HHS Public Access

Author manuscript

Nat Chem Biol. Author manuscript; available in PMC 2025 February 01.

Published in final edited form as:

Nat Chem Biol. 2024 February ; 20(2): 170–179. doi:10.1038/s41589-023-01458-4.

DNA-encoded library enabled discovery of proximity-inducing small molecules

Jeremy W. Mason[1,2], **Yuen Ting Chow**[1,3], **Liam Hudson**[1,2], **Antonin Tutter**[4], **Gregory Michaud**[4], **Matthias V. Westphal**[1,2], **Wei Shu**[5], **Xiaolei Ma**[5], **Zher Yin Tan**[1,3], **Connor W. Coley**[1], **Paul A. Clemons**[1], **Simone Bonazzi**[2], **Frédéric Berst**[6], **Karin Briner**[2], **Shuang Liu**[1,3,*], **Frédéric J. Zécri**[2,*], **Stuart L. Schreiber**[1,3,*]

[1]Chemical Biology and Therapeutics Science, Broad Institute, 415 Main Street, Cambridge, MA 02142, USA

[2]Global Discovery Chemistry, Novartis Institutes for BioMedical Research, 181 Massachusetts Avenue, Cambridge, MA 02139, USA

[3]Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA

[4]Chemical Biology and Therapeutics, Novartis Institutes for BioMedical Research, 181 Massachusetts Avenue, Cambridge, MA 02139, USA

[5]Structural and Biophysical Chemistry, Novartis Institutes for BioMedical Research, 5959 Horton Street, Emeryville, CA 94608, USA

[6]Global Discovery Chemistry, Novartis Institutes for BioMedical Research, Novartis Campus, Basel, CH-4002, Switzerland

Abstract

Small molecules that induce protein-protein associations represent powerful tools to modulate cell circuitry. We sought to develop a platform for the direct discovery of compounds able to induce association of any two pre-selected proteins, using the E3 ligase VHL and bromodomains as test

*Corresponding authors: stuart_schreiber@harvard.edu, frederic.zecri@novartis.com, sliu.broad@gmail.com.

Author contributions

J.W.M. conceived the project, designed and synthesized the CIP-DEL library, performed the BRD4 screening, analyzed the data, and produced screening hits off-DNA. Y.T.C. performed the BRD2/BRDT screening, synthesized hits off-DNA, and profiled the hits from these screens. L.H. and M.V.W. contributed to the design, screening, data analysis, and off-DNA hit synthesis. A.T. developed and conducted all NanoBiT, HiBiT, CTG, and TR-FRET experiments from the BRD4 screening and oversaw these experiments for BRD2/BRDT compound profiling. G.M. generated and analyzed all SPR data. W.S. and X.M. ran crystallization screens and solved the ternary complex structure. Z.Y.T. contributed to off-DNA compound synthesis. C.W.C. and P.A.C. developed the CIP-DEL data analysis pipeline. S.L. supervised and supported all BRD2/BRDT screening and profiling. S.B., F.B., K.B., F.J.Z., and S.L.S. provided context for the framing of the original goals, supervision, guidance, operational support, assisted in the interpretation of experimental outcomes, and made recommendations for a subset of the reported experiments. J.W.M. and S.L.S. wrote the manuscript and all authors read and edited the paper.

Competing interests

The authors declare the following competing financial interests: X.M. is a shareholder of Terremoto Biosciences. C.W.C. is an advisor to Anagenex. P.A.C. is an advisor to nference, Inc., Pfizer, Inc., and Belharra Therapeutics. S.L.S. is a shareholder and serves on the Board of Directors of Jnana Therapeutics and Kojin Therapeutics; is a shareholder and advises Kisbee Therapeutics, Belharra Therapeutics, Magnet Biomedicine, Exo Therapeutics, and Eikonizo Therapeutics; advises Vividian Therapeutics, Eisai Co., Ltd., Ono Pharma Foundation, F-Prime Capital Partners, and the Genomics Institute of the Novartis Research Foundation; and is a Novartis Faculty Scholar. The remaining authors declare no competing interests.

systems. Leveraging the screening power of DNA-encoded libraries (DELs), we synthesized ~one million DNA-encoded compounds that possess a VHL-targeting ligand, a variety of connectors, and a diversity element generated by split-and-pool combinatorial chemistry. By screening our DEL against bromodomains in the presence and absence of VHL, we could identify VHL-bound molecules that simultaneously bind bromodomains. For highly barcode-enriched library members, ternary complex formation leading to bromodomain degradation was confirmed in cells. Furthermore, a ternary complex crystal structure was obtained for our most enriched library member with $BRD4^{BD1}$ and a VHL complex. Our work provides a foundation for adapting DEL screening to the discovery of proximity-inducing small molecules.

Introduction

The discovery of small molecules that can modulate specific protein–protein interactions (PPIs) has long been of interest to the biomedical field and, although challenging, provides broad therapeutic opportunities.[1] While the discovery of PPI inhibitors has become an established paradigm in drug discovery, the identification of compounds that induce novel (neo)-protein associations has been gaining momentum. The therapeutic relevance of such compounds, referred to here as chemical inducers of proximity (CIPs), was established with the realization that the drugs cyclosporin (sandimmune),[2] FK506 (tacrolimus),[2] and later rapamycin (sirolimus)[3,4] work by such a mechanism. Continued identification of simpler synthetic CIPs, such as tool compounds synstab A[5] and indisulam,[6–10] and drugs thalidomide and lenalidomide (Revlimid),[11–17] have ignited widespread interest in the development of proximity-inducing molecules.

While the potential applications of CIPs continue to expand,[18–25] the available tools for identifying them remain limited. We set out to develop a platform that could accelerate the discovery of CIPs for any two pre-selected proteins, which we herein label as targets and presenters. For our initial implementation, we targeted the discovery of heterobifunctional degraders known as proteolysis-targeting chimeras (PROTACs).[26] Our hope was to develop an approach where both target and presenter proteins would be screened simultaneously to identify bifunctional compounds that induce optimal ternary complex formation in a single step, rather than requiring independent efforts to optimize separate binding and linking events.

We reasoned that DNA-encoded libraries (DELs)[27] might be adapted for the discovery of proximity-inducing small molecules. A library of DNA-encoded bifunctional compounds was designed that could in principle recruit any desired protein (the target) to the E3 ubiquitin ligase VHL (the presenter) for ubiquitination and subsequent degradation. Bromodomains from three separate proteins were chosen as our initial targets and we used affinity screens with and without the VHL–elongin C–elongin B (VCB) complex to identify compounds that bind the target protein in the presence of VHL. Here we present our first implementation of the “CIP-DEL” approach and report the successful validation of this platform through the identification of novel protein degraders.

Results

Library design and synthesis.

The established BRD4 degrader **MZ1** was used as a template for our initial library design and the available ternary complex structure (PDB: 5T35) was used to select an attachment vector for the requisite DNA barcodes (Fig. 1a and 1b).[28,29] The thiazole-methyl group of the VHL-targeting moiety of **MZ1** resides within a small channel directed towards solvent. For our library, the VHL-targeting portion of **MZ1** was synthesized with an alkyne functionality appended to the thiazole-methyl for copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) coupling to the DNA headpiece. This modified VHL ligand was functionalized with a panel of connectors having varied lengths and compositions. We focused primarily on shorter heterocyclic connectors to bias our screens towards CIPs that induce a close association between VHL and the protein of interest (POI), although traditional alkyl and PEG-based linkers were also included (Fig. 1d; Supplementary Fig. 1). To ensure that our VHL ligand–connector combinations did not disrupt VHL binding, a critical feature for our screens, we tested 22 ligand–connector combinations in a TR-FRET assay for VCB binding after the Fmoc-protected amines were converted to acetamides. We identified a collection of 15 connectors that displaced a HIF1α derived peptide from the VCB complex with EC_{50} values of ~1 μM (range: 0.75 – 1.51 μM). These connectors were selected for library inclusion, coupled to the DNA headpiece, and then encoded by unique DNA barcodes.

Each connector terminated with a primary or secondary amine that was used to construct a split-and-pool library. Numerous DEL library designs could be used for functionalizing the connector amines; here, we selected the well-established triazine-library concept for our initial experiments given the large diversity that can be generated with sequential S_NAr reactions on the triazine core.[30] Each connector amine was directly functionalized with the triazine scaffold, and two additional rounds of S_NAr were performed to complete the triazine substitution (Fig. 1e). To select amines for inclusion in our library, we performed validation experiments with a set of 264 primary amines and 169 secondary amines. Building blocks that gave a minimum of 75% conversion to product based on UPLC-MS analysis were selected for inclusion. This resulted in the inclusion of 290 amines (163 primary and 127 secondary) for the second triazine substitution and 249 amines (146 primary and 103 secondary) for the final substitution. Therefore, with 15 connectors (cycle 1), 290 amines included in the second substitution (cycle 2), and 249 amines included in the third substitution (cycle 3), the final DEL library contains 1,083,150 potential CIPs for targeting VHL. Since our library was designed using **MZ1** as a template, a DNA-tagged derivative of this degrader (**MZ1-DEL**) was synthesized separately from the library, encoded with a unique barcode, and spiked into the library at a theoretical equimolar ratio to each library member.

Affinity Screens and Sequencing Analysis

For our initial CIP-DEL screening, we targeted the first bromodomain (BD1) of BRD4 and screened the library using three parallel conditions: (i) a “beads-only” screen to control for matrix binders; (ii) a “BRD4 only” screen (denoted BRD4 (−) VHL) to identify

binders to $BRD4^{BD1}$; and (iii) a "BRD4 with VHL" screen (denoted BRD4 (+) VHL) that included excess VCB complex (~85 equivalents of VCB complex compared to CIP-DEL; ~43 equivalents of $BRD4^{BD1}$ compared to CIP-DEL) to identify library members that bind BRD4 in the presence of VCB. Our expectation was that BRD4 ligands identified in the BRD4 (−) VHL screen would show reduced enrichments in the screen with VCB if the library compound was unable to induce a stable ternary complex (*e.g.,* due to steric or electrostatic clashes between presenter and target proteins: Fig. 1c). In contrast, BRD4 ligand–connector combinations that induced a stable ternary complex would yield high enrichment levels in both (+/−) VHL screens. In the ideal situation, induction of neo-PPIs would lead to increases in experimental enrichment values in the presence of VCB, suggesting cooperative ternary complex formation.

Following the affinity-based screens, library members retained in each experimental condition were separated from the beads by heating and the output DNA was PCR-amplified and analyzed by next-generation sequencing (NGS). Our NGS data showed that we obtained roughly equal sequencing depth for each of the three screening conditions (Extended Data Fig. 1). However, in the BRD4 (+) VHL screen, we observed a marked increase in the number of high-count barcodes. We next looked at barcode counts for each individual building block across each screening condition (Extended Data Fig. 2). For each connector (denoted BB1), we saw roughly even distribution of barcode counts in the beads-only and BRD4 (−) VHL screens. Remarkably, in the screen with VHL present, we found a dramatic shift in NGS reads towards library members that possessed connector **C15** suggesting that this connector was a key component for successful ternary complex induction. Intriguingly, the second most prevalent connector was **C14**, which is a one methylene truncated analog of **C15** suggestive of a structure–activity relationship (SAR).

Barcode counts were then analyzed for preferred cycle-2 and −3 building blocks (denoted BB2 and BB3, respectively). In the BRD4 (+) VHL screen, the highest counts in cycle 2 came from **BB2–109**. Interestingly, the cycle 3 building block with the highest counts in the BRD4 (+) VHL screen was **BB3–114**, which has the same structure as **BB2–109** (Extended Data Fig. 2). Since a common set of amine-based building blocks was profiled for cycle-2 and −3 triazine substitutions, and amines that provided at least 75% UPLC-MS product yield were included in the library, some amines were included in both cycles 2 and 3. That **BB2–109** / **BB3–114** was the most enriched building block within each substitution cycle strongly suggests that this building block is a favored chemotype.

To refine our screening analysis further, we calculated enrichment values for each library member by comparing the barcode counts from the BRD4 +/− VHL screens to the counts from the beads-only screen. We modeled the NGS counts as a Poisson sampling of the barcodes and estimated an enrichment ratio, and the upper- and lower-bounds of the 95% confidence. For all subsequent analyses, we used the lower bound (lb) of the 95% confidence interval as our enrichment metric since this provided a more conservative estimate of enrichment. Cube plots were generated using the enrichment values to visualize SAR trends (Fig. 2a – c). We looked for library members with three different output profiles: (i) high-confidence enrichment in the BRD4 (−) VHL screen (Fig. 2a); (ii) high-

confidence enrichment in the BRD4 (+) VHL screen (Fig. 2b); and (iii) molecules that showed enrichment in both screens (Fig. 2c).

In the BRD4 (−) VHL screen, we found weaker overall enrichments compared to the BRD4 (+) VHL screen and greater dispersion of the compounds meeting the enrichment cutoff ≥ 7 (cutoff values were selected to give ~200 plotted data points). However, several line features were present and many of the enriched features contained **C15** and **BB2–109** / **BB3–114** (shown in green). When VCB complex was included in the screen, the preference for compounds containing the **C15** connector paired with **BB2–109** or **BB3–114** became strongly apparent (Fig. 2b; green). Strong enrichment values were seen across **BB2–109** and **BB3–114** line features, suggesting that a variety of amines may be tolerated at the final diversity element. Given the strong enrichments for **C15** and **BB2–109** / **BB3–114**, we focused our off-DNA follow-up efforts on compounds containing these features. In addition, we synthesized compounds containing **BB2–109** / **BB3–114** but with connectors that gave weaker enrichments to confirm that our screening results were predictive of off-DNA CIP activity.

As an additional query of our screening data, we plotted library members that met a nominal enrichment value in both BRD4 (+/−) VHL screens (cutoff ≥ 5; Fig. 2c). We observed similar trends to those seen when the BRD4 (+) VHL screen was analyzed independently (Fig. 2b) with some minor variations. Notably, we saw the appearance of an additional vertical line feature for **BB3–78**. However, this line feature displayed much lower enrichment values compared to **BB2–109** / **BB3–114** and as such, was deemed a second tier series.

SPR Validation of Off-DNA Library Members

A total of 21 library members were selected for off-DNA synthesis and evaluation (denoted **CIP-1** – **21**) in addition to the **MZ1** control. BRD4 binding affinities for the compounds were first measured by surface plasmon resonance (SPR), wherein biotinylated $BRD4^{BD1}$ was immobilized to the SPR chip. The control compound **MZ1** gave a binding affinity of 14 nM. In comparison, **CIP-1**, which corresponds to the most highly enriched barcode from the BRD4 (+) VHL screen, had a binary K_D of 45 nM (Fig. 2d; Supplementary Tables 1 and 2). Next, we sought to determine a ternary K_D by immobilizing $BRD4^{BD1}$ on the SPR chip, adding a pre-incubated mixture of compound and VCB to the flow cell, and measuring the binding affinity of the complex. A ternary K_D of 49 nM was obtained for **MZ1** and a ternary K_D of 158 nM was measured for **CIP-1**. From the binary and ternary K_D values, the cooperativity factor α (defined as $\alpha = K_D^{binary} / K_D^{ternary}$)[31] was calculated to be 0.29 for **CIP-1**, indicating negative cooperativity despite remaining a potent CIP.

In the BRD4 (−) VHL screen, the most highly enriched barcode corresponded to **CIP-2**, which was also the fourth most highly enriched barcode in the BRD4 (+) VHL screen (Supplementary Table 1). **CIP-2** contains the highly enriched **C15** and **BB3–114** pairing, with **BB2–289** as the final triazine substituent. **CIP-2** was evaluated by SPR and displayed a binary K_D of 13 nM and a ternary K_D of 34 nM, matching the potencies of **MZ1** with no off-DNA optimization.

Two additional representatives with the **C15** / **BB2–109** pairing that gave strong enrichment values in both screens are **CIP-3** and **CIP-4** (Fig. 2d). SPR analysis for **CIP-3** and **CIP-4** gave binary K_D values of 114 nM and 250 nM, and ternary K_D values of 636 nM and 276 nM, respectively. Interestingly, the weaker $BRD4^{BD1}$ ligand resulted in a stronger ternary K_D, indicating enhanced cooperativity.

For additional CIP-DEL validation, we synthesized **CIP-20** which contains the same cycle 2 and 3 building blocks as **CIP-1**, but with a **C6** connector (as used in **MZ1**) that did not show strong enrichment in our screening (Fig. 2d). Despite containing the same cycle 2 and 3 building blocks as **CIP-1**, **CIP-20** displayed substantially weaker binary and ternary K_D values, as suggested by our screening data, indicating that **C15** is indeed an important contributor to both binary and ternary binding affinities. The full collection of 21 library members selected for off-DNA synthesis and profiling can be found in Supplementary Table 1 and detailed SPR results are provided in Supplementary Table 3.

Confirmation of VHL Binding by TR-FRET

Each VHL ligand–connector pairing included in the CIP-DEL library had shown comparable binding to VCB when acetylated at the terminal amine of the connector (Fig. 1d). To verify that the addition of the triazine component did not disrupt binding, we tested our off-DNA validation compounds in the TR-FRET assay for VCB binding. All compounds were found to retain binding to VCB with an average EC_{50} of 35 nM (range: 6 to 153 nM; Supplementary Table 1). The TR-FRET assay was also performed with the addition of $BRD4^{BD1}$ to look for cooperative complex induction. Most compounds showed slight negative cooperativity in the presence of $BRD4^{BD1}$. Notably, **CIP-12** displayed strong negative cooperativity ($\alpha = 0.02$), whereas **CIP-7**, **−9**, and **−10** displayed slight positive cooperativity for VCB binding ($\alpha = 2.63$, 1.91, and 1.24, respectively) and also showed positive cooperativity for $BRD4^{BD1}$ binding in our SPR analysis ($\alpha = 6.47$, > 7.14, and > 5.65, respectively). Common to **CIP-7**, **−9**, and **−10** is **BB2–185**, which contains a constrained hydroxymethyl group. We speculate that this functionality may be a key contributor to the positive cooperativity.

Ternary Complex Induction and Protein Degradation

To complement our SPR and TR-FRET datasets, we assessed ternary complex induction in HEK293T cells using the high-sensitivity NanoBiT split-luciferase assay.[32] **CIP-1**, **−3**, and **−4** induced robust complex assembly with EC_{50} values of 163 nM, 625 nM, and 43 nM, respectively (Fig. 2e). Importantly, **CIP-20** induced minimal complex assembly, further validating our screening results. To confirm that the NanoBiT signal was dependent on VHL binding, we tested the *cis*-hydroxyproline derivative of **CIP-1** (***cis*-CIP-1**). As expected, ***cis*-CIP-1** produced negligible ternary complex induction. A truncated analog of **CIP-1** (denoted **BL-1**), which comprises only the $BRD4^{BD1}$ ligand with the **C15** connector, was synthesized and tested in the NanoBiT assay to confirm further that complex assembly required the VHL-targeting motif. Negligible activity was observed with **BL-1**, confirming that the complete bifunctional compound is necessary for complex induction. The complete set of NanoBiT EC_{50} and E_{max} values can be found in Supplementary Table 1.

Although our CIP-DEL screens were not coupled to protein degradation, the discovery of CIPs targeting BRD4 to VHL could result in functional degraders. To establish this, we profiled our DNA-free compounds in a HiBiT assay using HEK293T cells transfected with a HiBiT-tagged $BRD4^{BD1}$ plasmid.[33] We found that **CIP-1** and **CIP-4** produced potent $BRD4^{BD1}$ degradation with DC_{50} values of 95 nM and 112 nM and D_{max} values of 91% and 94% after 17 h of treatment (Fig. 2f). Interestingly, **CIP-3** was found to be ~10x less potent than **CIP-4** even though **CIP-3** was a more potent $BRD4^{BD1}$ ligand, reinforcing the idea that binary affinity to the POI is less consequential for degradation than cooperativity and ternary K_D. **CIP-2** provided a DC_{50} of 17 nM and a D_{max} of 93%, representing the most potent degrader we identified. Considering that **CIP-2** came directly from our screen with no ligand or connector optimization, this provides compelling validation of our CIP-DEL approach.

Consistent with our NanoBiT data, ***cis*-CIP-1** and **BL-1** induced no protein degradation in the HiBiT assay, confirming the importance of VHL binding for the biological effect (Fig. 2f). Treating cells with **CIP-20** produced no $BRD4^{BD1}$ degradation, suggesting that enrichment in the CIP-DEL screen was predictive of functional degraders. To confirm that the degradation observed with **CIP-1**, **−3**, and **−4** was cullin-RING ligase (CRL)-dependent, the HiBiT assay was repeated in the presence of the NEDDylation inhibitor (NAEi) **MLN4924**,[34] which fully reversed degradation (Fig. 2g). In fact, addition of the NAEi generated a dose-dependent increase in the HiBiT curves. A similar result was observed for ***cis*-CIP-1** and **BL-1** in the absence of NAEi. These data suggest that $BRD4^{BD1}$ binding promotes ligand-induced stabilization of the protein, a phenomenon recently exploited in the development of a HiBiT-based CETSA assay.[35]

Since an isolated protein domain was used in the HiBiT assay, we performed Western blot analysis of HEK293T cells treated with CIPs for 18 h to assess endogenous BRD4 degradation (Fig. 2h). A dose-dependent degradation was observed for **CIP-1**, **−3**, and **−4**, consistent with the HiBiT results; however, the relative potencies for **CIP-3** and **−4** appear to be different when examining endogenous BRD4 degradation. Co-treatment of cells with active CIPs and a NAEi completely reversed the observed degradation, and no degradation was observed for **CIP-20**, ***cis*-CIP-1,** or **BL-1** at 1 μM, whereas **MZ1** induced complete BRD4 removal at this concentration. The BRD4-binding element **(+)-JQ1**[36] and the VHL-binding element **VH032**[37] showed no degradation as monovalent ligands. Finally, cell viability was measured by CellTiter-Glo (CTG) to verify that the observed protein degradation was not due to compound toxicity. After 16 h of compound treatment, no toxicity was observed for the tested CIPs up to 20 μM (Supplementary Fig. 2).

CIP-DEL Correlations with Off-DNA Validation Data

Correlations between on-DNA screening results and off-DNA validation data were examined to assess the predictive capacity of the CIP-DEL platform (Fig. 3). The most direct comparisons can be drawn between on-DNA enrichment values and off-DNA binding affinities from our SPR data. Comparing BRD4 (−) VHL enrichments versus binary K_D (Fig. 3a), and BRD4 (+) VHL enrichments versus ternary K_D (Fig. 3b), we found that higher enrichment values were generally predictive of more potent binders in each of the screens. To assess the ability of CIP-DEL screening to identify compounds that promote

ternary complex formation and protein degradation in cells, we compared BRD4 (+) VHL enrichment values to our NanoBiT and HiBiT results (Fig. 3c and 3d). BRD4 (+) VHL enrichments correlated well with NanoBiT E_{max} and HiBiT D_{max} values, which represent important degrader attributes (Fig. 3d). HiBiT DC_{50} values showed reasonable correlation with BRD4 (+) VHL enrichments, but poor association was seen between NanoBiT EC_{50} data and enrichment values (Fig. 3c). Similar trends were observed when the NanoBiT and HiBiT data was compared to ternary K_D values from SPR (Fig. 3e and 3f), indicating that barcode enrichment and ternary complex binding by SPR were similarly predictive of performance in the NanoBiT and HiBiT assays. HiBiT data and NanoBiT data showed good inter-assay correlations (Fig. 3g and h).

Lastly, we hoped to determine if our screening system was predictive of CIP cooperativity. We calculated a ratio of the enrichment values from the screens +/− VHL and compared these ratios to the α values obtained by SPR (Fig. 3i). While the largest enrichment ratios from our screen did correspond to positively cooperative compounds, we identified few cooperative compounds and additional data will be needed to validate these trends.

Ternary Complex Crystal Structure

To understand better the striking preferences observed for **C15** and **BB2–109** / **BB3–114**, we solved the X-ray crystal structure of the $BRD4^{BD1}$–**CIP-1**–VHL–Elongin B–Elongin C complex (PDB: 8EWV, Fig. 4; Extended Data Fig. 3). In contrast to the “folds on itself” conformation of PEG-based degraders such as **MZ1**, the **C15** connector adopts a linear, extended conformation minimizing the entropic penalty incurred during complex induction. Assembly of the complex generates a short channel stretching between the two proteins that is occupied by the **C15** connector and surrounded primarily by hydrophobic residues from both VHL and $BRD4^{BD1}$ (Extended Data Fig. 3c – f). The connector structure generates an effective balance of close protein associations without inducing steric clashes and results in 216 $Å^2$ of solvent accessible surface area that is buried by the interface (Extended Data Fig. 4). The complex is further stabilized by ligand-induced PPIs. Most notably, Asp145 of $BRD4^{BD1}$ forms two hydrogen bonds with VHL, one to Tyr112 and the second to His110 (Fig. 4b). All these factors combine to explain the strong preference observed for the **C15** connector.

The VHL-targeting component of **CIP-1** maintains equivalent interactions with the E3 ligase as the parent ligand **VH032** (Fig. 4c).[37] In addition, the *tert*-butyl group of the VHL ligand interacts with hydrophobic residues Phe79 and Met149 of $BRD4^{BD1}$, further stabilizing the complex. The triazine component of **CIP-1** is positioned within the acetyl-lysine binding pocket of $BRD4^{BD1}$ and forms several key interactions with the WPF shelf residues of the protein. Phe83 of the WPF shelf, along with Met107 and Ile146, form a hydrophobic pocket occupied by the cyclopropyl group of the ligand. A hydrogen bond is formed between the cyclopropylamide carbonyl and Asn140 of $BRD4^{BD1}$ (Fig. 4b). Interestingly, **CIP-1** appears to form a weak intramolecular hydrogen bond between the cyclopropylamide N-H and a triazine nitrogen, which stabilizes the bound conformation of the ligand. Combined, these interactions provide a rationale for the strong enrichments observed for **BB2–109** / **BB3–**

114. The pyrazole substituent of the triazine forms a π-stacking interaction with Trp81 of the WPF shelf and hydrophobic interactions with Leu92 (Fig. 4b).

BRD2^{BD1} and BRDT^{BD1} Screening

To determine if the CIP-DEL library could be used to identify CIPs for additional targets, we screened the library against bromodomains from two additional proteins, BRD2^{BD1} and BRDT^{BD1}. From our results with BRD4^{BD1}, we found that high enrichments in the selection with VHL provided a good correlation with functional CIPs produced off-DNA. Therefore, we focused our analyses on these screens and the results of the BRD2^{BD1} (+) VHL and BRDT^{BD1} (+) VHL screens are shown in Fig. 5a and 5b, respectively. As may be expected from screens of highly related family members, we observed many of the same trends that were seen in the BRD4^{BD1} screen. For example, the same strong preferences were observed for the **C15** connector and **BB2–109** / **BB3–114**.

Across both screens, two library members stood out as common hits that showed the highest enrichments and we selected these two compounds for off-DNA synthesis (**CIP-22** and **CIP-23**, Fig. 5c). We again used the NanoBiT technology to confirm ternary complex induction (Fig. 5d) and HiBiT experiments to show protein degradation in cells (Fig. 5e). To confirm that the observed degradation was CRL-dependent, we co-treated cells with **MLN4924** which completely reversed degradation and instead resulted in ligand-induced protein stabilization, as seen with BRD4^{BD1} (Fig. 5f). Finally, we used the CTG assay to show that the observed protein degradation was not an artifact resulting from compound cytotoxicity (Fig. 5g). The full dataset for these experiments is summarized in Supplementary Table 4.

In both screens, **CIP-22** showed greater enrichment than **CIP-23** and this translated to increased potency for ternary complex induction and protein degradation, although direct correlations were modest, as seen with BRD4^{BD1}. To determine if the CIPs identified from our screens could induce endogenous BRD2 degradation, we used Western blotting to visualize protein levels in HEK293T cells (Fig. 5h). **CIP-22** produced robust BRD2 degradation although with lower potency than the HiBiT system reported. The weaker **CIP-23** did not induce endogenous BRD2 degradation up to 10 μM, consistent with the weaker degradation observed by HiBiT and the right-shift in potency seen for endogenous BRD2 degradation with **CIP-22**. Endogenous BRDT protein levels in the HEK293T cell line were too low for detection with Western blotting. Finally, we note that the BRD4 and BRD2/T screens were not run in parallel under identical conditions, so we are not able to reliable predict CIP selectivity from our screening data. Additional studies are underway to assess counter-screening for selectivity.

Discussion

The identification of heterobifunctional CIPs can be a tedious, multistage process entailing optimization of target- and presenter-protein binders, identification of exit vectors, and optimization of a connector. Our CIP-DEL approach streamlines this process by directly screening both target and presenter proteins using a large library of DNA-barcoded heterobifunctional compounds. As a testament to the success of this approach, we identified

a 17 nM degrader of BRD4^{BD1} directly from our library. Notably, this included the discovery of a novel BRD4 ligand and the identification of a preferred connector with no off-DNA optimization. Another benefit of the CIP-DEL approach is that it is not restricted to active-site ligands since the affinity-based screening interrogates all accessible protein pockets.

For our inaugural CIP-DEL library, we modified a VHL-targeting ligand to incorporate a vector for DNA barcode attachment. The native substrate of VHL is the hydroxyproline-containing HIF-1α peptide, which spans an extended groove on the protein surface and exits at both N- and C- termini.[38] The two accessible exit vectors from this binding site enabled the design of a presenter ligand with a DNA barcode at one end, and a connector and split-and-pool library at the other. However, other E3 ligands, such as thalidomide, do not provide a second exit vector for DNA barcode attachment. For these ligands, a modified library design might be adapted where other components, such as the connectors, provide the requisite barcode attachment sites as reported recently.[39] Nevertheless, we anticipate that the high-level library design and screening principles identified in this study should be applicable to diverse CIP-DEL implementations.

As a positive control for our BRD4^{BD1} screening, we generated a DNA-barcoded derivative of **MZ1**. The on-DNA compound **MZ1-DEL** appeared to be an outlier when compared to the off-DNA potencies for **MZ1** (*e.g.,* Fig. 3b). In fact, at the outset of our analysis we were surprised to find that **MZ1-DEL** was not a highly enriched library member given the known potency of **MZ1**. The DNA attachment point used for the CIP-DEL library was designed to minimize ternary complex disruption but was added to the VHL ligand near the PPI interface induced by **MZ1** (Fig. 1a). Our data suggests that this attachment vector interferes with ternary complex formation for this degrader.

Interestingly, our ternary complex structure with **CIP-1** places BRD4^{BD1} further from the barcode attachment point. Since our CIP-DEL design mimicked the configuration of **MZ1**, one potential outcome of our screening was that we identified a BRD4^{BD1} ligand that could substitute for the **JQ1** component of **MZ1** while maintaining the same ternary complex. However, we did not observe significant enrichments for the PEG connector of **MZ1** and the configuration of the ternary complex with **CIP-1** was dramatically different than the **MZ1** ternary complex (Extended Data Fig. 5).[29] We compared the orientations of additional VHL-containing ternary complexes and found substantial variability in the approach of the target proteins, highlighting the incredible influence the bifunctional small molecule has on ternary complex geometry.[40,41]

The field of small molecule-induced proximity induction is now in its growth stage and is expanding in many new directions. Preselecting targets and presenter proteins with robust discovery of proximity-inducing small molecules will enable many neo-protein and neo-substrate-related activities. Here we describe the development of a novel CIP discovery platform with vast opportunities for expansion.

Methods

NGS data processing and analysis

Reads observed during sequencing were processed into counts for each barcoded small molecule using Python scripting. FastQ files were read using the BioPython package and used to construct a sparse mapping describing the number of reads for each unique sequence. Fifty nucleotide-long reads were split into five regions corresponding to the cycle 1 tag, cycle 2 tag, cycle 3 tag, library tag, and unique molecular identifier (UMI). The barcode region (comprising the three cycle tags and the library tag) was used to aggregate the reads corresponding to each small molecule. To account for errors during PCR amplification and sequencing, up to two errors in the library tag and up to one error in the cycle 1 tag were allowed as measured by the Hamming distance between the expected and observed tags (no errors were allowed in cycle 2 and 3 tags). This error tolerance provided a marginal increase in the number of matched reads while still allowing unambiguous assignment of an observed sequence to a single expected barcode at that position. UMIs were used to deduplicate reads by constructing a directional graph that accounts for errors during PCR and sequencing, using the umi_tools package by Smith et al.[42] The number of unique UMI clusters corresponding to each barcode corresponds to that barcode's (integer) count.

Enrichment values were estimated by comparing count values across two distinct experimental conditions (i.e., a protein-containing selection versus the beads-only control), using a recently proposed Poisson enrichment metric.[43] We assume that each barcoded small molecule exists at a fixed but unknown concentration before PCR and NGS, λ. We also make the simplifying assumption that PCR amplification does not change the relative abundance of barcodes and that the stochastic process of NGS can be modeled using Poisson statistics. Poisson distributions are characterized by a single parameter λ that in this setting describes the relative abundance of a particular barcode compared to the total abundance of all other barcodes in the sample. If the (unknown) relative abundance of one compound following one experimental condition is λ_1 and following a second experimental condition is λ_2 (e.g., where the second experiment is a beads-only control), we define the enrichment ratio as $R = \lambda_1/\lambda_2$. Defining k_1 and k_2 as the integer numbers of counts for the particular barcode and n_1 and n_2 as the total integer counts for that experimental condition, naive estimates of λ_1 and λ_2 are simply k_1/n_1 and k_2/n_2, and hence a naïve estimate of R is $(k_1/n_1)/(k_2/n_2)$. However, this ratio does not consider uncertainty in the observed values of k, which is necessarily an integer and typically close to zero using standard sequencing depths; further, it fails to produce a finite estimate of the enrichment ratio when $k_2 = 0$.

As an uncertainty-aware enrichment metric, we directly estimate the value of $R = \lambda_1/\lambda_2$ and its confidence interval rather than the values of λ_1 and λ_2 separately. We therefore perform a Poisson ratio test that formally seeks to evaluate a hypothesis test about the value of the ratio of two Poisson metrics based on the observed integer counts and total exposures (total counts)[44]:

$$H_0: \frac{\lambda_1}{\lambda_2} \leq R_0 \qquad H_1: \frac{\lambda_1}{\lambda_2} > R_0$$

Using a variance stabilizing square root transformation, we calculate a z-score as:

$$z = 2\frac{\left(k_1 + \frac{3}{8}\right)^{1/2} - \left(k_2 + \frac{3}{8}\right)^{\frac{1}{2}}\left(\frac{n_1}{n_2}R_0\right)^{1/2}}{\left(1 + \frac{n_1}{n_2}R_0\right)^{1/2}}$$

As we expect this score to follow a standard normal distribution, we can calculate a *p*-value for different values of R_0. To estimate the lower and upper bounds for the enrichment ratio, we therefore invert this hypothesis test and calculate the values of R_0 corresponding to $z = \pm 2$. This procedure can be applied to the count data for individual compounds (single barcodes), as well as groups of compounds (sets of barcodes, e.g., those sharing a particular building block) as done for the linker analysis to identify **C15** as strongly enriched.

For analysis of our screening data, we use the lower bound (lb) of the 95% confidence interval as our enrichment metric since this provides a more conservative estimate of enrichment that may reduce the number of false positives. In effect, this approach accounts for the uncertainty present when barcode counts are low: if a barcode is observed three times in a screen and once in the beads-only control, we obtain a 3-fold enrichment but with low certainty; if a barcode is observed 150 times in the screen and 50 times in the beads-only control, we obtain a 3-fold enrichment with higher certainty. Finally, enrichment values for all compounds synthesized off-DNA were tabulate in Microsoft Excel, along with the associated biophysical, biochemical, and cell-based profiling data, to analyze correlations between on-DNA and off-DNA values.

Note regarding DNA barcode analysis

We profiled a common set of amine building blocks for both cycle 2 and cycle 3 triazine substitutions using a model reaction system. A cutoff of 75% UPLC-MS product yield was used for including amines in the CIP-DEL library. Since the cutoff was applied independently for cycle 2 and cycle 3 profiling, some amines met the inclusion criteria for cycle 2, some amines met the inclusion criteria for cycle 3, and some amines met the inclusion criteria for both cycles 2 and 3. Including a subset of identical amines in cycles 2 and 3 generated redundant library members due to symmetry of the triazine scaffold (e.g. BBx and BBy added in cycles 2 and 3 generates the same molecule as BBx and BBy added in cycles 3 and 2, respectively).

A total of 218 amines (121 primary amines and 97 secondary amines) were included in both cycle 2 and 3 substitutions, producing 354,795 duplicate library members (with distinct DNA barcodes) and 728,355 unique library members. For analysis of duplicate library members, we selected the component with the greatest enrichment value in the target protein (+) VHL selection condition. Enrichment values for the target protein (−) VHL selection were then matched with the selected library member. This analysis is based on

the assumption that higher enrichment values correspond to library members with greater synthetic yields and therefore, the more representative enrichment values. We note however that variability was observed in the enrichment values between identical library members encoded with distinct DNA barcodes and further studies on the implications of redundant DNA encoding are needed.

VHL TR-FRET binding assay

Binding competition reactions were created with a mixture of 10 nM purified recombinant VHL protein complex consisting of N-terminally Avi-tagged and biotinylated VHL, untagged Elongin-B and untagged Elongin-C, 2 nM terbium-conjugated Streptavidin (ThermoFisher P/N: PV3966) and 24 nM FITC-peptide probe: (bA)-DEALA(HyP)YIPD-OH (New England Peptide) in binding assay buffer (100 mM NaCl, 50 mM Tris pH 7.5, 0.1% Pluronic F127, 1 mM DTT). The mixture was dispensed into 1536-well plates (Greiner P/N: 782075) with 5 μL / well using a tip-based dispenser (GNF Systems One Tip Dispenser). 50 nL of CIP compounds in DMSO was added to target wells using an acoustic liquid handler (Labcyte Echo 500). The plates were incubated at RT for 60 min, and then read on a plate reader (Perkin Elmer Envision) using a TR-FRET protocol (terbium donor exc: 340 / 80 nm, ems: 492 / 7.5 nm; FITC acceptor ems: 520 / 7.5 nm). The TR-FRET signal was calculated as the ratio: signal 520 nm / signal 492 nm. AC_{50} values were calculated from curve fits generated by Spotfire analysis software (Tibco). For TR-FRET cooperativity assays, the experiments were run as described above except that 20 μM $BRD4^{BD1}$ was added to each well.

SPR binding measurements

Binary complexes: His-BRD4(44–168)-avi was immobilized (~3000 RUs) on a streptavidin chip using a Biacore 8K (GE Healthcare). To determine the kinetic affinity of the small molecules, compounds were diluted in DMSO such that final concentration gradients in DMSO were 50x. 3.1 μL of the small molecules were then added to 150 μL of buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.01% Tween 20, 1 mM EDTA, 1 mM DTT) in a 384- deep well plate (Greiner Bio-One P/N: 78127) and mixed using a Biomek FX. The small molecule solution gradients (6 concentrations / 2-fold dilutions) were then injected at 45 μL / min for 120 seconds of contact time followed by 1200 seconds of dissociation time in running buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.01% Tween 20, 1 mM EDTA, 1 mM DTT, 2% DMSO). The single-cycle kinetics data was fit to a 1:1 binding model or steady-state model to measure the association rate k_a (1 / M·s), the dissociation rate k_d (1 / s), and the affinity K_D (M).

Ternary complexes: His-BRD4(44–168)-avi was immobilized (~400 RUs) on a streptavidin chip using a Biacore 8K (GE Healthcare). Compounds were diluted in DMSO to give 50x final concentration gradients. To determine the apparent kinetic affinity for ternary complex formation for small molecules with immobilized BRD4 and soluble untagged VHL complex (VHL, Elongin-B, and Elongin-C), 3.1 μL of the small molecules were added to 150 μL of buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.01% Tween 20, 1 mM EDTA, 1 mM DTT) containing 2 or 4 μM VHL complex in a 384-deep well plate (Greiner Bio-One P/N: 78127) and mixed using a Biomek FX. The protein / small molecule solution gradients

(6 compound concentrations / 2-fold dilutions) were then injected at 45 μL / min for 120 seconds of contact time followed by a dissociation time of 1200 seconds in running buffer (50 mM Tris pH 8.0, 150 mM NaCl, 0.01% Tween 20, 1 mM EDTA, 1 mM DTT, 2% DMSO). The single-cycle kinetics data was fit to a 1:1 binding model or steady-state model to measure the association rate k_a (1 / M·s), the dissociation rate k_d (1 / s), and the affinity K_D (M).

Data Analysis: A detailed report of all kinetic and steady-state binding parameters for CIP compounds, along with calculated α-values, is provided in Supplementary Table 3. To simplify subsequent analyses, we averaged all experimentally determined K_D values for each compound and used this data (shown in Supplementary Table 1) to look for correlations across our data sets.

NanoBit cellular ternary complex complementation assay

BRD4^{BD1} assay protocol.—HEK293T cells were transfected by plating 15 mL of cells at 0.625 × 10^6 cells / mL in complete media (DMEM with 10% FBS) into a 75 cm^2 flask. 90 μL of Fugene HD (Promega P/N: E2312) was combined with 1500 μL of OptiMEM I (Gibco P/N: 31985–062) and added to 30 μg of total DNA (15 μg of HSVTK-driven VHL-LgBiT plasmid and 15 μg of CMV-driven SmBiT-BRD4^{BD1} plasmid). The transfection mixture was incubated for 15 min at RT before being added dropwise to the plated cells. After 24 h of incubation in a cell culture incubator (37°C, 5% CO_2), the transfected cells were trypsinized, normalized to 0.5 × 10^6 cells / mL in complete media and plated into 1536-well plates, 5 μL / well using a tip-based dispenser (GNF Systems One Tip Dispenser). The plates were incubated in a cell culture incubator (37°C, 5% CO_2) for 5 h. Next, 5 nL of 1 mM MLN4924 in DMSO was added to every well using an acoustic liquid handler (Labcyte Echo 500), and the plates were returned to the incubator for an additional 17 h. Next, 10 nL of CIP compounds in 100% DMSO, or DMSO alone as a control, were dispensed into target wells using the Echo, and the plates were returned to the incubator for another 2 h. The plates were allowed to cool to RT for 10 min, and 5 μL of NanoBiT NanoGlo reagent, consisting of NanoGlo subtrate (Promega P/N: N205B) diluted 1:25 in a mixture of 50% NanoGlo LCS buffer (Promega P/N: N206C) and 50% PBS, was dispensed into each well using the GNF dispenser. After a final incubation at RT for 10 min, the plates were read on a luminescence plate reader (Perkin Elmer ViewLux). Luminescence for each data point was normalized using the equation 100*(well value-NC)/NC, where NC is the neutral control well (DMSO). Curves were fitted using GraphPad Prism v9.0.0.

BRD2^{BD1} and BRDT^{BD1} assay protocol.—5 mL of HEK293T cells were plated at 0.625 × 10^6 cells / mL in complete media (DMEM with 10% FBS) with 10% FBS in a T25 flask. 30 μL of Fugene HD (Promega P/N: E2312) was combined with 500 μL of OptiMEM I (Gibco P/N: 31985062) and added to 5 μg of HSVTK-driven VHL-LgBiT plasmid and 5 μg of CMV-driven SmBiT-BRD2^{BD1} or SmBiT-BRDT^{BD1} plasmid. The transfection mixture was incubated for 10 min at RT before being added dropwise to the plated cells. After 16 h of incubation in a cell culture incubator (37°C, 5% CO_2), the transfected cells were trypsinized and normalized to 0.38 × 10^6 cells / mL in complete media. MLN4924 in DMSO was then added to the cells (2 μM final concentration). The cells were plated into

1536-well plates, 5 μL / well using a tip-based dispenser (GNF Systems One Tip Dispenser). The plates were incubated in a cell culture incubator (37°C, 5% CO_2) overnight for 22 h. Next, 15 nL of each compound in 100% DMSO, or DMSO alone as a control, were dispensed into target wells using an acoustic liquid handler (Labcyte Echo 655), and the plates were returned to the incubator for another 2 h. The plates were allowed to cool to RT for 10 min, and 5 μL of NanoBiT NanoGlo reagent, consisting of NanoGlo subtrate (Promega P/N: N205B) diluted 1:25 in a mixture of 50% NanoGlo LCS buffer (Promega P/N: N206B) and 50% PBS, was dispensed into each well using the GNF dispenser. After a final incubation at RT for 10 min, the plates were read on a luminescence plate reader (PHERAstar FS). Luminescence for each data point was normalized using the equation 100*(well value-NC)/NC, where NC is the neutral control well (DMSO). Curves were fitted using GraphPad Prism v9.3.1.

Cell-Titer Glo cell viability assay

The Cell-Titer Glo assay was performed identically to the HiBiT cellular protein degradation assay, except 5 μL of Cell-Titer Glo substrate (Promega P/N: G7571) was added instead of HiBiT lytic reagent.

BRD4 Western blotting

Degradation of endogenous BRD4 was validated by western blotting in HEK293T cells cultured in DMEM supplemented with 10% FBS and 100 U / mL penicillin-streptomycin. For each treatment condition, 350k cells in 2 mL of media were plated in 6-well culture plates and incubated at 37°C with 5% CO_2. After 4 h, 200 μL of media was removed and replaced with 200 μL of 10x drug treatment solution diluted in media. Final DMSO concentrations were 1%. Cells were treated with compounds for 18 h, then placed on ice and washed twice with ice cold PBS. The cells were lysed with 100 μL RIPA buffer (20 mM Tris pH 7.4, 150 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% NP-40, 0.25% deoxycholate, 1 mM Na_2VO_4, 1 mM PMSF, 1 mM NaF, and Roche protease inhibitor cocktail) and the lysates were clarified by centrifugation at 20,000 × g for 30 min at 4°C. Protein concentrations were determined for each lysate using the Pierce Detergent Compatible Bradford Assay (P/N: 1863028) with Bio-Rad BGG protein standards (P/N: 5000209) and 30 μg of each protein lysate was separated via electrophoresis on polyacrylamide gels. Proteins were transferred to nitrocellulose membranes using the Trans-Blot Turbo Transfer System (Bio-Rad) and the membranes were blocked with 5% BSA in TBST for 30 min. The blots were probed with an anti-BRD4 rabbit monoclonal antibody (Abcam, P/N: ab128874, 1:1000) at 4°C with gentle shaking overnight, then stained with IRDye 800CW donkey anti-rabbit secondary antibody (LiCOR, P/N: 926–32213, 1:10,000) and imaged using a LiCOR Odyssey infrared imaging system. Finally, blots were probed with an anti-β-actin primary antibody (Cell Signaling Technology, P/N: 4970S, 1:2000) and IRDye 800CW secondary (as above) and imaged to confirm uniform protein loading.

BRD2 Western blotting

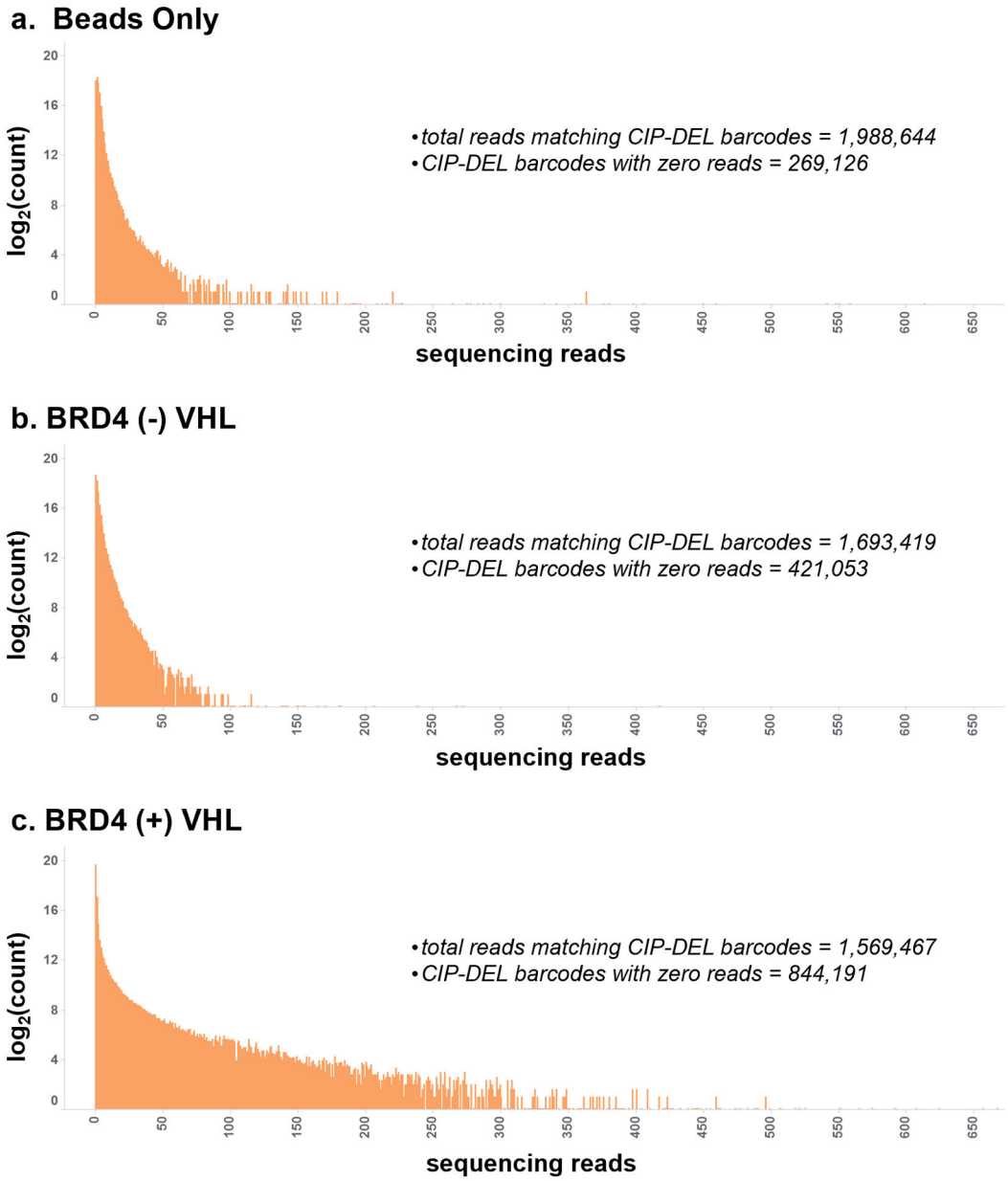
Degradation of endogenous BRD2 was validated by western blotting in HEK293T cells cultured in DMEM supplemented with 10% FBS and 100 U / mL penicillin-streptomycin. For each treatment condition, 300k cells in 1 mL of media were plated in 12-well culture

plates and incubated at 37°C with 5% CO_2. After 24 h, 500 μL of media was removed and replaced with 500 μL of 2x drug treatment solution diluted in media. Final DMSO concentrations were 0.5%. Cells were treated with compounds for 20 h. Media was removed, and the cells were lysed with 70 μL RIPA buffer with cOmplete™ EDTA-free Protease Inhibitor Cocktail (Sigma Aldrich, P/N: 04693159001) and incubated on ice for 30 min. 20 μL of 8 M urea was added, and the supernatant was mixed with 4x loading dye and run on a polyacrylamide gel (Thermo, P/N: NP0336BOX). Proteins were transferred to PVDF membranes (Thermo, P/N: IB24002) using the iBlot 2 Gel Transfer Device (Thermo) and the membranes were blocked with Intercept Blocking Buffer (LI-COR, P/N: 927–70001) for 1 h. The blots were probed with an anti-BRD2 rabbit monoclonal antibody (Abcam, P/N: ab139690, 1:500) and an anti-β actin mouse monoclonal antibody (Santa Cruz Biotechnology, P/N: sc-47778, 1:1000) at 4°C with gentle shaking overnight, then stained with goat anti-rabbit (IRDye 800CW, LI-COR P/N: 926–32211) and goat anti-mouse (IRDye 680RD, LI-COR P/N: 926–68070) secondary antibodies (1:10,000) and imaged using a LI-COR Odyssey infrared imaging system.

X-ray crystallography

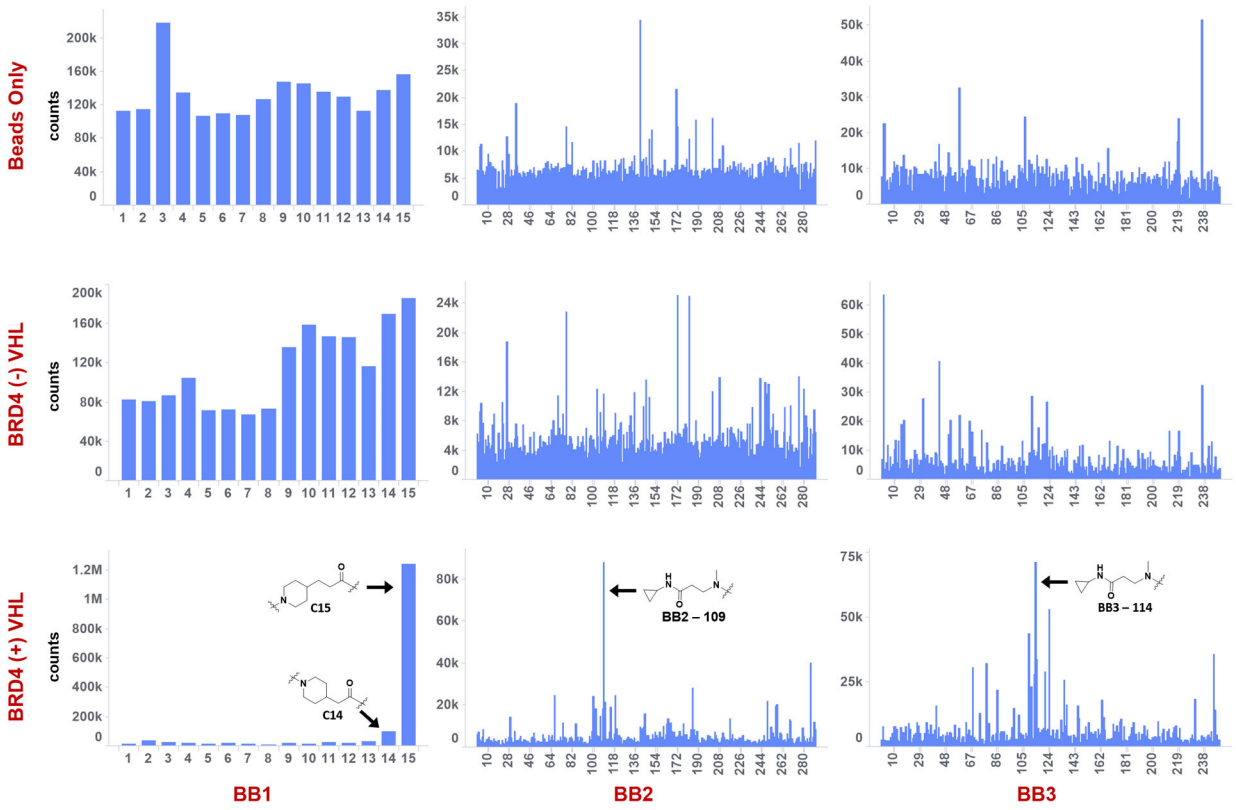
VCB complex (233 μM) and $BRD4^{BD1}$ (1280 μM) were dialyzed separately into buffer containing 5 mM Hepes pH 7.5, 100 mM NaCl, and 0.5 mM TCEP. After dialysis overnight, VCB complex and $BRD4^{BD1}$ were mixed at 1:1 molar ratio in the presence of 200 μM **CIP-1**. The VCB / $BRD4^{BD1}$ / **CIP-1** complex was concentrated to 10 mg / ml (180 μM) prior to crystallization screens. Multiple crystal hits were identified from related conditions such as 0.2 M sodium malonate, 20% w/v PEG 3350 (Hampton research PEG/Ion E8) and 0.2 M sodium citrate, 20% w/v PEG 3350 (Hampton research PEG/Ion D10), after 2 – 4 days of incubation. Single crystals suitable for data collection were obtained after several rounds of optimization with streak seeding in 0.2 M sodium malonate, 16 – 20% w/v PEG 3350. Crystals were cryo-protected in reservoir solution supplemented with 20% glycerol and flash-cooled in liquid nitrogen for data collection. Datasets were collected under cryogenic conditions (100 K) at the Advanced Light Source (ALS) beamline 5.0.2. All data were processed using CCP4[45], employing XDS[46] for data integration and AIMLESS[47] for scaling. Molecular replacement was carried out using PHASER[48] with coordinates from human VHL-ElonginC-ElonginB complex (PDB ID: 1VCB)[49] and the first bromodomain of human BRD4 (PDB ID: 2OSS).[50] Structure refinement was carried out in PHENIX[51] alternated with manual fitting in Coot.[52] Data collection and structure refinement statistics are included in Supplementary Table 5. Structural analysis was performed in PyMol (Schrodinger).

Extended Data

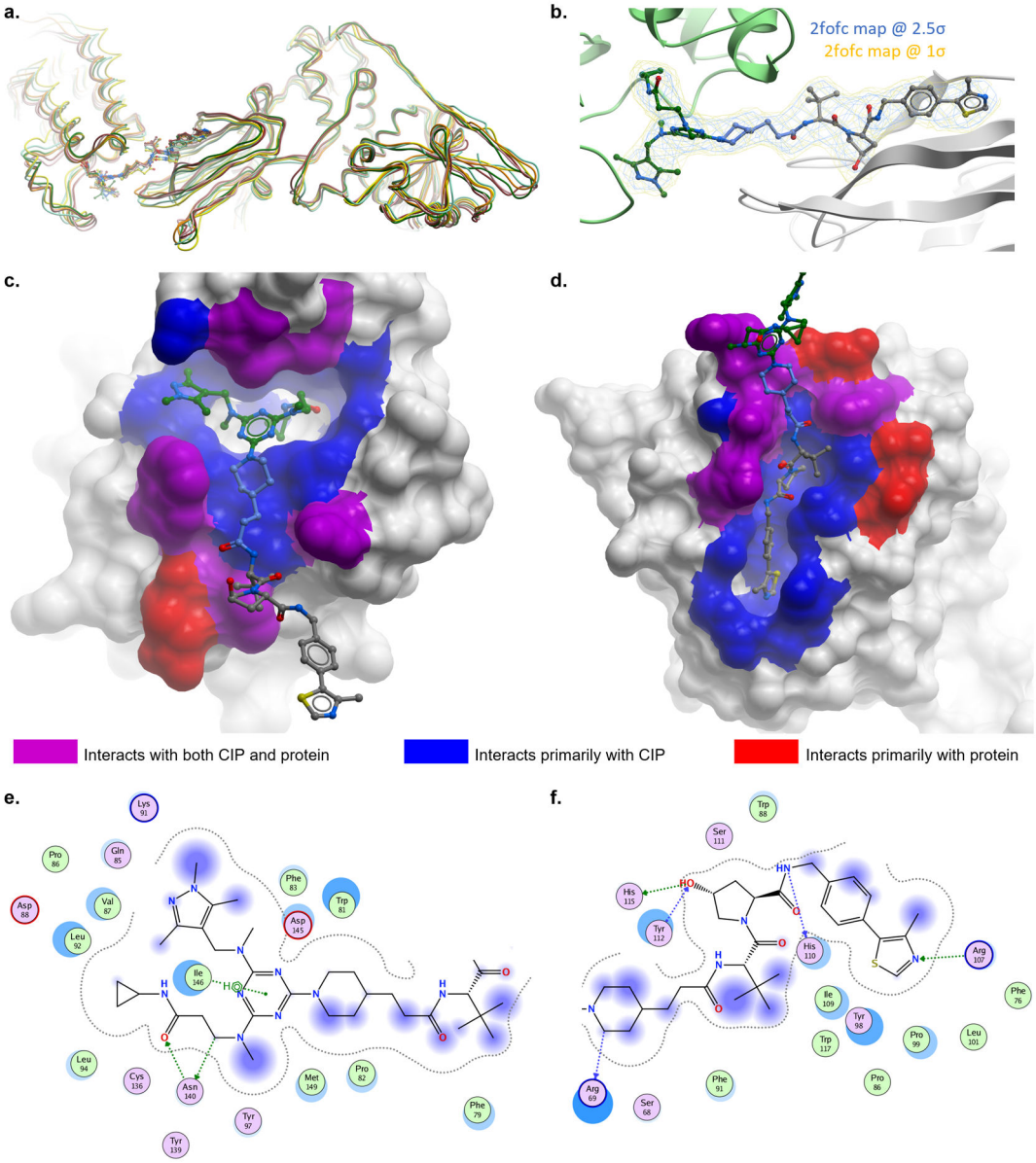


Extended Data Fig. 1.

Distribution of barcode counts for each CIP-DEL library member across each screening condition. (a) Barcode counts for the beads-only screen. (b) Barcode counts for the BRD4 (-) VHL screen. (c) Barcode counts for the BRD4 (+) VHL screen.



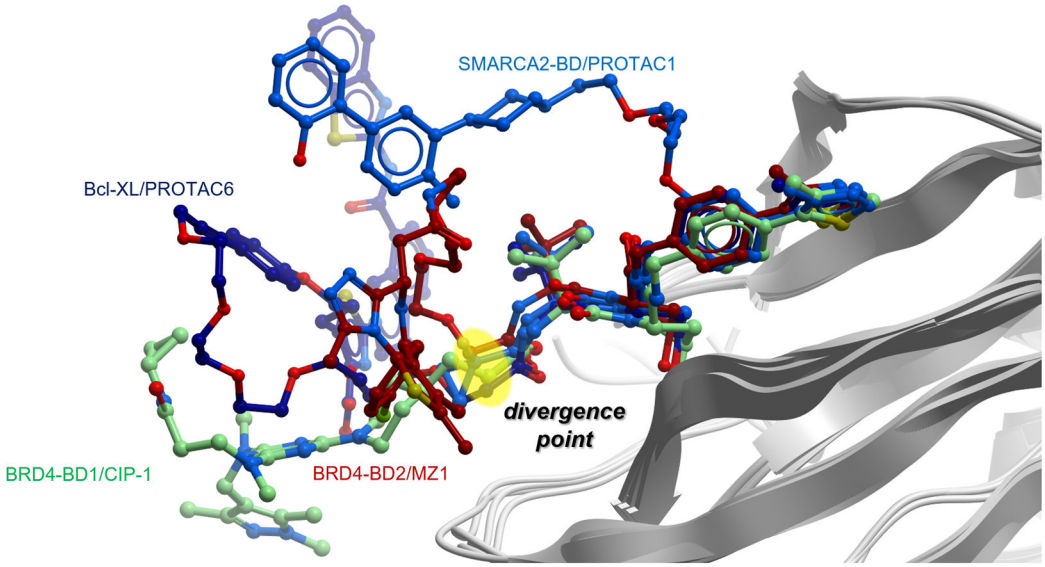
Extended Data Fig. 2.
Individual building block barcode counts for each screening condition.



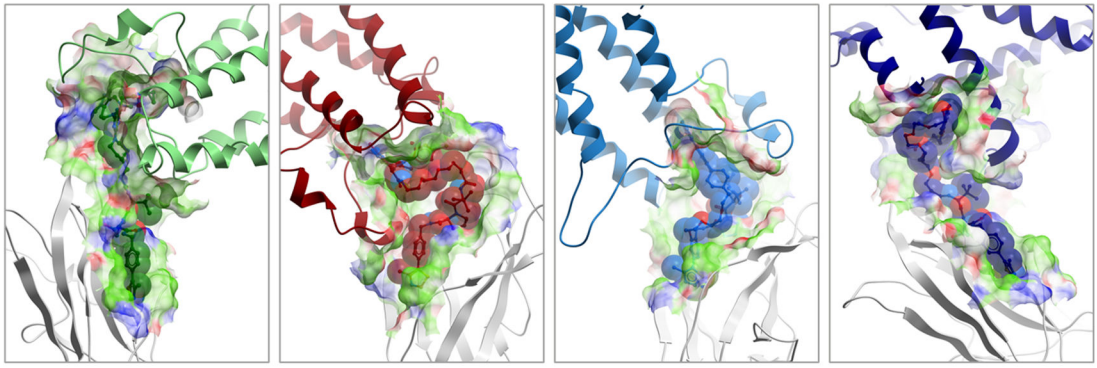
Extended Data Fig. 3.

CIP-1 ternary complex structural analysis. (a) Six copies of the ternary complex are present in each asymmetric unit. An overlay of the six ternary complexes shows the structures are highly similar with RMS deviations of 1.06 – 1.37 Å. (b) The electron density for CIP-1 is well defined in 4 of the 6 complexes of the asymmetric unit. $2F_o - F_c$ maps contoured at 2.5σ and 1σ are shown for CIP-1 in a ternary complex with well-defined density. (c) $BRD4^{BD1}$ protein surface with residues colored based on interactions with CIP-1 (blue), VHL (red), or both (purple). CIP-1 is shown in sticks. (d) VHL protein surface with residues colored based on interactions with CIP-1 (blue), $BRD4^{BD1}$ (red), or both (purple). (e) 2D diagram of CIP-1 interactions with $BRD4^{BD1}$. (f) 2D diagram of CIP-1 interactions with VHL.

a.



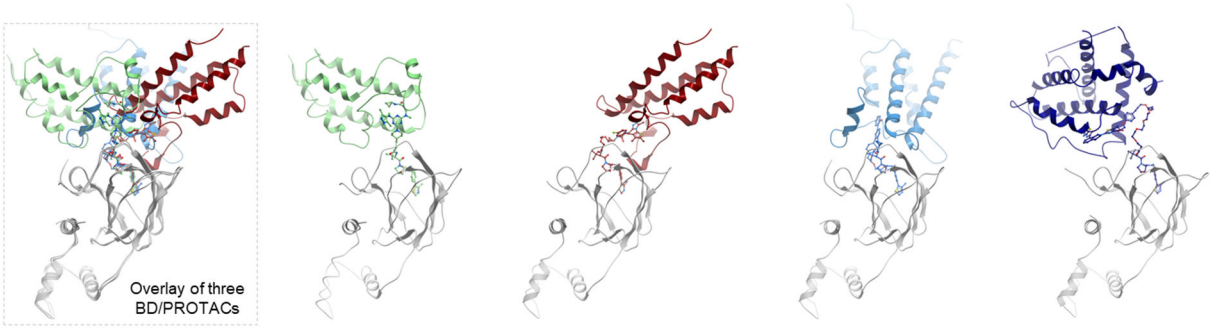
b.



LMW total area (Å²)	1412	1321	1271	1872
Percent buried	86.5%	82.5%	82.3%	81.5%
PPI interface area (neo-interactions) (Å²)	216	320	337	310

Extended Data Fig. 4.

Comparison of the small molecule components of VHL PROTAC ternary complex structures. (a) Overlaying the VHL binding moieties of ternary complex assemblies, the diverse projection vectors of the connectors and target protein ligands are shown. The initial point of divergence for the bifunctional molecules is the carbon atom following the terminal amide of the VH032 ligand. (b) Comparison of the buried surface area for four VHL-targeting CIPs.



	BRD4-BD1/CIP-1	BRD4-BD2/MZ1	SMARCA2-BD/PROTAC1	Bcl-XL/PROTAC6
Linker attachment to VH032	Left Hand Side (LHS)	LHS	RHS phenolic linkage	LHS side
Types of linker	short, rigid	flexible PEG-based linker	flexible PEG-based linker	flexible PEG-based linker
Linker conformation	extended	collapsed	collapsed	collapsed
Cooperativity (α)	0.3 by SPR	17.6 by ITC	4.8 by ITC, 4.4 by FP, 12 by TR-FRET	0.7 by SPR
PDB reference	8EWV	5T35	6HAY	6ZHC

Extended Data Fig. 5.

Comparison of the ternary complex assemblies for four VHL-targeting bifunctional degraders. The first three structures show bromodomains targets ($BRD4^{BD1}$, $BRD4^{BD2}$, and $SMARCA4^{BD}$) and the fourth complex is a non-bromodomain targeting degrader (Bcl-XL). VHL proteins are aligned at the bottom of each structure and the target protein orientation is shown at the top. Key details for each complex are provided in the table. Inset: the three bromodomain containing complexes are overlaid with VHL aligned at the bottom of the structures, highlighting the diverse target protein orientations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

John Capece, Jennifer Poirier, Philip Michaels, Carmelina Rakiec, Alicia Lindeman, Thomas Dice and Ritesh Tichkule are gratefully acknowledged for excellent technical and analytical support. Drs. Bruce Hua, Christopher Gerry, Wenyu Wang, Andrew Reidenbach, Katherine Lim, Shubhroz Gill, Bingqi Tong, Lawrence Chung, Christian Gampe, and Nichola Smith are gratefully acknowledged for their support, guidance, and valuable feedback during the preparation of this manuscript. We also thank Dustin Dovala and Michael Romanowski for supplying the protein reagents used in SPR. The research was supported in part by the National Institute of General Medical Sciences (R35GM127045 awarded to S.L.S.) and by the NIBR Scholar's Program.

Data availability

All data supporting the findings of this study are available within the article and Supplementary Information. The X-ray crystal structure for the VCB - CIP-1 - BRD4 complex is available on the PDB (ID: 8EWV). Source data for Figs. 2h and 5h are provided with the paper. The Fastq files, containing raw Illumina sequencing data are available at https://doi.org/10.5281/zenodo.8253891.

References

1. Schreiber SL The Rise of Molecular Glues. Cell 184, 3–9 (2021). [PubMed: 33417864]

2. Liu J et al. Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. Cell 66, 807–815 (1991). [PubMed: 1715244]
3. Brown EJ et al. A mammalian protein targeted by G1-arresting rapamycin–receptor complex. Nature 369, 756–758 (1994). [PubMed: 8008069]
4. Sabatini DM, Erdjument-Bromage H, Lui M, Tempst P & Snyder SH RAFT1: A mammalian protein that binds to FKBP12 in a rapamycin-dependent fashion and is homologous to yeast TORs. Cell 78, 35–43 (1994). [PubMed: 7518356]
5. Haggarty SJ et al. Dissecting cellular processes using small molecules: identification of colchicine-like, taxol-like and other small molecules that perturb mitosis. Chemistry & Biology 7, 275–286 (2000). [PubMed: 10780927]
6. Uehara T et al. Selective degradation of splicing factor CAPERα by anticancer sulfonamides. Nature Chemical Biology 13, 675–680 (2017). [PubMed: 28437394]
7. Han T et al. Anticancer sulfonamides target splicing by inducing RBM39 degradation via recruitment to DCAF15. Science 356, eaal3755 (2017). [PubMed: 28302793]
8. Du X et al. Structural Basis and Kinetic Pathway of RBM39 Recruitment to DCAF15 by a Sulfonamide Molecular Glue E7820. Structure 27, 1625–1633.e3 (2019). [PubMed: 31693911]
9. Bussiere DE et al. Structural basis of indisulam-mediated RBM39 recruitment to DCAF15 E3 ligase complex. Nature Chemical Biology 16, 15–23 (2020). [PubMed: 31819272]
10. Faust TB et al. Structural complementarity facilitates E7820-mediated degradation of RBM39 by DCAF15. Nature Chemical Biology 16, 7–14 (2020). [PubMed: 31686031]
11. Krönke J et al. Lenalidomide Causes Selective Degradation of IKZF1 and IKZF3 in Multiple Myeloma Cells. Science 343, 301–305 (2014). [PubMed: 24292625]
12. Lu G et al. The Myeloma Drug Lenalidomide Promotes the Cereblon-Dependent Destruction of Ikaros Proteins. Science 343, 305–309 (2014). [PubMed: 24292623]
13. Gandhi AK et al. Immunomodulatory agents lenalidomide and pomalidomide co-stimulate T cells by inducing degradation of T cell repressors Ikaros and Aiolos via modulation of the E3 ubiquitin ligase complex CRL4CRBN. British Journal of Haematology 164, 811–821 (2014). [PubMed: 24328678]
14. Fischer ES et al. Structure of the DDB1–CRBN E3 ubiquitin ligase in complex with thalidomide. Nature 512, 49–53 (2014). [PubMed: 25043012]
15. Petzold G, Fischer ES & Thomä NH Structural basis of lenalidomide-induced CK1α degradation by the CRL4CRBN ubiquitin ligase. Nature 532, 127–130 (2016). [PubMed: 26909574]
16. Donovan KA et al. Thalidomide promotes degradation of SALL4, a transcription factor implicated in Duane Radial Ray syndrome. eLife 7, e38430 (2018). [PubMed: 30067223]
17. Matyskiela ME et al. Crystal structure of the SALL4–pomalidomide–cereblon–DDB1 complex. Nature Structural & Molecular Biology 27, 319–322 (2020).
18. Banik SM et al. Lysosome-targeting chimaeras for degradation of extracellular proteins. Nature 584, 291–297 (2020). [PubMed: 32728216]
19. Caianiello DF et al. Bifunctional small molecules that mediate the degradation of extracellular proteins. Nature Chemical Biology 17, 947–953 (2021). [PubMed: 34413525]
20. Zhou Y, Teng P, Montgomery NT, Li X & Tang W Development of Triantennary N-Acetylgalactosamine Conjugates as Degraders for Extracellular Proteins. ACS Central Science 7, 499–506 (2021). [PubMed: 33791431]
21. Siriwardena SU et al. Phosphorylation-Inducing Chimeric Small Molecules. Journal of the American Chemical Society 142, 14052–14057 (2020). [PubMed: 32787262]
22. Henning NJ et al. Deubiquitinase-targeting chimeras for targeted protein stabilization. Nature Chemical Biology 18, 412–421 (2022). [PubMed: 35210618]
23. Wang WW et al. Targeted Protein Acetylation in Cells Using Heterobifunctional Molecules. Journal of the American Chemical Society 143, 16700–16708 (2021). [PubMed: 34592107]
24. Yamazoe S et al. Heterobifunctional Molecules Induce Dephosphorylation of Kinases–A Proof of Concept Study. Journal of Medicinal Chemistry 63, 2807–2813 (2020). [PubMed: 31874036]
25. Chen P-H et al. Modulation of Phosphoprotein Activity by Phosphorylation Targeting Chimeras (PhosTACs). ACS Chemical Biology (2021).

26. Sakamoto KM et al. Protacs: Chimeric molecules that target proteins to the Skp1–Cullin–F box complex for ubiquitination and degradation. Proceedings of the National Academy of Sciences 98, 8554–8559 (2001).
27. Brenner S & Lerner RA Encoded combinatorial chemistry. Proceedings of the National Academy of Sciences 89, 5381–5383 (1992).
28. Zengerle M, Chan K-H & Ciulli A Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. ACS Chemical Biology 10, 1770–1777 (2015). [PubMed: 26035625]
29. Gadd MS et al. Structural basis of PROTAC cooperative recognition for selective protein degradation. Nature Chemical Biology 13, 514–521 (2017). [PubMed: 28288108]
30. Clark MA et al. Design, synthesis and selection of DNA-encoded small-molecule libraries. Nature Chemical Biology 5, 647–654 (2009). [PubMed: 19648931]
31. Roy MJ et al. SPR-Measured Dissociation Kinetics of PROTAC Ternary Complexes Influence Target Degradation Rate. ACS Chemical Biology 14, 361–368 (2019). [PubMed: 30721025]
32. Dixon AS et al. NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. ACS Chemical Biology 11, 400–408 (2016). [PubMed: 26569370]
33. Schwinn MK et al. CRISPR-Mediated Tagging of Endogenous Proteins with a Luminescent Peptide. ACS Chemical Biology 13, 467–474 (2018). [PubMed: 28892606]
34. Soucy TA et al. An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. Nature 458, 732–736 (2009). [PubMed: 19360080]
35. Martinez NJ et al. A widely-applicable high-throughput cellular thermal shift assay (CETSA) using split Nano Luciferase. Scientific Reports 8, 9472 (2018). [PubMed: 29930256]
36. Filippakopoulos P et al. Selective inhibition of BET bromodomains. Nature 468, 1067–1073 (2010). [PubMed: 20871596]
37. Galdeano C et al. Structure-Guided Design and Optimization of Small Molecules Targeting the Protein–Protein Interaction between the von Hippel–Lindau (VHL) E3 Ubiquitin Ligase and the Hypoxia Inducible Factor (HIF) Alpha Subunit with in Vitro Nanomolar Affinities. Journal of Medicinal Chemistry 57, 8657–8663 (2014). [PubMed: 25166285]
38. Min J-H et al. Structure of an HIF-1α-pVHL complex: hydroxyproline recognition in signaling. Science 296, 1886–1889 (2002). [PubMed: 12004076]
39. Chen Q et al. Optimization of PROTAC Ternary Complex Using DNA Encoded Library Approach. ACS Chemical Biology 18, 25–33 (2023). [PubMed: 36606710]
40. Farnaby W et al. BAF complex vulnerabilities in cancer demonstrated via structure-based PROTAC design. Nature Chemical Biology 15, 672–680 (2019). [PubMed: 31178587]
41. Chung C. w. et al. Structural Insights into PROTAC-Mediated Degradation of Bcl-xL. ACS Chemical Biology 15, 2316–2323 (2020). [PubMed: 32697072]
42. Smith T, Heger A & Sudbery I UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Research 27, 491–499 (2017). [PubMed: 28100584]
43. Lim KS et al. Machine Learning on DNA-Encoded Library Count Data Using an Uncertainty-Aware Probabilistic Loss Function. Journal of Chemical Information and Modeling 62, 2316–2331 (2022). [PubMed: 35535861]
44. Gu K, Ng HKT, Tang ML & Schucany WR Testing the Ratio of Two Poisson Rates. Biometrical Journal 50, 283–298 (2008). [PubMed: 18311854]
45. Collaborative. The CCP4 suite: programs for protein crystallography. Acta Crystallographica Section D 50, 760–763 (1994).
46. Kabsch W XDS. Acta Crystallographica Section D 66, 125–132 (2010).
47. Evans PR & Murshudov GN How good are my data and what is the resolution? Acta Crystallographica Section D 69, 1204–1214 (2013).
48. McCoy AJ et al. Phaser crystallographic software. Journal of Applied Crystallography 40, 658–674 (2007). [PubMed: 19461840]

49. Stebbins CE, Kaelin WG Jr. & Pavletich NP Structure of the VHL-ElonginC-ElonginB complex: implications for VHL tumor suppressor function. Science 284, 455–61 (1999). [PubMed: 10205047]

50. Filippakopoulos P et al. Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family. Cell 149, 214–231 (2012). [PubMed: 22464331]

51. Adams PD et al. The Phenix software for automated determination of macromolecular structures. Methods (San Diego, Calif.) 55, 94–106 (2011). [PubMed: 21821126]

52. Emsley P & Cowtan K Coot: model-building tools for molecular graphics. Acta Crystallographica Section D 60, 2126–2132 (2004).

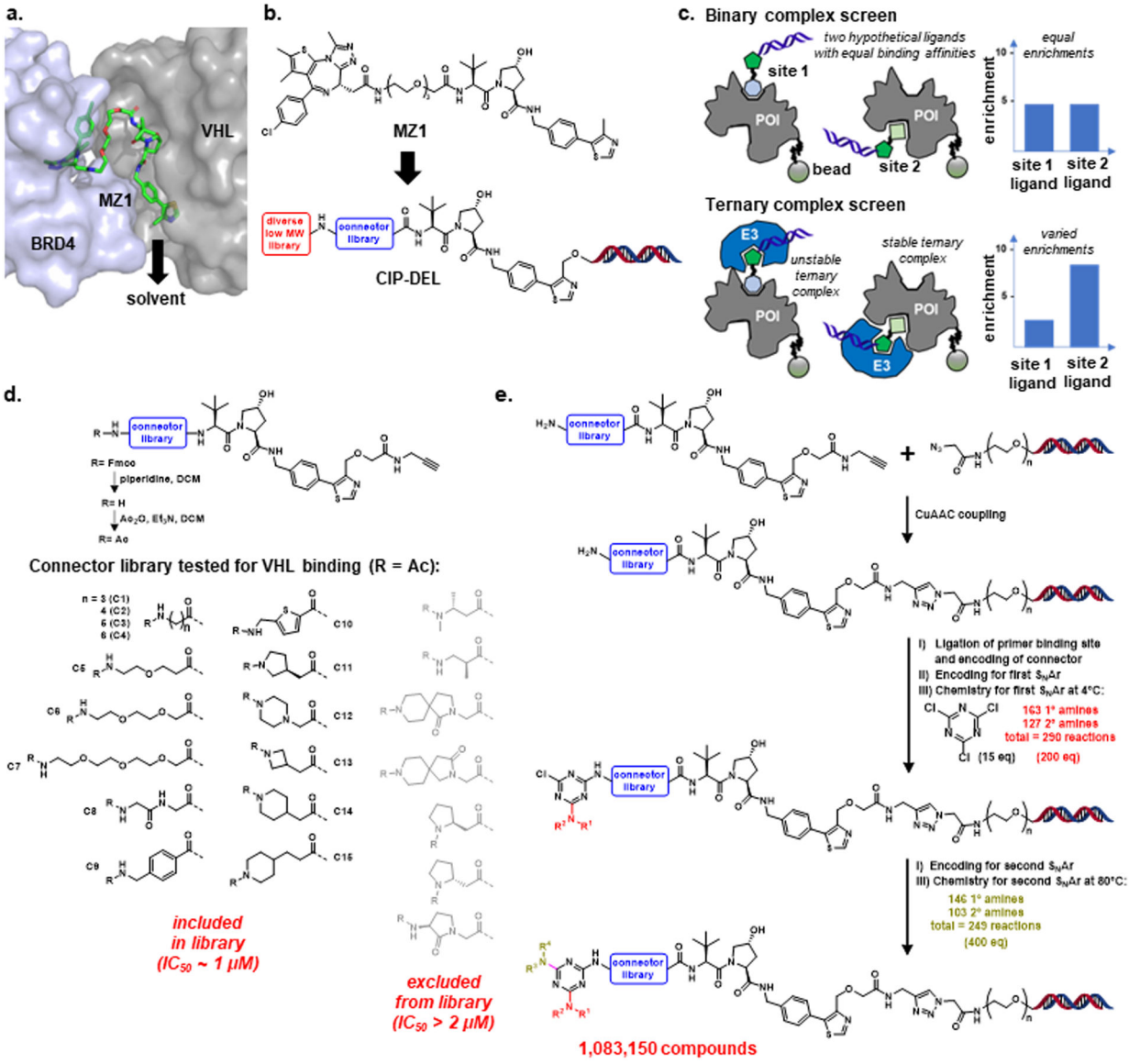


Figure 1. Library design and synthesis.

(**a**) Using the ternary complex crystal structure of **MZ1** with $BRD4^{BD2}$ and VHL (PDB: 5T35), the thiazole-methyl group of **MZ1** was selected as the site for DNA attachment as it projects towards a solvent exposed channel, potentially minimizing disruption of the ternary complex. (**b**) **MZ1** was used as a template for CIP-DEL library design. The VHL ligand was functionalized with a connector library and appended to DNA. Each connector terminated with an amine that was functionalized using split-and-pool DEL synthesis methods. (**c**) Affinity screening strategy for identifying complex-inducing library members. Parallel screens with the protein of interest +/− the E3 ubiquitin ligase are performed and enrichment signals that remain high in both screens suggest library members that are capable of stable ternary complex induction. (**d**) The VHL ligand was modified with an alkyne to enable DNA attachment and functionalized with 22 Fmoc-amino acid connectors. The Fmoc groups were replaced by acetyl groups, and each compound was assayed for binding to

VHL. We selected 15 connectors that retained binding to VHL to include in the CIP-DEL library. (**e**) Overview of the CIP-DEL library synthesis.

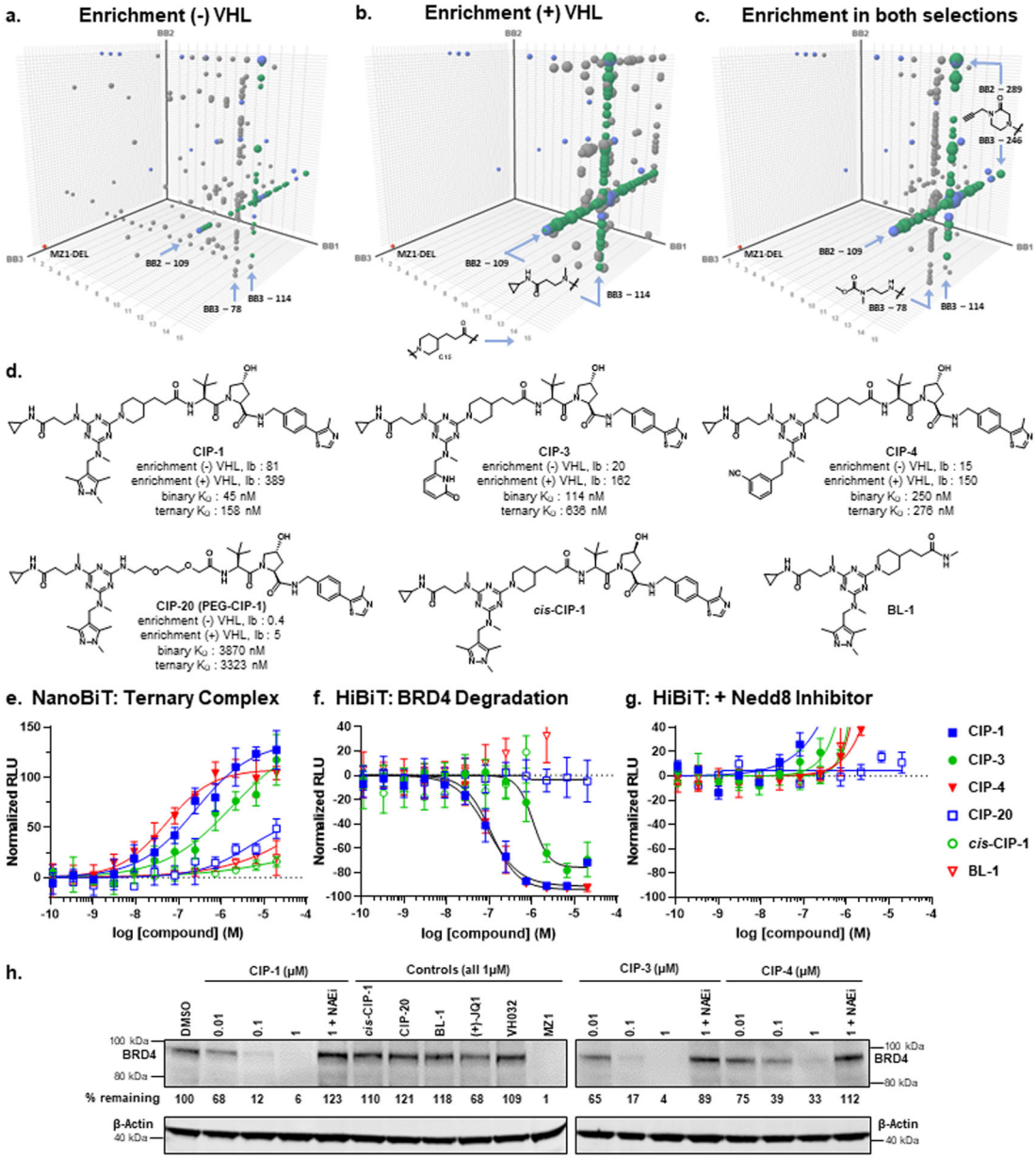


Figure 2. CIP-DEL screening analysis and evaluation of compounds synthesized off-DNA.

(**a - c**) Cube plots of enriched features in (**a**) the BRD4 (−) VHL screen (enrichment cutoff ≥ 7), (**b**) the BRD4 (+) VHL screen (enrichment cutoff ≥ 80), and (**c**) both the BRD4 (−) VHL and BRD4 (+) VHL screens (enrichment cutoff ≥ 5 in both). Points are sized by (−) VHL enrichment values in (**a**) and (+) VHL enrichment values in (**b**) and (**c**). Coloring: red = **MZ1-DEL** (shown at 1,1,1); grey = meets enrichment cutoff; green = contains **C15** and **BB2–109** or **BB3–114**; blue = synthesized off-DNA. Compounds selected for synthesis are shown in all plots and may not meet indicated enrichment cutoffs. (**d**) Structures of select library members synthesized off-DNA with binary and ternary K_D values from SPR (see Supplementary Table 1 for additional details). (**e**) Ternary complex induction by select CIP compounds using the NanoBiT assay. (**f**) Compound induced $BRD4^{BD1}$ degradation

measured by the HiBiT assay. (**g**) $BRD4^{BD1}$ degradation with CIP compounds is reversed in the presence of a Nedd8 inhibitor. Data in **e** - **g** represent the mean ± s.d. of $n = 2$ independent experiments. (**h**) Western blot analysis of endogenous BRD4 degradation with CIP compounds and controls ($n = 1$).

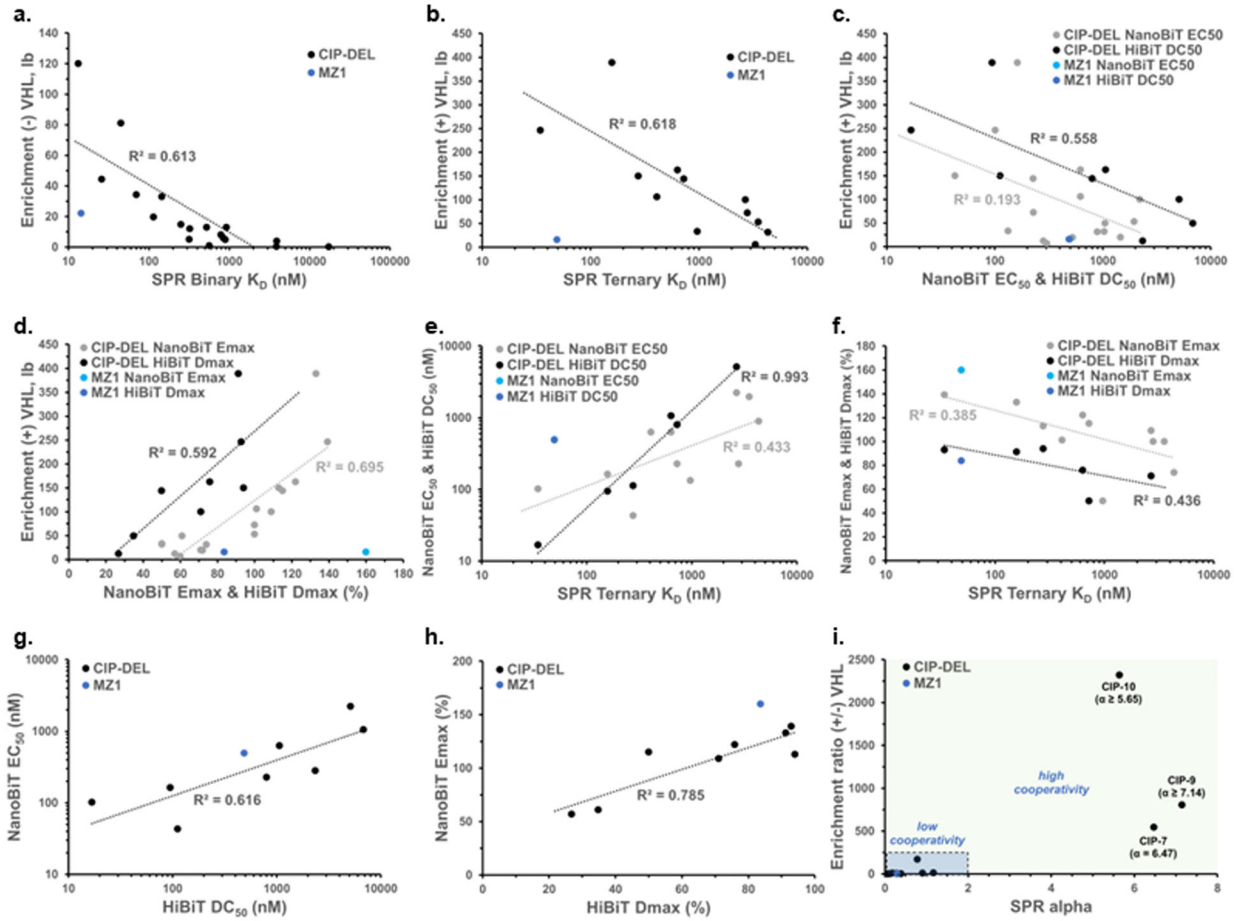


Figure 3. Correlations between CIP-DEL screening results and off-DNA validation data.

(**a**) Correlation between BRD4 (−) VHL enrichment values (lower-bound of 95% confidence interval) and binary K_D values obtain by SPR. (**b**) Relationship between BRD4 (+) VHL enrichments and ternary K_D values obtained by SPR. (**c**) BRD4 (+) VHL enrichment values compared to NanoBiT EC_{50} and HiBiT DC_{50} values. (**d**) BRD4 (+) VHL enrichments compared to NanoBiT Emax and HiBiT Dmax results. (**e**) Relationship between NanoBiT EC_{50} and HiBiT DC_{50} values, and the ternary K_D values obtained from SPR. (**f**) Correlation of NanoBiT Emax and HiBiT Dmax with the ternary K_D results obtained by SPR. (**g**) Association between the NanoBiT EC_{50} and HiBiT DC_{50} results obtained for each active off-DNA compound. (**h**) Correlation between NanoBiT Emax and HiBiT Dmax values for active DNA-free compounds. (**i**) Comparison of CIP-DEL enrichment ratios (+/−) VHL and the alpha values calculated from binary and ternary SPR K_D values. In all panels, data points for **MZ1** were excluded from the line fitting due to potential ternary complex disruptions caused by attachment of the DNA tag (see text).

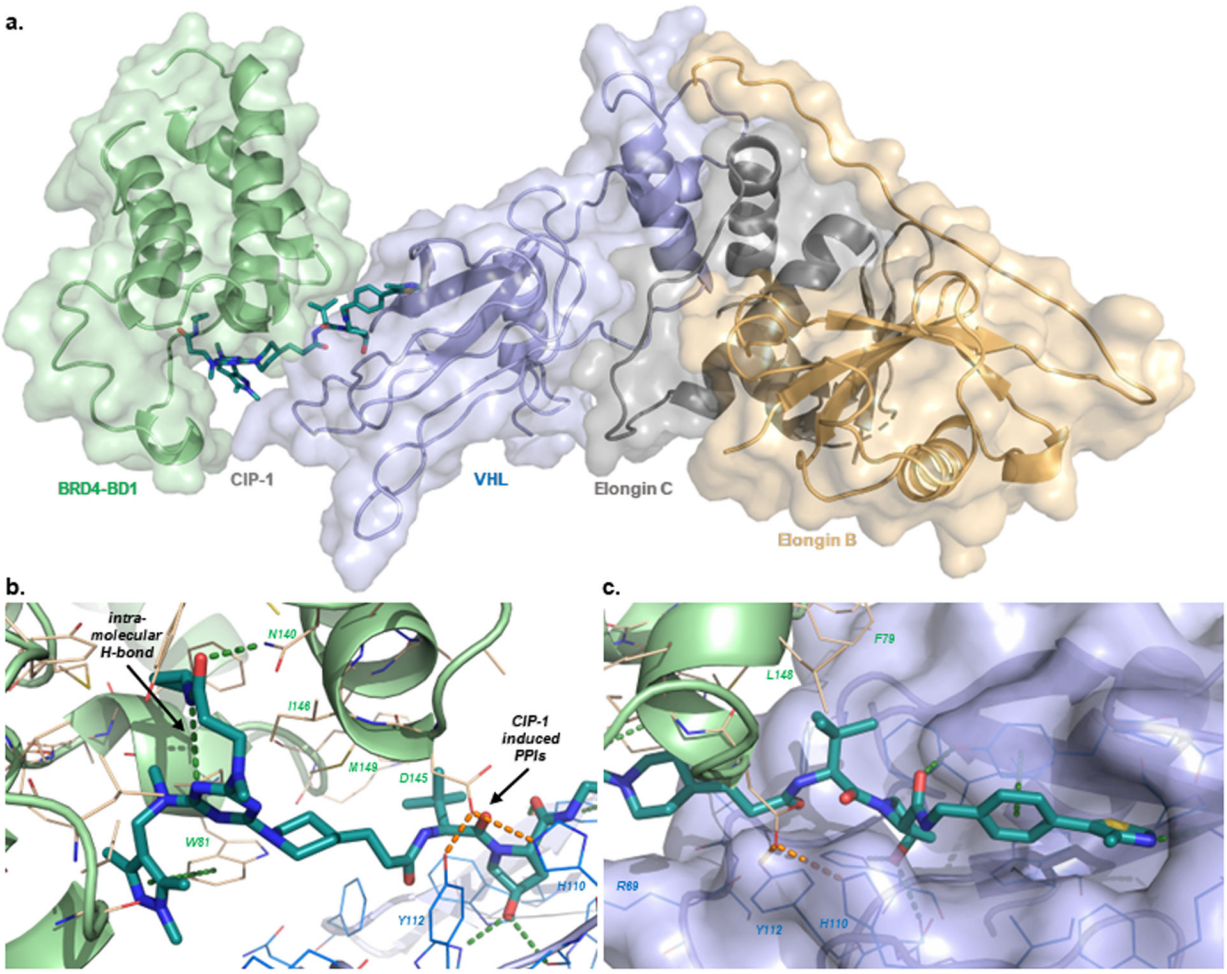


Figure 4.

Ternary complex crystal structure of CIP-1 with BRD4 and VHL-Elongin C-Elongin B. (a) Overall architecture of complex assembly. Protein components are shown as ribbons with semi-transparent surfaces and labeled by protein color. CIP-1 is shown as sticks. (b) Key interactions of CIP-1 with $BRD4^{BD1}$. An amide carbonyl from CIP-1 forms an H-bond with Asn140 of $BRD4^{BD1}$ and the pyrazole substituent of CIP-1 forms a pi-stacking interaction with Trp81. An intramolecular H-bond between an amide NH and a triazine nitrogen of CIP-1 reinforces the bound conformation of the BRD4 binding component. The ternary complex is stabilized by two H-bonds formed between Asp145 of $BRD4^{BD1}$ and Tyr112 and His110 of VHL. (c) Key interactions of CIP-1 with VHL. The VHL binding component of CIP-1 maintains equivalent interactions with the E3 ligase as the parent VH032 ligand.

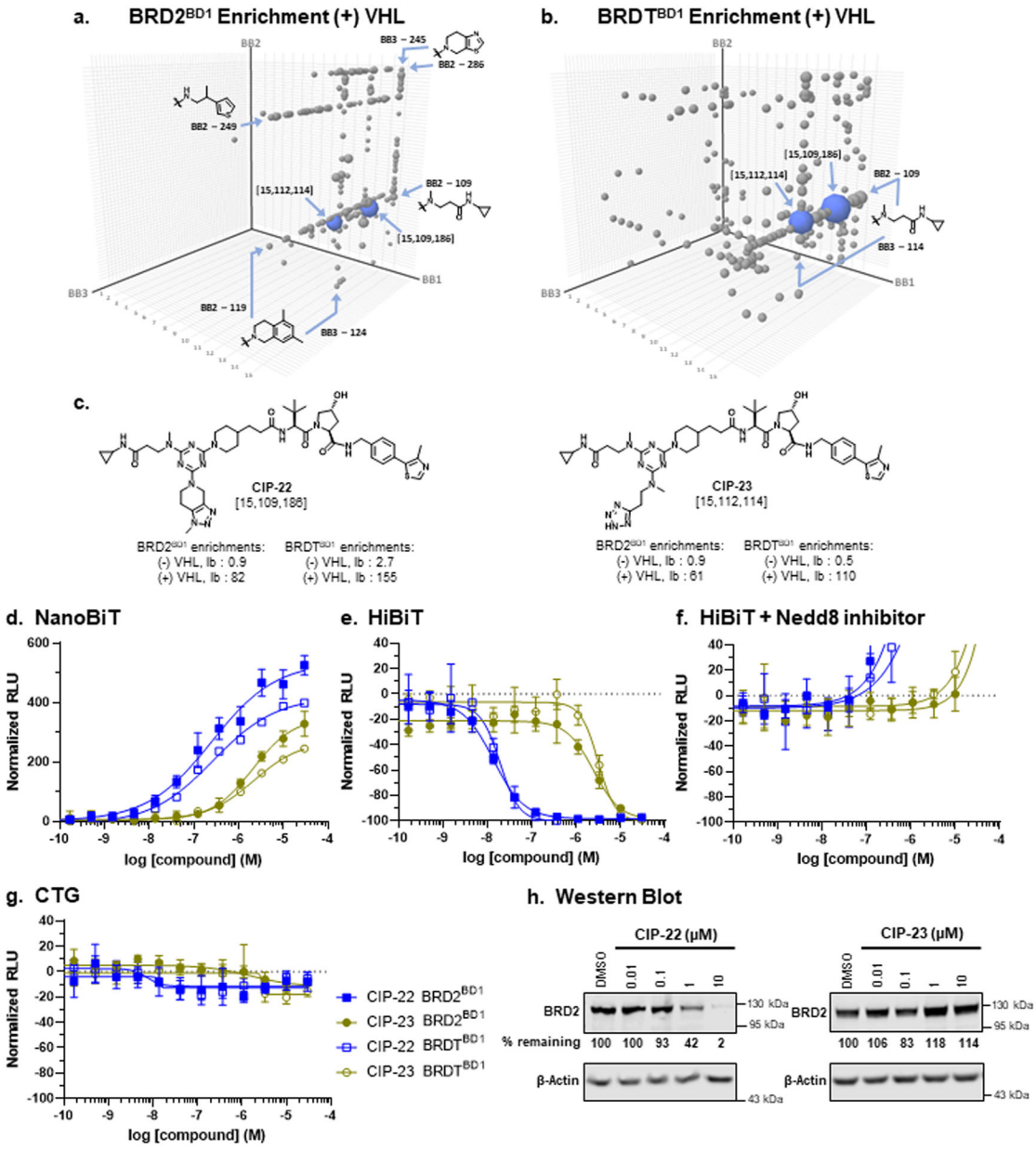


Figure 5. CIP-DEL screening against BRD2^{BD1} and BRDT^{BD1} with off-DNA profiling of two shared hits.

(a) Selection analysis of the BRD2^{BD1} (+) VHL screen (enrichment cutoff ≥ 7), and (**b**) the BRDT^{BD1} (+) VHL screen (enrichment cutoff ≥ 5). Points are sized by (+) VHL enrichment values and the two shared hits are labeled and colored blue. (**c**) Structures of hit compounds synthesized off-DNA with enrichment values observed in each screen. (**d**) Ternary complex induction between VHL and BRD2^{BD1} (filled) and BRDT^{BD1} (hollow) for **CIP-22** (blue) and **CIP-23** (green) measured using the NanoBiT assay. (**e**) CIP mediated degradation of BRD2^{BD1} and BRDT^{BD1} measured using the HiBiT assay. (**f**) The bromodomain degradation with CIPs is reversed in the presence of a Neddylation inhibitor. (**g**) Cell viability is not affected by the CIP compounds, shown by the CellTiter-Glo assay. Data in **d**

- **g** represent the mean ± s.d. of $n = 2$ independent experiments. (**h**) Western blot analysis of endogenous BRD2 degradation with CIP compounds ($n = 3$ independent experiments).