverse compounds against individual proteins (target-based) or cellular systems (phenotype-based)\(^7\),\(^8\). Hits from such libraries can often be challenging to optimize due to their structural complexity and suboptimal ligand efficiencies\(^9\). Further, many proteins are problematic to express, purify, and format for \emph{in vitro} HTS, especially if they are parts of large complexes and/or remain poorly characterized in terms of biochemical function. These challenges underscore the need for new methods that can more broadly assess the ‘ligandability’ (i.e., ability to bind small molecules) of the human proteome in native biological systems.

Fragment-based ligand discovery (FBLD) has emerged as a versatile approach for the discovery of atom-efficient, small-molecule binders for a wide range of proteins\(^10\)\(^\rightarrow\)\(^12\). However, due to the generally low affinity of fragment hits and the biophysical methods typically used for their discovery (e.g., NMR, surface plasmon resonance, isothermal calorimetry), FBLD has been mostly limited to the study of purified proteins \emph{in vitro}\(^11\). We recently introduced a strategy that integrates FBLD with chemical proteomics to globally assess small molecule-protein interactions in human cells\(^13\). Using a specialized library of fully functionalized fragment (FFF) probes, which possess variable fragment binding elements coupled to photoreactive and bioorthogonal reporter groups, we mapped >2000 reversible fragment-protein interactions in human cells and showed that these interactions can be advanced into more potent and selective compounds capable of modulating the activity of proteins in cells\(^13\).
The fragment binding elements in our initial studies were selected based on their representation in drug-like molecules and were accordingly diverse in structure and physicochemical properties. As a consequence, we found that individual FFF probes showed substantial differences in their overall proteomic interaction profiles, which made for complicated structure-activity relationships (SARs) requiring careful manual review to identify fragment-protein interactions that reflected authentic recognition events (versus simply correlating with the overall proteomic interaction profiles of the FFF probes). We describe herein a general strategy to address this confounding bottleneck in the form of a next-generation set of FFF probes consisting of physicochemically matched fragment pairs differing only in absolute stereochemistry. Using a set of eight pairs of enantiomeric FFF probes – or “enantioprobes” – we expeditiously identify >170 stereochemistry-dependent small molecule-protein interactions in human cells. The enantioprobe targets span diverse structural and functional classes and include proteins that lack chemical probes. We validate enantioprobe interactions for several recombinantly expressed proteins and show that the interactions occur at functionally relevant sites on these proteins. Finally, we describe a quantitative, multiplexed workflow capable of performing up to five enantioprobe pair comparisons in a single experiment, thereby greatly increasing the throughput and dimensionality of fragment-based ligand discovery in cells.

Results
Design and initial proteomic profiling of enantioprobes
Our original set of FFF probes were designed to contain: 1) a “variable” recognition element consisting of structurally diverse small-molecule fragments intended to promote interactions with distinct proteins in human cells; and 2) a structurally minimized “constant” region bearing a photoactivatable diazirine group and alkyne handle, which together enabled UV light-induced covalent modification and detection, enrichment, and identification of fragment-interacting protein targets (Fig. 1a). Here, we reasoned that the introduction of stereochemistry into FFF probe design could furnish pairs of compounds that display equivalent physicochemical
properties and gross overall protein binding in cells, but differ in their stereoselective interactions with authentic small molecule-binding pockets in proteins. The preferential enrichment of proteins by one member of an enantioprobe pair would then constitute instant evidence of ligandability for these proteins.

We synthesized a library of eight enantioprobe pairs, where members of each pair differ only in absolute stereochemistry of the fragment recognition element (Fig. 1a). We then qualitatively assessed the overall proteomic interaction profiles for enantioprobes using established SDS-PAGE methods. In brief, we treated HEK293T cells with each enantioprobe (20 µM, 30 min) followed by exposure to UV light (365 nm, 10 min), harvesting, lysis, coupling of probe-modified proteins to an azide-rhodamine reporter tag using copper-catalyzed azide-alkyne cycloaddition chemistry (CuAAC), and visualization of these proteins by SDS-PAGE and in-gel fluorescence scanning. As expected, we observed substantial differences in protein interactions across the enantioprobe pairs, with one probe pair ((R) and (S)-6) exhibiting much greater overall protein labeling compared to others (Fig. 1b). Encouragingly, however, the (R) and (S) members within each enantioprobe pair showed similar overall proteomic labeling with the exception of select proteins that exhibited stereochemistry-dependent (“stereoselective”) interactions (Fig. 1b and Supplementary Fig. 1a, red asterisks). The enantioprobe pairs also showed clear increases in protein labeling across a test concentration range of 5-100 µM (Supplementary Fig. 1b), indicating good cell permeability, and virtually all of these protein labeling events were dependent on UV light exposure (Supplementary Fig. 1c). We next turned our attention to mapping enantioprobe-protein interactions in human cells by quantitative mass spectrometry (MS)-based proteomics.

Global maps of enantioprobe-protein interactions in human cells

We evaluated two complementary cell types for enantioprobe interactions by quantitative MS-based proteomics: 1) primary human peripheral blood mononuclear cells (PBMCs); and 2) HEK239T cells. The selection of these cell types afforded an opportunity to directly compare
enantioprobe profiles to those generated with the original set of FFF probes (generated in HEK293T cells)\textsuperscript{13} and extend our understanding of fragment ligandability to primary human immune cells. Both cell types were treated with equal concentrations of (R)- or (S)-compounds from each enantioprobe pair (200 μM, 30 min) and then exposed to UV light to induce photocrosslinking of enantioprobe-bound proteins, lysed, and enantioprobe-labeled proteins conjugated to an azide-biotin tag by CuAAC chemistry, enriched by streptavidin, and analyzed by MS-based proteomics, where stereoselective interactions were quantified by isotopic labeling using either reductive dimethylation (ReDiMe) with heavy or light formaldehyde (PBMCs)\textsuperscript{16,17} or SILAC (stable isotope labeling by amino acids in cell culture; HEK293T cells)\textsuperscript{18} (Fig. 2a). For PBMCs, these experiments were performed in replicate in both isotopic directions (heavy vs light and light vs heavy) to furnish four independent experiments for each enantioprobe pair, and, for HEK293T cells, a subset of the enantioprobe pairs was examined. We designated proteins as showing stereoselective interactions if they were preferentially enriched by an average value of > 2.5-fold by one member of an enantioprobe pair in either PBMCs or HEK293T cells. For the SILAC studies in HEK293T cells, we also performed control experiments where the heavy- and light-labeled cells were treated with the same enantioprobe to ensure that, under these conditions, enantioprobe-enriched proteins showed a ratio of ~1.0. (Supplementary Fig. 2a, Supplementary Table 1, and Supplementary Table 2).

In total, 176 proteins showed stereoselective interactions with one or more enantioprobe pairs, which included 119 proteins identified in PBMCs (Fig. 2b, Supplementary Table 1, and Supplementary Table 2) and 108 proteins identified in HEK293T cells (Supplementary Fig. 2b, Supplementary Table 1, and Supplementary Table 2). We observed similar numbers of stereoselective protein interactions for each member of an enantioprobe pair (Fig. 2c and Supplementary Fig. 2c, d), indicating equivalent potential for the R or S-enantiomer to preferentially enrich proteins. Proteins identified in both PBMCs and HEK293T cells generally showed consistent profiles across the cell types, that is, stereoselective interactions identified in PBMCs were also observed in HEK293Ts and vice versa (Fig. 2d and Supplementary Fig. 2e).
The enantioprobe pairs displayed considerable differences in their total number of stereoselective interactions with the human proteome (Fig. 2e), and, notably, these profiles were unrelated to the extent of overall protein labeling displayed by the probes (Fig. 1b). This result suggests that stereoselective interactions are based on factors beyond the general protein binding potential of a given fragment structure.

The majority of proteins showing stereoselective interactions (>80%) did so with only one of the enantioprobe pairs (Fig. 2f and Supplementary Fig. 2f). Embedded within this specificity were multiple profiles, including proteins that were enriched by several enantioprobe pairs, but stereoselectively by one pair, as well as proteins that showed strong enrichment predominantly with a single enantioprobe across the entire probe set (Fig. 2g). Proteins showing stereoselective interactions with enantioprobes spanned diverse functional and structural classes (Fig. 3a, Table 1, and Supplementary Table 1). Perhaps unsurprisingly, many of the enantioprobe targets were enzymes, including kinases, methyltransferases, and various metabolic enzymes (Fig. 3a, Table 1, and Supplementary Table 1), likely reflecting the high potential for these proteins to specifically bind small molecules. We also, however, observed stereoselective interactions for various scaffolding/adaptor proteins and transcriptional regulators – classes that have been historically considered challenging to target with small molecules (Fig. 3a, Table 1, and Supplementary Table 1). To the extent that the magnitude of stereoselective enrichment is predictive of a robust small-molecule interaction, we further noted that high stereoselective enrichment (ratio values > 4.0 in pairwise comparisons of R vs S enantioprobes) was observed for several scaffolding/adaptor proteins and transcriptional regulators (Table 1 and Supplementary Table 1). Enantioprobe targets that were observed in PBMCs, but not HEK293T cells, tended to correspond to immune-enriched proteins (e.g., IRAK319, PARP1020) (Supplementary Table 1). We finally observed limited overlap of enantioprobe targets with proteins that demonstrated ligandability in previous chemical proteomic studies using cysteine21,22 or lysine-reactive23 electrophilic fragments (Fig. 3b), indicating that non-covalent and covalent fragments generally interact with distinct sets of
proteins in human cells. Moreover, while many of the enantioprobe targets were also enriched by members of the original FFF probe set\textsuperscript{13}, these previous profiles often did not provide useful SAR information, either reflecting substantial enrichment by all of the FFF probes or mirroring the respective global protein interaction footprints of these probes (Supplementary Fig. 2g).

Enantioprobes engage functionally relevant sites on proteins

We next sought to confirm stereoselective interactions for representative proteins targeted by diverse enantioprobes and originating from different functional classes, including enzymes (kinase, RPS6KA3; methyltransferase, SMYD3; and a metabolic enzyme; DCTPP1), a lipid-binding protein (UNC119B), a transporter (TSPO), a membrane-binding/adaptor protein (PACSIN2), a transcriptional regulator (HDGF), and an uncharacterized protein (TTC38). Each protein was recombinantly expressed with a FLAG epitope tag in HEK293T cells by transient transfection, and 48 h later, cells were treated with the indicated enantioprobe pair (5-80 µM, unless otherwise indicated) followed by photocrosslinking with UV light, CuAAC coupling to an azide-rhodamine tag, and visualization of protein labeling by SDS-PAGE and in-gel fluorescence scanning. All experiments also included mock-transfected cells as a control. Each recombinantly expressed protein displayed stereoselective interactions with the enantioprobes that mirrored the preferential labeling of the endogenous forms of these proteins in PBMCs and or HEK293T cells (Fig. 3c-f and Supplementary Fig. 3). Most of these stereoselective interactions could be detected with 5-10 µM of the preferred enantioprobe and were preserved across the entire enantioprobe concentration range (Fig. 3c-f and Supplementary Fig. 3).

Some of the protein targets have known ligands, which afforded an opportunity to test whether the enantioprobes and ligands share a common binding site on these proteins. Four representative proteins were selected for analysis – 1) the lysine methyltransferase SMYD3, a target of (R)-1 and (R)-5 that binds both cofactors (SAM, SAH) and the synthetic inhibitor EPZ031686\textsuperscript{24,25}; 2) the lipid-binding protein UNC119B, a target of (R)-1 and (R)-5 that binds the natural product squarunkin A\textsuperscript{26}; 3) the sterol transporter TSPO, a target of (R,R)-7 that binds the
synthetic ligand PK 11195\textsuperscript{27,28}, which is used to image brain injury and inflammation\textsuperscript{29,30}, and 4) the uncharacterized protein TTC38, a target of (S)-4 that has been found to bind the histone deacetylase (HDAC) inhibitor panobinostat\textsuperscript{31}. In all four cases, we found that enantioprobe interactions with both the endogenous and recombinantly expressed protein targets were blocked by increasing concentrations of ligand, as measured by MS-based (endogenous protein) and gel-based (recombinant protein) methods (Fig. 4a-c and Supplementary Fig. 4).

For SMYD3, only EPZ031686, but not SAM or SAH, blocked (R)-1 interactions (Fig. 4a). Structural studies have shown that EPZ031686 binds in the lysine substrate binding pocket of SMYD3 and noncompetitively with SAM\textsuperscript{24}. These data suggested that (R)-1 may also bind to the lysine substrate pocket, which we confirmed by mapping the site of (R)-1 photolabeling on SMYD3 by quantitative MS. Using previously described protocols\textsuperscript{13,32,33}, we identified a single tryptic peptide in SMYD3 that was photolabeled by (R)-1 – R.DQYCFECDCFR.C (amino acids (aa) 255-265) – with the predicted site(s) of photoreactivity being residues D255-Y257. These residues are located within 3.6 angstroms of an EPZ031686 analog in the SMYD3 co-crystal structure (Fig. 4d), and Y257 specifically has been found to interact with the methylated lysine in substrate co-crystal structures\textsuperscript{34}. Quantitative MS-based proteomics further demonstrated that photolabeling of the aa 255-265 peptide by (R)-1 was blocked by co-incubation with EPZ031686 and was not observed with the inactive enantioprobe (S)-1 (Fig. 4d).

We also mapped the primary sites of enantioprobe labeling for UNC119B (Fig. 4e) and TSPO (Fig. 4f) and confirmed the stereoselectivity of these labeling events and their blockade by treatment with competitive ligands (squarunkin A and PK 11195, respectively). For UNC119B, (R)-1 modified the tryptic peptide containing residues 227-236 (R.SDSFYFVDNK.L) with predicted sites of labeling spanning S227-Y231 (Fig. 4e). These residues represent a highly conserved stretch of amino acids in UNC119 proteins that are within 2.5 angstroms of a fatty acylated peptide in a co-crystal structure with the related protein UNC119A\textsuperscript{35}, and S229 is predicted to hydrogen bond with squarunkin A in a docking model of the natural product bound to UNC119A\textsuperscript{26} (Fig. 4e). For TSPO, (R,R)-7 modified the N-terminal peptide (aa 2-24,
M.APPWVPAMGFTLAPSLGCFVGSR.F with the principle site of labeling being C19 (Fig. 4f). In the solution structure of mouse TSPO, the corresponding residue (G19) is 2.8 angstroms away from the ligand PK 1119527.

Taken together, our follow-up studies on representative targets indicate that enantioprobes engage functionally relevant and druggable sites on diverse classes of proteins. We also noted that our chemical proteomic studies with enantioprobes identified additional, unanticipated targets for some of the tested small-molecule competitors. For instance, EPZ031686 blocked (R)-1 interactions with the solute carrier SLC35F2 and the peptidase PRCP (Fig. 4a), while PK 11195 decreased (R,R)-7 interactions with the lipid-binding protein ABHD5 (Fig. 4c). While ABHD5 also showed independent evidence of stereoselective interactions with enantioprobes (Supplementary Table 1), SLC35F2 and PRCP did not (Supplementary Table 1), suggesting that these latter proteins may specifically bind enantioprobes, but without stereochemical preference. Motivated to explore this general concept further, as well as to increase the throughput and information content of our chemical proteomic experiments, we set out to create a multiplexed platform for the streamlined analysis of enantioprobe-protein interactions in human cells.

**Multiplexed analysis of enantioprobe-protein interactions in cells**

While we were generally satisfied with the sensitivity and robustness of our chemical proteomic experiments using SILAC or ReDiMe as quantitative MS-based measurement protocols of enantioprobe-protein interactions, we also recognized that the pairwise nature of these comparisons had drawbacks. Prominently, the limited throughput prevented a deeper exploration of SAR both within and across enantioprobe pairs. For instance, a protein that interacts specifically, but without stereo-preference with both enantioprobes in a pairwise comparison is difficult to distinguish from a non-specific interaction, as both outcomes furnish an enrichment ratio of ~1.0. And, relatedly, the stochastic nature of protein identification events in untargeted MS-based proteomic experiments hindered confident assignment of proteins that
selectively interacted with one or a subset of enantioprobes across different experiments. Finally, the throughput of pairwise comparisons also becomes restrictive when attempting to compare the protein interaction profiles of several enantioprobes under various conditions (e.g., different cell types, probe concentrations, etc.).

We considered that many of the aforementioned challenges could be addressed by analyzing enantioprobes with a multiplexed approach for quantitative MS-based proteomics that uses isobaric tandem mass tags (TMT)\textsuperscript{36-38}. In this workflow, up to 10 separate populations of cells are each treated with an enantioprobe (200 µM, 30 min), photocrosslinked, lysed, conjugated to biotin azide via CuAAC, enriched and trypsinized as described above. Tryptic peptides stemming from each treatment group are then labeled with a TMT tag of equivalent parent mass, but differentiable by MS3-derived fragmentation products, combined, and analyzed in a single MS experiment (Fig. 5a)\textsuperscript{38}. Applying 10-plex TMT, we compared the protein interaction profiles of four enantioprobe pairs (Supplementary Table 3), alongside a previously described methyl control probe\textsuperscript{13} (in duplicate), in human PBMCs and HEK293T cells (Fig. 5b and Supplementary Fig. 5a). We required that at least three unique peptides were quantified for each protein to interpret stereoselective interactions with enantioprobes, and stereoselective interactions were defined as those displaying > 2.5-fold differential enrichment between (R) and (S) members of at least one enantioprobe pair, along with > 5-fold enrichment over the methyl control probe.

We observed a robust overall correlation between the enantioprobe profiles quantified by multiplexed (TMT-based) versus pairwise (ReDiMe/SILAC) comparative proteomic experiments (Fig. 5c, d and Supplementary Fig. 5b), and the vast majority (> 85%) of enantioprobe targets identified in pairwise comparisons showed consistent stereoselective interactions in multiplex experiments (Supplementary Tables 1-3). Another 115 stereoselective interactions were mapped by multiplexing, and these newly discovered events mostly corresponded to proteins that were not quantified in pairwise experiments performed with the relevant enantioprobe pair(s). In addition to recapitulating and extending the stereoselective enantioprobe-protein
interactions discovered in pairwise experiments, the multiplexed method also illuminated proteins that showed enrichment by one or more enantioprobe pairs, but without stereopreference. Examples included CYP27A1 and TLR8, which interacted preferentially with the (R)/(S)-2 and (R)/(S)-3 probe pairs, respectively, over the other enantioprobe pairs (Fig. 5c, lower panels). We interpret these enrichment profiles to also reflect specific probe-protein interactions, where the SAR across the enantioprobe set is driven by chemotype rather than stereotype.

We reasoned that the greater sample capacity afforded by multiplexing could also provide an efficient means of assessing the relative potency of enantioprobe-protein interactions by comparing protein enrichment profiles across several probe concentrations. We performed a proof-of-principle experiment with a representative enantioprobe pair – (S)-3 and (R)-3 – tested at five different concentrations (5, 20, 50, 100 and 200 µM) in human PBMCs (Supplementary Tables 1 and 3). The concentration-dependent profiles revealed that previously mapped enantioprobe targets maintained stereoselective interactions across the entire probe concentration range (e.g., see IRAK3 and PARP10 in Fig. 5e; also see Supplementary Fig. 6a). Some of these interactions showed evidence of saturated enrichment at lower concentrations of the preferred enantioprobe (e.g., TTC38, Fig. 5e), likely reflecting a higher affinity binding event. This type of saturated enrichment was also observed for certain proteins that did not show stereopreference between (S)-3 and (R)-3 (e.g., SLC25A20, Supplementary Fig. 6b). Finally, a third type of profile was observed, albeit rarely, where a protein displayed saturated enrichment with both (S)-3 and (R)-3, but the absolute signal plateaued at different values between the enantioprobes (e.g., PTGR2, Supplementary Fig. 6c). This outcome might reflect cases where equivalent binding is observed for both enantioprobes, but one of the probes generates a greater amount of photoadduct with the protein target (see Discussion below).

Taken together, these data indicate that the mapping of fragment-protein interactions in cells can be efficiently performed proteome-wide using multiplexing MS-based methods to
expedite the discovery of ligandable proteins with a rich body of integrated information on SAR and potency.

**Discussion**

Efforts to expand the proportion of the human proteome that can be targeted by chemical probes would benefit from methods capable of evaluating small molecule-protein interactions on a global scale in native biological systems. We previously described a chemical proteomic strategy to perform fragment-based ligand discovery experiments in human cells\(^{13}\). The initial set of fully functionalized (clickable, photoreactive) fragments (FFFs) uncovered many new small molecule-protein interactions, some of which were advanced to selective and cell-active chemical probes. Nonetheless, we also found that individual FFFs showed substantially different overall protein interaction profiles in human cells, which complicated the assignment of small molecule-protein interactions displaying authentic SARs (vs nonspecific binding to the FFFs). The enantioprobes described herein offer a general solution to this challenge by specifying ligandable proteins as those showing differential interactions with physicochemically matched compounds differing only in absolute stereochemistry. We identified many such stereoselective interactions across diverse functional and structural protein classes and verified several using recombinantly expressed proteins. Importantly, in each case where a protein target had an established ligand, we found that this ligand blocked enantioprobe binding. These results indicate that stereoselective interactions of enantioprobes often occur at functional sites on proteins. If this principle generalizes across the broader set of enantioprobe targets identified herein, it highlights the potential of fragment-based screening in cells to serve as a foundation for the pursuit of chemical probes that perturb the function of a wide range of proteins.

There are some important considerations when considering the broader implementation of enantioprobes for mapping protein ligandability in biological systems. First, we emphasize that a substantial fraction of enantioprobe targets showed stereoselective interactions with only a single \((R)/(S)\) probe pair. We interpret this result to indicate that, with our modest set of eight
enantioprobe pairs, we are vastly under-sampling the proportion of human proteins that have the capacity to show stereoselective interactions with small-molecule fragments. Future attention should thus be given to expanding the size and structural diversity of the enantioprobe library, as well as to applying these probes in more diverse cell types to survey a broader fraction of the human proteome. We also admit that, in most cases, we do not know with certainty whether stereoselective interactions between an enantioprobe and a protein reflect preferential binding versus photoreactivity (i.e., the extent of carbene adduct with a protein target following photoexcitation of the diazirine). In some cases, it is conceivable that both members of an enantioprobe pair bind equivalently to a protein target, but one probe produces a greater yield of photoadduct with the protein. While this SAR outcome would ultimately need to be clarified to guide efforts toward more advanced chemical probes that display higher affinity and selectivity for individual protein targets, we posit that stereoselective binding and stereoselective photoreactivity are equivalently useful parameters for identifying novel druggable sites in the proteome, as both would likely require specific interactions with a protein to discriminate between an enantioprobe pair. Finally, our data highlight the value of incorporating TMT-based multiplexing readouts into enantioprobe profiling experiments, which greatly expedited the discovery of stereoselective interactions without substantial losses in sensitivity or accuracy. Moreover, these multiplexing experiments provide additional SAR information by identifying proteins that interact in a chemoselective, rather than stereoselective, manner with the enantioprobe set.

Projecting forward, we envision several exciting pursuits with enantioprobes that should address fundamental questions about the ligandability of the human proteome. For instance, will the stereoselective interactions displayed by fragment enantioprobes be retained as these ligands are elaborated into more advanced chemical probes, or, alternatively, will the preferential interaction with a single stereocenter dissipate in importance as additional recognition elements are built into the probes? Toward this end, we note that the enantioprobes offer a convenient target engagement assay for assessing competitive binding of elaborated
analogues in cells, and that several of the enantiopure fragment recognition groups deployed herein are poised for direct modification using synthetic methodologies such as C-H bond activation chemistry\textsuperscript{39-41}. Second, would more structurally complex enantioprobes identify ligandable proteins that, for instance, do not display sufficient binding affinity to simple fragment probes? Finally, what fraction of stereoselective interactions observed proteome-wide occur at functional sites on proteins? Here, we admit that a complete answer is not likely to be soon forthcoming, as we are dependent on both mapping the sites of enantioprobe binding, a still technically challenging task, and the availability of protein structures to predict functional pockets. Consider TTC38, for instance, a poorly characterized protein that has been previously identified as an off-target of the HDAC inhibitor panobinostat\textsuperscript{31} and found herein to display stereoselective interaction with the (S)-4 probe that was blocked by panobinostat. We would presume that the site of binding of panobinostat and probe (S)-4 is relevant to TTC38 function, but absent a structure or, for that matter, even a biochemical activity for the protein, this conclusion is premature. Of course, ligands that are found to bind silent sites on proteins can still be converted into “functional” chemical probes that promote protein degradation using PROTAC-like technologies\textsuperscript{42-44}.

In summary, our findings demonstrate that enantioprobes offer a highly efficient way to discover small molecule-protein interactions in human cells. Differentiating proteins based on stereoselective interactions with otherwise physicochemically equivalent fragment probes offers instant evidence of authentic ligandability. These stereoselective interactions can then form the basis for pursuit of more advanced chemical probes targeting a diverse range of proteins for basic and translational research purposes.

**Methods**

A detailed Methods section is provided in the Supplementary Information.
Acknowledgements

We thank P.S. Baran for helpful discussions, PX. Shen for assistance in compounds characterization, and R. Park for technical advices on proteomics data analysis. The authors acknowledge the Scripps Normal Blood Donor Services, NMR, MS and Automated Synthesis core facilities. This work was supported by the NIH (CA231991, CA211526, and DK114785).

Author Contributions

Y.W., C.G.P. and B.F.C. conceived of the project, designed experiments and analyzed data. Y.W. synthesized and characterized compounds. Y.W. performed SILAC and ReDiMe quantitative proteomics experiments and data analysis. M.M.D. developed methods for tandem mass tag (TMT) labeling experiments. M.M.D. and Y.W. performed TMT-based proteomics experiments. Y.W. and J.R.R. conducted site of labeling experiments. Y.W. cloned, overexpressed proteins and conducted biological experiments. H.Y.L. assisted in biological experiments. Y.W. complied proteomics data and conducted computational analyses. S.P.G., M.K., R.M.L. and G.V. assisted in the design and analysis of mass spectrometry experiments. Y.W., C.G.P. and B.F.C. wrote the manuscript and all authors contributed to manuscript editing.

Competing Interests

The authors declare competing financial interests. B.F.C. is a founder and advisor to Vividion Therapeutics, a biotechnology company interested in using chemical proteomic methods to develop small-molecule drugs to treat human disease. C.G.P. serves as a consultant for Vividion Therapeutics.
References

1. Knight, Z. A. & Shokat, K. M. Chemical genetics: where genetics and pharmacology meet. *Cell* **128**, 425-430, doi:10.1016/j.cell.2007.01.021 (2007).

2. Belshaw, P. J., Ho, S. N., Crabtree, G. R. & Schreiber, S. L. Controlling protein association and subcellular localization with a synthetic ligand that induces heterodimerization of proteins. *P Natl Acad Sci USA* **93**, 4604-4607, doi:10.1073/pnas.93.10.4604 (1996).

3. Lai, A. C. & Crews, C. M. Induced protein degradation: an emerging drug discovery paradigm. *Nat Rev Drug Discov* **16**, 101-114, doi:10.1038/nrd.2016.211 (2017).

4. Lu, G. *et al.* The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. *Science* **343**, 305-309, doi:10.1126/science.1244917 (2014).

5. Schreiber, S. L. A Chemical Biology View of Bioactive Small Molecules and a Binder-Based Approach to Connect Biology to Precision Medicines. *Israel Journal of Chemistry* **0**, doi:10.1002/ijch.201800113 (2018).

6. Oprea, T. I. *et al.* Unexplored therapeutic opportunities in the human genome. *Nature Reviews Drug Discovery* **17**, 317-332, doi:10.1038/nrd.2018.14 (2018).

7. Schenone, M., Dancik, V., Wagner, B. K. & Clemons, P. A. Target identification and mechanism of action in chemical biology and drug discovery. *Nat Chem Biol* **9**, 232-240, doi:10.1038/Nchembio.1199 (2013).

8. Swinney, D. C. & Anthony, J. How were new medicines discovered? *Nature Reviews Drug Discovery* **10**, 507-519, doi:10.1038/nrd3480 (2011).

9. Hajduk, P. J. & Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat Rev Drug Discov* **6**, 211-219, doi:10.1038/nrd2220 (2007).

10. Bembenek, S. D., Touge, B. A. & Reynolds, C. H. Ligand efficiency and fragment-based drug discovery. *Drug Discov Today* **14**, 278-283, doi:10.1016/j.drudis.2008.11.007 (2009).

11. Scott, D. E., Coyne, A. G., Hudson, S. A. & Abell, C. Fragment-based approaches in drug discovery and chemical biology. *Biochemistry* **51**, 4990-5003, doi:10.1021/bi3005126 (2012).

12. Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W. & Jhoti, H. Twenty years on: the impact of fragments on drug discovery. *Nature Reviews Drug Discovery* **15**, 605-619, doi:10.1038/nrd.2016.109 (2016).

13. Parker, C. G. *et al.* Ligand and Target Discovery by Fragment-Based Screening in Human Cells. *Cell* **168**, 527-541 e529, doi:10.1016/j.cell.2016.12.029 (2017).

14. Welsch, M. E., Snyder, S. A. & Stockwell, B. R. Privileged scaffolds for library design and drug discovery. *Curr Opin Chem Biol* **14**, 347-361, doi:10.1016/j.cbpa.2010.02.018 (2010).
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 Engl* **41**, 2596-2599, doi:10.1002/1521-3773(20020715)41:14<2596::AID-ANIE2596>3.0.CO;2-4 (2002).

Boersema, P. J., Rajmakers, R., Lemeer, S., Mohammed, S. & Heck, A. J. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat Protoc* **4**, 484-494, doi:10.1038/nprot.2009.21 (2009).

Hsu, J. L., Huang, S. Y., Chow, N. H. & Chen, S. H. Stable-isotope dimethyl labeling for quantitative proteomics. *Anal Chem* **75**, 6843-6852, doi:10.1021/ac0348625 (2003).

Ong, S. E. *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**, 376-386, doi:10.1074/mcp.M200025-MCP200 (2002).

Hubbard, L. L. & Moore, B. B. IRAK-M regulation and function in host defense and immune homeostasis. *Infect Dis Rep* **2**, doi:10.4081/idr.2010.e9 (2010).

Verheugd, P. *et al.* Regulation of NF-κB signalling by the mono-ADP-ribosyltransferase ARTD10. *Nature Communications* **4**, 1683, doi:10.1038/ncomms2672 (2013).

Backus, K. M. *et al.* Proteome-wide covalent ligand discovery in native biological systems. *Nature* **534**, 570-574, doi:10.1038/nature18002 (2016).

Bar-Peled, L. *et al.* Chemical Proteomics Identifies Druggable Vulnerabilities in a Genetically Defined Cancer. *Cell* **171**, 696-709, doi:10.1016/j.cell.2017.08.051 (2017).

Hacker, S. M. *et al.* Global profiling of lysine reactivity and ligandability in the human proteome. *Nat Chem* **9**, 1181-1190, doi:10.1038/Nchem.2826 (2017).

Mitchell, L. H. *et al.* Novel Oxindole Sulfonamides and Sulfamides: EPZ031686, the First Orally Bioavailable Small Molecule SMYD3 Inhibitor. *ACS Med Chem Lett* **7**, 134-138, doi:10.1021/acsmedchemlett.5b00272 (2016).

Van Aller, G. S. *et al.* Structure-Based Design of a Novel SMYD3 Inhibitor that Bridges the SAM-and MEKK2-Binding Pockets. *Structure* **24**, 774-781, doi:10.1016/j.str.2016.03.010 (2016).

Mejuch, T. *et al.* Small-Molecule Inhibition of the UNC119-Cargo Interaction. *Angew Chem Int Ed Engl* **56**, 6181-6186, doi:10.1002/anie.201701905 (2017).

Jaremko, L., Jaremko, M., Giller, K., Becker, S. & Zweckstetter, M. Structure of the mitochondrial translocator protein in complex with a diagnostic ligand. *Science* **343**, 1363-1366, doi:10.1126/science.1248725 (2014).

Le Fur, G. *et al.* Peripheral benzodiazepine binding sites: effect of PK 11195, 1-(2-chlorophenyl)-N-methyl-N-(1-methylpropyl)-3-isoquinolinecarboxamide. I. In vitro studies. *Life Sci* **32**, 1839-1847 (1983).

Owen, D. R. J. & Matthews, P. M. Imaging Brain Microglial Activation Using Positron Emission Tomography and Translocator Protein-Specific Radioligands. *Int Rev Neurobiol* **101**, 19-39, doi:10.1016/B978-0-12-387718-5.00002-X (2011).

Rupprecht, R. *et al.* Translocator protein (18 kDa) (TSPO) as a therapeutic target for neurological and psychiatric disorders. *Nature Reviews Drug Discovery* **9**, 971-988,
31 Becher, I. et al. Thermal profiling reveals phenylalanine hydroxylase as an off-target of panobinostat. *Nat Chem Biol* **12**, 908-+, doi:10.1038/Nchembio.2185 (2016).

32 Niphakis, M. J. et al. A Global Map of Lipid-Binding Proteins and Their Ligandability in Cells. *Cell* **161**, 1668-1680, doi:10.1016/j.cell.2015.05.045 (2015).

33 Weerapana, E., Speers, A. E. & Cravatt, B. F. Tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP) - a general method for mapping sites of probe modification in proteomes. *Nature Protocols* **2**, 1414-1425, doi:10.1038/nprot.2007.194 (2007).

34 Fu, W. Q. et al. Structural Basis for Substrate Preference of SMYD3, a SET Domain-containing Protein Lysine Methyltransferase. *J Biol Chem* **291**, 9173-9180, doi:10.1074/jbc.M115.709832 (2016).

35 Jaiswal, M. et al. Novel Biochemical and Structural Insights into the Interaction of Myristoylated Cargo with Unc119 Protein and Their Release by Arl2/3. *J Biol Chem* **291**, 20766-20778, doi:10.1074/jbc.M116.741827 (2016).

36 Thompson, A. et al. Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem* **75**, 1895-1904, doi:10.1021/ac0262560 (2003).

37 Dayon, L. et al. Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Anal Chem* **80**, 2921-2931, doi:10.1021/ac702422x (2008).

38 McAlister, G. C. et al. MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Anal Chem* **86**, 7150-7158, doi:10.1021/ac502040v (2014).

39 Ming, S. et al. Modular, Stereocontrolled Cβ–H/Cα–C Activation of Alkyl Carboxylic Acids. (2019).

40 Tran, A. T. & Yu, J. Q. Practical Alkoxycarbonyl Auxiliaries for Iridium(I)-Catalyzed C-H Alkylation of Azacycles. *Angew Chem Int Ed Engl* **56**, 10530-10534, doi:10.1002/anie.201704755 (2017).

41 Jain, P., Verma, P., Xia, G. Q. & Yu, J. Q. Enantioselective amine alpha-functionalization via palladium-catalysed C-H arylation of thioamides. *Nat Chem* **9**, 140-144, doi:10.1038/Nchem.2619 (2017).

42 Deshaies, R. J. Protein degradation: Prime time for PROTACs (vol 11, pg 634, 2015). *Nat Chem Biol* **11**, 887-887, doi:10.1038/nchembio1115-887b (2015).

43 Neklesa, T. K., Winkler, J. D. & Crews, C. M. Targeted protein degradation by PROTACs. *Pharmacol Therapeut* **174**, 138-144, doi:10.1016/j.pharmthera.2017.02.027 (2017).

44 Winter, G. E. et al. Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science* **348**, 1376-1381, doi:10.1126/science.aab1433 (2015).
Table 1. Representative stereoselective protein targets of enantioprobes in human cells.

| UniPROT ID | Protein name | Protein type | Cell type           | Stereoselective enantioprobe(s) | Highest stereoselective ratio |
|------------|--------------|--------------|---------------------|---------------------------------|-------------------------------|
| Q9NUJ1     | ABHD10       | serine hydrolase | PBMCs              | (R)-7                            | 5.3                           |
| Q15018     | ABRAXAS2     | scaffolding protein | PBMCs, HEK293T | (R)-8                            | 20                            |
| P04083     | ANXA1        | inflammatory modulator | PBMCs, HEK293T | (R)-2                            | 3.9                           |
| Q13867     | BLMH         | cysteine endopeptidase | PBMCs, HEK293T | (R)-3                            | 9.1                           |
| P24941     | CDK2         | protein kinase   | PBMCs              | (R)-7                            | 3.9                           |
| O95628     | CNOT4        | E3 ubiquitin ligase | HEK293T          | (R)-3                            | 11                            |
| Q9H773     | DCTPP1       | pyrophosphatase  | PBMCs, HEK293T    | (S)-1, (S)-3, (S)-4               | 10                            |
| Q13045     | FLII         | nuclear receptor coactivator | HEK293T | (R)-1                            | 2.8                           |
| P51858     | HDGF         | transcriptional repressor | PBMCs, HEK293T | (S)-1                            | 6.3                           |
| Q5T447     | HECTD3       | E3 ubiquitin ligase | HEK293T          | (S)-8                            | 10                            |
| O95373     | IPO7         | nuclear transport receptor | HEK293T     | (R)-8                            | 3.2                           |
| Q9Y616     | IRAK3        | pseudokinase     | PBMCs              | (S)-3, (S)-4                      | 15                            |
| Q08J23     | NSUN2        | RNA methyltransferase | PBMCs          | (R)-3                            | 2.9                           |
| Q9UNF0     | PACSIN2      | membrane-binding protein/adaptor protein | PBMCs, HEK293T | (R)-3, (R)-6                      | 4.4                           |
| Q53GL7     | PARP10       | NAD+ ADP-ribosyltransferase | PBMCs      | (S)-3, (S)-4                      | 20                            |
| Q9UBV8     | PEF1         | calcium-dependent adapter | PBMCs, HEK293T | (S)-1                            | 4.7                           |
| Accession | Name   | Description                        | Tissue  | Enantiomers       | EC50  |
|-----------|--------|------------------------------------|---------|-------------------|-------|
| O00329    | PIK3CD | phosphoinositide-3-kinase          | PBMCs   | (R)-2             | 3.8   |
| O75475    | PSIP1  | transcriptional coactivator        | PBMCs   | (S)-1             | 3.7   |
| P63244    | RACK1  | scaffolding protein                | PBMCs, HEK293T | (R)-2, (R)-7     | 3.5   |
| P51812    | RPS6KA3| protein kinase                     | PBMCs, HEK293T | (R)-1, (R)-2, (S)-4 | 20    |
| Q8WWI5    | SLC44A1| choline transporter                | PBMCs   | (R)-2             | 14    |
| Q9H7B4    | SMYD3  | protein methyltransferase          | PBMCs, HEK293T | (R)-1, (R)-5     | 7.0   |
| Q5BJF2    | TMEM97 | orphan receptor                    | HEK293T | (S)-8             | 7.0   |
| P30536    | TSPO   | cholesterol transporter            | PBMCs   | (R)-4, (R)-7      | 7.5   |
| A6NIH7    | UNC119B| lipid-binding protein              | PBMCs   | (R)-1             | 2.9   |
| O95619    | YEATS4 | succinyl-lysine reader             | HEK293T | (R)-3             | 9.5   |
Figure 1. Enantioprobes for mapping stereoselective protein-small molecule fragment interactions in human cells. a, Structures of enantioprobes, which consist of a “variable” element of stereopure fragment pairs (enclosed box) and a “constant” region containing a diazirine photoreactive group and a clickable alkyne handle. b, Gel-based profiling of enantioprobe-protein interactions in human cells. HEK293T cells were treated with enantioprobes (20 µM) for 30 min, photocrosslinked, lysed, and proteomes conjugated to an azide-rhodamine tag using CuAAC chemistry and analyzed by SDS-PAGE and in-gel fluorescent scanning. Red asterisks mark representative stereoselective enantioprobe-protein interactions.
Figure 2. MS-based profiling of enantioprobe-protein interactions in human cells. 

a, Schematic workflow for identifying stereoselective enantioprobe-protein interactions in human cells. 

b, Heatmap showing relative protein enrichment ratios for pairwise comparisons of (R) and (S) enantioprobes (200 µM each) in both isotopic directions in human peripheral blood mononuclear cells (PBMCs). White signals in the heatmap either correspond to proteins with ratio values of ~ 1 or proteins that were not enriched and quantified with the indicated enantioprobe pair. (R)* and (S)* represent (R,R) and (S,S) for enantioprobe 7. 

c, Representative scatter plot showing protein enrichment ratios for (R)-1 versus (S)-1 in PBMCs. Proteins enriched > 2.5-fold by one enantiomer over the other are considered stereoselective targets. Red and blue protein targets show stereoselective interactions with (R)-1 and (S)-1,
respectively. Data reflect an average of at least two independent experiments for each isotopic direction. d, Similar stereoselective interactions are observed in different cell types. Plot depicts Log₂ values of protein enrichment ratios for (R)-1/(S)-1 in HEK293T cells (x-axis) versus PBMCs (y-axis). e, Number of stereoselective protein interactions found for each enantioprobe pair in PBMCs. f, Number of proteins showing stereoselective interactions with the indicated number of enantioprobe pairs in PBMCs. g, Quantity of aggregate spectral counts for PYGB (left graph) and ABRAXAS2 (right graph) enriched by each enantioprobe in PBMCs.
Figure 3. Characterization of stereoselective protein targets of enantioprobes. a, Functional classes of stereoselective protein targets of enantioprobes in PBMCs and HEK293T cells. b, Fraction of stereoselective protein targets of enantioprobes showing evidence of ligandability with cysteine and/or lysine-reactive fragments, as determined previously\textsuperscript{21-23}. The left graph includes all stereoselective targets; the right graph shows only those stereoselective targets with quantified cysteines and/or lysines in previous studies\textsuperscript{21-23}. c-f, Top: Confirmation of stereoselective enantioprobe-protein interactions with recombinantly expressed proteins. RPS6KA3 (c), PACSIN2 (d), SMYD3 (e), and UNC119B (f) were recombinantly expressed with FLAG epitope tags by transient transfection in HEK293T cells, and transfected cells were then treated with the indicated concentrations of enantioprobes, photocrosslinked, lysed, and
proteomes conjugated to an azide-rhodamine tag by CuAAC chemistry and analyze by SDS-PAGE and in-gel fluorescence scanning. Bottom left: Extracted MS1 chromatograms of representative tryptic peptides for endogenous forms of the protein targets in HEK293T cells or PBMCs treated with indicated enantioprobes (200 µM). Bottom right: quantification of protein labeling by the indicated enantioprobes derived from gel-based profiles show in Top section. Data represents average values ± SD for two independent experiments. Confirmation of additional stereoselective interactions shown in Supplementary Fig 3.
Figure 4. Stereoselective interactions occur at functional and druggable sites on protein targets of enantioprobes. a-c, Top left: structure of competitor ligands (a) EPZ031686; (b) Squarunkin A; (c) PK 11195. Top right (or middle for a): Waterfall plots of competitive blockade of enantioprobe interactions with endogenous protein targets for corresponding ligands in HEK293T cells. Bottom (or right for a): Gel-based profiles of competitive blockade of enantioprobe interactions with recombinantly expressed protein targets for corresponding ligands in transfected HEK293T cells. d, Structure of SMYD3 in complex with EPZ031686.
(shown as stick model; PDB 5CCM) highlighting (R)-1-modified tryptic peptide (aa 255-265, light red; predicted probe-modified residues D255-Y257, dark red). e, Structure of UNC119A in complex with a myristoylated peptide (yellow; PDB 5L7K) highlighting (R)-1-modified tryptic peptide (aa 227-235, light red, predicted probe-modified residues S227-Y230, dark red). f, Structure of TSPO in complex with PK 11195 (shown as stick model; PDB 2MGY) highlighting (R,R)-7-modified tryptic peptide (aa 2-24, light red; predicted probe-modified residue C19, dark red).
Figure 5. Multiplexed MS-based quantification for expedited discovery of stereoselective protein-enantioprobe interactions. a, Schematic of TMT-based workflow for mapping enantioprobe-protein interactions in a multi (10)-plex format. b, Heatmap depicting TMT quantification of stereoselective protein targets in PBMCs. Relative enrichment ratios are calculated as a percent of maximum signal per protein. c, Similar profiles are found for stereoselective protein targets of enantioprobes in pairwise (ReDiMe) versus multiplex (TMT) experiments (top panels). Multiplex experiments also enable the identification of proteins that interact with enantioprobes in a chemotype-selective manner (bottom panels). White signals in the heatmap either correspond to proteins with ratio values of ~ 1 or proteins that were not enriched and quantified with the indicated enantioprobe pair. d, Representative scatter plots showing the correlation between pairwise (x-axis) and multiplex (y-axis) experiments performed with enantioprobes (R/S)-1 and (R/S)-2. e, Concentration-dependent profiles for representative stereoselective enantioprobe-protein interactions as determined by multiplex experiments of PBMCs treated with 0, 5, 20, 50, 100 and 200 µM of the indicated enantioprobe pair.
Supplementary Information

Expedited mapping of the ligandable proteome using fully functionalized enantiomeric probe pairs

Yujia Wang¹, Melissa M. Dix¹, Jarrett Remsberg¹, Hsin-Yu Lee¹, Marian Kalocsay², Steven P. Gygi², Gregory Vite³, R. Michael Lawrence³, Christopher G. Parker¹⁴, Benjamin F. Cravatt¹*

¹ Department of Chemistry, The Scripps Research Institute, La Jolla, CA, USA
² Department of Cell Biology, Harvard Medical School, Boston, MA, USA
³ Research and Development, Bristol-Myers Squibb Company, Princeton, NJ, USA
⁴ Department of Chemistry, The Scripps Research Institute, Jupiter, FL, USA

To whom correspondence should be addressed: cparker@scripps.edu; cravatt@scripps.edu
Contents

(A) Supplementary Figures
(B) Supplementary Table Legends
(C) Biological Methods
(D) Synthetic Methods
(E) NMR spectra
(F) References
Supplementary Figs. 1. Gel-based profiling of enantioprobe-protein interactions in human cells. a, Particulate fraction of gel-based profiles shown in Fig. 1c. Red asterisks mark representative stereoselective protein-enantioprobe interactions. b, Enantioprobes show concentration-dependent increases in protein labeling in HEK293T cells. c, Enantioprobes show UV-dependent protein labeling in HEK293T cells.
Supplementary Fig 2. MS-based profiling of enantioprobe-protein interactions in human cells. **a**, Representative SILAC ratio plots for control experiments in which isotopically heavy and light amino acid-labeled HEK293T cells were treated with the same enantioprobe. **b**, Heatmap showing relative protein enrichment ratios in both isotopic directions for the indicated enantioprobe pairs in HEK293T cells. White signals in the heatmap either correspond to proteins with ratio values of ~ 1 or proteins that were not enriched and quantified with the indicated enantioprobe pair. **c, d**, Related to Fig. 2c. Scatter plots showing protein enrichment ratios for all enantioprobe pairs in primary human PBMCs (c) or HEK293T cells (d). Targets enriched > 2.5-fold by the (R) or (S)-member of an enantioprobe pair are shown in red and blue,
respectively. Data reflect an average of two independent experiments for each isotopic direction. (R)* and (S)*- represent (R,R) and (S,S)- for enantioprobe 7. e, Related to Fig. 2d. Similar stereoselective interactions are observed in different cell types. Plots depict Log2 values of protein enrichment ratios for indicated enantioprobe pairs in HEK293T cells (x-axis) versus PBMCs (y-axis). f, Related to Fig. 2f. Number of proteins showing stereoselective interactions with the indicated number of enantioprobe pairs in HEK293T. g, Heatmaps showing relative protein enrichment values for FFF probes (200 µM) versus a methyl control probe (200 µM) for stereoselective protein targets of enantioprob es. FFF profiles were taken from reference 1. White signals in the heatmap either correspond to proteins with ratio values of ~1 or proteins that were not enriched and quantified with the indicated enantioprobe pair or FFF probe.
Supplementary Fig. 3. Characterization of additional stereoselective protein targets of enantioprobes. Left: Confirmation of stereoselective enantioprobe-protein interactions with recombinantly expressed proteins. Experiments were performed in transfected HEK293T cells as described in the legend for Fig. 3c-f. Right: Extracted MS1 chromatograms of representative tryptic peptides for endogenous forms of the protein targets in HEK293T cells or PBMCs treated with indicated enantioprobes (200 µM).
Supplementary Fig. 4. Stereoselective interactions occur at functional and druggable sites on protein targets of enantioprobes. Left: structure of panobinostat. Middle: Waterfall plot of competitive blockade of enantioprobe (S)-4 interactions with endogenous protein targets by panobinostat in HEK293T cells. Inset shows representative MS1 chromatogram for TTC38. Right: Gel-based profile of competitive blockade of enantioprobe (S)-4 interactions with recombinant TTC38 by panobinostat in HEK293T cells.
Supplementary Fig. 5. Multiplex MS-based quantification for expedited discovery of stereoselective protein-enantioprobe interactions. a, Related to Fig. 5b. Heatmaps showing enantioprobe enrichment profiles for stereoselective protein targets in multiplex versus pairwise experiments in PBMCs. White signals in the pairwise heatmap either correspond to proteins with ratio values of ~ 1 or proteins that were not enriched and quantified with the indicated enantioprobe pair. (R)* and (S)*- represent (R,R) and (S,S)- for enantioprobe 7. b, Related to Fig. 5d. Scatter plots showing the correlation between pairwise and multiplex experiments performed with the indicated enantioprobe pairs in PBMCs.
Supplementary Fig. 6. Concentration-dependent profiles for enantioprobe pair (R)-3 and (S)-3 in PBMCs. a, Heatmap showing concentration-dependent profiles for stereoselective protein targets of (R)-3 and (S)-3 in PBMCs, at 5, 20, 50, 100, and 200 µM of the enantioprobes (right). The left heatmap shows the profiles for stereoselective protein targets in pairwise experiments performed with the (R)-3 and (S)-3 enantioprobes (200 µM each). White signals in the heatmap either correspond to proteins with ratio values of ~1 or proteins that were not enriched and quantified with the indicated enantioprobe pair. b, c, Concentration-dependent profiles for representative proteins SLC25A20 (b) and PTGR2 (c) displaying saturated enrichment signals across the test probe concentration range (5-200 µM).
(B) Supplementary Table Legends

**Supplementary Table 1.** Compiled stereoselective protein targets for enantioprobe pairs 1 – 8 in primary human peripheral blood mononuclear cells (PBMCs) and HEK293T cells. (Tab 1) Averaged (3 biological replicates for enantioprobe pair 1, 4 biological replicates for HEK293T cells at 200 µM of indicated enantioprobe) ReDiMe ratios of stereoselective protein targets obtained from pairwise experiments conducted in PBMCs at 200 µM of indicated enantioprobe. (Tab 2) Averaged (2 biological replicates) SILAC ratios of stereoselective protein targets obtained from pairwise experiments conducted in HEK293T cells at 200 µM of indicated enantioprobe. (Tab 3) Averaged (2 biological replicates) relative enrichment ratios of stereoselective protein targets obtained from multiplexed experiments conducted in PBMCs and HEK293T cells at 200 µM of indicated enantioprobe. Relative enrichment ratios are calculated as a percent of maximum signal per protein. (Tab 4) Stereoselective protein targets from pairwise experiments referenced with their relative enrichment ratios from multiplexed experiments. (Tab 5) Averaged (2 biological replicates) concentration-dependent relative enrichment values of high-engagement protein targets of (R)-3 and (S)-3 in PBMCs, at 5, 20, 50, 100 and 200 µM of the enantioprobe. High-engagement protein targets are defined as proteins with relative enrichment values greater than 50% at 50 µM and 75% at 100 µM for either (R)-3 or (S)-3. (Tab 6) Averaged (2 biological replicates) concentration-dependent relative enrichment values of stereoselective protein targets of (R)-3 and (S)-3 in PBMCs or HEK293T cells, at 5, 20, 50, 100 and 200 µM of the enantioprobe. (Tab 7) Tissue-specific patterns of mRNA expression of stereoselective protein targets discovered solely in PBMCs with enantioprobe 1, 2, 3 and 8. Protein targets with 3 × median mRNA expression amongst lymphoblast, B cells, T cells, dendritic cells, NK cells, monocytes or myeloid cells are considered as enriched in immune-related cells. (Tab 8) Averaged (2 biological replicates) SILAC ratios of proteins from HEK293T cells treated with the same indicated enantioprobe in HEK293T cells at 200 µM. The mean of median SILAC ratios was used as the normalization factor to correct incomplete incorporation of heavy amino acids in SILAC experiments.

**Supplementary Table 2.** Data sets from SILAC/ReDiMe experiments. (Tab 1) Summary of experimental design for pairwise experiments included in Supplementary Table 2. (Tab 2-31) Each tab displays the relative light/heavy ratios for all quantified tryptic peptides per protein from each replicate.

**Supplementary Table 3.** Data sets from TMT experiments. (Tab 1) Summary of experimental design for multiplexed experiments included in Supplementary Table 3. (Tab 2-5) Averaged (2 biological replicates) relative enrichment ratios of quantified proteins obtained from multiplexed experiments conducted in PBMCs and HEK293T cells at 200 µM of indicated enantioprobe. Relative enrichment ratios are calculated as a percent of maximum signal per protein. (Tab 6) Averaged (2 biological replicates) relative enrichment values of quantified protein targets of (R)-3 and (S)-3 in PBMCs, at 5, 20, 50, 100 and 200 µM of the enantioprobes. (Tab 7-16) Each tab displays the averaged reporter ions intensity for all quantified proteins from each replicate.

(C) Biological Methods
Cell lines and primary cells
HEK293T cells were maintained in high-glucose DMEM (Corning, 15-013-CV) supplemented with 10% (v/v) fetal bovine serum (FBS, Omega Scientific, FB-01), penicillin (100 U/mL) and streptomycin (100 µg/mL) and L-glutamine (2 mM, Corning, 25-005-CI). Cells were grown at 37 °C in a humidified 5% CO₂ atmosphere. For SILAC experiments, HEK293T cells was passaged at least six times in SILAC DMEM (Thermo), which lacks L-lysine and L-arginine, and supplemented with 10% (v/v) dialyzed FBS (Omega Scientific, FB-03), penicillin, streptomycin (as above), and either [13C₆, 15N₂]-L-lysine (Sigma Aldrich, 608041) and [13C₆, 15N₂]-L-arginine (Sigma Aldrich, 608033) (100 µg/mL each) or L-lysine, HCl and L-arginine, HCl (100 µg/mL each). Heavy and light cells were maintained in parallel and cell aliquots were frozen after six passages in SILAC media and stored in liquid N₂ until needed. Whenever thawed, cells were passaged at least twice before being used in experiments.

All studies using samples from human volunteers follow protocols approved by the TSRI institutional review board. Human blood from healthy donors were obtained after informed consent and peripheral blood mononuclear cells (PBMCs) were purified over Lymphoprep™ (STEMCELL, Catalog # 07861) density gradient medium according to manufacturer’s protocol. The purified PBMCs were washed with sterile DPBS (Corning, 21-031-CV) and resuspended in RPMI-1640 (Corning, 15-040-CV) supplemented with penicillin, streptomycin and L-glutamine (as above) before being used in experiments.

In situ labeling of live cells with enantioprobe
Experiments were performed similarly as previously reported. For gel-based experiments, HEK293T cells were grown in 6-well plates (Olympus, Cat #: 25-105) to near complete confluence before treatment. Cells were washed once with DPBS and replenished with 1mL fresh serum-free media before treated with enantioprobe, and, if applicable, competitors or DMSO vehicle. The treated cells were incubated at 37 °C for 30 min, exposed to 365 nm UV light (Stratagene, UV Stratalinker 1800) for 10 min in a cold room. For no-UV control experiments, cells were placed at 4 °C in the cold room for 10 min under ambient light. Following photocrosslinking, cells were harvested by scraping, centrifuged (1,400 g, 4 min, 4 °C), and washed with cold DPBS (2X). Pellets were either directly used for analysis or kept frozen at -80 °C until use. For MS-based experiments, HEK293T cells were treated similarly as above with minor modifications, including cells were grown in 6 cm plates (Olympus, Cat # 25-260) in ‘heavy’ or ‘light’ SILAC DMEM media, and treated in 1.5 mL fresh serum-free media. Human PBMCs were obtained as described above, resuspended in serum-free RPMI-1640 media and seeded in 6-well plates immediately before treated similarly as HEK293T cells.

Preparation of proteome for gel- and MS-based analysis
Cell pellets were lysed in cold DPBS (100 – 400 µL) using a Branson Ultrasonics Sonifier S-250A cell disruptor (~6 pluses, 35% duty cycle, output setting =3.5). For experiments with soluble and particulate fraction of proteomes, cell lysates were centrifuged (100,000 g, 20 min) to provide soluble (supernatant) and particulate (pellet) fractions. Particulate fractions were resuspended in cold DPBS with sonication. Protein concentration was determined with the DC Protein Assay (Bio-Rad) and the absorbance was measured using a CLARIOstar microplate reader following manufacturer’s instructions. Protein concentration was diluted to 1.5 mg/mL with cold DPBS for further use.

Gel-based analysis of probe labelled proteins in cells
To each sample (1.5 mg/mL, 50 µL), 6 µL of freshly prepared ‘click’ mixture (3 µL of 1.7 mM TBT A in 4:1 t-BuOH:DMSO, 1 µL of 50 mM CuSO₄ in H₂O, 1 µL of 1.25 mM Tetramethylrhodamine (TAMRA) azide in DMSO, 1 µL of freshly prepared 50 mM TCEP in DPBS) was added. The reaction mixture was mixed by pipetting and incubated at room temperature for 1 hour before quenching with 4X SDS gel loading buffer (17 µL). Proteins (30 µg total protein loaded per gel lane) were resolved by SDS-PAGE (10% acrylamide) made in-house and visualized by in-gel fluorescence on a Bio-Rad ChemiDoc MP Imaging System. The images were processed using Image Lab (version 5.2.1) software.

Preparation of isotopically labeled samples for pairwise MS-based analysis
For isotope dimethyl labeled experiments, samples were prepared as previous reported. To 0.5 mL cell lysates in a 1.5 mL Eppendorf tube, 55 µL freshly prepared ‘click’ mixture (30 µL of 1.7 mM TBT A in 4:1 t-BuOH:DMSO, 10 µL of 50 mM CuSO₄ in H₂O, 5 µL of 10 mM Biotin-PEG4-azide (ChemPep, cat # 271606) in DMSO, 10 µL of
freshly prepared 50 mM TCEP in DPBS) was added and the samples were rotated at room temperature for 1 hour. The mixtures were transferred to 15 mL falcon tubes and 3 mL cold methanol was added. The resulting cloudy mixtures were centrifuged (5,000 g, 10 min, 4 °C) to obtain protein pellets. After removing the supernatant, the protein pellets were washed with cold 1:1 MeOH:CHCl₃ (2 x 1 mL) and were resuspended in cold 4:1 MeOH:CHCl₃ (3 mL) by sonication. The cloudy mixtures were centrifuged (5,000 g, 10 min, 4 °C) to pellet the proteins. The pellets were solubilized in proteomics-grade urea (500 µL, 6 M in DPBS) containing 10 µL of 10% SDS by sonication. 50 µL 1:1 mixture of TCEP (200 mM in DPBS) and K₂CO₃ (600 mM in DPBS) was added to each sample and the mixture was incubated at 37 °C for 30 min to reduce the disulfides. Reduced thiols were then alkylated by adding 70 µL iodoacetamide (400 mM in DPBS) at room temperature protected from light for 30 min. To each solution, 130 µL of 10% SDS (in DPBS) was added and diluted to ~0.2% SDS with DPBS (5.5 mM) and incubated with pre-equilibrated streptavidin agarose beads (Thermo Fisher Scientific, Cat # 20347, 100 µL 1:1 slurry) for 1.5 hours at room temperature on a rotator. The beads were centrifuged (1,400 g, 2 min) and washed sequentially with 0.2% SDS in DPBS (1 x 5 mL), DPBS (2 x 5 mL) and H₂O (1 x 5 mL) to remove unbound proteins, excess detergent and reagents. The beads were transferred to a low-binding 1.5 mL microfuge tube (Axygen, CNT 20347, 100 µL) and the enriched proteins were on-beads digested overnight at 37 °C in ~200 µL of 2 M urea in DPBS containing 2 µg sequencing grade porcine trypsin (Promega, V5111) in the presence of CaCl₂ (1 mM). To each digested sample, 8 µL of 4% 'light' formaldehyde (Sigma Aldrich, 252549) or 8 µL of 4% 'heavy' formaldehyde-¹³C, d₂ (Sigma Aldrich, 596388) and 8 µL of sodium cyanoborohydride (0.6 M in H₂O) were added and the reaction was incubated at room temperature for 1 hour before quenching with 32 µL of 1% NH₄OH (in H₂O) followed by 16 µL of formic acid. The corresponding 'light' and 'heavy' sample were combined and centrifuged (1,400 g, 2 min). The supernatant was transferred to a low-binding microfuge tube and stored at -20 °C until analysis. For Stable Isotope Labeling with Amino Acids in Cell Culture (SILAC) samples, a procedure adapted from methods previously reported³ was followed.

Liquid chromatography-mass spectrometry (LC-MS) analysis of probe labeled proteins

Protein digests were pressure loaded onto a 250 µm (inner diameter) fused silica capillary column packed with 4 cm C18 resin (Phenomenex, Aqua 5 µm). Samples were analyzed on an LTQ-Orbitrap Elite mass spectrometer (Thermo Scientific) coupled to a UltiMate 3000 Series Nano/Cap Pump NCP-3200RS and a WPS-3000PL (RS) autosampler (Thermo Scientific Dionex). Peptides were eluted by two-dimensional separation on a column with a 5 µm tip [100 µm fused silica, packed with 10 cm C18 (Phenomenex, Aqua 5 µm) and 3 cm strong cation exchange resin (SCX, Phenomenex)] using a 5-step 'MudPIT' protocol that injects 5 µL 0%, 25%, 50%, 80%, 100% salt bumps of ammonium acetate (500mM) in Buffer A (95% H₂O, 5% acetonitrile, 0.1% formic acid) followed by an increasing gradient of buffer B (20% H₂O, 80% acetonitrile, 0.1% formic acid) in Buffer A in each step. The flow rate was 0.5 µL/min and the voltage applied to the nano-LC electrospray ionization source was 2.5 kV. Spectra were collected in a data-dependent acquisition mode with dynamic exclusion enabled (repeat count 2, exclusion duration 20 s). One full MS1 scan (400-1800 m/z) was followed by 30 MS2 scans (CID-ITMS) of the nth most abundant ions. Parent ions with unassigned or +1 charge state were excluded from fragmentation.

Peptide and protein identification and quantification

The MS2 spectra were extracted from the raw file using RAWconverter (available at http://fields.scripps.edu/rawconv/) with 'select monoisotopic m/z in DDA' enabled. MS2 spectra data were searched using ProLuCID algorithm against a reverse-concatenated, nonredundant variant of the Human UniProt database (2016-07) and filtered using DTASelect 2.0 within the Integrated Proteomics Pipeline (IP2). The precursor ion mass tolerance for a minimum envelop of three isotopic peaks was set to 50 ppm. All cysteines were specified with a static modification for carbamidomethylation (+57.02146) and up to one differential modification was allowed per peptide for either methionine oxidation (+15.994915) or glutamine/asparagine deamidation (+0.984016). For stable isotope dimethyl labeled samples, lysine and N-terminus were also specified with a static modification for demethylation (+28.0313). In addition, peptides were required to have at least one tryptic terminus. Each dataset was simultaneously searched for both 'light' and 'heavy' isotopic labeling by specifying the mass shifts on selected labeled amino acids, specifically, lysine (+8.0142) and arginine (+10.0083) for SILAC samples or lysine (+6.03182) and N-terminus (+6.03182) for dimethyl labeled samples. The minimum peptide length was set to six residues, at least 2 peptides per protein was required and the false-positive rate was set at 1% at spectrum level. Light/Heavy peptide ratios were quantified based on peak areas on MS1 chromatogram with the in-house software (CIMAGE) as previously described³. The co-elution correlation score filter was set to R² > 0.5 for stable isotope dimethyl labeled samples.
Proteomics data analysis of stereoselectively labeled proteins in pairwise comparisons

To minimize false stereoselective protein targets introduced by, for instance, altered protein expression levels in ‘light’ versus ‘heavy’ SILAC medium, probe treatment was performed in both ‘forward’ and ‘reverse’ experimental design. In ‘forward’ experiment, the isotopically ‘light’ and ‘heavy’ proteomes were treated with (R)- and (S)-probe, respectively; then the isotopically ‘light’ and ‘heavy’ proteomes were treated with (S)- and (R)-probe in ‘reverse’ experiment. Both isotopic directions were performed in biological duplicate to furnish four independent experiments for each enantioprobe pair. Within individual experiment dataset, unique peptides from the same protein were grouped together, and the median ratio from at least two unique peptides was calculated as the protein ratio if its standard deviation was less than 10; if the standard deviation was greater than 10, the peptide ratio closest to 1 was assigned to the protein ratio. Protein ratios processed as above were then averaged across replicate experiments if their standard deviation is less than 60% of the mean; otherwise the ratio closest to 1 was taken as the final ratio for a protein in certain experiment condition.

In addition to the above criteria, to qualify as stereoselective protein targets, proteins must fulfill the following additional criteria: (1) quantified in at least two datasets for any enantioprobe pair. Proteins quantified in only one dataset were considered for stereoselective protein target designation if the protein was stereoselectively labeled by other probes (Supplementary Table 1); (2) For proteins quantified in both ‘forward’ and ‘reverse’ isotopic directions, average stereoselectivity ratios greater than 2.5 were required; for proteins only quantified in one isotopic direction, stereoselectivity ratios greater than 4 were required; (3) have at least 2 unique quantified peptides (unique sequences, ignoring MudPIT salt pump steps, charge states, modifications, etc.) in individual dataset. But if a protein had evidences of being stereoselectively labeled by other probes or in the other cell type, single unique quantified peptide was also reported.

Meta-analysis of stereoselective protein targets

Stereoselective protein targets were queried against the DrugBank database (v. 5.1.1 released on 2018-7-3; group "All") and fractionated into DrugBank and non-DrugBank proteins. Protein targets were queried against the KEGG and UniProtKB/Swiss-Prot Protein Knowledge database for protein classification. Top level terms from KEGG BRITE and Gene Ontology terms were parsed to place each protein into a category (Transporter, Channel and Receptors; Enzymes; Gene Expression and Nucleic Acid Binding; Scaffolding, Modulators and Adaptors). If a protein can be classified into different categories, the abovementioned order of categories was used to prioritize the protein class. If no Gene Ontology term was available to assign protein class, the protein was sorted into the category ‘Uncategorized’. Stereoselective protein targets only discovered in human PBMCs are queried against GeneAtlas U133A, gcrma dataset on BioGPS for tissue-specific patterns of mRNA expression. Protein targets with 3 × median mRNA expression in any of lymphoblast, B cells, T cells, dendritic cells, NK cells, monocytes or myeloid cells are defined as enriched in immune-related cells.

Cloning and transient overexpression of proteins in HEK293T cells

Open reading frame (ORF) of genes of interest were cloned from the Human ORFome V8.1 Library (Dharmacon) into a pRK5-derived plasmid generated using the gateway vector conversion system (Invitrogen, cat # 11828029). Final plasmids contain a CMV promoter followed by a gateway cloning linker, a start methionine, the respective ORF without stop codon, a second gateway cloning linker and a C-terminal DDK tag with its own stop codon. All gene constructs were verified by DNA sequencing. To transiently overexpress proteins for in situ treatment, HEK293T cells were grown to ~60% confluence in 6-well plates (gel-based experiments) or 6-cm plates (MS-based experiments) in complete growth DMEM. 2 – 4 µg of desired plasmid and 6-12 µg of PEI (polyethylenimine, MW 40,000; Polysciences) were mixed in serum-free DMEM and incubated at room temperature for 30 min and added dropwise to the cells. Cells were grown for 48 – 72 hours before treatment. The pRK5 vector was a gift from David Sabatini lab (MIT).

Western blot analysis

After scanning for fluorescence, proteins were transferred to a nitrocellulose membrane (Amersham Protran, cat # 10600011) in Towbin buffer, the membrane was blocked for 1 hour at room temperature with 5% nonfat dry milk (w/v) in Tris-buffered saline with Tween 20 (TBST) and incubated with primary antibodies in the same
solution overnight at 4 °C or 1 hour at room temperature. Blots were washed (3 × 5 min, TBST), incubated with secondary antibodies (IRDye 800CW) in milk for 1 hour at room temperature, washed (3 × 5 min, TBST), rinsed in water and scanned with a LI-COR Odyssey Scanner.

**Preparation of tandem mass tag (TMT) labeled samples for multiplex MS-based analysis**

The procedure to prepare stable isotope dimethyl labeled samples as reported above was modified with the following modifications. After streptavidin enrichment, the beads were centrifuged (1,400 g, 2 min) and washed sequentially with 0.2% SDS in DPBS (1 × 5 mL), DPBS (2 × 5 mL) and 200mM EPDS (Sigma Aldrich, E9502, pH 8, 1 × 5 mL). The beads were transferred to a low-binding 1.5 mL microfuge tube (Axogen, CNT-1.5FL) and the enriched proteins were on-beads digested overnight at 37 °C in ~200 µL of 2 M urea in 200mM EPDS buffer containing 2 µg sequencing grade porcine trypsin (Promega, V5111) in the presence of CaCl₂ (1 mM). The samples were centrifuged to separate the beads and supernatant, then anhydrous acetonitrile was added to the supernatant to 30% final volume. 6 µL (20 µg/µL) of respective 10-plex TMT tag (Thermo Scientific, cat # 90110) was added and the reaction was incubated at room temperature for 1 hour with occasional vortex before quenching with 6 µL of 5% hydroxyamine for 15 min followed by 4 µL formic acid. All 10-plex samples were vacuum-centrifuged to near dryness, reconstituted in Buffer A (95% H₂O, 5% acetonitrile, 0.1% formic acid), combined and stored at -80°C until analysis.

**Mass spectrometry analysis of tandem mass tag (TMT) labeled peptides**

Labeled peptides were pressure loaded onto a 250 µm (inner diameter) fused silica capillary column packed with 4 cm C18 resin (Phenomenex, Aqua 5 µm). Samples were analyzed on an Orbitrap Fusion mass spectrometer (Thermo Scientific) coupled to an UltiMate 3000 Series Rapid Separation LC system and autosampler (Thermo Scientific Dionex). Peptides were separated on a 100 µm inner diameter capillary column with a 5 µm tip packed with 10 cm C18 (Phenomenex, Aqua 5 µm) and 3 cm strong cation exchange resin (SCX, Phenomenex) using a 10-step ‘MudPIT’ protocol that injects 5 µL 5%, 10%, 15%, 20%, 25%, 30%, 40%, 50%, 80%, 100% salt bumps of ammonium acetate (500mM) in buffer A (100% H₂O, 0.1% formic acid) followed by an increasing gradient of buffer B (100% acetonitrile, 0.1% formic acid) in Buffer A in each step. The flow rate was 0.6 µL/min and the voltage applied to the nano-LC electrospray ionization source was 1.9 kV. A MS3-based TMT method was used for data acquisition. The scan sequence began with a MS1 master scan (Orbitrap analysis, resolution 120, 000, 400–1700 m/z, RF lens 60%, automatic gain control [AGC] target 2E5, maximum injection time 50 ms, centroid mode) with dynamic exclusion enabled (repeat count 1, duration 15s). The top ten precursors were then selected for MS2/MS3 analysis. MS2 analysis consisted of: collision-induced dissociation (CID), quadrupole ion trap analysis, AGC 1.8E4, CID collision energy 35%, Activation Q 0.25, maximum injection time 120 ms, and isolation window at 0.7. Following acquisition of each MS2 spectrum, Synchronous Precursor Selection (SPS) was enabled to include up to 10 MS2 fragment ions for the MS3 spectrum. MS3 precursors were fragmented by HCD and analyzed using the Orbitrap (collision energy 55%, AGC 1.5E5, maximum injection time 120 ms, resolution was 50, 000). For MS3 analysis, we used charge state–dependent isolation windows. For charge state z = 2, the MS isolation window was set at 1.2; for z = 3-6, the MS isolation window was set at 0.7.

The spectra files were uploaded to Integrated Proteomics Pipeline (IP2) and searched using ProLuCID algorithm against a reverse-concatenated, nonredundant variant of the Human UniProt database (2012-11). The precursor ion mass tolerance for a minimum envelop of three isotopic peaks was set to 50 ppm. All cysteines were specified with a static modification for carbamidomethylation (+57.02146) and up to two differential modification was allowed per peptide for methionine oxidation (+15.994915). Lysine and N-terminus were also specified with a static modification for TMT tag (+229.1629). The minimum peptide length was set to six residues, at least 1 peptide per protein was required and the false-positive rate was set at 1% at spectrum level. MS3-based peptide quantification was performed with reporter ion mass tolerance set to 20 ppm and intensity threshold set to 5000 for sum of all reporter ions. Proteins were required to have at least three unique peptides (unique sequences, ignoring salt pump steps, charge states, modifications, etc.) and quantified by summing reporter ion intensities across all matching PSMs and normalized to the highest signal channel per protein. All TMT-based experiments were performed in duplicate, and the averaged values were reported.

**Preparation for MS-based analysis of probe-modified peptides**

A procedure previously reported¹ was followed. In brief, HEK293T cells were grown in 6 cm plates and in situ labeled with probes (200 – 250 µM) as described above. The harvested proteomes were adjusted to 1.8 mg/mL (500 µL), to which a freshly prepared ‘click’ mixture (30 µL of 1.7 mM TBTA in 4:1 t-BuOH:DMSO, 10 µL of 50
mM CuSO$_4$ in H$_2$O, 10 µL of 5 mM either ‘light’ or ‘heavy’ isotopically labeled Biotin-TEV-azide$^7$ in DMSO, 10 µL of freshly prepared 50 mM TCEP in DPBS) was added. The samples were rotated at room temperature for 1 hour and centrifuged (16,000 g, 5 min, 4°C). The resulting protein pellets were sonicated in ice-cold methanol (500 µL) and the corresponding ‘light’ and ‘heavy’ samples were then combined and centrifuged. The pellets were sonicated in 1mL of 1.2% SDS in DPBS and heated to 95 °C for 5 min. Then the samples were transferred to 15 mL falcon tubes and diluted with 5 mL DPBS, to which 100 µL of pre-equilibrated streptavidin-agarose beads (Thermo Fisher Scientific, Cat # 20347, 1:1 slurry) was added. After 3 hours incubated on a rotator, the beads were centrifuged (1,400 g, 2 min) and washed sequentially with DPBS (2 × 1 mL) and H$_2$O (2 × 1 mL). The beads were transferred to Eppendorf tubes and resuspended in 500 µL DPBS containing 6 M proteomics-grade urea. DTT (25 µL of 200 mM stock in DPBS) was added and the samples were incubated at 65 °C for 15 min. Then iodoacetamide (25 µL of 400 mM stock in DPBS) was added and allowed to react at 37 °C for 30 min. The beads mixture was diluted with 900 µL DPBS, centrifuged and resuspended in 200 µL of 2 M urea in DPBS containing 2 µg sequencing grade porcine trypsin (Promega, V5111) in the presence of CaCl$_2$ (1 mM). In the next day, the beads were centrifuged, washed with DPBS (2 × 1 mL), H$_2$O (2 × 1 mL), transferred to low-binding Eppendorf tubes, then washed with 150 µL of TEV buffer (50 mM Tris, pH 8, 0.5 mM EDTA, 1 mM DTT). To the washed beads, 150 µL of fresh TEV buffer with 80 µM TEV protease was added and the samples were rotated overnight at 29 °C. The TEV digests were separated from the beads by centrifugation (1,400 g, 3 min) and the beads were washed once with water (100 µL). The samples were then acidified with formic acid (~5% final v/v) and stored at -80°C until analysis. The MS data were collected on a LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific) as described above with differences in the salt bumps applied in the LC gradients, which in this case were 0%, 30%, 60%, 90% and 100% NH$_4$OAc (500 mM). The spectra searches were performed with the following changes applied to identify the peptides modified with the corresponding probe and the cleaved TEV tag. All amino acids were considered as possible residues for differential modification. The mass of the modification used to search was +651.38567 m/z for (R)-1 and (S)-1, +695.41188 m/z for (R,R)-7 and (S,S)-7, which are the monoisotopic masses for the corresponding probe plus the light TEV-tag and an additional +6.0138 m/z for the heavy counterpart. To account for potential diazirine insertion at cysteine thiols, an additional search only considering differential modification on cysteine without static carbamidomethylation (+57.02146) was also performed. The isoTOP ratios for probe labeled peptides were quantified using the in-house software CIMAGE. All MS1 chromatograms and MS2 fragment spectra were manually reviewed to ensure correct assignments.

(D) Synthetic Methods
Chemicals and reagents were purchased from commercial vendors, including Sigma-Aldrich, Fisher Scientific, CombiBlocks, Alfa Aesar and AstaTech, and were used as received without further purification, unless otherwise noted. Anhydrous solvents were purchased from Sigma-Aldrich in Sure/Seal™ formulations. All reactions were monitored by thin-layer chromatography (TLC, Merck silica gel 60 F-254 plates). The plates were stained either with potassium permanganate, anisaldehyde, iodine or directly visualized with UV light. Reaction purification was carried out using Flash chromatography (230 – 400 mesh silica gel) or preparative thin layer chromatography (pTLC, Analtech, 500-2000 µm thickness). NMR spectra were recorded on Bruker DPX-400 or Bruker AV-500 spectrometers in the indicated solvent. Multiplicities are reported with the following abbreviations: s singlet; d doublet; t triplet; q quartet; p pentet; m multiplet; br broad; dd doublet of doublets; dt doublet of triplets; td triplet of doublets; Chemical shifts are reported in ppm relative to the residual solvent peak and J values are reported in Hz. Mass spectrometry data were collected on an Agilent 6120 single-quadrupole LC/MS instrument (ESI, low resolution) or an Agilent ESI-TOF instrument (ESI-TOF, HRMS).

3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)propanoic acid was prepared according to literature procedures¹. Analytical data are in agreement with previously reported data.

General Procedure: coupling procedure for the synthesis of chiral fragment probes. 3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)propanoic acid (0.1 mmol, 1 equiv.) and the commercially-available chiral amine (0.11 mmol, 1.1 equiv.) were dissolved in anhydrous dichloromethane (1.5 mL) in a 4 mL vial. DIPEA (2.2 equiv.), EDCI (1.1 equiv.) and HOAt (1.1 equiv.) were added. The reaction mixture was stirred overnight at room temperature or when TLC indicated reaction completed. The crude mixture was concentrated by blowing a constant stream of nitrogen over it to remove the excess solvent. The resulting mixture was diluted with ethyl acetate (5 mL) and washed with saturated aqueous NH₄Cl (5 mL), saturated aqueous NaHCO₃ (5 mL) and brine (5 mL). The organic phase was dried over anhydrous Na₂SO₄, concentrated by removing solvent under reduced pressure and purified by flash chromatography or preparative TLC. The chromatography was run with the indicated solvent and the product was eluted from the silica. Evaporation of the solvent resulted in the desired product.
(R)-3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)-1-(2-phenylpyrrolidin-1-yl)propan-1-one ((R)-1). General procedure. The flash chromatography was run with n-hexane/ethyl acetate 3:1. 21.4 mg (72%) of the product were obtained. $^1$H NMR (500 MHz, Chloroform-$d$) (70:30 mixture of rotamers, peaks corresponding to minor rotamer starred) δ 7.35 (m, 1.35H), 7.27 (m, 1.35H), 7.21 – 7.16* (m, 0.3H), 7.16 – 7.10 (m, 2H), 5.19* (dd, $J = 8.1, 2.7$ Hz, 0.3H), 4.88 (dd, $J = 8.0, 2.4$ Hz, 0.7H), 3.79 – 3.64 (m, 1.7H), 3.54* (ddd, $J = 9.9, 8.6, 7.2$ Hz, 0.3H), 2.45 – 2.32 (m, 0.7H), 2.29 – 2.18* (m, 0.3H), 2.09 (t, $J = 7.6$ Hz, 0.6H), 2.04 – 1.95 (m, 1.4H), 1.95 – 1.82 (m, 5.7H), 1.80 – 1.62 (m, 2.3H), 1.62 – 1.55 (m, 0.7H), 1.55 – 1.41 (m, 1.3H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 171.10, 169.95, 143.53, 143.34, 129.28, 128.77, 127.82, 127.09, 125.80, 125.77, 83.22, 83.16, 77.69, 77.43, 77.18, 69.46, 61.99, 60.97, 47.92, 47.58, 36.76, 34.35, 33.03, 32.78, 28.85, 28.79, 28.23, 27.88, 24.07, 22.16, 13.73, 13.61. HRMS (m/z) calculated for C$_{18}$H$_{22}$N$_3$O [M+H]$^+$: 296.1757; found: 296.1760.

(S)-3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)-1-(2-phenylpyrrolidin-1-yl)propan-1-one ((S)-1). General procedure. The flash chromatography was run with n-hexane/ethyl acetate 3:1. 22.3 mg (75%) of the product were obtained. $^1$H NMR (500 MHz, Chloroform-$d$) (70:30 mixture of rotamers, peaks corresponding to minor rotamer starred) δ 7.38 – 7.32 (m, 1.35H), 7.31 – 7.25 (m, 1.35H), 7.21 – 7.16* (m, 0.3H), 7.16 – 7.10 (m, 2H), 5.19* (dd, $J = 8.1, 2.7$ Hz, 0.3H), 4.88 (dd, $J = 8.0, 2.4$ Hz, 0.7H), 3.79 – 3.64 (m, 1.7H), 3.54* (ddd, $J = 9.9, 8.6, 7.3$ Hz, 0.3H), 2.44 – 2.34 (m, 0.7H), 2.29 – 2.19* (m, 0.3H), 2.09 (t, $J = 7.6$ Hz, 0.6H), 2.04 – 1.95 (m, 1.4H), 1.94 – 1.84 (m, 5.7H), 1.79 – 1.62 (m, 2.3H), 1.61 – 1.55 (m, 0.7H), 1.55 – 1.41 (m, 1.3H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 170.76, 169.61, 143.12, 142.94, 128.90, 128.39, 127.44, 126.71, 125.42, 125.39, 125.37, 82.84, 82.78, 77.29, 77.04, 76.78, 69.08, 68.98, 61.61, 61.59, 60.59, 60.57, 47.54, 47.20, 47.18, 36.36, 33.96, 32.62, 32.38, 29.71, 28.46, 28.40, 28.38, 28.03, 27.83, 27.81, 27.48, 23.66, 21.76, 21.74, 13.34, 13.22. HRMS (m/z) calculated for C$_{18}$H$_{22}$N$_3$O [M+H]$^+$: 296.1757; found: 296.1761.

Enantiopurity was determined by SFC (see below). $T_{rac} = 1.925$ and 2.293 min.

|        | Retention (major) | Retention (minor) | Area (major) | Area (minor) | e.e. |
|--------|------------------|------------------|-------------|-------------|-----|
| (R)-1  | 1.911            | 2.277            | 1049882     | 12758       | 97.6|
| (S)-1  | 2.280            | 1.966            | 1015497     | 4775        | 99.1|
(R)-3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)-N-(1-hydroxy-3-(1H-indol-3-yl)propan-2-yl)propenamide (([R]-2)). General procedure. The flash chromatography was run with n-hexane/ethyl acetate 2:3. 24.7 mg (73%) of the product were obtained. $^1$H NMR (500 MHz, Chloroform-d) $\delta$ 8.25 (s, 1H), 7.64 (d, $J = 7.9$ Hz, 1H), 7.36 (d, $J = 8.1$ Hz, 1H), 7.24 – 7.17 (m, 1H), 7.17 – 7.09 (m, 1H), 7.04 (d, $J = 2.4$ Hz, 1H), 5.83 (d, $J = 7.6$ Hz, 1H), 4.35 – 4.21 (m, 1H), 3.78 – 3.55 (m, 2H), 3.01 (d, $J = 6.9$ Hz, 2H), 2.90 (d, $J = 5.1$ Hz, 1H), 2.01 – 1.93 (m, 3H), 1.87 (t, $J = 7.1$ Hz, 2H), 1.81 – 1.73 (m, 2H), 1.57 (t, $J = 7.4$ Hz, 2H). $^{13}$C NMR (126 MHz, CDCl$_3$) $\delta$ 171.91, 136.30, 127.62, 122.76, 122.32, 119.74, 118.68, 111.47, 111.34, 82.77, 77.31, 77.05, 76.80, 69.31, 64.47, 60.45, 52.26, 32.27, 30.52, 28.30, 27.91, 26.51, 14.21, 13.26. HRMS (m/z) calculated for C$_{19}$H$_{23}$N$_4$O$_2$ [M+H]$^+$: 339.1815; found: 339.1817.

Enantiopurity was determined by SFC (see below). $T_{ Rac}$ = 1.721 and 1.975 min.

| Retention (major) | Retention (minor) | Area (major) | Area (minor) | e.e. (%) |
|------------------|------------------|-------------|-------------|----------|
| 1.331            | 1.538            | 1.533       | 1.832       | 18.27    |
| 1.540            | 2.070            | 1.657       | 2.280       | 18.42    |

(S)-3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)-N-(1-hydroxy-3-(1H-indol-3-yl)propan-2-yl)propenamide (([S]-2)). General procedure. The flash chromatography was run with n-hexane/ethyl acetate 2:3. 27.8 mg (82%) of the product were obtained. $^1$H NMR (500 MHz, Chloroform-d) $\delta$ 8.17 (s, 1H), 7.65 (d, $J = 7.9$ Hz, 1H), 7.37 (d, $J = 8.1$ Hz, 1H), 7.24 – 7.18 (m, 1H), 7.14 (t, $J = 7.5$ Hz, 1H), 7.06 (d, $J = 2.4$ Hz, 1H), 5.77 (d, $J = 7.5$ Hz, 1H), 4.35 – 4.22 (m, 1H), 3.77 – 3.61 (m, 2H), 3.02 (d, $J = 6.9$ Hz, 2H), 2.72 (s, 1H), 2.03 – 1.92 (m, 3H), 1.88 (t, $J = 7.2$ Hz, 2H), 1.83 – 1.72 (m, 2H), 1.58 (t, $J = 7.2$ Hz, 2H). $^{13}$C NMR (126 MHz, CDCl$_3$) $\delta$ 171.85, 136.30, 127.60, 122.70, 119.79, 118.69, 111.53, 111.32, 82.75, 77.29, 77.03, 76.78, 69.27, 64.60, 52.27, 32.30, 30.95, 30.52, 28.29, 27.89, 26.52, 13.26. HRMS (m/z) calculated for C$_{19}$H$_{23}$N$_4$O$_2$ [M+H]$^+$: 339.1815; found: 339.1813.

Enantiopurity was determined by SFC (see below). $T_{ Rac}$ = 1.721 and 1.975 min.
(R)-3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)-N-(1,2,3,4-tetrahydronaphthalen-1-yl)propenamide ((R)-3). General procedure. The preparative TLC was run with n-hexane/ethyl acetate 3:1. 19.6 mg (66%) of the product were obtained. $^1$H NMR (500 MHz, Chloroform-d) δ 7.25 (d, J = 2.5 Hz, 1H), 7.18 (td, J = 6.8, 6.0, 3.9 Hz, 2H), 7.13 – 7.07 (m, 1H), 5.67 (d, J = 8.5 Hz, 1H), 5.17 (dd, J = 8.9, 5.2 Hz, 1H), 2.78 (qt, J = 17.0, 5.9 Hz, 2H), 2.03 (td, J = 7.6, 2.9 Hz, 3H), 1.98 – 1.92 (m, 3H), 1.89 (m, 2H). 13C NMR (126 MHz, CDCl₃) δ 170.20, 137.63, 136.45, 129.23, 128.76, 127.39, 126.31, 82.72, 77.29, 77.04, 76.78, 69.22, 47.53, 32.46, 30.61, 30.04, 29.71, 29.21, 28.40, 27.88, 19.85, 13.32. HRMS (m/z) calculated for C₁₈H₂₂N₃O [M+H]^+: 296.1757; found: 296.1756.

Enantiopurity was determined by SFC (see below). T_{rac} = 1.441 and 1.854 min.

(S)-3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)-N-(1,2,3,4-tetrahydronaphthalen-1-yl)propenamide ((S)-3). General procedure. The preparative TLC was run with n-hexane/ethyl acetate 3:1. 22.6 mg (77%) of the product were obtained. $^1$H NMR (500 MHz, Chloroform-d) δ 7.25 (d, J = 2.6 Hz, 1H), 7.18 (td, J = 6.8, 5.9, 3.8 Hz, 2H), 7.12 – 7.06 (m, 1H), 5.67 (d, J = 8.5 Hz, 1H), 5.17 (dt, J = 9.4, 5.4 Hz, 1H), 2.78 (qt, J = 17.0, 6.0 Hz, 2H), 2.03 (ddd, J = 10.4, 8.3, 4.5 Hz, 3H), 1.95 (dt, J = 12.2, 3.3 Hz, 3H), 1.89 (m, 2H), 1.83 (m, 3H), 1.66 (t, J = 7.1 Hz, 2H). 13C NMR (126 MHz, CDCl₃) δ 170.20, 137.63, 136.45, 129.23, 128.76, 127.39, 126.31, 82.72, 77.29, 77.04, 76.78, 69.22, 47.53, 32.46, 30.61, 30.04, 29.71, 29.21, 28.40, 27.88, 19.85, 13.32. HRMS (m/z) calculated for C₁₈H₂₂N₃O [M+H]^+: 296.1757; found: 296.1756.
|     | Retention (major) | Retention (minor) | Area (major) | Area (minor) | e.e. (%) |
|-----|-------------------|-------------------|--------------|--------------|---------|
| (R)-3 | 1.427            | 1.833             | 915936       | 11204        | 97.6    |
| (S)-3 | 1.826            | 1.430             | 1113421      | 2175         | 99.6    |

(R)-3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)-N-(2,3-dihydro-1H-inden-1-yl)propenamide ((R)-4). General procedure. The preparative TLC was run with n-hexane/ethyl acetate 3:1. 21.0 mg (75%) of the product were obtained. $^1$H NMR (500 MHz, Chloroform-d) $\delta$ 7.30 – 7.27 (m, 1H), 7.25 – 7.18 (m, 3H), 5.69 (d, $J$ = 8.5 Hz, 1H), 5.46 (q, $J$ = 7.7 Hz, 1H), 2.98 (m, 1H), 2.58 (m, 1H), 2.03 (td, $J$ = 7.5, 2.6 Hz, 2H), 1.98 – 1.93 (m, 3H), 1.89 (m, 2H), 1.80 (m, 1H), 1.66 (t, $J$ = 7.5 Hz, 2H). $^{13}$C NMR (126 MHz, CDCl$_3$) $\delta$ 170.81, 143.47, 142.96, 128.09, 126.82, 124.03, 82.72, 77.30, 77.04, 76.79, 69.24, 54.77, 33.99, 32.45, 30.48, 30.23, 28.40, 27.89, 13.32. HRMS (m/z) calculated for C$_{17}$H$_{20}$N$_3$O [M+H]$^+$: 282.1601; found: 282.1598.

Enantiopurity was determined by SFC (see below). $T_{rac} = 1.400$ and 1.596 min.
|      | Retention (major) | Retention (minor) | Area (major) | Area (minor) | e.e. (%) |
|------|------------------|------------------|-------------|-------------|----------|
| (R)-4 | 1.384            | n.d.             | 940583      | n/a         | 100      |
| (S)-4 | 1.576            | n.d.             | 1082102     | n/a         | 100      |

(R)-N-(1-benzylpyrrolidin-3-yl)-3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)-N-methylpropanamide (((R)-5)). General procedure. The flash chromatography was run with n-hexane/ethyl acetate 3:2 to 2:3. 27.8 mg (82%) of the product were obtained. $^1$H NMR (500 MHz, Chloroform-d) (60:40 mixture of rotamers, peaks corresponding to minor rotamer starred) $\delta$ 7.30 (m, 4H), 7.27 – 7.21 (m, 1H), 5.22 (m, 0.6H), 4.37* (m, 0.4H), 3.64 (d, $J = 12.9$ Hz, 1H), 3.50 (dd, $J = 18.2, 12.9$ Hz, 1H), 2.92 (m, 3.6H), 2.79* (td, $J = 8.7, 4.1$ Hz, 0.4H), 2.62 – 2.57 (m, 1H), 2.54* (m, 0.4H), 2.39 (q, $J = 8.6, 7.9$ Hz, 1H), 2.27 – 2.10 (m, 2H), 2.10 – 1.94 (m, 5H), 1.85 – 1.80 (m, 2H), 1.65 (t, $J = 7.5$ Hz, 2H). $^{13}$C NMR (126 MHz, CDCl$_3$) $\delta$ 171.05, 170.50, 139.07, 138.65, 128.56, 128.53, 128.37, 128.29, 127.17, 127.01, 82.81, 77.31, 77.05, 76.80, 69.13, 69.10, 60.18, 60.11, 57.19, 57.11, 55.54, 55.83, 53.83, 53.58, 51.82, 32.58, 32.55, 30.05, 29.71, 29.06, 28.96, 28.30, 28.17, 27.98, 27.91, 27.62, 27.24, 13.33. HRMS (m/z) calculated for C$_{20}$H$_{27}$N$_4$O [M+H]$^+$: 339.2179; found: 339.2185.

(S)-N-(1-benzylpyrrolidin-3-yl)-3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)-N-methylpropanamide (((S)-5)). General procedure. The flash chromatography was run with n-hexane/ethyl acetate 3:2 to 2:3. 25.7 mg (76%) of the product were obtained. $^1$H NMR (500 MHz, Chloroform-d) (60:40 mixture of rotamers, peaks corresponding to minor rotamer starred) $\delta$ 7.36 – 7.30 (m, 4H), 7.26 (m, 1H), 5.25 (m, 0.6H), 4.40* (m, 0.4H), 3.67 (d, $J = 12.9$ Hz, 1H), 3.53 (dd, $J = 17.8, 12.9$ Hz, 1H), 2.95 (m, 3.6H), 2.82* (td, $J = 8.7, 4.3$ Hz, 0.4H), 2.62 (dd, $J = 10.1, 3.8$ Hz,
1H), 2.56* (dd, J = 10.1, 8.0 Hz, 0.4H), 2.46 – 2.38 (m, 1H), 2.29 – 2.12 (m, 2H), 2.12 – 1.96 (m, 5H), 1.89 – 1.82 (m, 2H), 1.68 (t, J = 7.6 Hz, 2H). 13C NMR (126 MHz, CDCl3) δ 171.05, 170.50, 138.64, 128.56, 128.53, 128.37, 128.29, 127.17, 127.01, 82.80, 77.30, 77.04, 76.79, 69.12, 69.10, 60.17, 60.11, 57.18, 57.11, 55.54, 53.83, 53.57, 51.82, 32.58, 32.55, 30.05, 29.71, 29.06, 28.95, 28.30, 28.17, 27.98, 27.91, 27.63, 27.24, 13.33. HRMS (m/z) calculated for C20H27N4O [M+H]+: 339.2179; found: 339.2184.

Enantiopurity was determined by SFC (see below). T_{rac} = 1.296 and 1.474 min.

|        | Retention (major) | Area (major) | Retention (minor) | Area (minor) | e.e. (%) |
|--------|------------------|--------------|------------------|--------------|----------|
| (R)-5  | 1.293            | 1152767      | 1.474            | 64955        | 89.3     |
| (S)-5  | 1.274            | 1164051      | 1.471            | 10419        | 98.2     |

(R)-N-benzyl-3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)-N-(1-phenylethyl)propenamide ((R)-6). General procedure. The preparative TLC was run with n-hexane/ethyl acetate 13:7. 26.0 mg (72%) of the product were obtained. 1H NMR (500 MHz, Chloroform-d) (70:30 mixture of rotamers, peaks corresponding to minor rotamer starred) δ 7.33 – 7.16 (m, 8H), 7.12 (m, 0.7H), 7.05 (m, 1.3H), 6.17 (q, J = 7.2 Hz, 0.7H), 5.13 (q, J = 7.2 Hz, 0.3H), 4.94 (d, J = 15.5 Hz, 0.3H), 4.35 (d, J = 17.9 Hz, 0.7H), 4.18 (d, J = 17.9 Hz, 0.7H), 3.98 (d, J = 15.5 Hz, 0.3H), 2.23 (m, 0.7H), 2.06 – 1.77 (m, 6.3H), 1.66 (m, 0.7H), 1.58 (m, 1.3H), 1.46 (d, J = 7.3 Hz, 3H). 13C NMR (126 MHz, CDCl3) δ 172.42, 140.93, 138.09, 128.83, 128.72, 128.53, 128.28, 127.55, 127.51, 127.27, 127.20, 126.58, 125.83, 82.80, 77.29, 77.04, 76.78, 69.06, 51.80, 47.19, 46.55, 32.55, 28.11, 27.97, 16.96, 13.29. HRMS (m/z) calculated for C23H28N3O [M+H]+: 360.2070; found: 360.2070.
(S)-N-benzyl-3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)-N-(1-phenylethyl)propenamide ((S)-6). General procedure. The preparative TLC was run with n-hexane/ethyl acetate 13:7. 29.2 mg (81%) of the product were obtained. $^1$H NMR (500 MHz, Chloroform-d) (70:30 mixture of rotamers, peaks corresponding to minor rotamer starred) δ 7.39 – 7.17 (m, 8H), 7.14 (m, 0.7H), 7.07 (m, 1.3H), 6.20 (q, J = 7.2 Hz, 0.7H), 5.16 (q, J = 7.0 Hz, 0.3H), 4.96 (d, J = 15.5 Hz, 0.3H), 4.37 (d, J = 17.9 Hz, 0.3H), 4.21 (d, J = 17.9 Hz, 0.7H), 4.00 (d, J = 15.5 Hz, 0.3H), 2.25 (m, 0.7H), 2.08 – 1.79 (m, 6.3H), 1.68 (m, 0.7H), 1.60 (m, 1.3H), 1.48 (d, J = 7.3 Hz, 3H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 172.42, 140.93, 138.09, 128.83, 128.72, 128.53, 128.28, 127.55, 127.51, 127.28, 127.20, 126.58, 125.83, 82.79, 77.29, 77.04, 76.79, 69.06, 51.80, 47.19, 32.55, 28.11, 27.97, 16.96, 13.29. HRMS (m/z) calculated for C$_{23}$H$_{26}$N$_3$O [M+H]$^+$: 360.2070; found: 360.2070.

Enantiopurity was determined by SFC (see below). T$_{rac}$ = 1.837 and 2.139 min.

| (R)-6  | Retention (major) | Retention (minor) | Area (major) | Area (minor) | e.e. (%) |
|--------|------------------|------------------|--------------|--------------|---------|
| 1.832  | 2.131            |                  | 1581571      | 50450        | 93.8    |

| (S)-6  | Retention (major) | Retention (minor) | Area (major) | Area (minor) | e.e. (%) |
|--------|------------------|------------------|--------------|--------------|---------|
| 1.831  | 2.124            |                  | 1624433      | 5316         | 99.4    |

N-((1R,2R)-2-(benzyloxy)cyclopentyl)-3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)propenamide ((R,R)-7). The preparative TLC was run with n-hexane/ethyl acetate 13:7. 24.1 mg (71%) of the product were obtained. $^1$H NMR (500 MHz, Chloroform-d) δ 7.36 – 7.30 (m, 4H), 7.29 – 7.26 (m, 1H), 5.30 (d, J = 7.3 Hz, 1H), 4.60 (q, J = 12.0 Hz, 2H), 4.20 (m, 1H), 3.77 (dt, J = 6.3, 3.9 Hz, 1H), 2.18 (m, 1H), 2.04 – 1.94 (m, 3H), 1.91 – 1.66 (m, 7H), 1.66 – 1.60 (m, 3H), 1.39 (m, 1H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 170.60, 138.67, 128.36, 127.69, 127.51, 84.59, 82.72,
77.29, 77.03, 76.78, 71.09, 69.22, 55.93, 32.45, 30.57, 30.48, 30.46, 28.31, 27.88, 21.68, 13.29. HRMS (m/z) calculated for C\textsubscript{20}H\textsubscript{26}N\textsubscript{3}O\textsubscript{2} [M+H]\textsuperscript{+}: 340.2019; found: 340.2027.

\textbf{N-((1S,2S)-2-(benzylloxy)cyclopentyl)-3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)propenamide ((R,S)-7).} General procedure. The preparative TLC was run with n-hexane/ethyl acetate 13:7. 21.5 mg (63\%) of the product were obtained. \textsuperscript{1}H NMR (500 MHz, Chloroform-d) \(\delta\) 7.35 – 7.31 (m, 4H), 7.29 – 7.25 (m, 1H), 5.37 – 5.20 (m, 1H), 4.60 (q, \(J = 12.1\) Hz, 2H), 4.20 (m, 2H), 3.77 (dt, \(J = 6.3, 3.9\) Hz, 1H), 2.18 (m, 1H), 1.93 – 1.76 (m, 6H), 1.76 – 1.69 (m, 1H), 1.68 – 1.57 (m, 4H), 1.39 (m, 1H). \textsuperscript{13}C NMR (126 MHz, CDCl\textsubscript{3}) \(\delta\) 170.59, 138.67, 128.36, 127.69, 127.51, 84.59, 82.72, 77.28, 77.03, 76.78, 71.09, 69.22, 55.93, 32.46, 30.57, 30.48, 30.46, 28.31, 27.88, 21.68, 13.29. HRMS (m/z) calculated for C\textsubscript{20}H\textsubscript{26}N\textsubscript{3}O\textsubscript{2} [M+H]\textsuperscript{+}: 340.2019; found: 340.2025.

Enantiopurity was determined by SFC (see below). \(T_{90} = 1.428\) and 1.530 min.

|       | Retention (major) | Retention (minor) | Area (major) | Area (minor) | e.e. (%) |
|-------|-------------------|-------------------|--------------|--------------|----------|
| (R,R)-7 | 1.528             | 1.426             | 753576       | 4809         | 98.7     |
| (S,S)-7 | 1.427             | n.d.              | 805126       | n/a          | 100      |

(R)-3-(3-(but-3-yn-1-yl)-3H-diazirin-3-yl)-N-(1-hydroxy-3-methyl-1,1-diphenylbutan-2-yl)propenamide ((R)-8). The preparative TLC was run with n-hexane/ethyl acetate 5:3. 38.7 mg (96\%) of the product were obtained. \textsuperscript{1}H NMR (500 MHz, Chloroform-d) \(\delta\) 7.46 (m, 4H), 7.31 (dd, \(J = 8.5, 7.1\) Hz, 2H), 7.28 – 7.24 (m, 2H), 7.23 – 7.14 (m, 2H), 5.97 (d, \(J = 9.9\) Hz, 1H), 4.95 (dd, \(J = 9.9, 2.3\) Hz, 1H), 2.87 (s, 1H), 1.96 – 1.92 (m, 3H), 1.90 – 1.78 (m, 2H), 1.76 – 1.62 (m, 2H), 1.56 – 1.46 (m, 2H), 0.90 (dd, \(J = 7.7, 6.8\) Hz, 6H). \textsuperscript{13}C NMR (126 MHz, CDCl\textsubscript{3}) \(\delta\) 171.25, 146.02, 145.32, 128.49, 128.41, 126.99, 126.97, 125.33,
The preparative TLC was run with n-hexane/ethyl acetate 5:3. 38.4 mg (95%) of the product were obtained. \(^1\)H NMR (500 MHz, Chloroform-\(d\)) \(\delta\) 7.46 (m, 4H), 7.31 (dd, \(J = 8.5, 7.1\) Hz, 2H), 7.26 (t, \(J = 7.8\) Hz, 2H), 7.23 – 7.14 (m, 2H), 5.98 (d, \(J = 9.9\) Hz, 1H), 4.95 (dd, \(J = 10.0, 2.3\) Hz, 1H), 2.88 (s, 1H), 1.97 – 1.91 (m, 3H), 1.90 – 1.77 (m, 2H), 1.77 – 1.62 (m, 2H), 1.56 – 1.49 (m, 2H), 0.90 (dd, \(J = 8.5, 6.8\) Hz, 6H). \(^13\)C NMR (126 MHz, CDCl\(_3\)) \(\delta\) 171.26, 146.03, 145.32, 128.49, 128.41, 126.99, 126.97, 125.33, 125.20, 82.73, 82.31, 77.29, 77.04, 76.78, 69.19, 57.86, 32.22, 30.60, 28.82, 28.39, 27.79, 22.88, 17.81, 13.30. HRMS (m/z) calculated for C\(_{25}\)H\(_{29}\)N\(_3\)O\(_2\)Na \([\text{M+Na}]^+\): 426.2152; found: 426.2159.

Enantiopurity was determined by SFC (see below). \(T_{\text{rac}} = 1.237\) and 1.367 min.

|     | Retention (major) | Area (major) | Retention (minor) | Area (minor) | e.e. (%) |
|-----|------------------|--------------|-------------------|--------------|---------|
| (R)-8 | 1.236            | 1447134      | n.d.              | n/a          | 100     |
| (S)-8 | 1.366            | 1446498      | 1.225             | 26237        | 96.4    |
(E) NMR spectra

$^1$H-NMR for ($R$)-1

$^{13}$C-NMR for ($R$)-1
$^1$H-NMR for (S)-1

$^{13}$C-NMR for (S)-1
$^1$H-NMR for (R)-2

$^{13}$C-NMR for (R)-2
$^1$H-NMR for (S)-2

$^{13}$C-NMR for (S)-2
$^1\text{H-NMR for } (R)-3$

$^{13}\text{C-NMR for } (R)-3$
$^1$H-NMR for (S)-3

$^{13}$C-NMR for (S)-3
$^1$H-NMR for (R)-4

$^{13}$C-NMR for (R)-4
$^1$H-NMR for (S)-4

$^{13}$C-NMR for (S)-4
$^{1}H$-NMR for (R)-5

$^{13}$C-NMR for (R)-5
$^1$H-NMR for (R)-6

$^{13}$C-NMR for (R)-6
$^1\text{H-NMR for (S)-6}$

$^{13}\text{C-NMR for (S)-6}$
$^1$H-NMR for (R)-7

$^{13}$C-NMR for (R)-7
$^1$H-NMR for (R)-8

$^{13}$C-NMR for (R)-8
$^1$H-NMR for (S)-8

$^{13}$C-NMR for (S)-8
(F) References

1. Parker, C. G. et al. Ligand and Target Discovery by Fragment-Based Screening in Human Cells. *Cell* **168**, 527-541 e529, doi:10.1016/j.cell.2016.12.029 (2017).
2. Parker, C. G. et al. Chemical Proteomics Identifies SLC25A20 as a Functional Target of the Ingenol Class of Actinic Keratosis Drugs. *ACS Central Science* **3**, 1276-1285, doi:10.1021/acscentsci.7b00420 (2017).
3. Weerapana, E. et al. Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790, doi:10.1038/nature09472 (2010).
4. Maciejewski, A. et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Research* **46**, D1074-D1082, doi:10.1093/nar/gkx1037 %J Nucleic Acids Research (2017).
5. Su, A. I. et al. A gene atlas of the mouse and human protein-encoding transcriptomes. *Science* **320**, 560-565, doi:10.1126/science.1135821 %J Proceedings of the National Academy of Sciences of the United States of America (2004).
6. Afrasiabi, C., Tsueng, G., Su, A. I., Wu, C. & Jin, X. BioGPS: building your own mash-up of gene annotations and expression profiles. *Nucleic Acids Research* **44**, D313-D316, doi:10.1093/nar/gkv1104 %J Nucleic Acids Research (2015).
7. Backus, K. M. et al. Proteome-wide covalent ligand discovery in native biological systems. *Nature* **534**, 570-574, doi:10.1038/nature18002 (2016).
| File Name                                      | Size       | Download | View on ChemRxiv |
|-----------------------------------------------|------------|----------|------------------|
| Wang_etal_2019_SupplementaryTable3.xlsx       | 10.18 MiB  | download | view on ChemRxiv |
| Wang_etal_2019_SupplementaryTable2.xlsx       | 30.62 MiB  | download | view on ChemRxiv |
| Wang_etal_2019_SupplementaryTable1.xlsx       | 329.95 KiB | download | view on ChemRxiv |