

A chemoproteomic platform for reactive fragment profiling against the deubiquitinases

Rosa Cookson^{1,2,3†}, Aini Vuorinen^{1,2,3†}, Jonathan Pettinger¹, Cassandra R. Kennedy², Joanna M. Kirkpatrick³, Rachel E. Peltier-Heap^{1,3}, Andrew Powell¹, Ambrosius P. Snijders³, Mark Skehel³, David House¹, Katrin Rittinger^{2*}, Jacob T. Bush^{1*}

¹ Crick-GSK Biomedical LinkLabs, GSK, Gunnels Wood Road, Stevenage, Hertfordshire, SG1 2NY, UK.

² Molecular Structure of Cell Signalling Laboratory, The Francis Crick Institute, 1 Midland Road, London, NW1 1AT, UK.

³ Proteomics Science Technology Platform, The Francis Crick Institute, 1 Midland Road, London, NW1 1AT, UK.

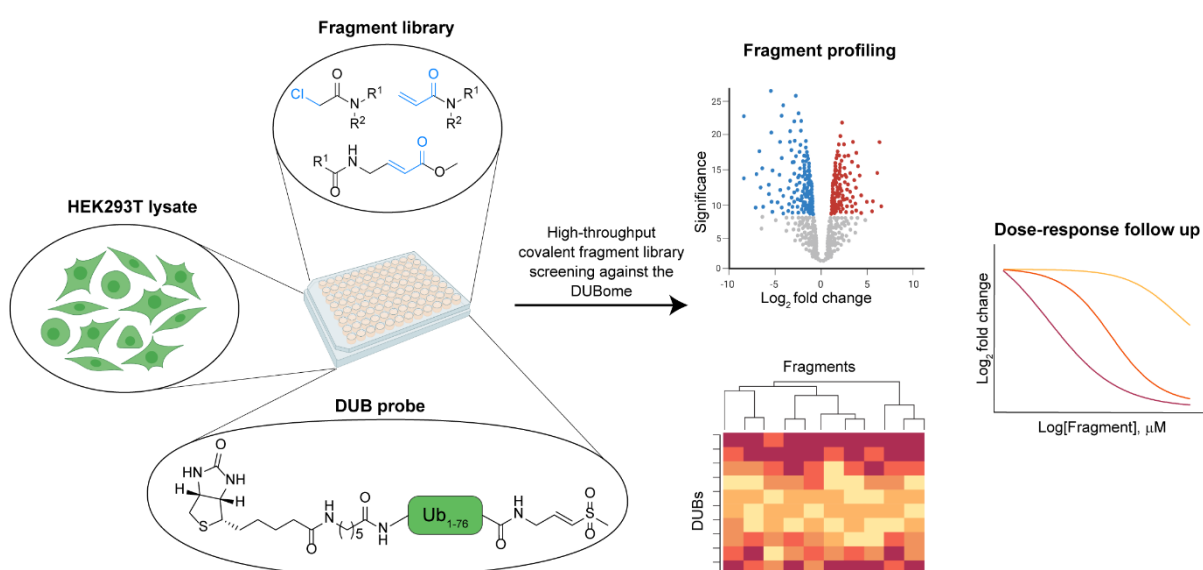
† Authors contributed equally

* Corresponding author

jacob.x.bush@gsk.com

katrin.rittinger@crick.ac.uk

Graphical abstract



Summary

Chemoproteomics is a powerful method capable of detecting interactions between small molecules and the proteome, however its use as a high-throughput screening method for chemical libraries has so far been limited. To address this need, we have further developed a chemoproteomics workflow to screen cysteine reactive covalent fragments in cell lysates against the deubiquitinating (DUB) enzymes using activity-based protein profiling. By using targeted ubiquitin probes, we have addressed sensitivity and affinity limitations, enabling target identification and covalent fragment library profiling in a 96-well plate format. The use of data independent acquisition (DIA) methods for MS analysis combined with automated Evosep liquid chromatography (LC) reduced instrument runtimes to 21 minutes per sample and simplified the workflow. In this proof-of-concept study, we have profiled 138 covalent fragments against 57 DUB proteins and validated four hit fragments against OTUD7B and UCHL3 through site identification experiments and orthogonal biochemical activity assays.

Key words

Deubiquitinases, reactive fragments, chemoproteomics, activity-based protein profiling, mass spectrometry, fragment-based drug discovery

Introduction

The development of chemical probes for the entire human proteome is an ambitious goal of the research community that will have profound impact on our understanding of protein function and human disease, ultimately translating into new innovative medicines.¹⁻³ Currently, ligands exist for approximately 10% of the proteome, and just 22-40% is estimated to be druggable by traditional small molecules.^{4,5} Recent advances in genomic approaches are highlighting novel, understudied targets from the human genome, and chemical probes provide an essential method to enable rapid validation and prioritisation of these targets.⁶⁻¹¹

Traditional small molecule screening has so far been well utilised for chemical probe discovery, however, in order to accelerate this discovery, new orthogonal technologies are required.¹²⁻¹⁶ One such opportunity is the screening of reactive fragment libraries through chemoproteomic workflows.¹⁷⁻¹⁹ Fragments allow efficient coverage of chemical space with small libraries, while the functionalisation of the fragment with a reactive group affords increased potency and selectivity, permitting robust detection of fragment-protein interactions in cells (Figure 1A). Reactive fragment screening is typically performed with purified protein and has had significant impact in this field,¹⁹⁻²⁶ even enabling targeting of shallow protein pockets. Examples of fragment-derived medicines include inhibitors of KRas^{G12C} (Sotorasib),²⁷ tumour target Pin1,²⁸ and the SARS-CoV-2 target M^{pro}.²⁹

Forthcoming advances will likely take in fragment-based drug discovery from an *in vitro* screening method to an *in cellulo* platform capable of simultaneously detecting and studying fragment-protein interactions across the proteome.^{17, 18, 30, 31} Importantly, this would enable profiling of target engagement of endogenous proteins in their native context, where protein-protein interactions, post-translational modifications and subcellular localisation alter their activities and ligandability. Furthermore, such a platform allows the study of proteins that are difficult to isolate recombinantly for *in vitro* screening and assays.

So far, a key limitation to developing this reactive fragment screening platform in cells has been the sensitivity of mass spectrometry (MS) analysis, which makes achieving sufficient coverage of functionally relevant protein sites a challenge. Despite rapid advances in profiling cysteine reactive fragments against the ‘cysteinome’ using iodoacetamide based probes,^{32, 33} only ~10% of the cysteinome can currently be liganded with fragments. One approach to circumvent this is to use targeted probes against a protein family of interest, for example activity-based probes (ABPs) which target a conserved pocket or protein surface. Employing these ABPs as reporter molecules in a chemoproteomics increases the depth of protein coverage

within the desired protein family when compared to a cysteinome-wide approach to a level where fragment screening can be employed effectively.

A second barrier for chemoproteomic screening of chemical libraries is the time required for liquid chromatography (LC), which can typically be up to three hours per sample. Significant improvements have been achieved through isotope labelling, enabling multiplexing and analysis of 10-15 samples in one run.^{32, 34, 35} However, isotope labelling reagents are expensive and therefore not ideally suited to screening large numbers of compounds. In addition, data-dependent acquisition (DDA) is the most frequently utilised MS method in chemoproteomics, but this can limit the comparability between MS runs, a necessity for robust profiling of a library. Recent advances in data independent acquisition (DIA) and label free quantification methods have enabled shorter LC run times, and have reduced the number of missing values, relative to DDA,³⁶⁻³⁸ offering great potential for library profiling.^{36, 37, 39}

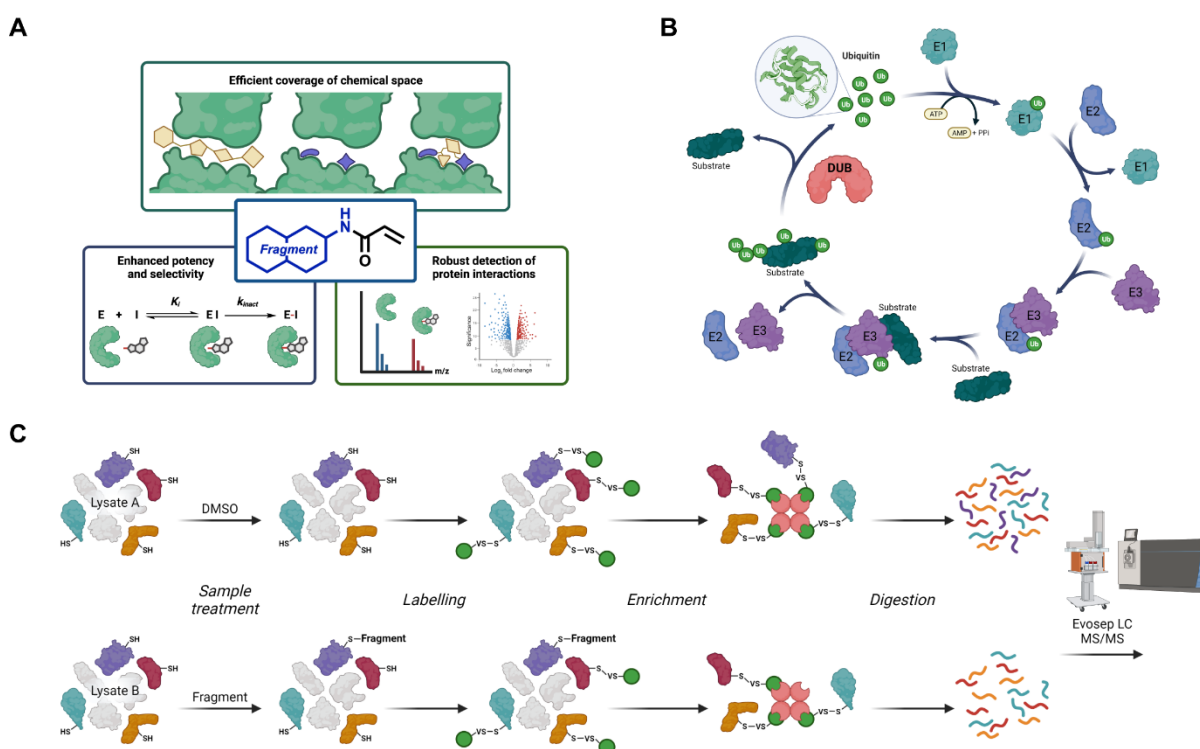


Figure 1. Fragment-based drug discovery of DUBs using chemoproteomics

A) Key advantages of screening with reactive fragments.

B) The ubiquitin cycle, with ubiquitin (green), E1-activating enzymes (turquoise), E2-conjugating enzymes (blue), E3 ligases (purple), substrates (dark green) and deubiquitinase enzymes (pink).

C) Schematic of the activity-based protein profiling workflow. Lysates are treated with covalent fragments or DMSO, followed by treatment with an activity-based DUB probe. Labelled proteins are enriched, digested and analysed by LC-MS/MS.

Herein we describe a high throughput chemoproteomics platform to screen cysteine reactive fragment libraries against deubiquitinases (DUBs). Ubiquitination is a critical and complex post-translational modification which controls and regulates most cellular processes. The ubiquitin cycle is carefully maintained by ubiquitin system proteins: substrate ubiquitination is a result of an enzymatic cascade involving E1-activating enzymes, E2-conjugating enzymes, and E3 ligases; substrate deubiquitination is controlled by DUBs (Figure 1B).²³ In humans, DUBs can be classified into two categories: the metalloproteases (JAMM family) and cysteine proteases (USP, OUT, MJD, UCH, MINDY and ZUFSP families). Both categories of DUBs have been implicated in many diseases including central nervous system (CNS) disorders, inflammation, immunity and infectious diseases,⁴⁰ and have been ardently studied to decipher their mechanisms and substrate specificities. Despite this, new chemical probes are still needed to explore the biology of DUB proteins and aid development of therapeutics.⁴¹

In this study we describe the development of a workflow that was optimised in 96-well plate format and employs DIA-MS analysis to enable fast analysis times while profiling >50% of the native DUBome (Figure 1C). This platform was used to screen a library of 138 fragments, identifying functionally relevant hits for numerous members of the DUB family and demonstrating the utility of the approach in expanding the liganded proteome.

Results

Activity-based probe selection and chemoproteomic platform development

To establish a platform for screening cysteine-reactive fragments against the endogenous DUBome, a DUB specific ABP was first selected. DUB-targeting ABPs mimic native ubiquitin but are modified at the ubiquitin C-terminus with an electrophilic warhead. When bound to cysteine protease DUBs, this warhead irreversibly labels the active site cysteines.^{42, 43} An affinity handle, such as biotin, at the N-terminus of ubiquitin enables downstream enrichment of labelled proteins. We selected three DUB ABPs with different electrophiles to identify which would give the best coverage and enrichment of the DUBome: biotinylated ubiquitin-vinylsulfone (Biotin-Ahx-Ub-VS), biotinylated ubiquitin-propargylamide (Biotin-Ahx-Ub-PA) and biotinylated ubiquitin-vinyl methyl ester (Biotin-Ahx-Ub-VME) (Figure 2A).

We first tested the DUB ABPs in a proteomics workflow that was designed to be fully applicable for subsequent downstream fragment library screening. To reduce running costs and negate the need for multiplexing, simplifying the workflow, we chose to use DIA MS analysis. Furthermore, we aimed to reduce the MS instrument runtime as much as possible. HEK293T lysates were treated with each of the three probes at two concentrations (0.5 and 2 μ M) and DMSO for one hour, before labelled proteins were enriched using avidin and digested with Trypsin. Peptides were analysed by LC-MS/MS using an Evosep One – Orbitrap Fusion Lumos with a 44 minute LC gradient and a DIA method.³⁹

Each probe afforded similar protein intensities, coefficients of variation (CV) and high sequence coverage of the DUBs (Figure 2B and SI Figure 1B and E). All three probes were found to enrich 51-57 DUBs compared to the DMSO control (SI Figure 1A), representing 55-61% of the catalytic cysteine-containing DUB family. Of these, 48 were enriched across the three probes, and Biotin-Ahx-Ub-VS showed the best coverage (Figure 2C) and was selected for future experiments. Furthermore, Biotin-Ahx-Ub-VS showed coverage across each DUB subfamily (Figure 2D), with the exception of the metalloprotease JAMM DUB family which are out of scope for these probes. In addition to selective enrichment compared to the DMSO control, DUB proteins were ranked highly in absolute intensity, highlighting the robust levels of enrichment over the proteome background (Figure 2E). Biotin-Ahx-Ub-VS exhibited high selectivity for the identified cysteine protease DUBs (median avg \log_2 ratio = 4.86) over the 2502 non-DUB enriched proteins (median avg \log_2 ratio = -0.04) (Figure 2F). A lower concentration of Biotin-Ub-VS probe at 0.2 μ M was explored, however, while the CVs remained similar, this led to a decrease in DUB protein intensities (SI Figure 1C and D).

We subsequently sought to increase the throughput of the analysis by decreasing the LC runtimes. Three LC gradients were compared, via 44-, 21- and 11-minute methods. Minimal reduction in the intensities and CVs of DUBs was observed upon reduction of run-time, suggesting all methods would be suitable for library screening (SI Figure 2A, B).

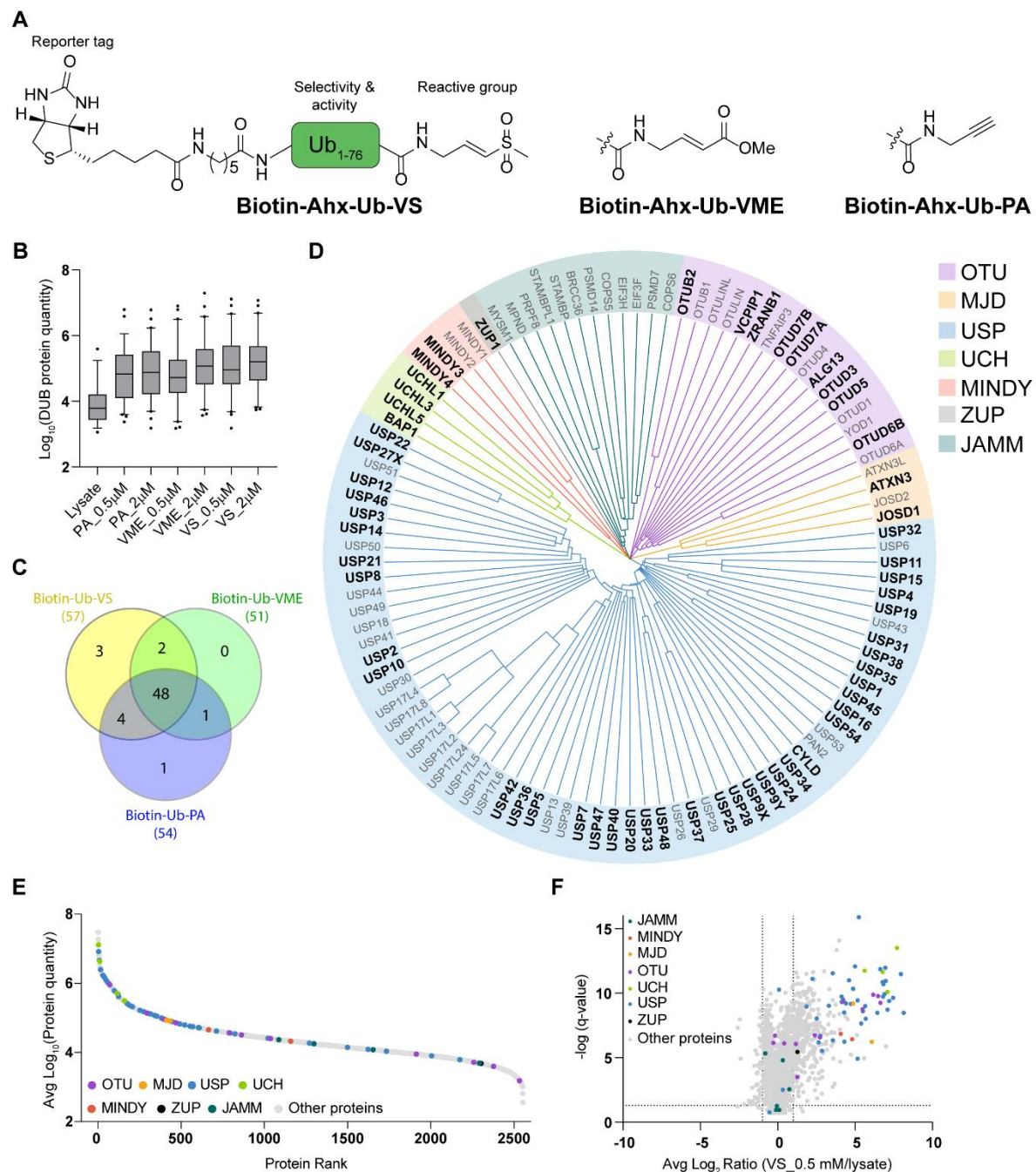


Figure 2. Method development and DUBome coverage using automated Evosep One LC and Orbitrap Fusion Lumos MS alongside DIA for MS analysis

A) Structures of activity-based DUB probes Biotin-Ahx-Ub-VS, Biotin-Ahx-Ub-VME and Biotin-Ahx-Ub-PA.

B) DUB protein quantities identified by each activity-based DUB probe (VS = Biotin-Ahx-Ub-VS; VME = Biotin-Ahx-Ub-VME; PA = Biotin-Ahx-Ub-PA). Box plot shows the median central line and extends from 25th to 75th percentiles. Whiskers represent protein quantity within 5th to 95th percentiles.

C) Venn diagram showing overlap of significantly enriched DUBs for Biotin-Ahx-Ub-VS, Biotin-Ahx-Ub-VME and Biotin-Ahx-Ub-PA probes at 0.5 μ M when compared to untreated lysate.

D) DUB phylogenetic tree highlighting in bold the significantly enriched DUBs when Biotin-Ahx-Ub-VS (0.5 μ M) treatment is compared to untreated lysate (avg \log_2 ratio ≤ -1 , q-value ≤ 0.05). Full length DUB sequences were aligned with COBALT⁴⁴ and subsequently visualised with iTOL.⁴⁵

E) Protein rank plot for Biotin-Ahx-Ub-VS treated sample (0.5 μ M) listing an average protein quantity for all quantified proteins in descending values. DUBs are coloured based on their subfamilies.

F) Volcano plot showing significantly enriched proteins (avg \log_2 ratio ≤ -1 , q-value ≤ 0.05) when Biotin-Ahx-Ub-VS (0.5 μ M) treated sample is compared to untreated lysate. DUBs are coloured based on subfamilies.

Validation of chemoproteomics workflow with known inhibitors

To validate our protocol in a displacement assay, we selected three reported DUB ligands for initial profiling; the parent of Biotin-Ahx-Ub-VS, Ub-VS,⁴⁶ the pan-DUB inhibitor PR619,⁴⁷ and a Ubiquitin Specific Protease (USP) inhibitor (herein referred to as 'USP probe') that targets 12 USPs.⁴⁸ HEK293T lysates were incubated with inhibitor (2.5 μ M Ub-VS, 50 μ M PR619, or 50 μ M USP probe) or DMSO vehicle for three hours prior to treatment with Biotin-Ub-VS probe (0.5 μ M) and subsequent enrichment and MS-analysis as described above. This was performed with six replicates to thoroughly determine reproducibility.

Ub-VS outcompeted 42 out of 53 DUBs enriched by the Biotin-Ahx-Ub-VS ABP (q-value ≤ 0.05 and avg \log_2 ratio ≤ -1 , Figure 3A), indicative of mechanism-based capture. PR619 and USP probe outcompeted 27 and 35 DUBs respectively (q-value ≤ 0.05 and average \log_2 ratio ≤ -1 , Figures 3B and 3C). Of these, 21 DUBs were competed by both compounds. The reproducibility of the platform was found to be excellent, with 49 out of 51 DUBs quantified in all six replicates (DMSO control) and median CVs for the DUB proteins were between 24-29% (SI Figures 3A and B). As a result of the high reproducibility, the number of biological replicates was reduced to three in subsequent experiments.

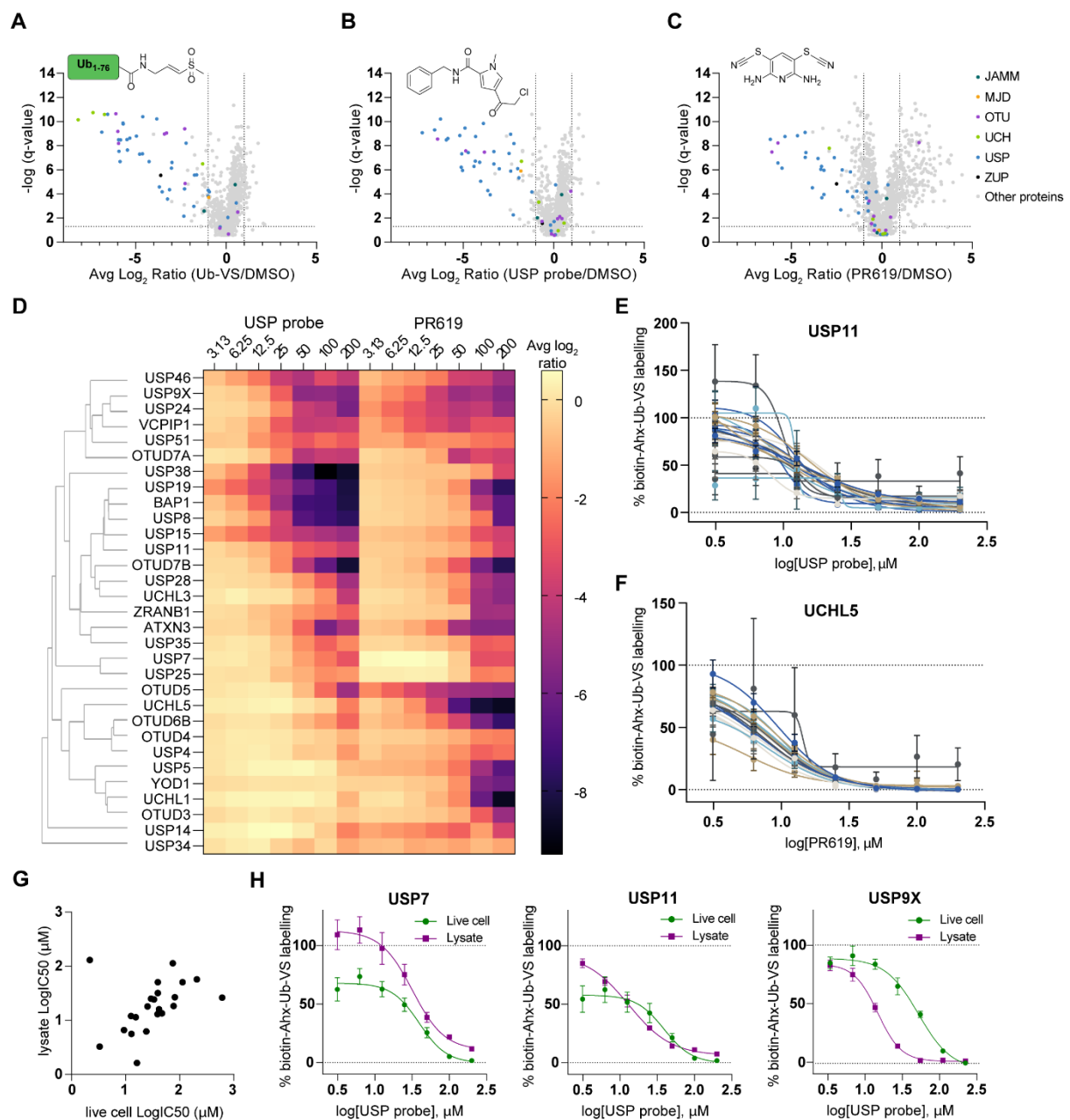


Figure 3. Validation of the chemoproteomics workflow with known DUB inhibitors.

A) Volcano plot showing significantly competed proteins (avg log₂ ratio ≤ -1, q-value ≤ 0.05) when Ub-VS (2.5 μM) treated sample is compared to DMSO (Biotin-Ahx-Ub-VS, 0.5 μM). DUBs are coloured based on subfamilies.

B) Volcano plot showing significantly competed proteins (avg log₂ ratio ≤ -1, q-value ≤ 0.05) when USP probe (50 μM) treated lysate is compared to DMSO (Biotin-Ahx-Ub-VS, 0.5 μM). DUBs are coloured based on subfamilies.

- C) Volcano plot showing significantly competed proteins (avg log₂ ratio ≤ -1, q-value ≤ 0.05) when PR619 (50 μM) treated lysate is compared to DMSO (Biotin-Ahx-Ub-VS, 0.5 μM). DUBs are coloured based on subfamilies.
- D) Heatmap showing dose-response chemoproteomics data for a panel of DUBs (lysate treatment). Avg log₂ ratios of inhibitor/DMSO are plotted for each DUB. DUBs are clustered on the y-axis by single linkage clustering method using cosine correlation distance measure.
- E) Dose-response chemoproteomics data for USP11 peptides. Data is presented as mean ± SEM, *n* = 3. The curves were fitted with GraphPad Prism 9 using four parameter nonlinear regression.
- F) Dose-response chemoproteomics data for UCHL5 peptides. Data is presented as mean ± SEM, *n* = 3. The curves were fitted with GraphPad Prism 9 using four parameter nonlinear regression.
- G) Scatter plot showing logIC₅₀ values for a panel of DUBs comparing USP probe treated lysate and live cell dose-response chemoproteomics data. logIC₅₀ values were extracted from dose-response curves that were fitted with GraphPad Prism 9 using four parameter nonlinear regression.
- H) Selected examples of dose-response chemoproteomics data to compare lysate and live cells treated with USP probe. Data is presented as mean ± SEM, *n* = 3. The curves were fitted with GraphPad Prism 9 using four parameter nonlinear regression. All dose-response curves are presented in SI Figure 4.

To further validate the workflow for screening compound libraries against the DUBome, the two DUB inhibitors, PR619 and USP probe, were profiled in a seven-point dose-response experiment, with concentration responses observed at both the protein and peptide level (Figure 3D, E and F). Clear differences were observed for the DUB interaction profiles of the two compounds. A greater number of IC₅₀ values were determined for the USP probe (32) compared to PR619 (26), consistent with the initial single concentration experiment (SI Table 1). As expected, the pan-DUB inhibitor PR619 exhibited activity across a broad spectrum of DUBs: IC₅₀ values below 10 μM were observed for OTUD5, USP46, VCPIP1, USP9X and UCHL5. The USP inhibitor showed increased selectivity for the USP proteins at lower concentrations, with IC₅₀ values below 10 μM for USP19, USP15, USP38, USP46, BAP1 and USP8.⁴⁸

Comparison of inhibitor treatments in lysates and in live cells

Following workflow optimisations in cell lysates, we profiled the USP probe in a live cell dose response experiment. HEK293T cells were incubated for three hours with USP probe (0 - 200 μ M), and cells were lysed prior to addition of Biotin-Ahx-Ub-VS probe (0.5 μ M). Labelled proteins were enriched and MS-analysis was performed as described previously. Dose response curves were obtained for 37 DUBs, consistent with the lysate-based protocol. However, there were some notable differences in which DUBs were detected. We speculate that this is a consequence of the respective stability of these DUBs in the two different workflows. Importantly, there was a good correlation between IC₅₀ values observed from the cell and lysate treatments, providing confidence that lysate-based screening is a reasonable proxy for engagement of proteins in cells, while enabling higher throughput (Figure 3G, H and SI Figure 4). Nevertheless, cell-based experiments would always be recommended for validation of any interactions identified in lysate-based screening.

Cysteine reactive fragment library screening against the DUBome

Following the establishment and validation of this new and powerful high throughput chemoproteomics platform for the DUBome, we next wanted to perform fragment screening with our methodology. A cysteine reactive fragment library was screened to identify new starting points for the development of covalent DUB inhibitors (Figure 4A). A library of 138 cysteine reactive fragments was selected, comprised of three electrophilic groups (chloroacetamide, acrylamide and methyl acrylate) linked to a chemically diverse set of fragments to assess a range of reactivities, geometries, and functionality (Figure 4A, SI Figure 5A). The library of fragments had molecular weights of between 180 and 400 g/mol (SI figure 5B). The fragments were incubated in cell lysates at 200 μ M for three hours and enriched DUBs were analysed using a 21 minute LC method. This enabled a throughput of 60 samples per day, with the entire screen requiring just 11 days of instrument time. These 11 days included a number of quality controls (QCs) and blanks to ensure the quality of the data. Excellent consistency in the performance of the protocol and instrument performance was observed across all plates and the multiple days of data acquisition (SI Figure 6).

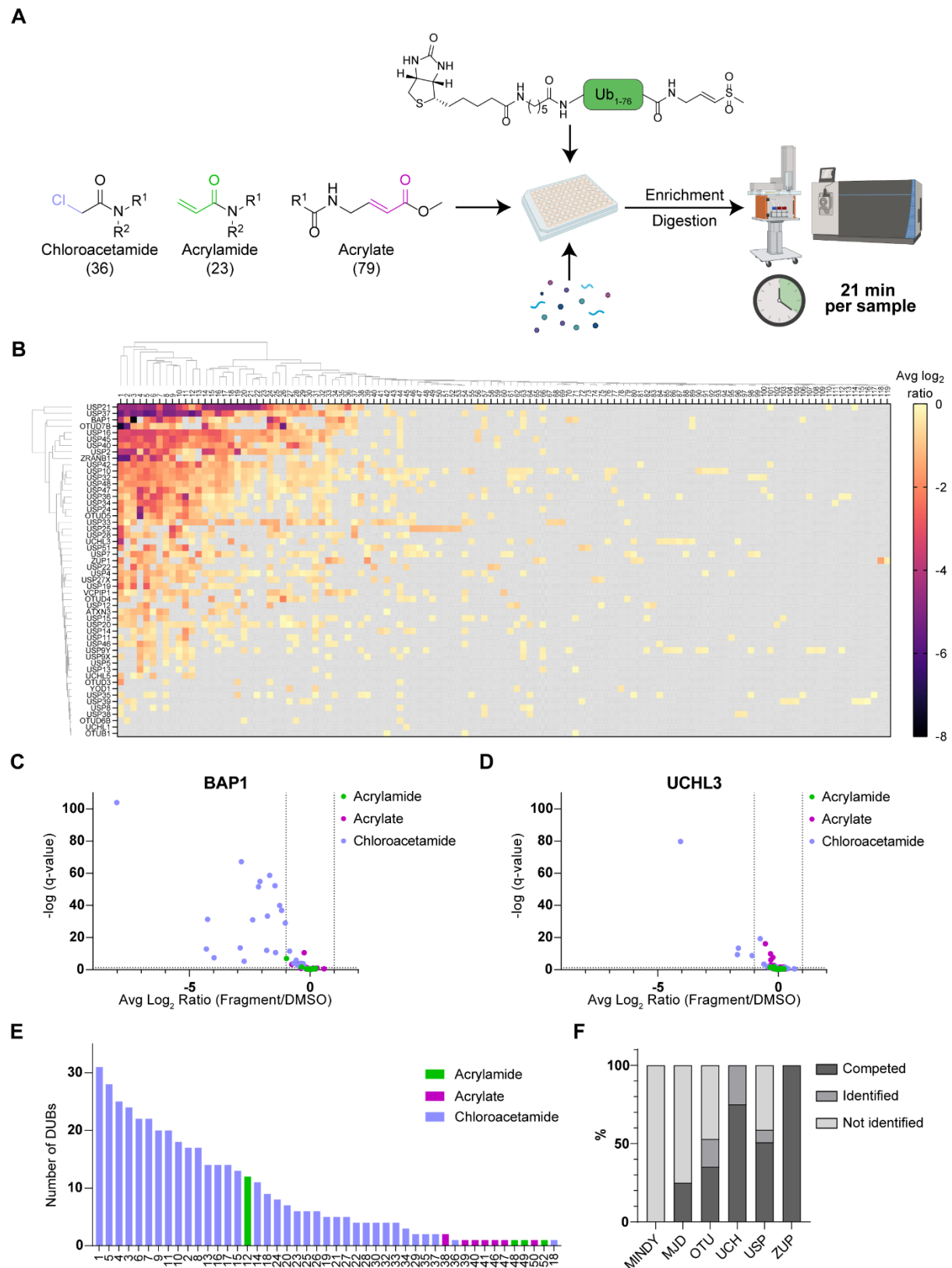


Figure 4. Fragment library screening.

A) Schematic overview of reactive fragment library screening by chemoproteomics. Lysates were treated with the reactive fragments, followed by Ub-VS incubation, enrichment of labelled proteins and digestion for LC-MS/MS analysis.

B) Heatmap showing fragment library screening results. Avg \log_2 ratio of fragment/DMSO ≤ 0 is plotted for each DUB when q-value ≤ 0.05 . Fragments and DUBs are clustered by UPGMA clustering method using Euclidean distance measure.

C) Volcano plot showing all fragments avg \log_2 ratio and q-value for all fragments (200 μM) against BAP1.

D) Volcano plot showing all fragments avg \log_2 ratio and q-value for all fragments (200 μM) against UCHL3.

E) Number of DUBs competed by each fragment when avg \log_2 ratio ≤ -1 and q-value ≤ 0.05 . Bars are coloured based on electrophile.

F) Percentage of DUBs in each family that were competed (avg \log_2 ratio ≤ -1 , q-value ≤ 0.05), identified and not identified.

The resulting full-matrix data set gave competition ratios to describe the interaction between 119 fragments and 52 endogenous DUBs (average \log_2 ratio ≤ 0 , q-value ≤ 0.05 and peptides ≥ 2) (Figure 4B). The chloroacetamides were found to exhibit increased promiscuity (as seen by clustering to the lefthand side of the heatmap) in line with their reactivity, while the acrylates and acrylamides were found to interact with fewer DUBs (as seen by clustering to the righthand side of the heatmap). Fragments interactions were filtered for with average $\log_2 \leq -1$ ($> 50\%$ competition). Some proteins were found to interact with multiple fragments, such as BAP1 and USP21 (Figure 4C and SI Figure 7), whereas other DUBs interacted with few fragments, such as UCHL3 and USP24 (Figure 4D and SI Figure 7). In total, 43 DUBs were competed by 47 fragments (Figure 4E). At least one DUB interaction was found for 35 of the 36 chloroacetamides (97%) screened. This is in comparison to 17% of acrylamides and 10% of acrylates. *p*-Nitrophenyl acrylamide 12 sits as an outlier of the acrylamides as it significantly competes 12 DUB proteins, which is consistent with the enhanced reactivity induced by the electron deficient phenyl ring. There was good coverage across all DUB sub-families, with covalent fragments found to compete USP, MJD and OTU families (59, 25 and 53% respectively, Figure 4F). The highest proportion of covalent hits were observed for the smaller sub-families UCH and ZUP, with covalent fragments found to compete 75% and 100% members of the families, respectively.

Validation of fragment-DUB interactions

Four hit fragments were selected for dose-response validation based on high significance scores, competition ratios and a range of selectivity profiles (Figure 5A, B and SI Figure 8A, B). Fragments 1, 2 and 3 labelled multiple DUB proteins, and clustered together to the lefthand side of the heat map. Fragment 26 was selected for further analysis based on its selective labelling profile, even at 200 μ M. Two of the DUBs significantly competed by fragment 26 (BAP1 and OTUD7B) were also significantly competed by fragments 1-3.

The four fragments were further profiled in the chemoproteomics DUB screening platform at a range of concentrations from 3.1 - 200 μ M (Figure 5C and SI Figure 8). For each of these four fragments, the most significantly competed DUB protein in the single concentration experiment corresponded with the highest potency in the dose response. For fragments 1, 2 and 26 this was OTUD7B, and for fragment 3 it was BAP1. Each of the fragments showed engagement of BAP1, confirming that BAP1 and OTUD7B are both promiscuous proteins. Fragment 3 showed a high selectivity window for BAP1, which was not predicted from the single concentration experiment. Three DUBs significantly competed by fragment 26 (OTUD7B, BAP1 and OTUD4) in the original single concentration experiment were engaged in the dose-response with the same rank order selectivity (Figure 5B, C).

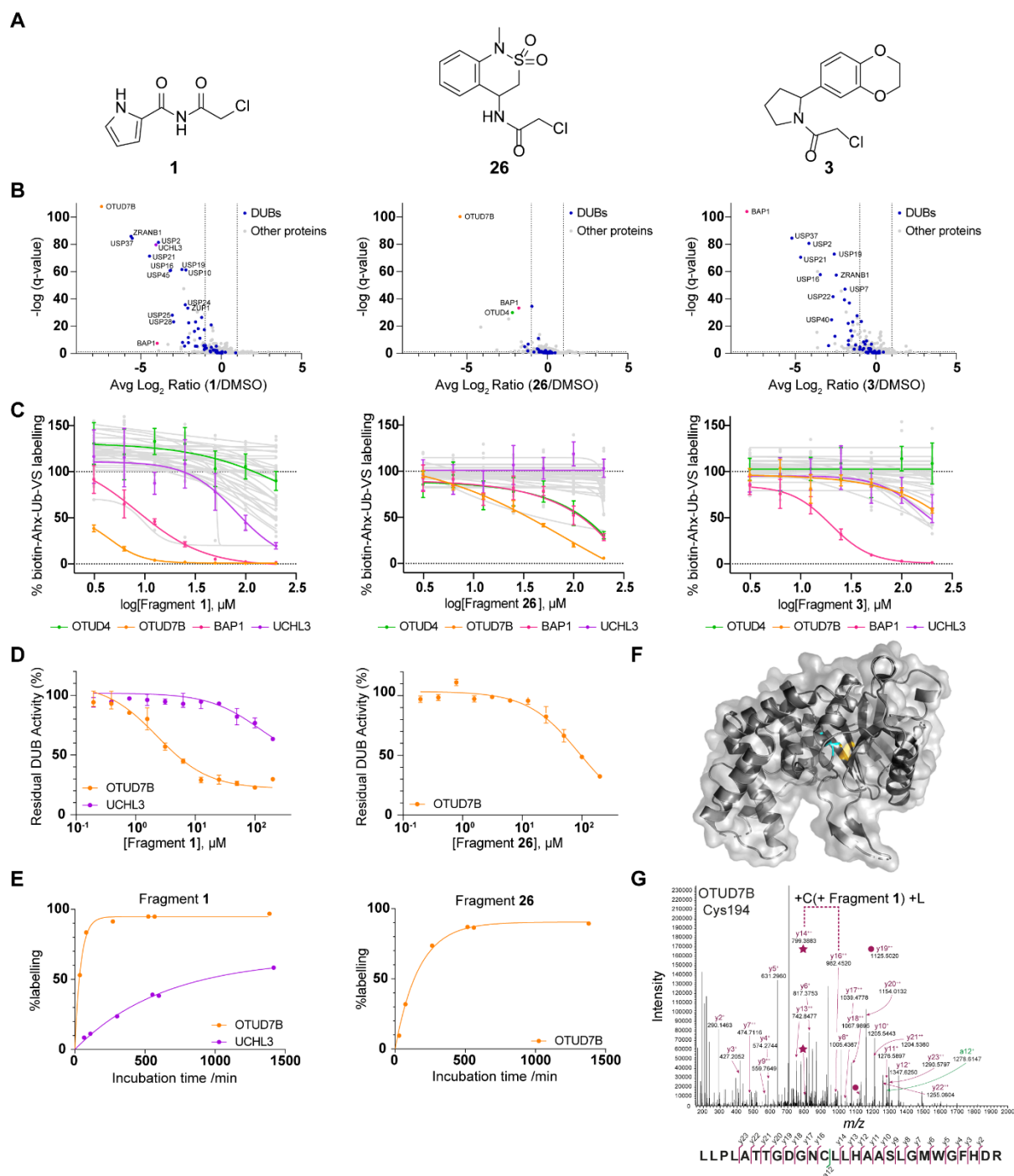


Figure 5. Hit fragment validation.

A) Structures of selected hit fragments **1**, **3** and **26**.

B) Volcano plots showing significantly competed proteins (avg log₂ ratio ≤ -1, q-value ≤ 0.05) when hit fragment (200 μM) treated samples are compared to DMSO (Biotin-Ahx-Ub-VS, 0.5 μM). DUBs are highlighted in blue.

C) Dose-response chemoproteomics data for DUB proteins. Data is presented as mean ± SEM, *n* = 3. The curves were fitted with GraphPad Prism 9 using four parameter nonlinear regression.

D) Enzymatic inhibition assay data for hit fragments **1** and **26**. Data is presented as mean \pm SD, $n = 3$. The curves were fitted with GraphPad Prism 9 using three parameter nonlinear regression.

E) Intact-protein LCMS time-course percentage labelling data for fragments **1** and **26**. The curves were fitted with GraphPad Prism 9 using one phase association.

F) Structure of OTUD7B highlighting the active site Cys194 (PDB: 5LRU).⁴⁹

G) LC-MS/MS spectra of the peptide ₁₈₃LLPLATTGDGNCLLHAASLGMWGFHDR₂₀₉ crosslinked to fragment **1** indicating OTUD7B Cys194 as the fragment binding site.

Orthogonal assays for DUB target validation

To gain further confidence in the fragment-target pairings, the fragments were profiled in orthogonal assays. To assess inhibition of DUB activity by covalent modification, an activity-based enzymatic assay that follows cleavage of ubiquitin-rhodamine by recombinant protein was employed. Recombinant OTUD7B was incubated with fragments **1** and **26** (Figure 5D) and fragment **2** (SI Figure 8D) for three hours prior to the addition of Ub-rhodamine. Fluorescence intensities were plotted as a percent activity compared to DMSO controls. Consistent with chemoproteomics dose experiments, each fragment showed a dose dependent response, with the most potent interaction seen for fragment **1**. The assay was repeated with recombinant UCHL3 and fragment **1** (Figure 5D), which was less potent in agreement with the chemoproteomic experiment. Unfortunately, BAP1 could not be assessed in this assay due to instability of the recombinant protein. To investigate kinetics of protein labelling with fragments **1**, **2** and **26** we carried out a time-course of fragment engagement with recombinant OTUD7B and UCHL3 over 24 hours followed by LCMS analysis (Figure 5E and SI Figure 8E). The time dependent labelling of OTUD7B and UCHL3 correlated well with the enzymatic assays and chemoproteomic experiments. Labelling of UCHL3 by fragment **1** was significantly slower than of OTUD7B, explaining the less potent inhibition in the enzymatic assay.

Given the mechanism-based capture of the DUB activity-based probe it was expected that covalent modification of the fragment was occurring at the active site cysteine. In order to confirm this, site ID MS experiments were performed with UCHL3, OTUD7B and fragments **1**, **2** and **26** (Figure 5F, G and SI Figure 8F, G and H). For all fragments, the predominant site of labelling was identified to be the catalytic cysteine of OTUD7B (Cys 194) across all fragments and UCHL3 (Cys 95) with fragment **1**.

Discussion

The development and discovery of chemical probes for the entire human proteome is recognised as a key objective that will accelerate our understanding of biology and disease. Ligand discovery traditionally relies on biochemical screening with purified protein, which can be time and cost intensive, and often does not effectively reflect *in vivo* ligandability. Covalent fragment-based ligand discovery *in cellulo* provides a powerful alternative due to the small size of libraries required and convenient readout of target engagement by LC-MS. However, to efficiently employ covalent fragment screening of entire libraries in a cellular environment requires the development of high throughput chemoproteomic workflows that are reliable and cost effective. Here we have reported the development of a powerful and efficient workflow and have used it to screen a library of 138 cysteine-reactive fragments against DUBs in an endogenous setting, enabling the identification of numerous fragment-protein interactions.

Efficient chemoproteomic profiling of reactive fragment against the DUBs was achieved through optimisation of multiple steps. Typically, chemoproteomic workflows employ a purification or protein precipitation step, which complicates the workflow and limits parallel sample preparation. Through the use of a high affinity probe that was pre-conjugated to biotin, we removed the requirement for protein precipitation and established the protocol in 96-well plate format. The throughput of mass spectrometry analysis was optimised through the use of an Evosep LC system combined with label-free DIA methods to achieve instrument run times of 21 minutes per sample and a throughput of 60 samples per day, which we envisage could be further reduced to 11 minutes per sample. Advantages of this label free approach include facilitating comparison across hundreds of samples and reduction of costs associated with isotope labelling reagents. The workflow described here enabled reliable quantification of 57 DUBs with high reproducibility and low coefficients of variation.

Validation of the workflow as a screening platform was performed using three established DUB inhibitors, Ub-VS,⁴⁶ PR619,⁴⁷ and a Ubiquitin Specific Protease (USP) inhibitor, USP probe. Interactions with multiple endogenous DUBs were reliably detected and concentration response experiments enabled differentiation of the potencies of each inhibitor-DUB interaction. Furthermore, screening of fragments in live cells afforded IC₅₀ values that correlated well with those measured in lysates, supporting our approach of lysate screening as a suitable proxy for cells to improve the throughput of compound treatment.

Finally, we employed the platform to profile a reactive fragment library to identify ligands for proteins of the DUB family. A library of 138 reactive fragments was profiled against 52 endogenous DUBs, using just 11 days of instrument time. Hits were identified for 43 of the

DUBs (>50% competition at 200 μ M), and these may serve as starting points for the optimisation of more potent and selective inhibitors. Four hits were selected for profiling at a range of concentrations, with many interactions found to drop off quickly at lower concentration, but revealing relatively selective interactions for OTUB7B ($IC_{50} < 3 \mu$ M) and BAP1 ($IC_{50} \sim 10 \mu$ M). The potency and selectivity of these interactions correlated with covalent modification of purified proteins, as determined by intact mass spectrometry. Finally, inhibition of DUB activity was measured for OTUD7B and UCHL3, which was found to match the profile measured by chemoproteomics.

This chemoproteomics platform demonstrates a cost and time effective method for identifying fragments which can be further developed into potent and specific chemical probes against the DUBome. We envisage this platform could be used alongside AI/ML-based approaches for rapid compound optimisation. Furthermore, this adaptable workflow provides opportunities for screening additional protein classes and could be coupled with alternative reactive fragment warheads to target alternative amino acids.

Acknowledgments

This work was supported by the Francis Crick Institute which receives its core funding from Cancer Research United Kingdom (CC 2075), the United Kingdom Medical Research Council (CC 2075), and the Wellcome Trust (CC 2075), by the Biotechnology and Biological Research Council, BB/T014547/1 to KR and DH, and by the Engineering and Physical Sciences Research Council, EP/V038028/1, to DH, JB and KR. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. Figures were created in BioRender.com under the institutional license belonging to the Francis Crick Institute.

Author Contributions

K.R., D.H., A.P.S., and J.T.B. conceived the study. R.C., A.V., J.P., C.R.K., J.M.K., and R.E.P.-H. designed and performed experiments. M.S. and A.P. supported proteomics and experimental work. R.C., A.V., C.R.K., J.P., K.R., and J.T.B. wrote the manuscript, and all authors approved the final version.

References

- (1) Frye, S. V. The art of the chemical probe. *Nature Chemical Biology* **2010**, 6 (3), 159-161.
- (2) Bunnage, M. E.; Chekler, E. L. P.; Jones, L. H. Target validation using chemical probes. *Nature Chemical Biology* **2013**, 9 (4), 195-199.
- (3) Carter, A. J.; Kraemer, O.; Zwick, M.; Mueller-Fahrnow, A.; Arrowsmith, C. H.; Edwards, A. M. Target 2035: probing the human proteome. *Drug Discovery Today* **2019**, 24 (11), 2111-2115.
- (4) Antolin, A. A.; Tym, J. E.; Komianou, A.; Collins, I.; Workman, P.; Al-Lazikani, B. Objective, Quantitative, Data-Driven Assessment of Chemical Probes. *Cell Chemical Biology* **2018**, 25 (2), 194-205.e195.
- (5) Finan, C.; Gaulton, A.; Kruger Felix, A.; Lumbers, R. T.; Shah, T.; Engmann, J.; Galver, L.; Kelley, R.; Karlsson, A.; Santos, R.; et al. The druggable genome and support for target identification and validation in drug development. *Science Translational Medicine* **2017**, 9 (383), eaag1166.
- (6) Plenge, R. M.; Scolnick, E. M.; Altshuler, D. Validating therapeutic targets through human genetics. *Nature Reviews Drug Discovery* **2013**, 12 (8), 581-594.
- (7) Heilbron, K.; Mozaffari, S. V.; Vacic, V.; Yue, P.; Wang, W.; Shi, J.; Jubbs, A. M.; Pitts, S. J.; Wang, X. Advancing drug discovery using the power of the human genome. *The Journal of Pathology* **2021**, 254 (4), 418-429, <https://doi.org/10.1002/path.5664>.
- (8) Lander, E. S.; Linton, L. M.; Birren, B.; Nusbaum, C.; Zody, M. C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; et al. Initial sequencing and analysis of the human genome. *Nature* **2001**, 409 (6822), 860-921.
- (9) Tachmazidou, I.; Hatzikotoulas, K.; Southam, L.; Esparza-Gordillo, J.; Haberland, V.; Zheng, J.; Johnson, T.; Koprulu, M.; Zengini, E.; Steinberg, J.; et al. Identification of new therapeutic targets for osteoarthritis through genome-wide analyses of UK Biobank data. *Nature Genetics* **2019**, 51 (2), 230-236.
- (10) Ochoa, D.; Hercules, A.; Carmona, M.; Suveges, D.; Gonzalez-Urriarte, A.; Malangone, C.; Miranda, A.; Fumis, L.; Carvalho-Silva, D.; Spitzer, M.; et al. Open Targets Platform: supporting systematic drug–target identification and prioritisation. *Nucleic Acids Research* **2021**, 49 (D1), D1302-D1310.
- (11) Hyde, C. L.; Nagle, M. W.; Tian, C.; Chen, X.; Paciga, S. A.; Wendland, J. R.; Tung, J. Y.; Hinds, D. A.; Perlis, R. H.; Winslow, A. R. Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nature Genetics* **2016**, 48 (9), 1031-1036.
- (12) Baker, L. M.; Aimon, A.; Murray, J. B.; Surgenor, A. E.; Matassova, N.; Roughley, S. D.; Collins, P. M.; Krojer, T.; von Delft, F.; Hubbard, R. E. Rapid optimisation of fragments and hits to lead compounds from screening of crude reaction mixtures. *Communications Chemistry* **2020**, 3 (1), 122.
- (13) Machutta, C. A.; Kollmann, C. S.; Lind, K. E.; Bai, X.; Chan, P. F.; Huang, J.; Ballell, L.; Belyanskaya, S.; Besra, G. S.; Barros-Aguirre, D.; et al. Prioritizing multiple therapeutic targets in parallel using automated DNA-encoded library screening. *Nature Communications* **2017**, 8 (1), 16081.
- (14) Resnick, E.; Bradley, A.; Gan, J.; Douangamath, A.; Krojer, T.; Sethi, R.; Geurink, P. P.; Aimon, A.; Amitai, G.; Bellini, D.; et al. Rapid Covalent-Probe Discovery by Electrophile-Fragment Screening. *Journal of the American Chemical Society* **2019**, 141 (22), 8951-8968.
- (15) O'Connell, T. N.; Ramsay, J.; Rieth, S. F.; Shapiro, M. J.; Stroh, J. G. Solution-based indirect affinity selection mass spectrometry--a general tool for high-throughput screening of pharmaceutical compound libraries. *Analytical chemistry* **2014**, 86 (15), 7413-7420.
- (16) Grant, E. K.; Fallon, D. J.; Hann, M. M.; Fantom, K. G. M.; Quinn, C.; Zappacosta, F.; Annan, R. S.; Chung, C.-w.; Bamborough, P.; Dixon, D. P.; et al. A Photoaffinity-Based Fragment-Screening Platform for Efficient Identification of Protein Ligands. *Angewandte Chemie International Edition* **2020**, 59 (47), 21096-21105, <https://doi.org/10.1002/anie.202008361>.
- (17) Backus, K. M.; Correia, B. E.; Lum, K. M.; Forli, S.; Horning, B. D.; González-Páez, G. E.; Chatterjee, S.; Lanning, B. R.; Teijaro, J. R.; Olson, A. J.; et al. Proteome-wide covalent ligand discovery in native biological systems. *Nature* **2016**, 534, 570.
- (18) Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B. D.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F. Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **2010**, 468 (7325), 790-795.
- (19) Zanon, P. R. A.; Lewald, L.; Hacker, S. M. Isotopically Labeled Desthiobiotin Azide (isoDTB) Tags Enable Global Profiling of the Bacterial Cysteinome. *Angewandte Chemie International Edition* **2020**, 59 (7), 2829-2836.
- (20) Parker, C. G.; Galmozzi, A.; Wang, Y.; Correia, B. E.; Sasaki, K.; Joslyn, C. M.; Kim, A. S.; Cavallaro, C. L.; Lawrence, R. M.; Johnson, S. R.; et al. Ligand and Target Discovery by Fragment-Based Screening in Human Cells. *Cell* **2017**, 168 (3), 527.
- (21) Varca, A. C.; Casalena, D.; Chan, W. C.; Hu, B.; Magin, R. S.; Roberts, R. M.; Liu, X.; Zhu, H.; Seo, H.-S.; Dhe-Paganon, S.; et al. Identification and validation of selective deubiquitinase inhibitors. *Cell Chemical Biology* **2021**.
- (22) Johansson, H.; Isabella Tsai, Y.-C.; Fantom, K.; Chung, C.-W.; Kümper, S.; Martino, L.; Thomas, D. A.; Eberl, H. C.; Muelbauer, M.; House, D.; et al. Fragment-Based Covalent Ligand Screening Enables Rapid Discovery of Inhibitors for the RBR E3 Ubiquitin Ligase HOIP. *Journal of the American Chemical Society* **2019**, 141 (6), 2703-2712.
- (23) Kennedy, C.; McPhie, K.; Rittinger, K. Targeting the ubiquitin system by fragment-based drug discovery. *Frontiers in Molecular Biosciences* **2022**, 9, Mini Review.
- (24) Gilbert, K. E.; Vuorinen, A.; Aatkar, A.; Pogány, P.; Pettinger, J.; Grant, E. K.; Kirkpatrick, J. M.; Rittinger, K.; House, D.; Burley, G. A.; et al. Profiling Sulfur(VI) Fluorides as Reactive Functionalities for Chemical Biology Tools and Expansion of the Ligandable Proteome. *ACS Chemical Biology* **2023**.

- (25) Thomas, R. P.; Heap, R. E.; Zappacosta, F.; Grant, E. K.; Pogány, P.; Besley, S.; Fallon, D. J.; Hann, M. M.; House, D.; Tomkinson, N. C. O.; et al. A direct-to-biology high-throughput chemistry approach to reactive fragment screening. *Chemical Science* **2021**, *12* (36), 12098-12106, 10.1039/D1SC03551G.
- (26) Aatkar, A.; Vuorinen, A.; Longfield, O.; Gilbert, K. E.; Peltier-Heap, R.; Wagner, C.; Zappacosta, F.; Rittinger, K.; Chung, C.-w.; House, D.; et al. Efficient Ligand Discovery Using Sulfur(VI) Fluoride Reactive Fragments. *ChemRxiv* **2022**.
- (27) Shin, Y.; Jeong, J. W.; Wurz, R. P.; Achanta, P.; Arvedson, T.; Bartberger, M. D.; Campuzano, I. D. G.; Fucini, R.; Hansen, S. K.; Ingersoll, J.; et al. Discovery of N-(1-Acryloylazetidin-3-yl)-2-(1H-indol-1-yl)acetamides as Covalent Inhibitors of KRASG12C. *ACS Medicinal Chemistry Letters* **2019**, *10* (9), 1302-1308.
- (28) Dubiella, C.; Pinch, B. J.; Koikawa, K.; Zaidman, D.; Poon, E.; Manz, T. D.; Nabet, B.; He, S.; Resnick, E.; Rogel, A.; et al. Sulfopin is a covalent inhibitor of Pin1 that blocks Myc-driven tumors in vivo. *Nature Chemical Biology* **2021**, *17* (9), 954-963.
- (29) Douangamath, A.; Fearon, D.; Gehrtz, P.; Krojer, T.; Lukacik, P.; Owen, C. D.; Resnick, E.; Strain-Damerell, C.; Aimon, A.; Ábrányi-Balogh, P.; et al. Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. *Nature Communications* **2020**, *11* (1), 5047.
- (30) Mullard, A. Fragment-based screening sees the light. *Nature Reviews Drug Discovery* **2020**, *19*, 742-743.
- (31) Jones, H. B. L.; Heilig, R.; Fischer, R.; Kessler, B. M.; Pinto-Fernández, A. ABPP-HT - High-Throughput Activity-Based Profiling of Deubiquitylating Enzyme Inhibitors in a Cellular Context. *Frontiers in Chemistry* **2021**, *9* (44), Original Research.
- (32) Kuljanin, M.; Mitchell, D. C.; Schweppe, D. K.; Gikandi, A. S.; Nusinow, D. P.; Bulloch, N. J.; Vinogradova, E. V.; Wilson, D. L.; Kool, E. T.; Mancias, J. D.; et al. Reimagining high-throughput profiling of reactive cysteines for cell-based screening of large electrophile libraries. *Nature Biotechnology* **2021**.
- (33) Maurais, A. J.; Weerapana, E. Reactive-cysteine profiling for drug discovery. *Current Opinion in Chemical Biology* **2019**, *50*, 29-36.
- (34) Pappireddi, N.; Martin, L.; Wühr, M. A Review on Quantitative Multiplexed Proteomics. *ChemBioChem* **2019**, *20* (10), 1210-1224, <https://doi.org/10.1002/cbic.201800650>.
- (35) Wühr, M.; Haas, W.; McAlister, G. C.; Peshkin, L.; Rad, R.; Kirschner, M. W.; Gygi, S. P. Accurate Multiplexed Proteomics at the MS2 Level Using the Complement Reporter Ion Cluster. *Analytical Chemistry* **2012**, *84* (21), 9214-9221.
- (36) Yang, F.; Jia, G.; Guo, J.; Liu, Y.; Wang, C. Quantitative Chemoproteomic Profiling with Data-Independent Acquisition-Based Mass Spectrometry. *Journal of the American Chemical Society* **2022**, *144* (2), 901-911.
- (37) Jones, H. B. L.; Heilig, R.; Davis, S.; Fischer, R.; Kessler, B. M.; Pinto-Fernández, A. ABPP-HT*—Deep Meets Fast for Activity-Based Profiling of Deubiquitylating Enzymes Using Advanced DIA Mass Spectrometry Methods. *International Journal of Molecular Sciences* **2022**, *23* (6), 3263.
- (38) Doerr, A. DIA mass spectrometry. *Nature Methods* **2015**, *12* (1), 35-35.
- (39) Bache, N.; Geyer, P. E.; Bekker-Jensen, D. B.; Hoerning, O.; Falkenby, L.; Treit, P. V.; Doll, S.; Paron, I.; Müller, J. B.; Meier, F.; et al. A Novel LC System Embeds Analytes in Pre-formed Gradients for Rapid, Ultra-robust Proteomics. *Molecular & Cellular Proteomics* **2018**, *17* (11), 2284-2296.
- (40) Harrigan, J. A.; Jacq, X.; Martin, N. M.; Jackson, S. P. Deubiquitylating enzymes and drug discovery: emerging opportunities. *Nature Reviews Drug Discovery* **2018**, *17* (1), 57-78.
- (41) Lange, S. M.; Armstrong, L. A.; Kulathu, Y. Deubiquitinases: From mechanisms to their inhibition by small molecules. *Molecular Cell* **2021**.
- (42) Hewings, D. S.; Flygare, J. A.; Bogoy, M.; Wertz, I. E. Activity-based probes for the ubiquitin conjugation–deconjugation machinery: new chemistries, new tools, and new insights. *The FEBS Journal* **2017**, *284* (10), 1555-1576.
- (43) Borodovsky, A.; Ovaa, H.; Kolli, N.; Gan-Erdene, T.; Wilkinson, K. D.; Ploegh, H. L.; Kessler, B. M. Chemistry-Based Functional Proteomics Reveals Novel Members of the Deubiquitinating Enzyme Family. *Chemistry & Biology* **2002**, *9* (10), 1149-1159.
- (44) Papadopoulos, J. S.; Agarwala, R. COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics* **2007**, *23* (9), 1073-1079.
- (45) Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research* **2021**, *49* (W1), W293-W296.
- (46) Borodovsky, A.; Kessler, B. M.; Casagrande, R.; Overkleeft, H. S.; Wilkinson, K. D.; Ploegh, H. L. A novel active site-directed probe specific for deubiquitylating enzymes reveals proteasome association of USP14. *The EMBO Journal* **2001**, *20* (18), 5187-5196.
- (47) Altun, M.; Kramer, Holger B.; Willems, Lianne I.; McDermott, Jeffrey L.; Leach, Craig A.; Goldenberg, Seth J.; Kumar, K. G. S.; Konietzny, R.; Fischer, R.; Kogan, E.; et al. Activity-Based Chemical Proteomics Accelerates Inhibitor Development for Deubiquitylating Enzymes. *Chemistry & Biology* **2011**, *18* (11), 1401-1412.
- (48) Ward, J. A.; McLellan, L.; Stockley, M.; Gibson, K. R.; Whitlock, G. A.; Knights, C.; Harrigan, J. A.; Jacq, X.; Tate, E. W. Quantitative Chemical Proteomic Profiling of Ubiquitin Specific Proteases in Intact Cancer Cells. *ACS Chemical Biology* **2016**, *11* (12), 3268-3272.
- (49) Mevissen, T. E. T.; Kulathu, Y.; Mulder, M. P. C.; Geurink, P. P.; Maslen, S. L.; Gersch, M.; Elliott, P. R.; Burke, J. E.; van Tol, B. D. M.; Akutsu, M.; et al. Molecular basis of Lys11-polyubiquitin specificity in the deubiquitinase Cezanne. *Nature* **2016**, *538* (7625), 402-405.