Structural basis of PROTAC cooperative recognition for selective protein degradation

Gadd, Morgan; Testa, Andrea; Lucas, Xavier; Chan, Kwok Ho; Chen, Wenzhang; Lamont, Douglas; Zengerle, Michael; Ciulli, Alessio

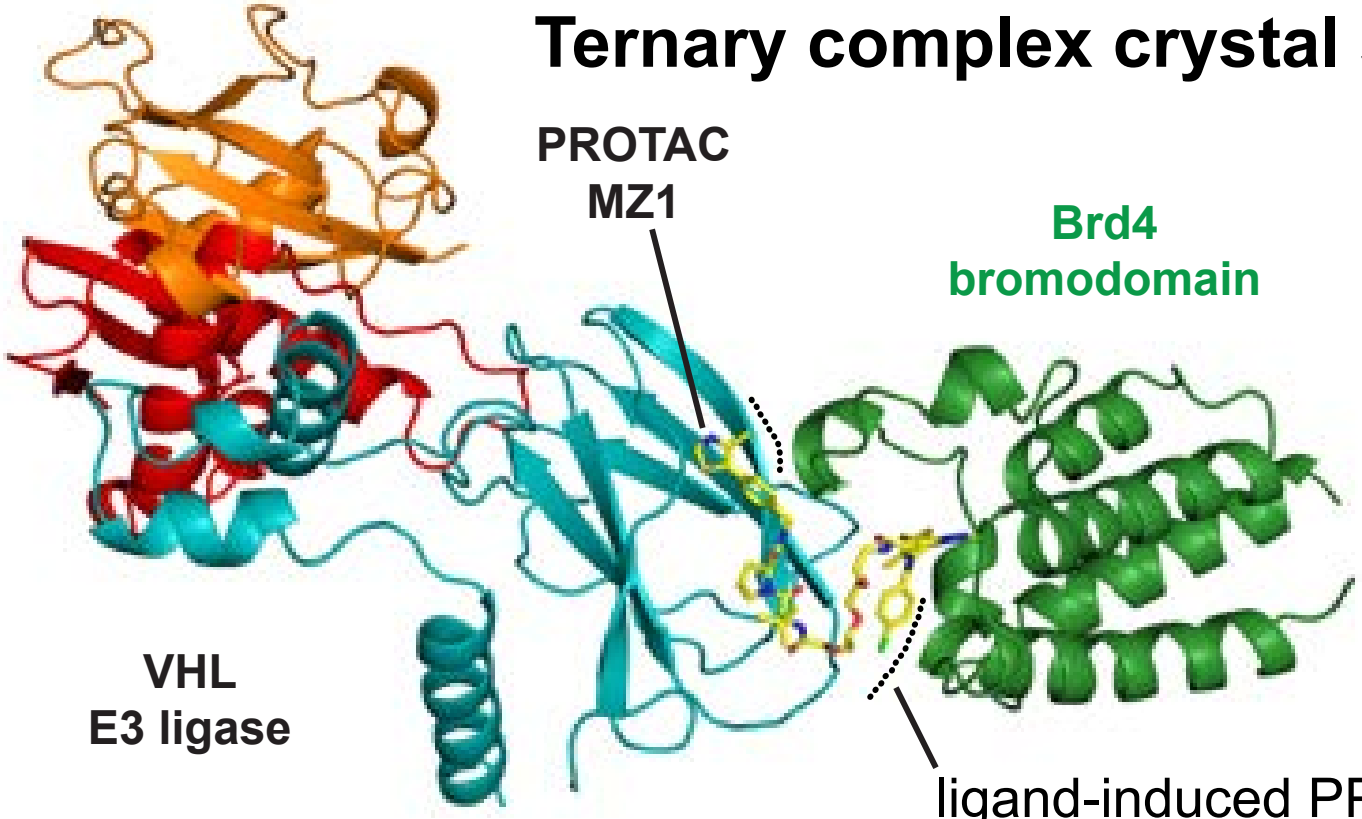
Published in:
Nature Chemical Biology

DOI:
10.1038/nchembio.2329

Publication date:
2017

Document Version
Peer reviewed version

Ternary complex crystal structure



- Cooperative and stable
- Highly populated
- **Selective degradation**

Title: Structural basis of PROTAC cooperative recognition for selective protein degradation

Authors: Morgan S. Gadd[1], Andrea Testa[1], Xavier Lucas[1], Kwok-Ho Chan, Wenzhang Chen, Douglas J. Lamont, Michael Zengerle, Alessio Ciulli[*]

Affiliation: Division of Biological Chemistry and Drug Discovery, School of Life Sciences, University of Dundee, Dow Street, Dundee, DD1 5EH, Scotland, UK.

[*]To whom correspondence should be addressed. E-mail: a.ciulli@dundee.ac.uk

[1] These authors contributed equally to this work

Abstract:

Inducing macromolecular interactions with small molecules to activate cellular signaling is a challenging goal. PROTACs (proteolysis-targeting chimaeras) are bifunctional molecules that recruit a target protein in proximity to an E3 ubiquitin ligase to trigger protein degradation. Structural elucidation of the key ternary species ligase:PROTAC:target and how this impacts target degradation selectivity remains elusive. We solved the crystal structure of Brd4-degrader MZ1 in complex with human VHL and the Brd4 bromodomain ($Brd4^{BD2}$). The ligand folds into itself to allow formation of specific intermolecular interactions in the ternary complex. Isothermal titration calorimetry studies, supported by surface mutagenesis and proximity assays, are consistent with pronounced cooperative formation of ternary complexes with $Brd4^{BD2}$. Structure-based designed compound AT1 exhibits highly selective depletion of Brd4 in cells. Our results elucidate how PROTAC-induced *de novo* contacts dictate preferential recruitment of a target protein into stable and cooperative complex with an E3 ligase for selective degradation.

Introduction

Regulating protein function through targeted degradation as opposed to more conventional target inhibition has emerged as a new modality of discovery chemistry with attractive potential both as tools for target validation and for the development of novel therapeutics [1-4]. PROTACs (proteolysis-targeting chimaeras) are bifunctional molecules that bring a target protein into spatial proximity with an E3 ubiquitin ligase to trigger target ubiquitination and subsequent proteasomal degradation [5-11]. Recent developments in the field have led to PROTACs being designed with increasingly "drug-like" molecular properties, and remarkable activities both in cells and *in vivo* [10]. Furthermore, we and others have shown that target depletion selectivity by PROTACs can significantly exceed the binding selectivity of their constitutive warhead ligands [7,12]. This realization provides proof-of-concept for turning non-selective or promiscuous ligands into more selective degraders, which can be highly desirable for both chemical probes and drug leads. A characteristic feature of PROTACs mode of action is their sub-stoichiometric catalytic action that alleviates the requirement for target engagement and occupancy of traditional inhibitors [9]. Effective redirection of a ligase poly-ubiquitination activity toward a new substrate protein requires formation of a ternary complex ligase:PROTAC:target, an intermediate species that is crucial to the cellular activity of degrader molecules. However, structural elucidation of such ternary species and how it may influence selectivity of target degradation remain elusive since the PROTAC concept was first incepted and demonstrated in 2001 (ref. [5]).

In 2015 we and others reported the first examples of small-molecule PROTACs that target the bromo- and extra-terminal (BET) family proteins for degradation by recruiting substrate-recognition subunits von Hippel–Lindau protein (VHL) [7] and cereblon (CRBN) [6,8] of the respective cullin RING ligases (CRLs), $CRL2^{VHL}$ and $CRL4^{CRBN}$. Compound MZ1 (ref. [7]) conjugates the pan-BET inhibitor JQ1 (ref. [13]) to VH032, a potent and specific VHL ligand [14,15], via a 3-unit PEG linker (**Fig. 1a**). MZ1 and its analogues, MZ2 and MZ3 (**Supplementary Results**, **Supplementary Fig. 1**), induce more effective depletion of a single BET member, Brd4 (a validated drug target against cancer and other diseases [16]), over its family paralogues Brd2 and Brd3 (ref. [7]). These observations led us to hypothesize a structural basis for target selectivity, imparted as a result of PROTAC-induced recruitment of the ligase and bromodomain together in a ternary complex. To reveal the molecular details of complex formation, we pursued the crystal structure of MZ1 in complex with VHL and a BET bromodomain. Here we present for the first time a structure of a PROTAC bound to both E3 ligase and target protein. The structure reveals MZ1 is "sandwiched" between the two proteins, inducing extensive new protein-protein and protein-ligand contacts of both hydrophobic and electrostatic nature. Biophysical binding studies in solution allowed measurement of full thermodynamics parameters of complex formation, which revealed marked isoform-specific cooperativity of ternary complexes. Surface mutagenesis swap and proximity binding assays data support the induced PPI contacts drive specificity of the cooperative recognition, impacting on the relative population of ternary complexes. Furthermore, new PROTAC molecules designed guided by the crystal structure showed exquisite selectivity for inducing cellular depletion of Brd4 over its BET family members Brd2 and Brd3.

Results

Ternary complex crystal structure. To elucidate the structural details of PROTAC-induced substrate recruitment to an E3 ligase, we solved the crystal structure of MZ1 bound in a

ternary complex with the second bromodomain (BD) of Brd4 ($Brd4^{BD2}$) and VHL to 2.7 Å resolution (**Fig. 1a, Supplementary Table 1**). The asymmetric unit of the crystal contained two ternary $Brd4^{BD2}$:MZ1:VCB (VHL, ElonginB and ElonginC) complexes of overall identical quaternary architecture (**Supplementary Fig. 2a**) and only minor deviations at either end when superposed over the central VHL subunit (**Supplementary Fig. 2b,c**). The first complex (chains A, B, C and D) had lower average *B* factors (**Supplementary Fig. 2a**) so we refer to this in all subsequent analyses. The electron density around MZ1 was fully defined (see inset panel in **Fig 1a** and **Supplementary Fig. 2d,e** for each protomer in the asymmetric unit). MZ1 is bound within a bowl-shaped interface formed by extensive protein-protein interactions (PPIs) between $Brd4^{BD2}$ and VHL (**Supplementary Fig. 3a**). The bowl has a hydrophobic "base" which is formed by two key points of contact (**Fig. 1b** and **Supplementary Fig. 3a**). Firstly, Trp374 from the characteristic hydrophobic region named "WPF shelf" (ref. [17]) of $Brd4^{BD2}$ interacts with residues Arg69, Pro71 and Tyr112 of VHL (**Fig. 1b**). Pro71 provides an additional stack to the WPF, forming an extended "PWPF" shelf (**Fig. 1b**). Secondly, Ala384 and Leu385 from the second helical turn of the ZA loop of $Brd4^{BD2}$ contact the hydrophobic side chains of Arg108, Ile109 and His110 in β4 of VHL (**Fig. 1b**). Two electrostatic "arms" complete the rim of the bowl. At one end, Asp381 and Glu383 in the ZA loop of $Brd4^{BD2}$ form a tight zipper structure of complementary charges with Arg107 and Arg108 (**Fig. 1c**). At the opposite end $Brd4^{BD2}$ residue Glu438, residing in the BC loop, contacts Arg69 from VHL (**Fig. 1d**). In the induced interface between the two proteins, $Brd4^{BD2}$ recapitulates some of the interactions made by the HIF-1α CODD segment peptide with VHL [18,19], as the electrostatic zipper structure generated by Asp381 and Glu383 contacts the same VHL residues (Arg107 and Arg108) as Asp569 and Asp571 of HIF-1α (**Fig. 1c** and **Supplementary Fig. 3b,c**). In contrast, VHL does not contact the surface of Brd4 bromodomain bound to acetyllysine histone H4 peptide [20] (**Supplementary Fig. 3d,e**). In total the PPIs induced in the complex bury a surface area of 688 $Å^2$ (**Supplementary Table 2**).

MZ1 is cupped within the bowl structure in such a way that its two heads recapitulate the binding modes of the respective ligands individually – JQ1 in the acetyllysine-binding pocket of $Brd4^{BD2}$ [13,20], and VH032 in the hydroxyproline-binding site of VHL [14] (**Supplementary Fig. 3f,g**). In addition to the expected binary protein-ligand interactions, MZ1 forms additional protein-ligand interactions within the ternary complex. The PEG linker makes van der Waals interactions with the BC loop of $Brd4^{BD2}$ and a hydrogen bond between the ether oxygen adjacent to the amide linkage to JQ1 and the BD2-specific residue His437 (**Fig. 1d**). The same hydrogen bond between His437 and a PEG oxygen is also observed in a recent crystal structure of $Brd4^{BD2}$ bound to MT1, a bivalent BET inhibitor comprised of two JQ1 moieties linked by a PEG unit in a similar fashion as in MZ1 (**Supplementary Fig. 4**) [21]. VHL and $Brd4^{BD2}$ come together to sandwich MZ1 against their respective binding surfaces, burying otherwise solvent-exposed regions of the JQ1 and VH032 ligands. Specifically, Ala384, Leu385 and Gly386 from the ZA loop of $Brd4^{BD2}$ contact the $–CH_2–$phenyl portion of VH032, whilst His110 and Tyr112 of β4 of VHL contact one of the two thiophene methyl groups and the *para*-chlorophenyl ring of JQ1 (**Fig. 1b**). In addition, MZ1 folds onto itself in such a way that its PEG linker is packed between the *tert*-butyl group of the VH032 moiety and the *para*-chlorophenyl ring of JQ1 (**Fig. 1a,d**). Throughout a 100 ns molecular dynamics simulation, favorable intermolecular contacts were observed from JQ1 atoms to VHL, from VH032 atoms to $Brd4^{BD2}$, and from the PEG linker atoms to both VHL and $Brd4^{BD2}$, as well as intramolecular contacts within MZ1 involving the PEG linker and the JQ1 and VH032 moieties (**Fig. 2a** and **Supplementary Fig. 5)**. The surface area further buried by the ligand folding within the bowl-shaped interface was 1,933 $Å^2$, resulting in a total extended buried

surface area of 2,621 Å^2 for the ternary complex (see **Supplementary Table 2** for a comparison with cereblon:phthalimides:target ternary complexes [2,3]).

Isoform-specific cooperativity of ternary complexes. The extensive new contacts observed in the crystal structures suggested the possibility that isoform-specific PPIs could play a role in the "cooperativity" of the ternary complex equilibria [22,23]. To assess the thermodynamics of PROTAC-induced complex formation with VHL and BET BDs, we employed isothermal titration calorimetry (ITC) in solution (**Table 1** and **Supplementary Fig. 6**). To disentangle contributions from binary and ternary complex formation equilibria we performed reverse titrations i.e. protein in syringe and ligand in cell. This experimental strategy avoids the characteristic "hook effect" observed with increasing concentrations of PROTACs, as formation of binary complexes competes with and eventually surpasses formation of ternary ones [23]. First, we titrated a solution of BET BD against MZ1, ensuring no excess unbound PROTAC compound would be present at the end of the titration. This was followed by a titration of VCB into the saturated MZ1:BD complex, forming the ternary VCB:MZ1:BD complex (**Fig. 2b**, right panel). Titration of VCB into MZ1 alone (**Fig. 2b**, left panel) was then performed and used as reference, allowing potential cooperativity of ternary complex formation to be accurately quantified. By definition, a ternary system is considered *positively* cooperative if interactions enhance formation of the ternary complex ($\alpha = \frac{K_d\ (Binary)}{K_d\ (Ternary)} > 1$; $\Delta pK_d = pK_d^T$ (*ternary*) – pK_d (*binary*) >0). Conversely, a system is termed *negatively* cooperative when formation of the ternary complex is diminished ($\alpha < 1$; $\Delta pK_d < 0$), for example because of repulsive interactions or steric hindrance between the two components in the ternary complex. Non-cooperative equilibria would instead show unchanged K_d for the two steps ($\alpha = 1$; $\Delta pK_d = 0$), suggesting no interactions (**Fig. 2c**). With all BET BDs used, we observed significant positively cooperative ternary complex formation (see **Table 1**, and plots of ΔpK_d in **Fig. 2d**). Strikingly, the strongest cooperativity was observed for Brd4^{BD2} (α = 18), followed by Brd3^{BD2} (α = 11; **Table 1** and **Fig. 2d**). The large cooperativity observed led to steep transition of the sigmoidal binding curve (**Fig. 2b**, right panel), suggesting that the fitted K_d value could potentially be underestimating the real binding affinity of this ternary complex. All BD1s also exhibited positive cooperativity, albeit to a much lesser extent (α between 2 and 3). Brd4^{BD2} and Brd3^{BD2} not only exhibited the greatest cooperativity amongst all BET BDs, they also formed the most stable ternary complexes overall (ΔG (*binary* + *ternary*) = –22.2 ± 0.1 and 22.0 ± 0.2 kcal/mol, respectively), ~2 kcal/mol more stable than e.g. Brd2^{BD1} (ΔG = –20.3 ± 0.2 kcal/mol).

To understand the impact of cooperativities of different BET BDs ternary complexes on their relative population, we applied a mathematical model of ternary equilibrium [23]. We simulated the fraction of ternary complex formation for VCB, MZ1 and six individual BET BDs, using our measured binary K_ds and cooperativities α (**Table 1**) and protein concentrations of 40 nM (to be around the K_d values and to match the concentrations used later in AlphaLISA). Overlay of simulations showed that the relative populations of each ternary complex vary significantly, with Brd4^{BD2} being the most populated, and, as an example, ~2.5 fold greater than Brd2^{BD1} at any given concentration of PROTAC (**Fig. 2e**). To interrogate this relative trend experimentally, we employed a proximity AlphaLISA assay that can achieve high signal amplification in response to formation of ternary complexes over an energy transfer distance of up to 200 nm [24,25]. At every fixed component concentration, the relative trend observed in AlphaLISA signal was broadly consistent with the cooperativity trends measured by ITC, with Brd4^{BD2} and Brd3^{BD2} giving greater signal, while Brd2^{BD1} giving the lowest response (**Fig. 2f**). A similar trend was observed with the analogous MZ2 (PEG_4), MZ3

(PEG_3-Phe) and MZ4 (PEG_2) (**Supplementary Fig. 7a–h**). Taken together the data are consistent with target-specific cooperativities and stabilities of ternary complexes impacting on the relative population of this key intermediate species.

Specificity of MZ1-induced protein-protein interactions. To evaluate to what extent the cooperativity of ternary complex formation is dictated by surface complementarity between VHL and the $Brd4^{BD2}$ bromodomain, we mutated semi-conserved or non-conserved bromodomain residues forming key induced PPI contacts, but not directly involved in binding of MZ1 (**Fig. 3a** and **Supplementary Fig. 5**). Inspection of sequence alignments (**Fig. 3a**) and the crystal structure (**Fig. 3b**) guided us to select residues Glu383 and Ala384 in $Brd4^{BD2}$ (the most cooperative of bromodomains), for site directed mutagenesis. These residues are Val106 and Lys107 in the corresponding positions in $Brd2^{BD1}$, one of the least cooperative BET domains. In addition, the MD simulation evidenced extensive movement of loop 7 of VHL (ref. [26]) bringing it in close contact with Lys378 of $Brd4^{BD2}$ (**Supplementary Fig. 5b–d**), which corresponds to Gln101 in $Brd2^{BD1}$. Based on these considerations, triple mutant $Brd4^{BD2}$ K378Q/E383V/A384K (named QVK for simplicity) was designed. The mutations would make the PPI surface of $Brd4^{BD2}$ closer to that of $Brd2^{BD1}$, albeit with the caveat of introducing an extra charge overall. Conversely, a triple mutant of $Brd2^{BD1}$ was designed in which the corresponding residues are switched to those of $Brd4^{BD2}$ (Q101K/V106E/K107A, named KEA for simplicity). In ITC, the QVK mutant exhibited significantly weakened cooperative complex formation relative to $Brd4^{BD2}$ wild-type (WT) (α = 4; **Table 1** and $\Delta p K_d = 0.64 \pm 0.04$, see plots in **Fig. 3c**). Conversely, the cooperativity of the KEA mutant increased relative to $Brd2^{BD1}$ WT (α = 8; **Table 1** and **Fig. 3c**). Crucially, these mutations did not affect the K_d of each domain for MZ1 (**Table 1**), suggesting that the switch of cooperativity is independent of binary target engagement and instead is dictated by the induced PPIs. Consistent with the cooperativity switch measured by ITC, QVK reduced ternary complex formation in AlphaLISA compared to WT, whereas KEA displayed the opposite effect (**Fig. 3d** with MZ1 and **Supplementary Fig. 7i** with MZ2). Taken together these data validate the ternary complex structure in solution and elucidate how specific PPIs influence cooperative recruitment of two target proteins to each other by a bifunctional molecule.

Structure-designed AT1 is highly selective Brd4 degrader. In chemical probe and drug development knowledge of ligand-bound structures can guide the design of next-generation compounds. We therefore sought to create new PROTACs based on our crystal structure that could exhibit enhanced target depletion selectivity in cells toward Brd4. We noted that the side chain of the key *tert*-Leu group of VH032 projected an attractive vector to link directly to the JQ1 moiety (**Fig. 4a**), which we hypothesized could better discriminate against the relative binding orientation observed in the crystal. We therefore replaced *tert*-Leu with penicillamine and synthesized **1** (AT1, **Fig. 4b**) and other analogues **2–6** (AT2–AT6) bearing thioether linkages of varying length to JQ1 (**Supplementary Fig. 1** and **Methods**). The modified VHL ligand within AT1 retained binding to VHL (K_d 330 nM, **Supplementary Table 3**, and **Supplementary Fig. 8**), a less than two-fold loss of potency relative to VH032 (ref. [14]) but approximately five-fold less compared to MZ1. ITC data comparing binary and ternary complexes revealed $Brd4^{BD2}$ as the BET bromodomain forming the most cooperative (α = 7; $\Delta p K_d = 0.84 \pm 0.07$) and most stable ($\Delta G = -20.2 \pm 0.2$ kcal/mol) of all ternary complexes with AT1 (**Fig. 4c**, **Supplementary Table 3**, and **Supplementary Fig. 8**). The same trend of cooperativity observed for MZ1 and the QVK and KEA mutants relative to WT was also observed with AT1 (**Fig. 4c**, **Supplementary Table 3**), suggesting AT1 recruits VHL and $Brd4^{BD2}$ in the same relative orientation as does MZ1. We consistently observed by

AlphaLISA preferential recruitment of Brd4^{BD2} over the other BDs by AT1–6 (**Fig. 4d** and **Supplementary Fig. 9**). We next tested the activity of the new structure-designed molecules to induce degradation of BET proteins in cells, and observed remarkable Brd4-selective depletion at all concentrations tested, with depletion of Brd4 after 24 h treatment with 1–3 μM of AT1 and negligible activity against Brd2 and Brd3 (**Fig. 4e,** see **Supplementary Fig. 9** for AT2–6, and full uncut gel images in **Supplementary Fig. 11**). Specificity for Brd4 degradation was not due to differences in protein synthesis rates, as shown by control treatments with cycloheximide, which blocks protein translation (**Supplementary Fig. 10**). To assess whether ubiquitination of lysine residues could play a role in the observed selectivity, we combined our EloBC–VHL–MZ1–Brd4^{BD2} complex with existing whole CRL structural information into a model of the entire CRL2^{VHL}–MZ1–bromodomain assembly (**Supplementary Fig. 12a**). Several lysine residues are surface exposed and accessible to the E2-Ub in this model at distances between 50–60 Å, consistent with known CRL substrates [27]. Mapping MZ1-induced ubiquitination sites *in vitro* identified Lys346 on Brd4^{BD2} and several sites on the other BET-BDs (**Supplementary Fig. 12b-d**). Unbiased and quantitative isobaric tagging mass spectrometry proteomics confirmed Brd4 as the sole protein markedly depleted (to ~40%) upon treatment with AT1, amongst the 5,674 detected proteins that passed filtering criteria (**Fig. 4f** and **Supplementary Data Set 1**). Crucially, no effect on protein levels of Brd2 and Brd3 was observed with AT1, in contrast with MZ1 that exerted a broader and more profound effect across all BET proteins, albeit still preferential toward Brd4 (**Fig. 4g**). Together these data qualify AT1 as a new highly selective degrader of Brd4 in cells.

Discussion

We put forth a model for how cooperative recruitment of a target close to an E3 ligase in a ternary complex by a PROTAC molecule can impact on the effectiveness and selectivity of target degradation (**Fig. 5**). Our work shines structural insights into how bifunctional molecules can induce target-specific interactions in the 'enzyme-substrate' ternary complex species key for PROTAC catalytic activity. These cooperative molecular recognition features contribute to how tightly and stably the '*neo*-substrate' can be bound to the ligase, impacting on the relative population of the complex and consequently on the catalytic efficiency of the process. For homologous targets as is the case with BET bromodomains, we show how these features add a level of target depletion selectivity independently of binary target engagement.

We illustrate the relevance of measuring cooperativities of ternary complex formation in solution using ITC. One of the main advantages of the assay set-up as described here is that it is designed to avoid issues associated with the hook effect. While the assay may be used to characterize in full a handful of compounds and systems, it requires large quantities of material and notably lacks throughput. Alternative bioassays to quantify formation of ternary complexes could circumvent this limitation, and we provide evidence that proximity AlphaLISA assay can be used for these purposes. However, AlphaLISA data should be interpreted with caution, even when comparing highly conserved domains as done here, because in this bead-based technology the multiplicity of binding sites and relative linkage and orientation of components immobilized to the beads may influence the measured signal [24,28]. Moreover, it can be difficult to deconvolute individual binding parameters from dose–response curves monitoring ternary complex formation, because these are often bell–shaped curves complicated by the hook effect [23]. We anticipate that future assay developments in this direction will help prioritize complexes for structural studies and to drive drug development programs. The results of our study are of particular relevance because VHL-based BET

degraders similar to the ones described here have proven to be bioavailable and active *in vivo* [29] and could potentially enter clinical trials as early as 2017 (ref. [4]).

Our study points to the importance that the bifunctional molecule "folds in itself" in such a way that its two heads can recruit the respective targets into productive proximity. The result of this process is the burial of extensive surface area, the formation of new PPIs, contributing to the high stability and cooperativity of ternary complexes. Our work has therefore major implications for future PROTAC drug design, which has so far been somewhat empirical and has largely adopted a combinatorial "plug-and-play" strategy [11]. The proposed model suggests that bifunctional molecules should be rationally designed and prioritized based on their ability to induce favorable contacts and allow forming a stable complex between the E3 ligase and the target. While ligand-induced proximity is expected to strengthen potential PPIs because of the reduced entropic cost, we show that the level of surface complementarity between the two proteins in their relative orientation imposed by the bifunctional molecule dictates cooperative complexation. While the exact relative orientation between VHL and Brd4 observed in the crystal may not be the only one that the system can adopt in solution once free from potential constraints of crystal packing, our data suggests that it captures a significant species underlying MZ1 function. Maximizing the diversity of E3s recruited [30], and linking positions and vectors from the E3 and target ligands, will thus be important to achieve target-specific degraders.

In an example of first layer of this rational design, we show how new PROTACs designed based on our ternary structure can lead to enhanced selectivity of depletion in cells for the crystallized target Brd4. The efficiency and selectivity of cellular protein knockdown will inevitably depend on other factors, including compound permeability and stability, the expression level of the hijacked CRL and its relative activity and flexibility [27,31], as well as target abundance and re-synthesis rates. Differing ubiquitination rates could also in principle influence target degradation selectivity. To this end, in addition to increasing their relative population, cooperative and stable complexes would be expected to exhibit slower dissociation rates and longer half-lives, potentially aiding the efficiency of target ubiquitination by the hijacked ligase. Differing availability and access of surface lysine residues between alternate substrates could also play a role. However, based on our data, the presence of many surface Lys on BET-BDs, and the flexibility and large ubiquitination zone of CRLs [2,27], we view it unlikely that target ubiquitination plays a role in the observed selectivity of Brd4 degradation. Obtaining a more detailed biochemical picture of target ubiquitination in a cellular context will be of clear importance for future investigation.

For targeted protein degradation, converting a pan-selective or promiscuous probe ligand into a more selective degrader probe provides new opportunities to improve target validation and could minimize off-target effects. In addition to dictate selectivity of target degradation, highly cooperative ternary E3:PROTAC:target systems would be anticipated to unlock the possibility to effectively degrade hitherto "undruggable" targets using ligands with inherently weak binary binding affinities. A more general implication of this study is the feasibility to induce *de novo* protein-protein interactions, or stabilize weakened ones, using bifunctional small molecules, a feature previously established with mono-functional 'molecular glues' [32] such as the plant hormones auxin [33] and jasmonate [34], the phthalimide immunomodulatory drugs (IMiDs) [2,3,31,35-38], and macrocyclic natural products such as rapamycin and cyclosporine [39,40]. We envision that extensions of PPI-stabilizing capabilities to hetero- or homo-bifunctional small molecules [21,41,42] beyond PROTACs as highlighted here could

expand the target spectrum accessible to PPI stabilizers, and provide a new paradigm of selective chemical intervention for structural chemical biology and drug discovery.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes

Atomic coordinates and structure factors for hsBrd4^{BD2}–MZ1–hsVHL–hsEloC–hsEloB have been deposited in the Protein Data Bank (PDB) under accession number 5T35.

Data availability

Any supplementary information, chemical compound information and source data are available in the online version of the paper. Correspondence and requests for materials should be addressed to A.C. (a.ciulli@dundee.ac.uk).

Acknowledgements

This work was supported by the European Research Council (ERC-2012-StG-311460 DrugE3CRLs Starting Grant to A.C.); the UK Biotechnology and Biological Sciences Research Council (BBSRC grant BB/J001201/2 to A.C.); the European Commission (H2020-MSCA-IF-2014-655516 Marie Skłodowska-Curie Actions Individual Fellowship to K-H.C., and H2020-MSCA-IF-2015-806323 Marie Skłodowska-Curie Actions Individual Fellowship to X.L.); and the Wellcome Trust (Strategic Awards 100476/Z/12/Z for biophysics and drug discovery and 094090/Z/10/Z for structural biology and X-ray crystallography to the Division of Biological Chemistry and Drug Discovery). We are thankful to P. Fyfe for support with the in-house X-ray facility; L. Finn for support with tissue culture facility (MRC-PPU); the Ferguson lab for access to LI-COR equipment; T. Cardote for the gift of full-length Cul2-Rbx1 and A. Knebel (MRC-PPU/DSTT) for the gift of E1 and E2 enzymes; the Division of Computational Biology for support with computational cluster; and to Diamond Light Source for beamtime (BAG proposal MX10071) and beamline support at beamline I04-1.

Author Contributions

A.C. conceived the idea and directed the project.
M.S.G., X.L., A.T., K.-H.C. and A.C. designed the experiments and interpreted results
M.S.G., X.L., A.T., and K.-H.C. performed experiments
A.T. and M.Z. contributed to compound design and synthesized compounds
W.C. performed MS proteomics experiments under the supervision of D.J.L.
M.S.G., X.L. and A.C. wrote the manuscript with input from all other authors.

Competing financial interests

The authors declare no competing financial interests.

References

1. Huang, X. & Dixit, V. M. Drugging the undruggables: exploring the ubiquitin system for drug development. *Cell Res* **26,** 484–498 (2016).
2. Petzold, G., Fischer, E. S. & Thomä, N. H. Structural basis of lenalidomide-induced

CK1α degradation by the CRL4(CRBN) ubiquitin ligase. *Nature* **532,** 127–130 (2016).
3. Matyskiela, M. E. *et al.* A novel cereblon modulator recruits GSPT1 to the CRL4(CRBN) ubiquitin ligase. *Nature* **535,** 252–257 (2016).
4. Lai, A. C. & Crews, C. M. Induced protein degradation: an emerging drug discovery paradigm. *Nat Rev Drug Discov* (2016). doi:10.1038/nrd.2016.211
5. Sakamoto, K. M. *et al.* Protacs: chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation. *P Natl Acad Sci Usa* **98,** 8554–8559 (2001).
6. Winter, G. E. *et al.* Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science* **348,** 1376–1381 (2015).
7. Zengerle, M., Chan, K.-H. & Ciulli, A. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. *ACS Chem Biol* **10,** 1770–1777 (2015).
8. Lu, J. *et al.* Hijacking the E3 Ubiquitin Ligase Cereblon to Efficiently Target BRD4. *Chem Biol* **22,** 755–763 (2015).
9. Bondeson, D. P. *et al.* Catalytic in vivo protein knockdown by small-molecule PROTACs. *Nat Chem Biol* **11,** 611–617 (2015).
10. Deshaies, R. J. Protein degradation: Prime time for PROTACs. *Nat Chem Biol* **11,** 634–635 (2015).
11. Toure, M. & Crews, C. M. Small-Molecule PROTACS: New Approaches to Protein Degradation. *Angew Chem Int Ed Engl* **55,** 1966–1973 (2016).
12. Lai, A. C. *et al.* Modular PROTAC Design for the Degradation of Oncogenic BCR-ABL. *Angew Chem Int Ed Engl* **55,** 807–810 (2016).
13. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468,** 1067–1073 (2010).
14. 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. *J Med Chem* **57,** 8657–8663 (2014).
15. Frost, J. *et al.* Potent and selective chemical probe of hypoxic signalling downstream of HIF-α hydroxylation via VHL inhibition. *Nat Commun* **7,** 13312 (2016).
16. Zuber, J. *et al.* RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **478,** 524–528 (2011).
17. Nicodeme, E. *et al.* Suppression of inflammation by a synthetic histone mimic. *Nature* **468,** 1119–1123 (2010).
18. Hon, W.-C. *et al.* Structural basis for the recognition of hydroxyproline in HIF-1 alpha by pVHL. *Nature* **417,** 975–978 (2002).
19. Min, J.-H. *et al.* Structure of an HIF-1alpha -pVHL complex: hydroxyproline recognition in signaling. *Science* **296,** 1886–1889 (2002).
20. Filippakopoulos, P. *et al.* Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* **149,** 214–231 (2012).
21. Tanaka, M. *et al.* Design and characterization of bivalent BET inhibitors. *Nat Chem Biol* **12,** 1089–1096 (2016).
22. Whitty, A. Cooperativity and biological complexity. *Nat Chem Biol* **4,** 435–439 (2008).
23. Douglass, E. F., Miller, C. J., Sparer, G., Shapiro, H. & Spiegel, D. A. A comprehensive mathematical model for three-body binding equilibria. *J Am Chem Soc* **135,** 6092–6099 (2013).
24. Eglen, R. M. *et al.* The use of AlphaScreen technology in HTS: current status. *Curr Chem Genomics* **1,** 2–10 (2008).
25. Roberts, J. M. & Bradner, J. E. A Bead-Based Proximity Assay for BRD4 Ligand

Discovery. *Curr Protoc Chem Biol* **7,** 263–278 (2015).
26. Stebbins, C. E., Kaelin, W. G. & Pavletich, N. P. Structure of the VHL-ElonginC-ElonginB complex: implications for VHL tumor suppressor function. *Science* **284,** 455–461 (1999).
27. Duda, D. M. *et al.* Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation. *Cell* **134,** 995–1006 (2008).
28. Zhou, M., Li, Q. & Wang, R. Current Experimental Methods for Characterizing Protein-Protein Interactions. *ChemMedChem* **11,** 738–756 (2016).
29. Raina, K. *et al.* PROTAC-induced BET protein degradation as a therapy for castration-resistant prostate cancer. *P Natl Acad Sci Usa* **113,** 7124–7129 (2016).
30. Bulatov, E. & Ciulli, A. Targeting Cullin-RING E3 ubiquitin ligases for drug discovery: structure, assembly and small-molecule modulation. *Biochem J* **467,** 365–386 (2015).
31. Fischer, E. S. *et al.* Structure of the DDB1-CRBN E3 ubiquitin ligase in complex with thalidomide. *Nature* **512,** 49–53 (2014).
32. Fischer, E. S., Park, E., Eck, M. J. & Thomä, N. H. SPLINTS: small-molecule protein ligand interface stabilizers. *Curr Opin Struct Biol* **37,** 115–122 (2016).
33. Tan, X. *et al.* Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* **446,** 640–645 (2007).
34. Sheard, L. B. *et al.* Jasmonate perception by inositol-phosphate-potentiated COI1-JAZ co-receptor. *Nature* **468,** 400–405 (2010).
35. Ito, T. *et al.* Identification of a primary target of thalidomide teratogenicity. *Science* **327,** 1345–1350 (2010).
36. Lu, G. *et al.* The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. *Science* **343,** 305–309 (2014).
37. Krönke, J. *et al.* Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. *Science* **343,** 301–305 (2014).
38. Chamberlain, P. P. *et al.* Structure of the human Cereblon-DDB1-lenalidomide complex reveals basis for responsiveness to thalidomide analogs. *Nat Struct Mol Biol* **21,** 803–809 (2014).
39. Pommier, Y. & Marchand, C. Interfacial inhibitors: targeting macromolecular complexes. *Nat Rev Drug Discov* **11,** 25–36 (2012).
40. Thiel, P., Kaiser, M. & Ottmann, C. Small-Molecule Stabilization of Protein-Protein Interactions: An Underestimated Concept in Drug Discovery? *Angew Chem Int Ed Engl* **51,** 2012–2018 (2012).
41. Illendula, A. *et al.* A small-molecule inhibitor of the aberrant transcription factor CBFβ-SMMHC delays leukemia in mice. *Science* **347,** 779–784 (2015).
42. Waring, M. J. *et al.* Potent and selective bivalent inhibitors of BET bromodomains. *Nat Chem Biol* **12,** 1097–1104 (2016).

Figure Legends

Figure 1. The crystal structure of the Brd4^{BD2}:MZ1:VHL-ElonginC-ElonginB complex. a, Overall structure of Brd4^{BD2}:MZ1:VHL-ElonginC-ElonginB in ribbon representation. Top middle, chemical structure of bifunctional PROTAC molecule MZ1. Top right, F_o−F_c omit map generated prior to ligand modelling contoured at 3.0σ around bound MZ1. **b**, Key residues forming the hydrophobic "base" of the induced Brd4^{BD2}:VHL interface. The "WPF" shelf of Brd4^{BD2} and extended "PWPF" stack are outlined in black. The JQ1 and VH032 elements of MZ1 are labelled in yellow. **c**, Electrostatic potential map showing the charged zipper contacts between Brd4^{BD2} residues D381 and E383 with VHL residues R107 and R108. **d**, Electrostatic potential map showing the interaction

between Brd4^{BD2} residue E438 with VHL residue R69. The hydrogen bond between H437 of Brd4^{BD2} and the PEG linker of MZ1 is also shown. Dashed lines indicate hydrogen bonds with shown distance in angstroms (Å).

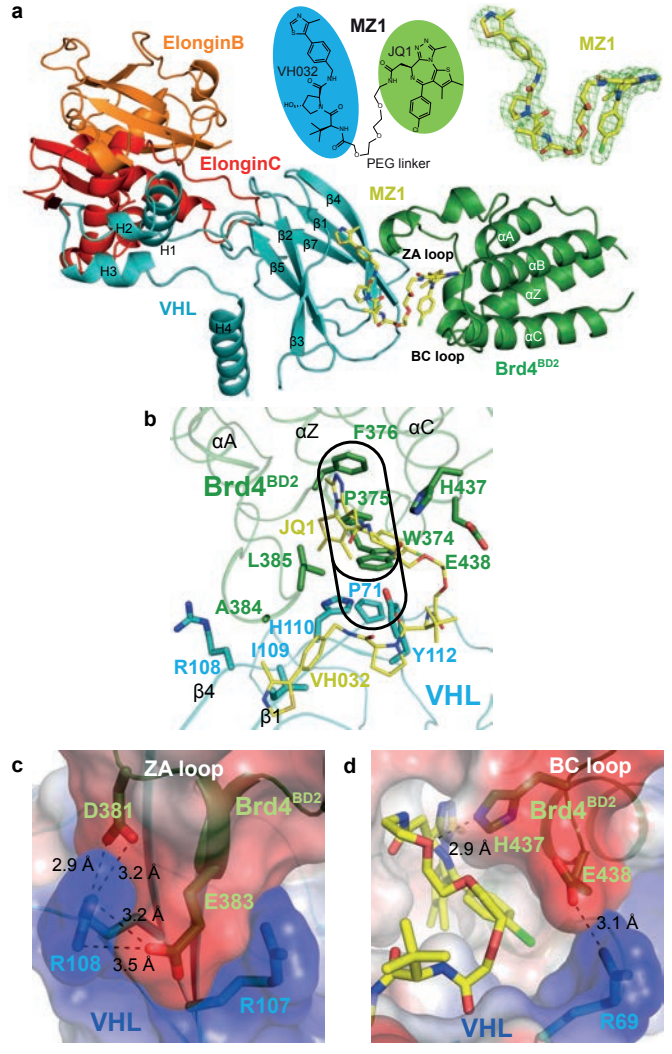
Figure 2. Brd4^{BD2} and VHL form a stable, cooperative complex in the presence of MZ1. a, Novel ligand contacts are induced by ternary complex formation. Colour strength (from white to red) indicates the mean enthalpic energies of individual MZ1 atoms in contacting Brd4^{BD2} (left) or VHL (right), as well as intra-ligand contacts within MZ1 (centre) in a 100 ns MD simulation. **b,** Inverse ITC titrations of VCB into MZ1 (left, representative of eight replicates) and VCB into the pre-formed MZ1:Brd4^{BD2} (right, performed in duplicate) **c**, Ternary complex equilibria and definition of cooperativities. **d**, ΔpK_d measured for VCB with MZ1 and the indicated BET-BDs, reported as difference (± uncertainty), from pK_d values measured as mean (± 1 s.e.m.) as described in Online Methods. Statistical significance of pK_d values for ternary titrations compared to the corresponding binary titrations was assessed by two-tailed *t*-test assuming equal variances, and is indicated as * (*p*-value < 0.05), ** (*p*-value < 0.01) or *** (*p*-value < 0.001). **e**, Simulated fraction of ternary complexes based on mathematical model described in ref. [23]. **f**, AlphaLISA intensity values titrating VCB against BET-BDs with MZ1. AlphaLISA intensities represent mean (± 1 s.d.) of intensity values from four technical replicates. The hook effect observed on these curves is due to biotinylated-VCB oversaturating the donor beads, resulting in a progressive decrease in signal.

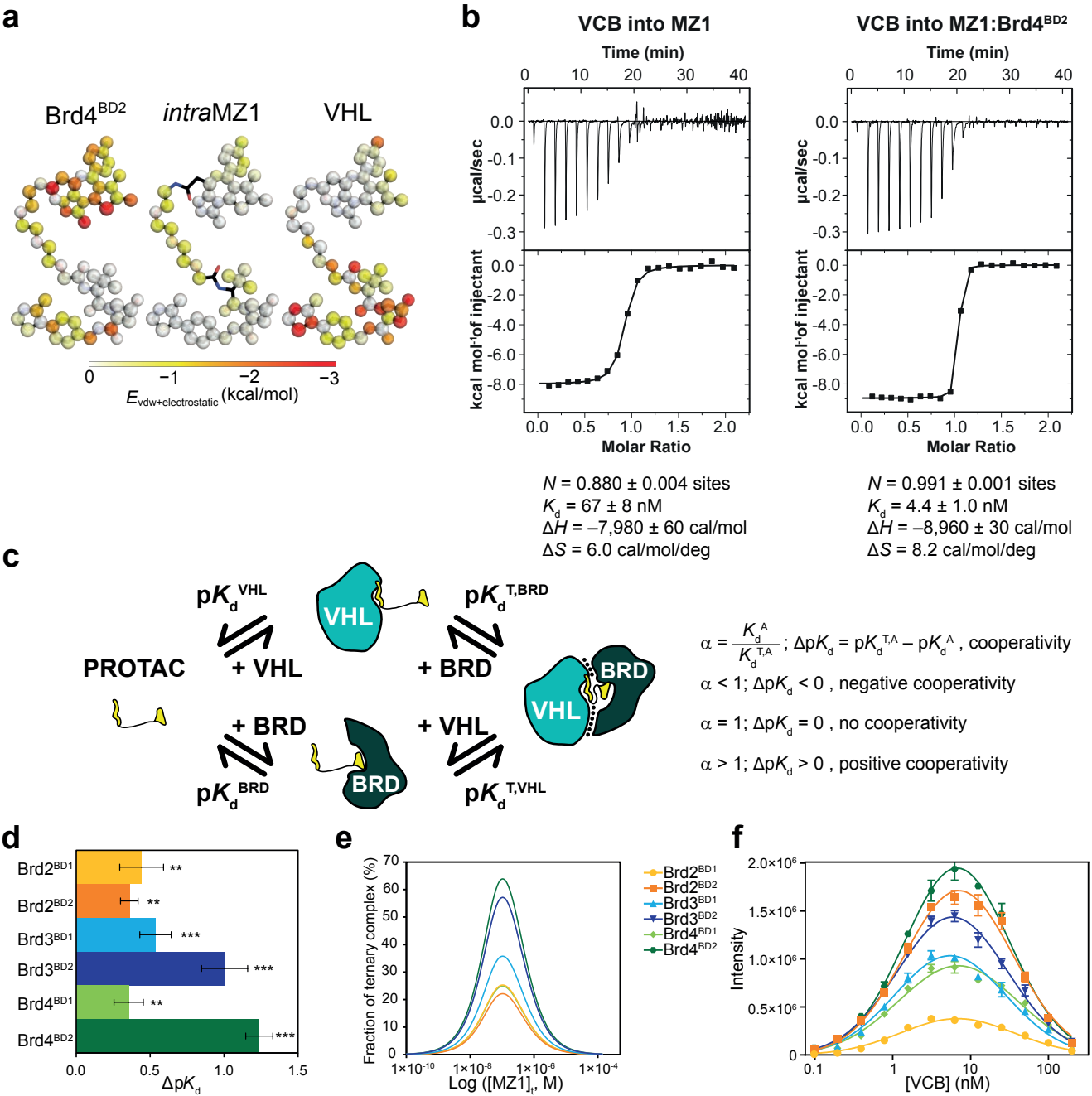
Figure 3. The molecular basis of MZ1-induced compact complex formation between Brd4^{BD2} and VHL. a, Sequence alignment of BET bromodomains. Residues of Brd4^{BD2} in contact with MZ1 and/or VHL are highlighted. **b**, Structural alignment of Brd2^{BD1} (yellow) superposed on Brd4^{BD2} (dark green) in the ternary structure with MZ1 and VHL. Key, non-conserved interacting residues are shown in sticks. **c**, ΔpK_d measured for VCB with MZ1 and the indicated BET-BDs, reported as difference (± uncertainty), from pK_d values measured as mean (± 1 s.e.m.) as described in Online Methods. Statistical significance of pK_d values for ternary titrations compared to the corresponding binary titrations (in black) and for ternary titrations of WT compared to the corresponding triple-mutant (in red) was assessed by two-tailed *t*-test assuming equal variances, and is indicated as * (*p*-value < 0.05), ** (*p*-value < 0.01) or *** (*p*-value < 0.001). **d**, AlphaLISA intensity values titrating Brd2^{BD1}, Brd4^{BD2} and corresponding mutants against VCB with MZ1. AlphaLISA intensities are the mean (± 1 s.d.) of intensity values from four technical replicates.

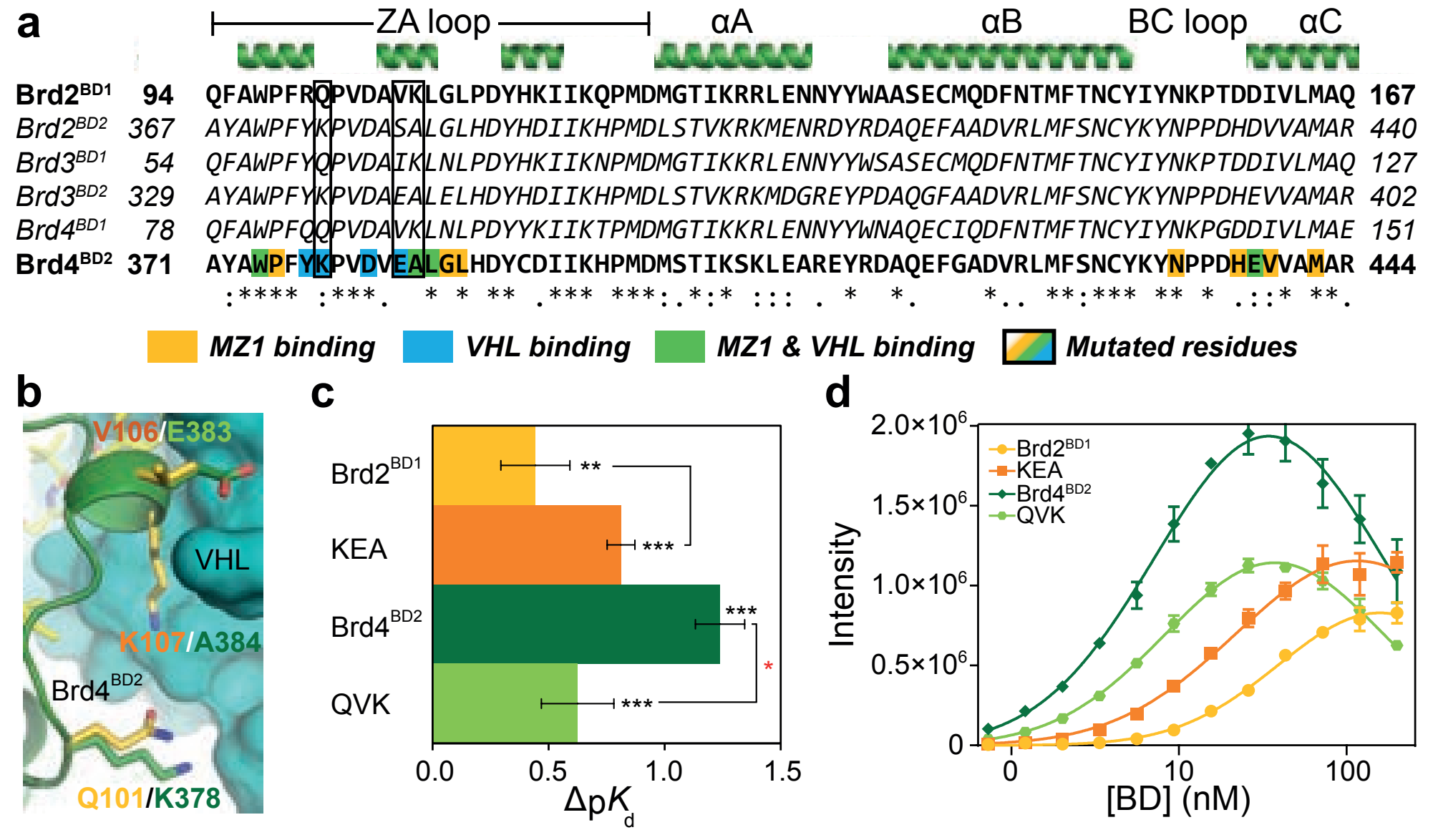
Figure 4. Structure-guided design and characterization of Brd4-selective degrader AT1. a, A vector linking VH032 to JQ1 that maintains the relative binding orientation. **b**, Chemical structure of AT1. **c**, ΔpK_d measured for VCB with AT1 and the indicated BET-BDs, reported as difference (± uncertainty), from pK_d values measured as mean (± 1 s.e.m.). Statistical significance of pK_d values for ternary titrations compared to corresponding binary (black) and for ternary WT compared to corresponding mutant (red) was assessed by two-tailed *t*-test assuming equal variances. **d**, AlphaLISA intensity values titrating VCB against BET-BDs with AT1 (left) and BET-BDs against VCB with AT1 (right). AlphaLISA intensities are the mean (± 1 s.d.) of intensity values from four technical replicates. **e**–**g**, Highly selective degradation of Brd4 by AT1 in HeLa cells after 24 h. **e**, Protein levels are shown from one representative of three biological replicates, visualized by immunoblot (left) and quantified relative to DMSO (right). Intensity values were measured as described in Online Methods. Full gels are provided in Supplementary Fig. 11. **f**, Impact of AT1 (1 μM, 24 h) on the cellular proteome. Data plotted as fold change (log_2) of replicate 1 vs replicate 2, for a total of 5,674 proteins quantified (see Online Methods). **g**, Quantified levels of BET proteins shown are mean (± 1 s.e.m.)

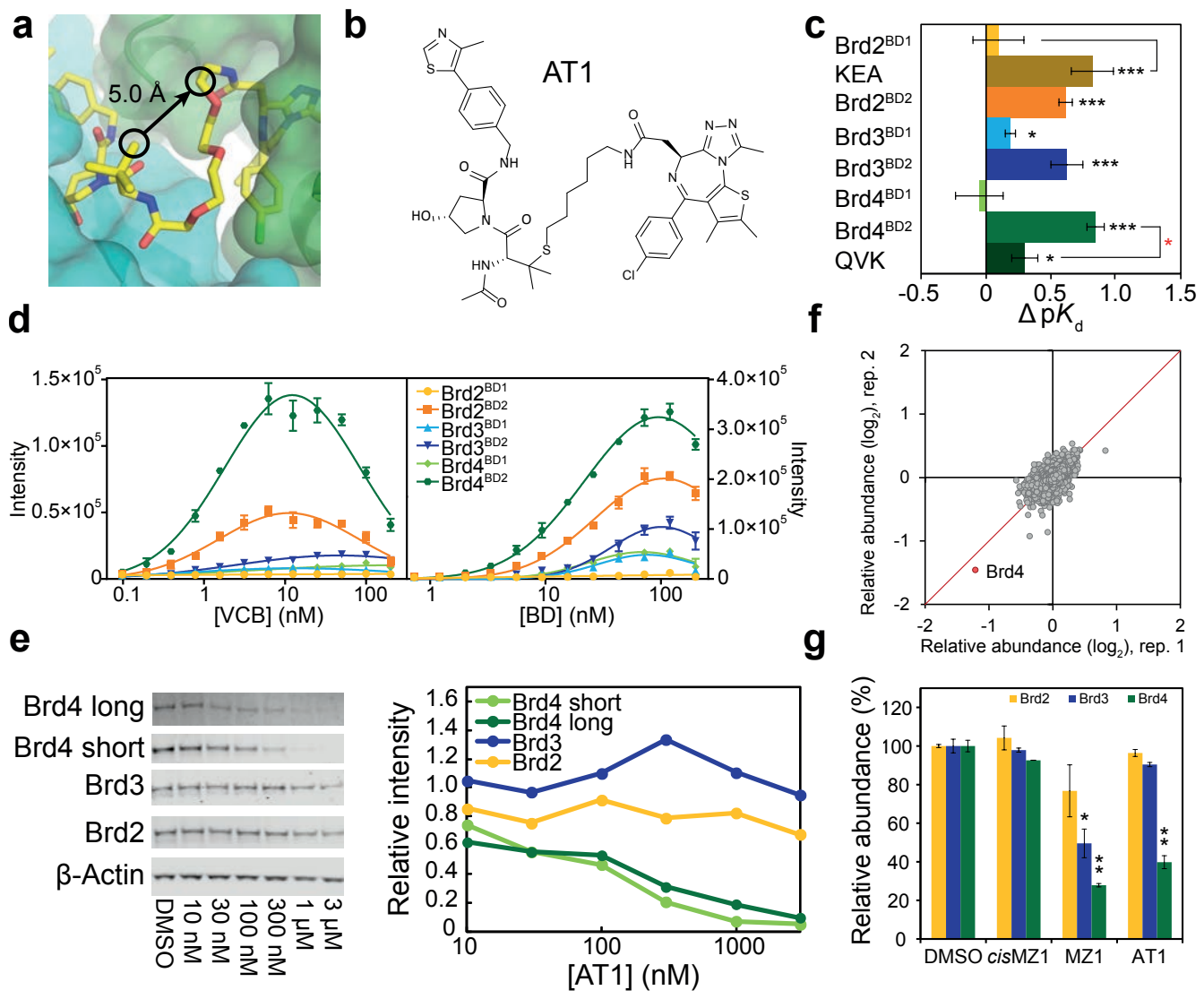
from two replicates relative to mean of vehicle. Statistical significance of relative protein abundance compared to DMSO was assessed by two-tailed *t*-test assuming equal variances. Statistical significance indicated as * (p-value < 0.05), ** (p-value < 0.01) or *** (p-value < 0.001).

Figure 5. Schematic model of selective PROTAC-induced target degradation. A target is preferentially recruited in a stable and positively cooperative ternary complex with the E3 ubiquitin ligase upon folding of the bifunctional probe to induce formation of specific PPIs.


a
ElonginB
ElonginC
VHL
MZ1
VH032
JQ1
PEG linker
MZ1
MZ1
H2
H1
H3
H4
β1
β2
β3
β4
β5
β7
ZA loop
BC loop
αA
αB
αZ
αC
Brd4BD2
b
αA
αZ
F376
αC
Brd4BD2
P375
H437
JQ1
W374
L385
E438
P71
A384
H110
I109
Y112
R108
β4
VH032
VHL
β1
c
ZA loop
D381
Brd4BD2
2.9 Å
3.2 Å
3.2 Å
E383
R108
3.5 Å
R107
VHL
d
BC loop
Brd4BD2
2.9 Å
H437
E438
3.1 Å
VHL
R69


a
Brd4BD2
intraMZ1
VHL
0
−1
−2
−3
Evdw+electrostatic (kcal/mol)
b
VCB into MZ1
VCB into MZ1:Brd4BD2
Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio
N = 0.880 ± 0.004 sites
Kd = 67 ± 8 nM
ΔH = –7,980 ± 60 cal/mol
ΔS = 6.0 cal/mol/deg
N = 0.991 ± 0.001 sites
Kd = 4.4 ± 1.0 nM
ΔH = –8,960 ± 30 cal/mol
ΔS = 8.2 cal/mol/deg
c
PROTAC
+ VHL
+ BRD
VHL
BRD
pKd VHL
pKd T,BRD
pKd BRD
pKd T,VHL
α = KdA / KdT,A ; ΔpKd = pKdT,A – pKdA , cooperativity
α < 1; ΔpKd < 0 , negative cooperativity
α = 1; ΔpKd = 0 , no cooperativity
α > 1; ΔpKd > 0 , positive cooperativity
d
Brd2BD1
Brd2BD2
Brd3BD1
Brd3BD2
Brd4BD1
Brd4BD2
ΔpKd
e
Fraction of ternary complex (%)
Log ([MZ1]t, M)
f
Intensity
[VCB] (nM)


a
ZA loop
αA
αB
BC loop
αC
Brd2BD1 94 QFAWPFRQPVDAVKLGLPDYHKIIKQPMDMGTIKRRLENNYYWAASECMQDFNTMFTNCYIYNKPTDDIVLMAQ 167
Brd2BD2 367 AYAWPFYKPVDASALGLHDYHDIIKHPMDLSTVKRKMENRDYRDAQEFAADVRLMFSNCYKYNPPDHDVVAMAR 440
Brd3BD1 54 QFAWPFYQPVDAIKLNLPDYHKIIKNPMDMGTIKKRLENNYYWSASECMQDFNTMFTNCYIYNKPTDDIVLMAQ 127
Brd3BD2 329 AYAWPFYKPVDAEALELHDYHDIIKHPMDLSTVKRKMDGREYPDAQGFAADVRLMFSNCYKYNPPDHEVVAMAR 402
Brd4BD1 78 QFAWPFQQPVDAVKLNLPDYYKIIKTPMDMGTIKKRLENNYYWNAQECIQDFNTMFTNCYIYNKPGDDIVLMAE 151
Brd4BD2 371 AYAWPFYKPVDVEALGLHDYCDIIKHPMDMSTIKSKLEAREYRDAQEFGADVRLMFSNCYKYNPPDHEVVAMAR 444
MZ1 binding
VHL binding
MZ1 & VHL binding
Mutated residues
b
V106/E383
VHL
K107/A384
Brd4BD2
Q101/K378
c
Brd2BD1
KEA
Brd4BD2
QVK
**
***
***
***
*
0.0
0.5
1.0
1.5
ΔpKd
d
2.0×10^6
1.5×10^6
1.0×10^6
0.5×10^6
0
Intensity
Brd2BD1
KEA
Brd4BD2
QVK
0
10
100
[BD] (nM)


a
5.0 Å
b
AT1
c
Brd2BD1
KEA
Brd2BD2
Brd3BD1
Brd3BD2
Brd4BD1
Brd4BD2
QVK
***
***
*
***
***
*
*
-0.5
0
0.5
1.0
1.5
ΔpKd
d
1.5×10^5
1.0×10^5
0.5×10^5
0
Intensity
Brd2BD1
Brd2BD2
Brd3BD1
Brd3BD2
Brd4BD1
Brd4BD2
4.0×10^5
3.0×10^5
2.0×10^5
1.0×10^5
0
Intensity
0.1
1
10
100
[VCB] (nM)
1
10
100
[BD] (nM)
f
Relative abundance (log2), rep. 2
Brd4
Relative abundance (log2), rep. 1
e
Brd4 long
Brd4 short
Brd3
Brd2
β-Actin
DMSO
10 nM
30 nM
100 nM
300 nM
1 μM
3 μM
Relative intensity
Brd4 short
Brd4 long
Brd3
Brd2
[AT1] (nM)
g
Relative abundance (%)
Brd2
Brd3
Brd4
DMSO
cisMZ1
MZ1
AT1

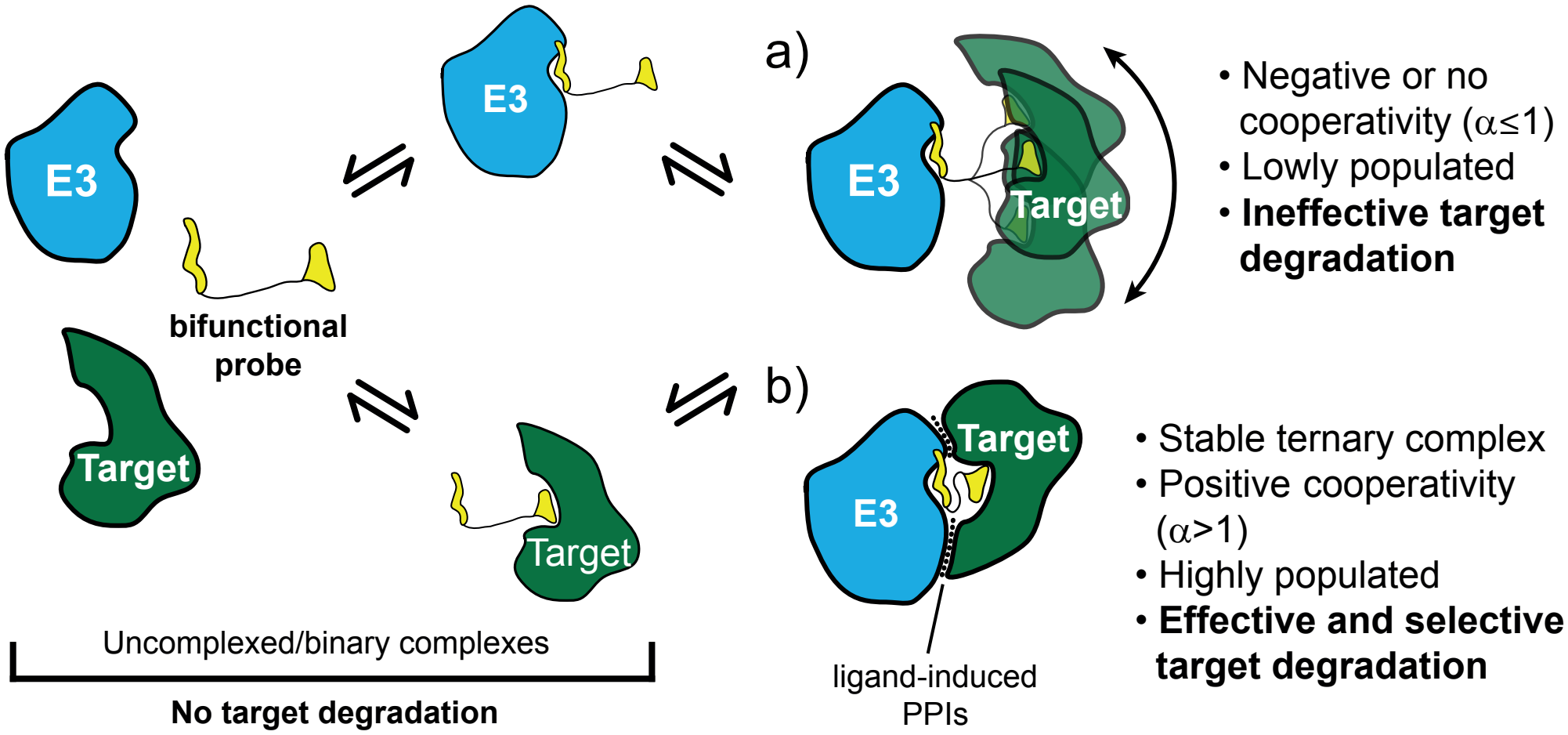

E3
E3
bifunctional probe
Target
Target
a)
E3
Target
• Negative or no cooperativity (α≤1)
• Lowly populated
• Ineffective target degradation
b)
Target
E3
• Stable ternary complex
• Positive cooperativity (α>1)
• Highly populated
• Effective and selective target degradation
ligand-induced PPIs
Uncomplexed/binary complexes
No target degradation

Table 1. Thermodynamic parameters of formation of binary and ternary complexes between MZ1, VCB and BET bromodomains measured by isothermal titration calorimetry.

Protein in syringe	Species in cell	K_d (nM)	ΔG (kcal × mol^{-1})	ΔH (kcal × mol^{-1})	$-T\Delta S$ (kcal × mol^{-1})	N	α	ΔpK_d
Brd2^{BD1}	MZ1	62 ± 6	–9.84 ± 0.06	–12.8 ± 0.7	3.0 ± 0.8	1.1 ± 0.1		
Brd2^{BD2}		60 ± 3	–9.85 ± 0.03	–9.8 ± 0.3	–0.1 ± 0.3	1.2 ± 0.1		
Brd3^{BD1}		21 ± 5	–10.2 ± 0.1	–14.7 ± 0.8	4.2 ± 0.9	1.1 ± 0.1		
Brd3^{BD2}		13 ± 3	–10.8 ± 0.1	–14.0 ± 0.9	3.3 ± 0.7	1.05 ± 0.02		
Brd4^{BD1}		39 ± 9	–10.1 ± 0.1	–14.7 ± 0.4	4.6 ± 0.5	0.95 ± 0.03		
Brd4^{BD2}		15 ± 1	–10.68 ± 0.04	–10.9 ± 0.4	0.2 ± 0.4	1.08 ± 0.07		
Brd2^{BD1} KEA		69 ± 9	–9.78 ± 0.08	–14 ± 1	4 ± 1	0.83 ± 0.08		
Brd4^{BD2} QVK		22 ± 8	–10.5 ± 0.2	–12.4 ± 0.5	1.9 ± 0.2	1.0 ± 0.1		
VCB[a]	MZ1[a]	66 ± 6	–9.81 ± 0.05	–7.7 ± 0.3	–2.1 ± 0.3	0.93 ± 0.04		
VCB	MZ1:Brd2^{BD1}	24 ± 8	–10.4 ± 0.2	–7.3 ± 0.2	–3.1 ± 0.4	1.1 ± 0.2	2.9	0.4 ± 0.2
	MZ1:Brd2^{BD2}	28 ± 3	–10.3 ± 0.1	–10.5 ± 0.1	0.2 ± 0.2	1.07 ± 0.02	2.3	0.36 ± 0.06
	MZ1:Brd3^{BD1}	19 ± 4	–10.6 ± 0.1	–8.8 ± 0.5	–1.8 ± 0.7	1.01 ± 0.01	3.5	0.5 ± 0.1
	MZ1:Brd3^{BD2}	7 ± 2	–11.2 ± 0.2	–6.3 ± 0.1	–4.9 ± 0.3	0.99 ± 0.04	10.7	1.0 ± 0.2
	MZ1:Brd4^{BD1}	28 ± 6	–10.3 ± 0.1	–9.1 ± 0.9	–1 ± 1	0.97 ± 0.06	2.3	0.4 ± 0.1
	MZ1:Brd4^{BD2}	3.7 ± 0.7	–11.5 ± 0.1	–8.9 ± 0.1	–2.6 ± 0.2	1.02 ± 0.02	17.6	1.24 ± 0.09
	MZ1:Brd2^{BD1} KEA	12 ± 7	–10.9 ± 0.4	–5.7 ± 0.2	–5.2 ± 0.2	0.8 ± 0.1	7.9	0.8 ± 0.3
	MZ1:Brd4^{BD2} QVK	14.9 ± 0.1	–10.68 ± 0.03	–6.2 ± 0.3	–4.5 ± 0.3	0.9 ± 0.1	4.2	0.62 ± 0.04

All ITC titrations were performed at 25 °C. Values reported are the mean ± S.E.M. from two independent measurements, except for VCB titration into MZ1 (line [a]) for which values reported are the mean ± S.E.M. from eight independent measurements.

Online Methods

Chemical synthesis: Synthesis of compounds described in this paper and their intermediates are described in the **Supplementary Note**.

Constructs, protein expression and purification. Wild-type and mutant versions of human proteins VHL (UniProt accession number: P40337), ElonginC (Q15369), ElonginB (Q15370), Brd2 (P25440), Brd3 (Q15059) and Brd4 (O60885) were used for all protein expression. For expression of VBC, N-terminally His_6-tagged VHL (54–213), ElonginC (17–112) and ElonginB (1–104) were co-expressed in *Escherichia coli* BL21(DE3) at 24 °C for 16 h using 3 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). *E. coli* cells were lysed using a pressure cell homogenizer (Stansted Fluid Power) and lysate clarified by centrifugation. His_6-tagged VCB was purified on a HisTrapFF affinity column (GE Healthcare) by elution with an imidazole gradient. The His_6 tag was removed using TEV protease and the untagged complex dialysed into low concentration imidazole buffer. VCB was then flowed through the HisTrapFF column a second time, allowing impurities to bind as the complex eluted without binding. VCB was then additionally purified by anion exchange and size-exclusion chromatography using MonoQ and Superdex-75 columns (GE Healthcare), respectively. The final purified complex was stored in 20 mM Bis Tris, pH 7, 150 mM sodium chloride and 1 mM dithiothreitol (DTT). $Brd2^{BD1}$ (71–194), $Brd2^{BD2}$ (344–455), $Brd3^{BD1}$ (24–146), $Brd3^{BD2}$ (306–416), $Brd4^{BD1}$ (44–178) and $Brd4^{BD2}$ (333–460) as well as equivalent mutant constructs were expressed with an N-terminal His_6 tag in *E. coli* BL21(DE3) at 18 °C for 20 h using 0.2 mM IPTG. His_6-tagged BDs were purified on nickel Sepharose™ 6 fast flow beads (GE Healthcare) by elution with increasing concentrations of imidazole. For crystallography the His_6-tagged BD was cleaved with TEV protease and dialysed into low concentration imidazole buffer. The BD was then flowed over the nickel beads a second time to remove impurities and protease. BDs were then additionally purified by size-exclusion chromatography using a Superdex-75 column. For AlphaLISA, ITC and ubiquitination reactions, following elution of His_6-tagged BDs from the nickel beads, the BDs were purified by size-exclusion chromatography using a Superdex-75 column. The final purified proteins were stored in 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.5, 150 mM sodium chloride and 1 mM DTT. All chromatography purification steps were performed using Äkta FPLC purification systems (GE Healthcare) or glass econo-columns (Bio-Rad) at room temperature.

Crystallography. VCB, MZ1 and $Brd4^{BD2}$ were mixed as a 1:1:1 stoichiometric ternary complex with a final concentration of 10 mg/mL. Drops of the ternary complex were mixed 1:1 with 13% (w/v) PEG 8000 and 0.1 M sodium citrate (pH 6.3) in the hanging-drop vapour diffusion format. Crystals appeared within minutes and were fully grown after one day. A crystal was dehydrated in a solution containing 18% (w/v) PEG 8000 for a few minutes and flash-cooled in liquid nitrogen using 20% 2-methyl-2,4-pentanediol in liquor solution as a cryoprotectant. Diffraction data were collected at Diamond Light Source beamline I04-1 using a Pilatus 6M-F detector at a wavelength of 0.92819 Å. Indexing and integration of reflections was performed using XDS with the XDSGUI interface[43], and scaling and merging with AIMLESS[44] in CCP4i[45]. The Wilson *B* factor was estimated at 47.2 $Å^2$. To solve the phase problem the molecular replacement method was used with the program PHASER[46] using search models derived from the coordinates of VCB (PDB entry 1VCB[26]) and $Brd4^{BD2}$ (PDB entry 2OUO[20]). Two instances of the ternary complex were found in the asymmetric unit, indicating a final solvent content of 68% as calculated from the Matthews coefficient. The initial model was refined iteratively using REFMAC[47] and COOT[48]. Ligand structures

and restraints were generated using the PRODRG server[49]. The MOLPROBITY server[50] was used to validate the geometry and steric clashes in the structures; the distribution of backbone torsion angles in the Ramachandran plot are 98.3% in the favored region and 1.7% in the allowed region. The structure has been deposited in the protein data bank (PDB) with accession code 5T35 and data collection and refinement statistics are presented in **Supplementary Table 1**. Interfaces and contacts observed in the crystal structure were calculated with PISA[51] and LIGPLOT[52]. All figures were generated using PyMOL.

Molecular dynamics simulations. MD simulations were carried out using the NAMD program[53] and the CHARMM 36 force field[54]. We attempted to derive *ab initio* topology and parameter files for MZ1 using Jaguar v. 9.0 (Schrödinger Inc., LLC, New York, NY, US). However, characterization of the minimized structure as a minimum by vibrational analysis proved unsuccessful (number of imaginary frequencies > 0) using several approaches and initial structures, probably due to the large number of atoms. Therefore topology and parameter files were generated using the CGenFF server[55].
To simulate the $Brd4^{BD2}$:MZ1:VHL ternary complex in solution, the coordinates of the X-ray crystal structure of the complex (chains A and D) were used as starting structure for simulation. ElonginB and ElonginC, which are sufficiently far from the hydroxyproline recognition site of VHL (> 20 Å), were excluded to increase computational efficiency throughout the simulation. The model was solvated in a TIP3P water box with a padding of 12 Å from the edge of the box to any protein atom. The system charges were neutralized with sodium or chloride ions as appropriate. The solvated system was minimized for 3,000 steps with all protein and MZ1 atoms restrained to eliminate residual unfavorable interactions between each other and the solvent, followed by another 5,000 steps with all atoms free to move. Heating of the system from 0 to 300 K was achieved in 100 ps (time step of 1 fs), with fixed protein backbone atoms to allow relaxation of the solvent. The system was subsequently equilibrated for 600 ps (time step of 2 fs) with all atoms free to move. The NPT ensemble was used during production simulation of 100 ns (time step of 2 fs). The temperature was controlled with a Langevin thermostat at 300 K, and the pressure with a Nose-Hoover Langevin piston barostat at 1 bar. A SHAKE constraint was applied to all bonds containing hydrogen atoms. Short-range nonbonded interactions were switched at 10 Å and cut off at 12 Å, and particle mesh Ewald summation was employed for long-range non-bonded interactions.
The trajectory was analysed using VMD v. 1.9.2[56] and taking snapshots every 10 ps of simulation, unless otherwise stated. To calculate *root-mean-square deviations (RMSD)* throughout the simulation, ternary complexes were superposed to the crystallographic complex using an in-house PyMOL script considering only Cα atoms of the VHL protein within a shell of 10 Å from MZ1. This was implemented in order to diminish the effect of structural rearrangements occurring far from the hydroxyproline recognition site of VHL during the simulation arising from the absence of ElonginB and ElonginC. *Radius of gyration (R_g)* of the ternary complex, *i.e*. the radius of a sphere with equivalent moment of inertia, was computed using Carma[57] at each snapshot considering the protein backbone. *Buried surface area (BSA)* upon complex formation, *i.e.* the difference in surface-accessible surface area (SASA) between the formed complex and the unbound partners in the system, was computed considering all protein atoms and a spherical probe of radius 1.4 Å. *Intermolecular contacts*, *i.e.* pair-to-pair contacts between an amino acid in $Brd4^{BD2}$ and VHL, were considered formed if more than five atoms of the amino acid were at a distance closer than 4.0 Å from the partner protein. Intermolecular contacts were computed using the Timeline plugin v. 2.3 as implemented in VMD. Per-residue and per-atom inter- and intramolecular interaction energies *($E_{vdW + electrostatic}$)* were computed using an in-house automated routine of the NAMD

Energy plugin v. 1.4 as implemented in VMD. Interaction energies were estimated by adding the pair-wise van der Waals and electrostatic contributions between individual amino acids or atoms and the corresponding partner. In the case of per-atom calculations, the interaction energies of hydrogen atoms were added to their corresponding heavy atom. For intramolecular interactions analysis, MZ1 was divided into three sections (JQ1, PEG linker, and VH032) and pair-wise energetic contributions from the atoms of each section to the rest of the molecule (excluding 1–4 bonded atoms) were calculated. In order to obtain comparable and interpretable results, the following scaling factor and cutoff value were applied to the electrostatic contribution:

$$\begin{cases} if\ |E_{\mathrm{vdW}}| \geq 0.1\text{: } E_{\mathrm{electrostatic}} = 0.07 * E_{\mathrm{raw\ electrostatics}} \\ if\ |E_{\mathrm{vdW}}| < 0.1\text{: } E_{\mathrm{electrostatic}} = 0 \end{cases}$$

Isothermal titration calorimetry (ITC) Titrations were performed on an ITC200 micro-calorimeter (GE Healthcare). The titrations were all performed as reverse mode (protein in syringe, ligand in cell) and consisted of 19 injections of 2 µL protein solution (20 mM Bis-tris propane, 150 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP), 0.02 % DMSO, pH 7.4) at a rate of 2 sec/µL at 120 s time intervals. An initial injection of protein (0.4 µL) was made and discarded during data analysis. All experiments were performed at 25 ºC, whilst stirring at 600 rpm. PROTACs (MZ1 or AT1) were diluted from a 10 mM DMSO stock solution to 20 µM in buffer containing 20 mM Bis-tris propane, 150 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP), pH 7.4. The final DMSO concentration was 0.01 or 0.02 %. Bromodomain (100 or 200 µM, in the syringe) was titrated into the PROTAC (10 or 20 µM, in the cell). At the end of the titration, the excess of solution was removed from the cell, the syringe was washed and dried, VCB complex (84 or 168 µM, in the same buffer) was loaded in the syringe and titrated into the complex PROTAC:bromodomain. The concentration of the complex in the cell (*C)* after the first titration (8.4 or 16.8 µM), was calculated as follow:

$$C = \frac{C_0 \cdot V_{cell}}{V_{cell} + V_{inj}}$$

where: C_0 is the initial concentration of the PROTAC in the cell (20 µM), V_{cell} is the volume of the sample cell (200.12 µL) and V_{inj} is the volume of titrant injected during the first titration (38.4 µL). Titrations for the binary complex PROTAC:VCB were performed as follow: to the solution of PROTAC (10 or 20 µM, in the cell), buffer (38.4 µL) was added by means of a single ITC injection. The excess of solution was removed from the cell, the syringe was washed and dried, VCB complex (84 or 168 µM, in the same buffer) was loaded in the syringe and titrated into the diluted PROTAC solution. The data were fitted to a single binding site model to obtain the stoichiometry n, the dissociation constant K_d and the enthalpy of binding ΔH using the Microcal LLC ITC200 Origin software provided by the manufacturer. The reported values are the mean ± s.e.m. from independent measurements (eight for VCB into MZ1; seven for VCB into AT1; two for each BD into VCB:PROTAC).

Simulations of ternary complex fractions. Fractions of ternary complexes were calculated by applying the ternary equilibria model in the excel spreadsheet provided in ref. [23]. Input parameters were $[VHL]_t = [BD]_t = 40$ nM; K_d (VHL) = 66 nM; K_d and α (BD) were as measured by ITC (**Table 1**).

Biotinylation of VCB. To biotinylate VCB the complex was mixed in a 1:1 stoichiometry with EZ-Link NHS-PEG_4-Biotin (Thermo Scientific) and incubated at room temperature for 1 h. To remove any unreacted NHS-biotin the sample was run over a PD-10 desalting column (GE Healthcare) into 20 mM HEPES, pH 7.5, 150 mM sodium chloride and 1 mM DTT.

AlphaLISA assays. All assays were performed at room temperature and plates sealed with transparent film between addition of reagents. All reagents were diluted in 50 mM HEPES, pH 7.5, 100 mM NaCl, 0.1% (w/v) bovine serum albumin and 0.02% (w/v) 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) and each solution was prepared as a 5× stock and mixed into a final volume of 25 µL per well. Each protein sample (biotinylated VCB and His_6-BD) and PROTAC were mixed and incubated for 1 h. Ni-coated acceptor beads were added and plates incubated another 1 h. Streptavidin-coated donor beads were added and plates incubated for a final 1 h. Plates were read on a PHERAstar FS (BMG Labtech) using an optic module with an excitation wavelength of 680 nm and emission wavelength of 615 nm. Within each read there was a settling time of 0.1 s, an excitation time of 0.3 s and an integration time of 0.6 s. For BET-BD titration experiments, VCB and PROTACs or Biotin-JQ1 alone were kept constant at a final concentration of 20 nM and His_6-BD was serially diluted three-in-five from 200 nM. For VCB/Biotin-JQ1 titration experiments, His_6-BDs and PROTACs were kept constant at a final concentration of 40 nM and VCB or Biotin-JQ1 was serially diluted one-in-two from 200 nM. The intensity values were plotted with concentration values on a $\log_{10}$ scale.

Tissue culture. Human HeLa cell lines were obtained from ATCC and were kept in DMEM medium (Gibco) supplemented with 10% FBS (Gibco), L-glutamate (Gibco), penicillin and streptomycin. Cells were kept in incubator at 37 °C, 5% CO_2. All cell lines were tested for mycoplasma contamination every month using MycoAlert™ Mycoplasma detection kit (Lonza).

Cell treatment and lysis. HeLa cells were seeded at 2.5×10^5 per well on a standard six-well plate. After a day, cells were treated with test compounds for the desired period of time. Cells were washed with PBS twice before lysis. Lysis was achieved by applying RIPA buffer (Sigma), supplemented with 1× protease inhibitor cocktail (Roche), Benzonase (Merck) and 0.5 mM $MgCl_2$ to the cells on ice. Lysate was briefly sonicated and then centrifuged at 20,000 × *g* for 10 min at 4 °C. Supernatant was collected as cell extract and protein concentration was measured by BCA assay. The extract was snap frozen in liquid nitrogen for storage before being used for Western blot analysis. Cycloheximide (C7698, Sigma Aldrich) was used at 100 µg/mL for the indicated times.

Western blot. Blots were probed with antibodies for Brd4 (AbCam ab128874, 1:1,000 dilution), Brd3 (AbCam ab50818, 1:500 dilution), Brd2 (AbCam ab139690, 1:2,000 dilution), β-actin (AbCam ab8227, 1:2,000 dilution) and cMyc (AbCam ab32072, 1:1,000 dilution) antibodies. Blots were developed with anti-Mouse or anti-Rabbit IRDye® 800CW secondary antibody from Licor (1:10,000 dilution) and bands visualized using Licor Odyssey Sa imaging system. Image processing and band intensity quantification were done using Licor Image Studio software Version 5.2.5. Ubiquitination blots were probed with anti-6×His antibody (AbCam ab18184, 1:2,000 dilution) and then with Anti-mouse IgG, HRP-linked Antibody (Cell Signaling technology #7076, 1:2,000 dilution). Probed blots were visualised with ECL Western Blotting Substrate (Pierce #32106) on film.

Sample preparation for MS proteomics. HeLa cells were seeded at 2×10^6 on a 100mm plate 24 h before treatment. To treat the cells, culture medium was replaced with 12 mL of medium containing the test compound. After 24 h, medium was removed and cells were washed with 12 mL of cold PBS twice. Samples were kept on ice from this point onwards. Cells were lysed in 0.5 mL of 100 mM Tris pH 8.0, 4% (w/v) SDS supplemented with protease inhibitor cocktail (Roche). The lysate was pulse sonicated briefly and then centrifuged at 17,000 × *g* for 15 min at 4 °C. The supernatant fraction of cell extract was snap-frozen in liquid nitrogen and stored in −80 °C freezer before further processing.

Sample preparation for MS proteomics. Samples were quantified using a micro BCA protein assay kit (Thermo Fisher Scientific) and 200µg of each sample was processed and digested using the Filter Aided Sample Preparation (FASP) method [58]. The samples were then desalted using a 7 mm, 3 ml C18 SPE cartridge column (Empore™, 3M) and labelled with TMT [59] 10plex™ Isobaric Label Reagent Set (Thermo Fisher Scientific) as per manufacturers instructions. After labelling, the peptides from the 10 samples were pooled together in equal proportion. The pooled sample was fractionated into 20 discrete fractions using high pH reverse phase chromatography [60] on a XBridge peptide BEH column (130 Å, 3.5 µm 2.1 X 150 mm, Waters) using an Ultimate 3000 HPLC system (Thermo Scientific/Dionex). Column temperature was set to 20 ^{o}C. The peptides were separated using a mix of buffers A (10 mM ammonium formate in water, pH 10) and B (10 mM ammonium formate in 90% CH_3CN, pH 10). The peptides were eluted from the column using a flow rate of 200 µl/min and a linear gradient from 5% to 60% buffer B over 60 min. The peptides eluted from the column were separated into 40 fractions prior to concatenation into 20 fractions based on the UV signal of each fraction. All the fractions were dried in a speedvac concentrator and resuspended in 10 µl 5% formic acid, then diluted to 1% prior to MS analysis.

nLC-MS/MS analysis. The fractions were analysed sequentially on a Q Exactive™ HF Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (Thermo Scientific) coupled to an Ultimate 3000 RSLC nano UHPLC system (Thermo Scientific) and Easyspray column (75 µm × 50 cm, PepMap RSLC C18 column, 2 µm, 100 Å, Thermo Scientific). The peptides from each fraction were separated using a mix of buffer A (0.1% formic acid in MS grade water) and B (0.08% formic acid in 80% MS grade CH_3CN). The peptides from each fraction were eluted from the column using a flow rate of 300 nl/min and a linear gradient from 5% to 40% buffer B over 122 min. The column temperature was set at 50 ^{o}C. The Q Exactive™ HF Hybrid Quadrupole-Orbitrap™ Mass Spectrometer was operated in data dependent mode with a single MS survey scan followed by 10 sequential m/z dependant MS2 scans. The 10 most intense precursor ions were sequencially fragmented by Higher energy Collision Dissociation (HCD). The MS1 isolation window was set to 0.4 Da and the resolution set 120,000. MS2 resolution was set as 60,000. The maximum ion injection time for MS1 and MS2 were set at 50 msec and 200 msec, respectively.

Peptide and protein identification. The raw ms data files for all 20 fractions were merged and searched against the Sprot database with taxonomy set to *Homo sapiens* by Proteome Discoverer Version 1.4 (Thermo Scientific) using the Mascot v.2.4.1 (Matrix Science) search engine for protein identification and TMT reporter ion quantitation. The identification was based on the following database search criteria: enzyme used Trypsin/P; maximum number of missed cleavages equal to 2; precursor mass tolerance equal to 10 ppm; fragment mass tolerance equal to 0.06 Da; dynamic modifications: Oxidation (M), Dioxidation (M), Acetyl (N-term), Gln->pyro-Glu (N-term Q), Pro->Hyp (P), Deamidation(N,Q); static modifications:

Carbamidomethyl (C), TMT10plex (K), TMT10plex (N-term). For protein identification the mascot ion score threshold was set at 30 and a minimum of 2 peptides was required.

Peptide and protein quantitation. The ratios of TMT reporter ion abundances in MS/MS spectra generated by HCD from raw data sets were used for TMT quantification. Isotopic correction factors were applied for the batch of TMT reagents used in this experiment as per manufacturers recommendation. A minimum of two unique peptides was used for quantitation and the resultant ratios were normalized based on protein median. Quantified proteins were filtered if the absolute fold change difference between the two DMSO replicates was ≥ 1.3.

Model construction of the multisubunit CRL2^{VHL}–MZ1–Brd4 complex. A structural model of the CRL2^{VHL} (VHL–EloC–EloB–Cul2–Rbx1) with bound MZ1–Brd4^{BD2} at one end and E2–Ubiquitin at the other end was constructed in PyMOL by aligning our Brd4^{BD2}–MZ1–VHL–EloC–EloB on to the quaternary structure VHL–EloC–EloB–Cul2^{NTD} (PDB entry 4WQO). Cul2^{NTD} and Cul2^{CTD} were modelled based on the structures of Cul5^{NTD} (PDB entry 2WZK) and Cul1^{CTD}–Rbx1 (PDB entry 3RTR) and superposed onto full-length Cul1 from PDB entry 1LDK. Finally, the Rbx1–E2–Ub arm was modelled based on the crystal structure of Rbx1–Ubc12~NEDD8–Cul1–Dcn1 (PDB entry 4P5O) superposed via the cullin subunit.

Recombinant ubiquitination experiments and ubiquitination site identification. His$_6$-tagged BET-BDs (2 μM) were ubiquitinated in the presence of E1 Ube1 (19 nM), E2 Ube2d1 (145 nM), ubiquitin (Ubiquigent, 1 mg/mL), recombinant VHL-ElonginC-ElonginB-Cullin2-Rbx1 complex (330 nM) and MZ1 (2 μM) standing for 24 h at room temperature in a buffer of 25 mM HEPES, pH 7.5, 5 mM $MgCl_2$, 100 mM NaCl, 2 mM ATP, 0.1 mg/mL BSA and 1 mM TCEP. Reactions were terminated by the addition of 1× NuPAGE LDS sample buffer (Invitrogen).

Sample preparation for MS. Samples were run 1–2 cm into a pre-cast 4–12% Bis-Tris NuPAGE gel and the entire protein content of each lane excised, washed and dried. Proteins were reduced with 10 mM DTT and 20 mM ammonium carbonate at 56 °C for 60 min and then alkylated with 50 mM N-ethylmaleimide and 20 mM ammonium carbonate at room temperature for 30 min. Proteins were trypsinized overnight at 30 °C and the resulting peptides extracted and dried down.

nLC-MS/MS analysis. Each peptide sample was reconstituted in 10 μl 5% formic acid then diluted to 1% prior to MS analysis. Peptide samples were injected onto a C18 PepMap 100 (300 μm x 5 mm, Thermo Scientific) trap column with buffer A (0.1% formic acid in MS grade water) using an Ultimate 3000 RSLC nano UHPLC system. After a 5 min wash at 5 μl/min the sample was then eluted onto an EasySpray PepMap RSLC C18 column (75 μm x 50 cm, Thermo Scientific) into a LTQ Orbitrap Velos Pro via an EasySpray ion source. The peptides were eluted from the column using a flow rate of 300 nl/min and a linear gradient from 2% to 40% buffer B (0.08% formic acid in 80% MS grade CH_3CN) over 124 min. The column temperature was set at 50 ^oC. The Orbitrap Velos Pro ms system was operated in data dependant acquisition mode using a Top 15 method with Lockmass = 445.120024. A MS1 survey scan with a range of 335–1800 m/z and a resolution of 60,000 was followed by 15 sequential MS2 scans at the normal scan rate using the LTQ Velos ion trap. The FTMS and ITMS AGC targets were set to $1e^6$ ions and $5e^3$ ions respectively. The FTMS and ITMS maximum fill times were set to 500 msec and 100 msec respectively. ITMS isolation width

was set at 2 Da with a normalised collision energy of 35, a default charge state of 2, an Activation Q of 0.250 and Activation Time of 10 msec.

Peptide and protein identification. The resultant raw data was searched against the Sprot database with a taxonomy filter set to *H. sapiens* using the Mascot v. 2.4.1 (Matrix Science) search engine to identify peptides containing Lysines with εN-linked di-glycine modifications. Peptide mass tolerance was set to 10 ppm and the fragment mass tolerance set to 0.6 Da. The number of maximum miss-cleavages was set to 2. The enzyme was set to Trypsin/P and the following variable modifications were considered: Acetyl (N-term), Deamidated (NQ), Dioxidation (M), Gln->pyro-Glu (N-term Q), GlyGly (K), Oxidation (M). A fixed modification for Cysteine was set to N-ethylmaleimide. A mascot ion score threshold was set at 37 to filter non-significant peptide identifications.

Statistical methods. No statistical methods were used to predetermine sample size. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment. For all experiments, number of replicates (*n*), mean value, error value and P value cutoffs are described in the respective figure legends. Error bars are shown for all data points with replicates as a measure of variation with the group. All *t*-tests performed were two-tailed *t*-tests assuming equal variances.

43. Kabsch, W. XDS. *Acta Crystallogr D Biol Crystallogr* **66,** 125–132 (2010).
44. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69,** 1204–1214 (2013).
45. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67,** 235–242 (2011).
46. McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40,** 658–674 (2007).
47. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **67,** 355–367 (2011).
48. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66,** 486–501 (2010).
49. Schüttelkopf, A. W. & van Aalten, D. M. F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* **60,** 1355–1363 (2004).
50. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66,** 12–21 (2010).
51. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372,** 774–797 (2007).
52. Laskowski, R. A. & Swindells, M. B. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51,** 2778–2786 (2011).
53. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J.Comput.Chem.* **26,** 1781–1802 (2005).
54. Brooks, B. R. *et al.* CHARMM: the biomolecular simulation program. *J.Comput.Chem.* **30,** 1545–1614 (2009).
55. Vanommeslaeghe, K. & Mackerell, A. D. Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *J Chem Inf Model* **52,** 3144–3154 (2012).
56. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics.

J.Mol.Graphics **14,** 33–8– 27–8 (1996).
57. Glykos, N. M. Software news and updates. Carma: a molecular dynamics analysis program. *J.Comput.Chem.* **27,** 1765–1768 (2006).
58. Manza, L. L., Stamer, S. L., Ham, A.-J. L., Codreanu, S. G. & Liebler, D. C. Sample preparation and digestion for proteomic analyses using spin filters. *Proteomics* **5,** 1742–1745 (2005).
59. Thompson, A. *et al.* Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal.Chem.* **75,** 1895–1904 (2003).
60. Gilar, M., Olivova, P., Daly, A. E. & Gebler, J. C. Orthogonality of separation in two-dimensional liquid chromatography. *Anal.Chem.* **77,** 6426–6434 (2005).

SUPPLEMENTARY INFORMATION

Structural basis of PROTAC-induced cooperative recognition for selective protein degradation

Morgan S. Gadd[1], Andrea Testa[1], Xavier Lucas[1], Kwok-Ho Chan, Wenzhang Chen, Douglas J. Lamont, Michael Zengerle, Alessio Ciulli*

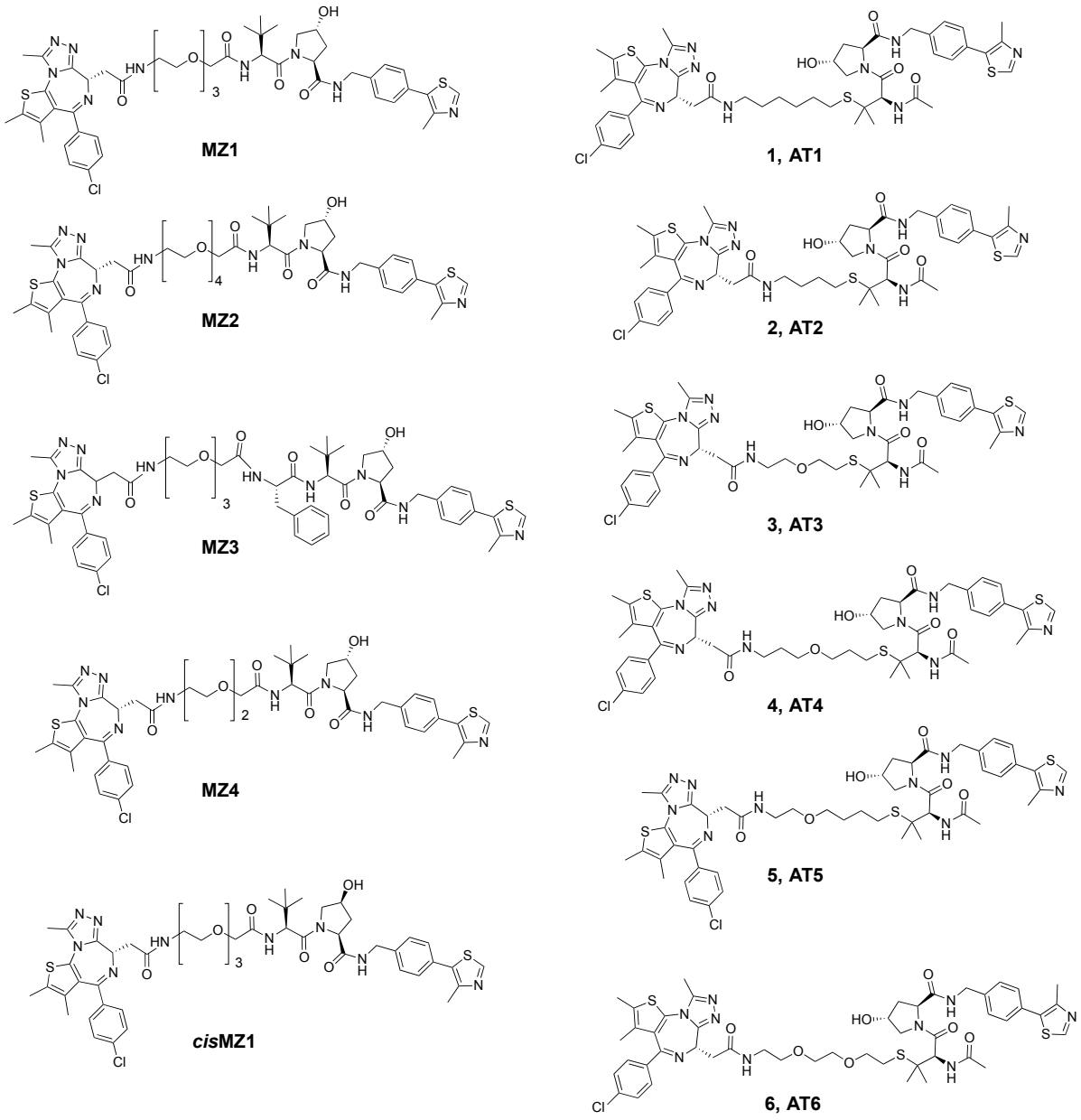
Division of Biological Chemistry and Drug Discovery, School of Life Sciences, University of Dundee, Dow Street, Dundee, DD1 5EH, Scotland, UK.

*To whom correspondence should be addressed: E-mail: a.ciulli@dundee.ac.uk

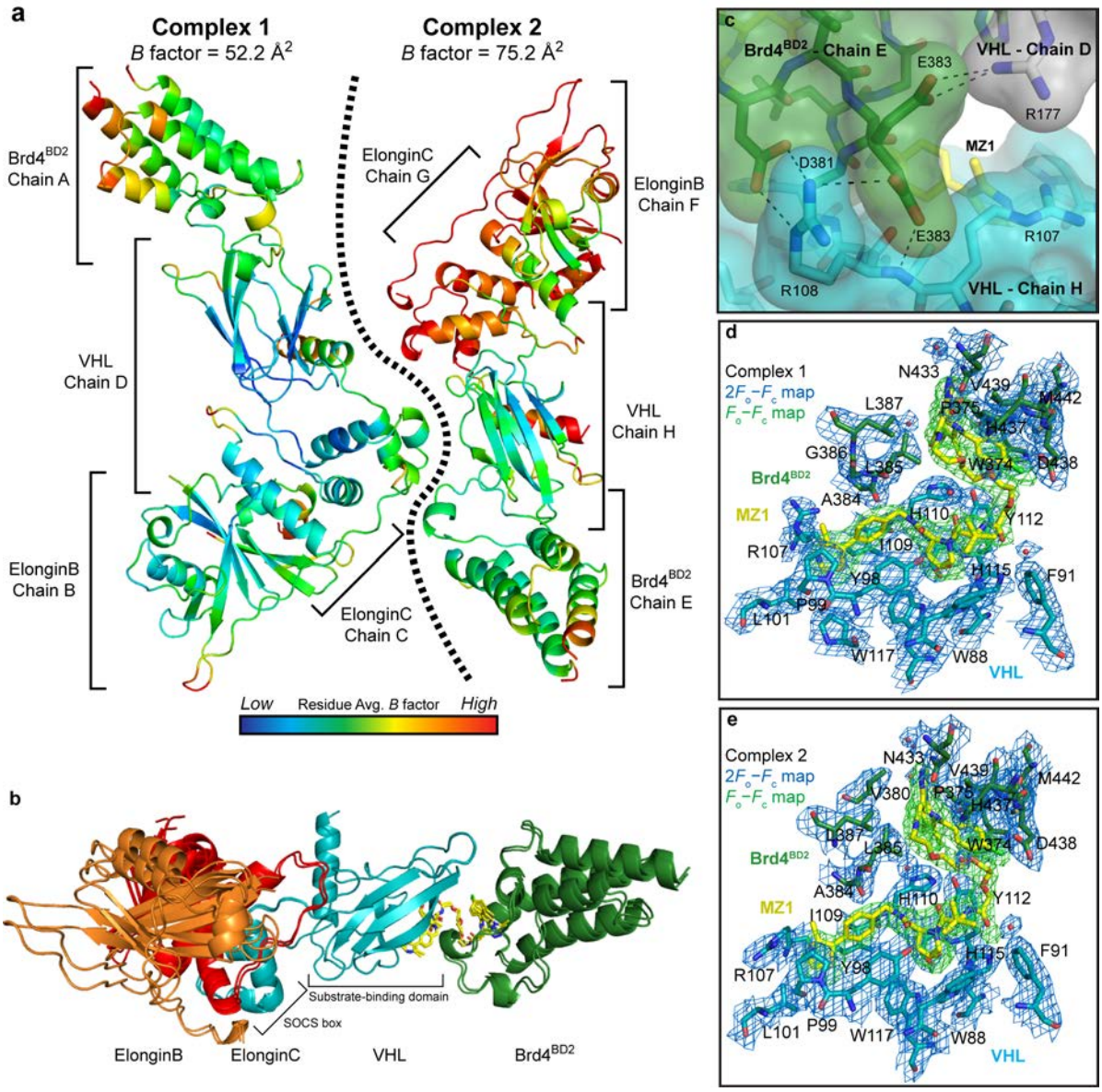
[1] These authors contributed equally to this work

- **Supplementary Figures 1-12**
- **Supplementary Tables 1-3**

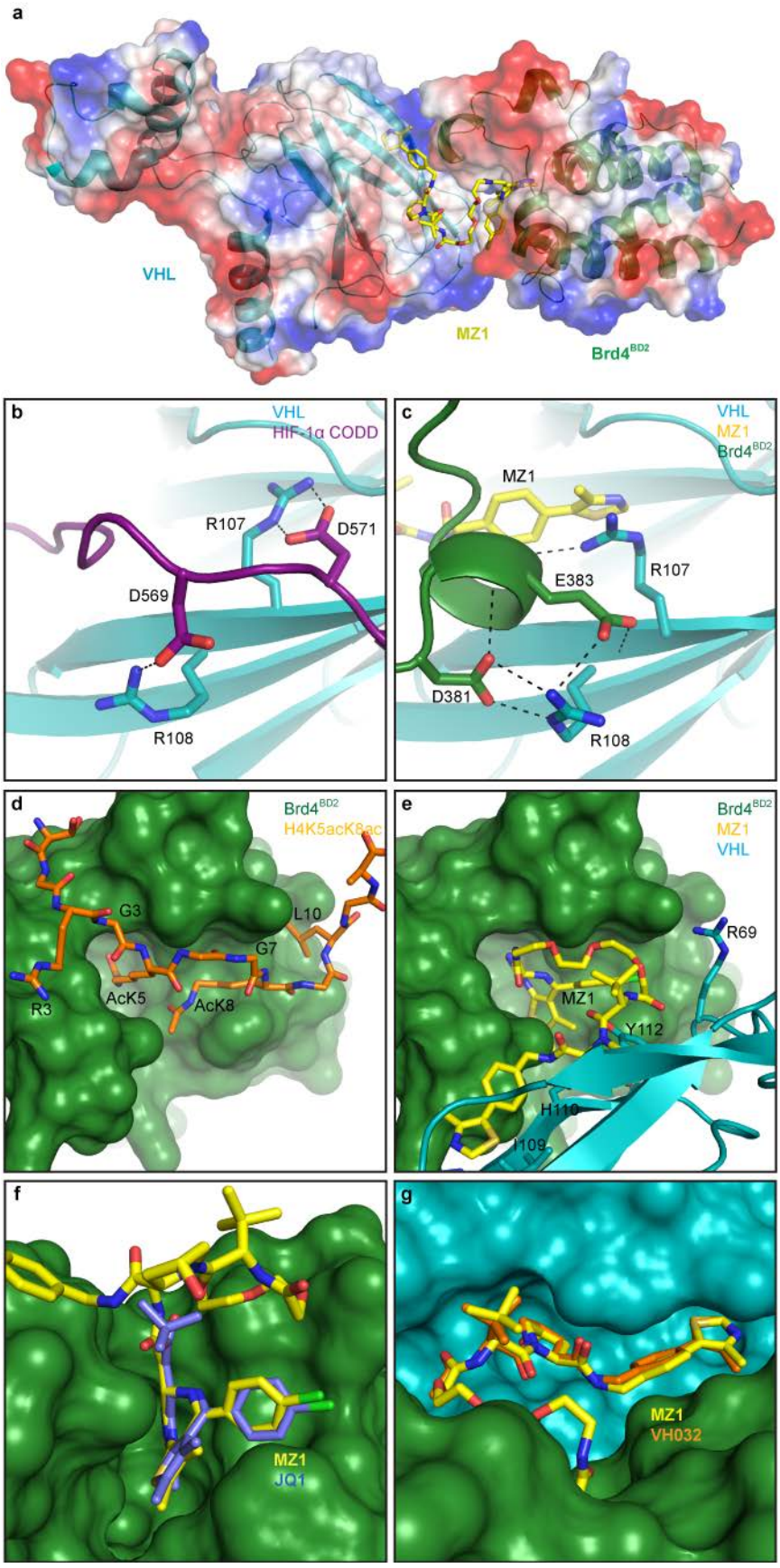
Supplementary Results



Supplementary Figure 1. Chemical structures of the PROTAC compounds used in this work.

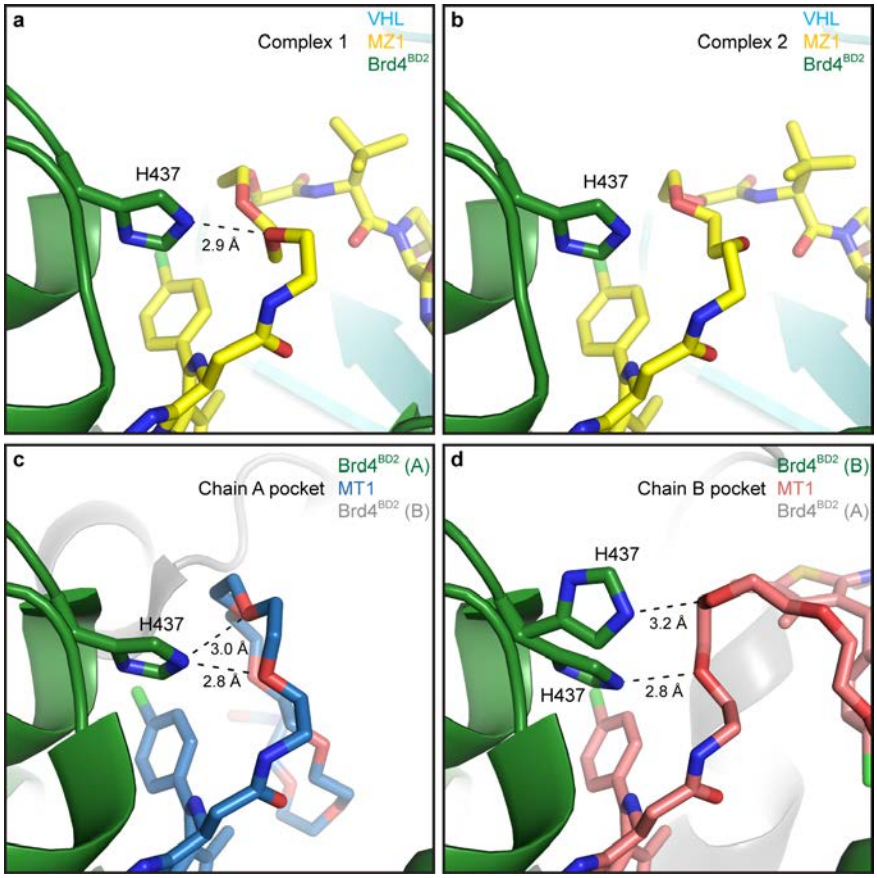


Supplementary Figure 2. Analysis of the asymmetric unit of the Brd4^{BD2}:MZ1:VCB crystal structure. a, Two complexes are present in the asymmetric unit. The eight protein chains (A–H) are shown as ribbons coloured according to average *B* factor per residue. Complex 1 is comprised of chains A–D and Complex 2 is comprised of chains E–H. **b**, Superposition of the two Brd4^{BD2}:MZ1:VCB ternary complexes in the asymmetric unit. The complexes were superposed via the backbone atoms of the VHL substrate-binding domain (residues 61–153). **c**, Analysis of crystal contact imposing on the induced Brd4^{BD2}:VHL interface in Complex 2. Arg177 from VHL in Complex 1 (chain D) interacts with Glu383 from Brd4^{BD2} in Complex 2 (chain E), resulting in two observable conformations of the side chain and displacing VHL residue Arg107 in Complex 2 (chain H) from the interface. Due to this and lower average *B* factors we used the Complex 1 (chains A, B, C and D) for all subsequent analyses. **d**, **e**, F_o−F_c omit maps (green meshes) of MZ1 prior to ligand modelling in Complexes 1 (**d**) and 2 (**e**) contoured at 3.0σ with a carve radius of 2.2 Å and 2F_o−F_c maps (blue meshes) covering residues of Brd4^{BD2} and VHL within 4 Å of MZ1 contoured at 1.5σ with a carve radius of 1.8 Å.

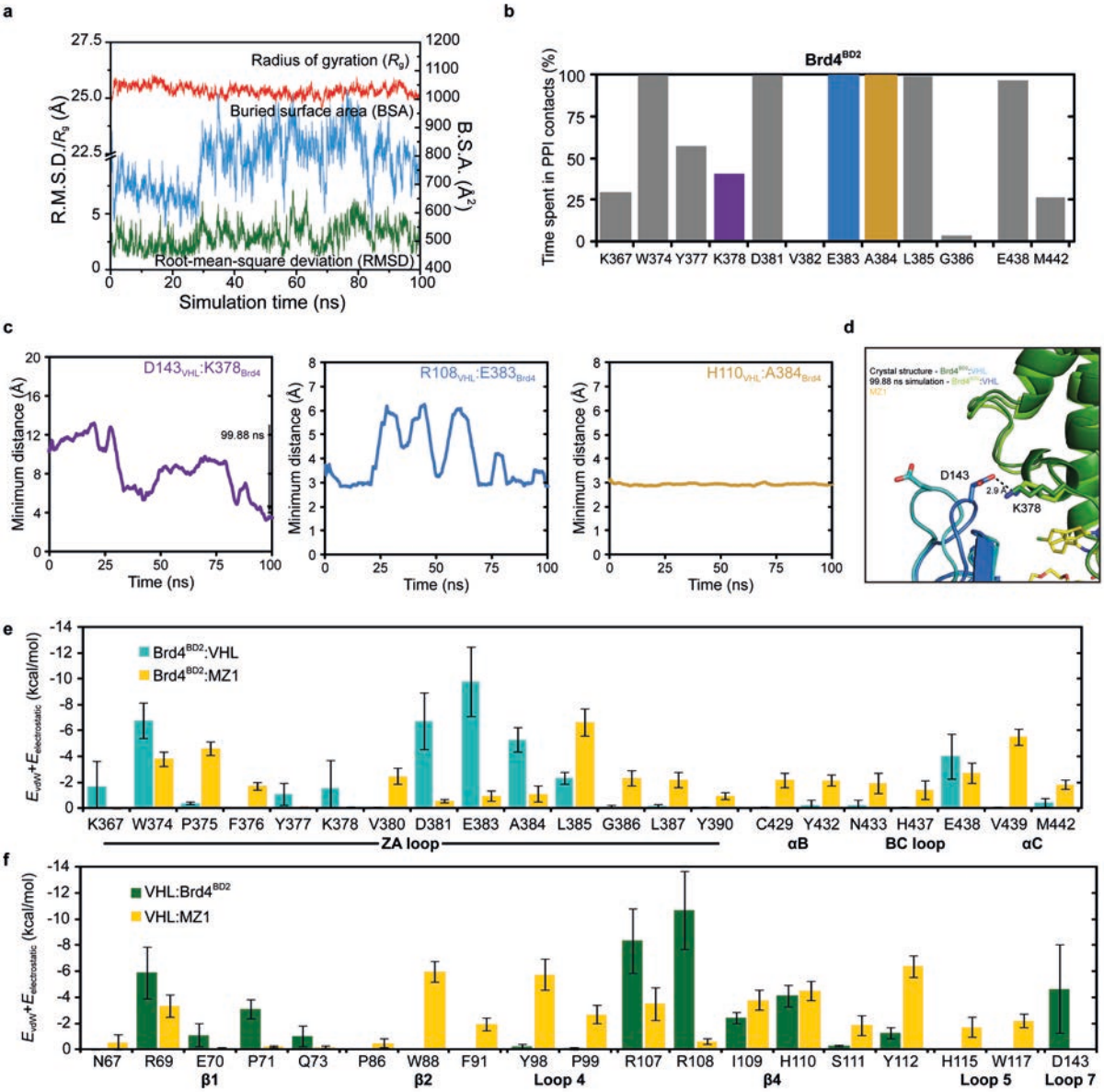


Supplementary Figure 3. Analysis and comparison of ligand binding at the Brd4^{BD2}:MZ1:VHL interface. a, The Brd4^{BD2}:MZ1:VHL interface forms a bowl shape. Overall view of MZ1 and the induced interacting proteins. Brd4^{BD2} and VHL are shown as surfaces coloured according to electrostatic potential and the underlying ribbon illustrating secondary structure. MZ1 sits on the

hydrophobic base (white surface) of the bowl and is surrounded by the binding sites of $Brd4^{BD2}$ and VHL and two complementary electrostatic arms forming the induced protein-protein interactions (Fig. 1c,d). **b**, Structure of HIF-1α (dark purple) with VHL (PDB entry 1LM8; teal) reveals contacts made by HIF-1α residues Asp569 and Asp571 with VHL residues Arg108 and Arg107, respectively. **c**, Ternary structure of $Brd4^{BD2}$ (green), MZ1 (yellow carbons) and VHL reveals similar contacts made by $Brd4^{BD2}$ residues Asp381 and Glu383 with the same VHL residues. **d**, Structure of diacetylated histone 4 tail (H4K5acK8ac; orange sticks) bound to $Brd4^{BD1}$ (PDB entry 3UVW; not shown) superposed with $Brd4^{BD2}$ from the ternary structure (green surface). **e**. Ternary structure of $Brd4^{BD2}$, MZ1 and VHL reveals that VHL does not make contacts with $Brd4^{BD2}$ that are similar to those observed made by the bound histone tail. The PEG linker of MZ1 traces a similar line to the peptide backbone of histone 4 between residues Gly4 and Gly7. **f**, Superposition of JQ1 (purple carbons) bound in $Brd2^{BD2}$ (PDB entry 3ONI; not shown) with MZ1 (yellow carbons) bound in $Brd4^{BD2}$ (green surface). **g**, Superposition of VH032 (orange carbons) bound in VHL (PDB entry 4W9H; teal surface) with MZ1 (yellow carbons).

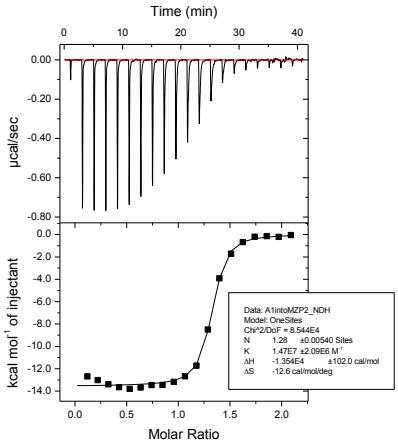


Supplementary Figure 4. Comparison of protein-ligand interactions involving PEG linkers of MZ1 and MT1. a,**b**, Structure of MZ1 (yellow sticks) bound to Brd4^{BD2} (green) from Complex 1 (**a**, chain A) and Complex 2 (**b**, chain E) and VHL (teal). A hydrogen bond (black dashes) between His437 of Brd4^{BD2} from Complex 1 and the proximal ether oxygen to JQ1 in MZ1 is shown. **c**,**d**, Structure of MT1 (PDB entry 5JWM; yellow sticks) bound to Brd4^{BD2} (green) from chain A (**c**) and chain B (**d**). Hydrogen bonds (black dashes) between His437 and the ether oxygens of MT1 are shown. The complementary Brd4^{BD2} molecule in the MT1-induced dimer is shown in grey in each case.

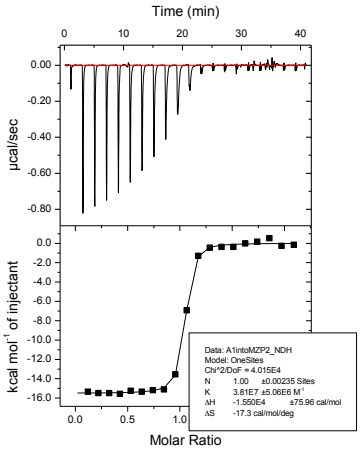


Supplementary Figure 5. Computational study of the intermolecular interactions in the Brd4^{BD2}:MZ1:VHL complex. a, Radius of gyration (R_g), buried surface area (BSA) of the PPI and root-mean-square deviation (RMSD) for the 100 ns MD simulation of the complex starting from the crystal structure. **b**, Brd4^{BD2} and VHL maintain an extensive interface throughout the MD simulation. Residues of Brd4^{BD2} are plotted according to the proportion of the time during the 100 ns of the simulation they spend making an intermolecular contact with VHL. Residues selected for side-directed mutagenesis are colored distinctly. **c**, Persistence of individual protein–protein interaction contacts in the Brd4^{BD2}:MZ1:VHL complex along the MD simulation. For the sake of clarity, value shown is the mean averaged over 5 ns. **d**, Superposition of MD simulations (light green and blue) at 99.88 ns with Brd4^{BD2}:MZ1:VHL crystal structure (dark green and teal) through Cα atoms of VHL. **e**, **f**, Per-residue intermolecular interaction energies derived from MD simulations for Brd4^{BD2} (**e**) or VHL (**f**) residues contacting MZ1 (yellow) and VHL (cyan) or Brd4^{BD2} (green), respectively. Energies are the mean (± 1 s.d.) of energy values calculated as described in Online Methods from one MD run of 100 ns and collected every 10 ps.

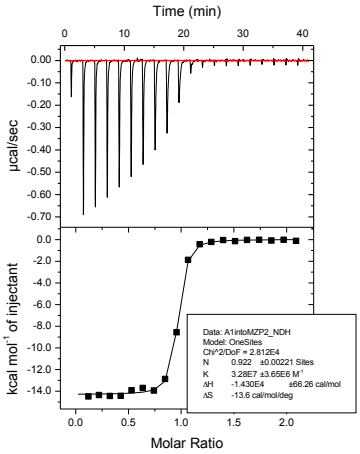
ITC titrations for binary complex MZ1:protein



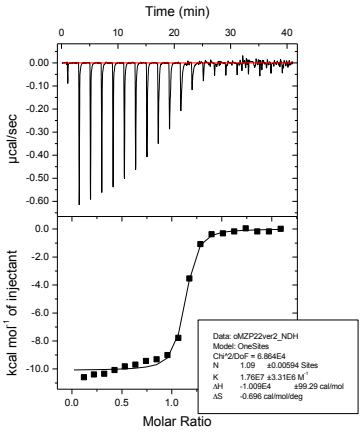
Brd2^{BD1}



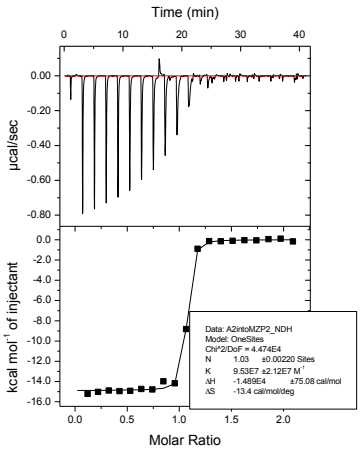
Brd3^{BD1}



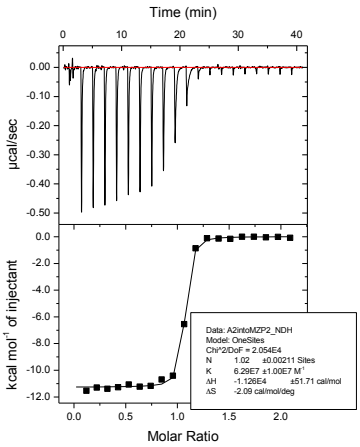
Brd4^{BD1}



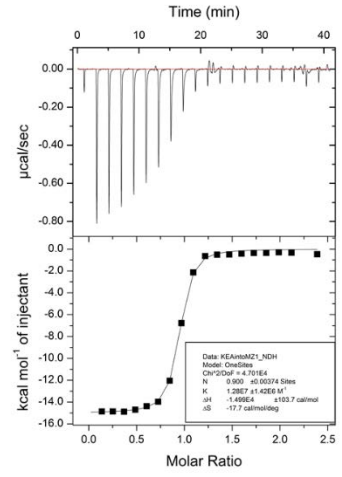
Brd2^{BD2}



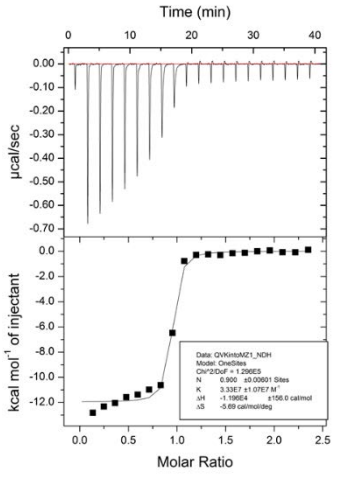
Brd3^{BD2}



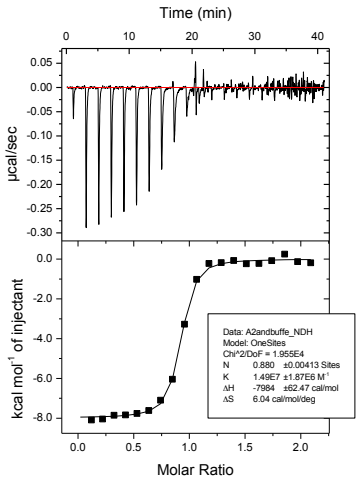
Brd4^{BD2}



Brd2^{BD1} (KEA)

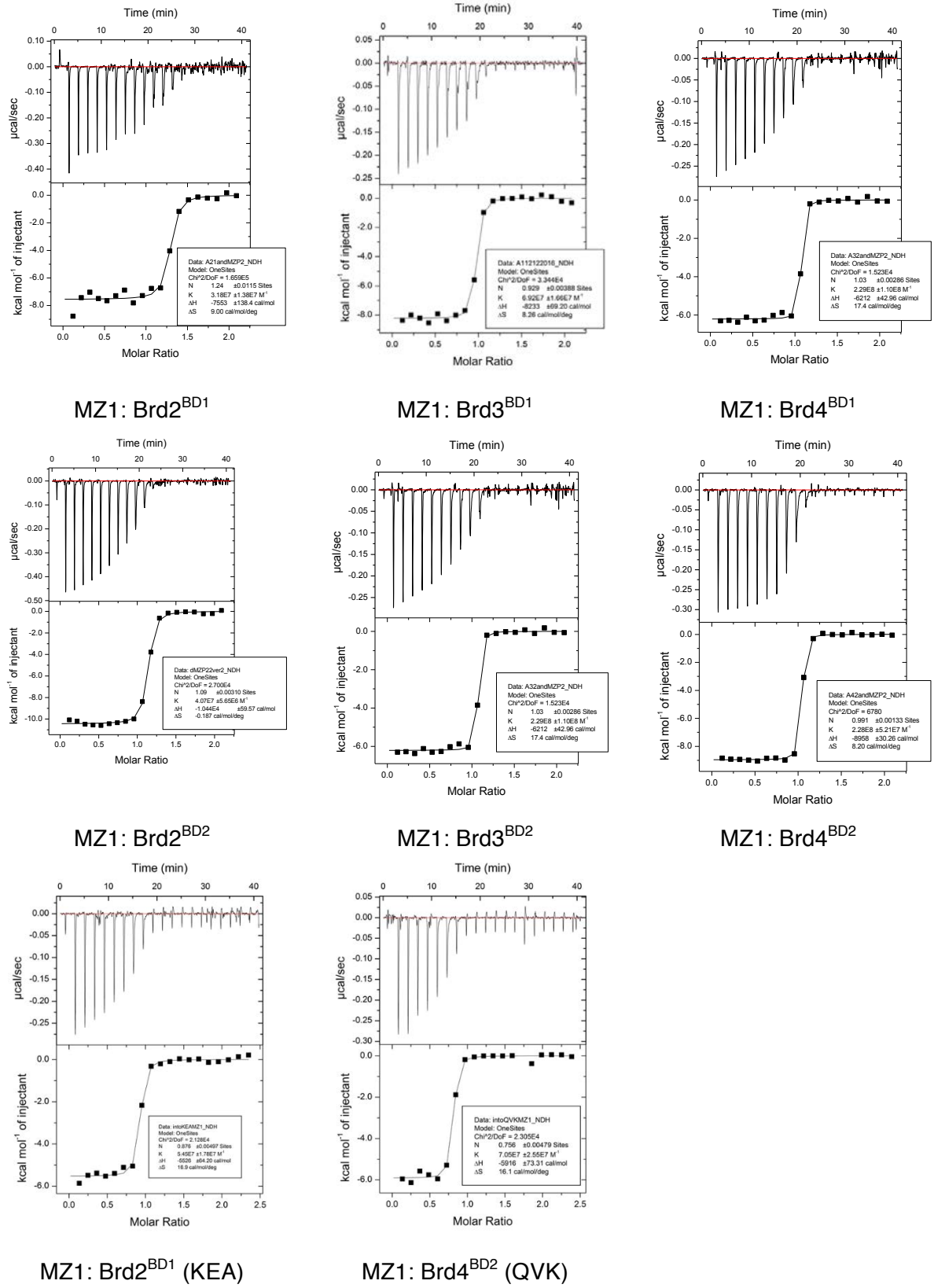


Brd4^{BD2} (QVK)

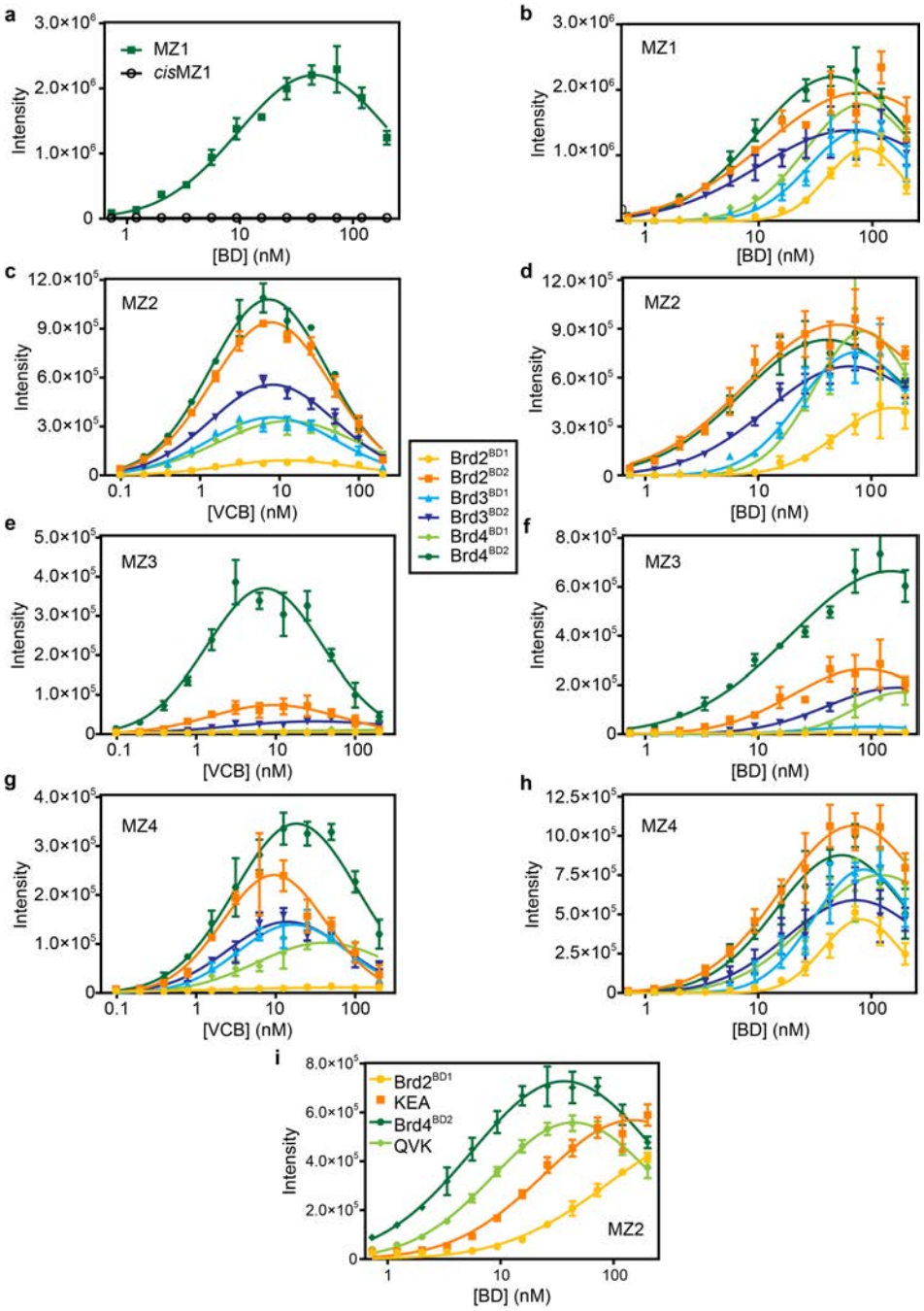


VCB

ITC titrations for ternary complex VCB:MZ1:BD

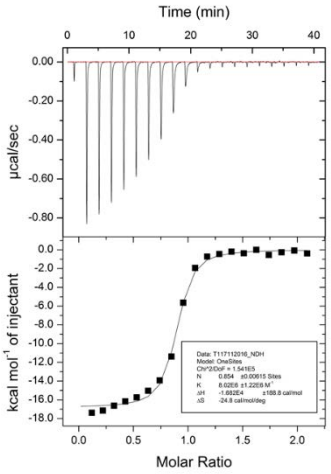


Supplementary Figure 6. Representative ITC titrations to form binary and ternary complexes. Binary complexes: MZ1 (20 µM in the cell), and BET BD (200 µM in the syringe). For VCB titration, MZ1 (16.8 µM in the cell), and VCB (168 µM) in the syringe. Ternary complexes: MZ1:BD (16.8 µM) in the cell, and VCB (168 µM) in the syringe. All titrations were performed at 25 °C.

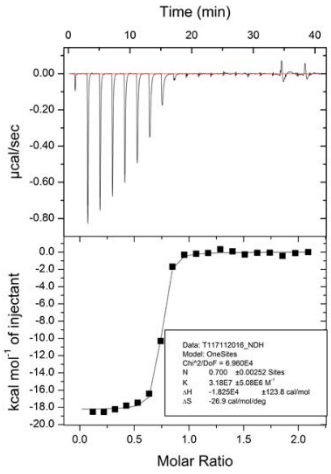


Supplementary Figure 7. Analysis of binary and ternary complex formation by AlphaLISA. a, AlphaLISA intensity values titrating $Brd4^{BD2}$ against VCB with MZ1 (green) or *cis*MZ1 (black circles). **b**, AlphaLISA intensity values titrating each BET-BD against VCB with MZ1. **c**, **e**, **g**, AlphaLISA intensity values titrating VCB against BET-BDs with MZ2 (**c**), MZ3 (**e**), and MZ4 (**g**). **d**, **f**, **h**, AlphaLISA intensity values titrating each BET-BD against VCB with MZ2 (**d**), MZ3 (**f**) and MZ4 (**h**). **i**, AlphaLISA intensity values titrating wild-type and mutant BET-BDs against VCB with MZ2. AlphaLISA intensities are the mean (± 1 s.d.) of intensity values from four technical replicates.

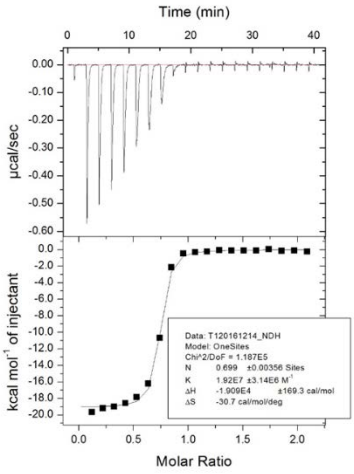
ITC titrations for binary complex AT1:protein



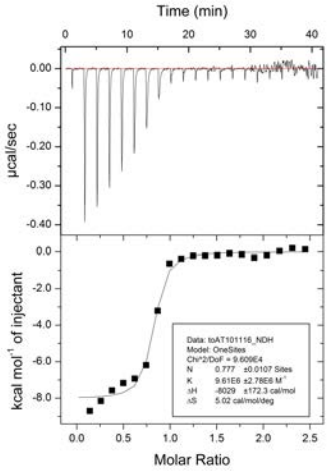
Brd2^{BD1}



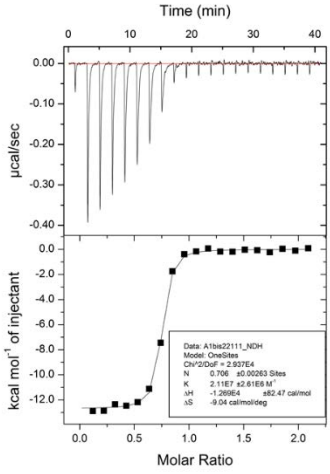
Brd3^{BD1}



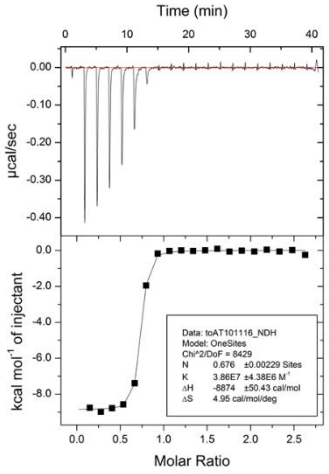
Brd4^{BD1}



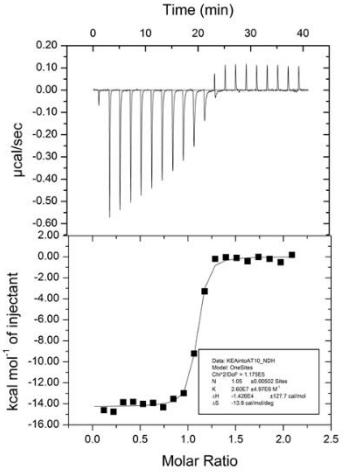
Brd2^{BD2}



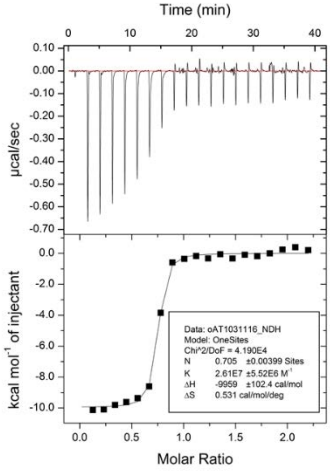
Brd3^{BD2}



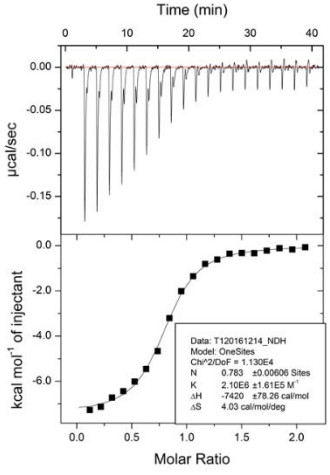
Brd4^{BD2}



Brd2^{BD1} (KEA)

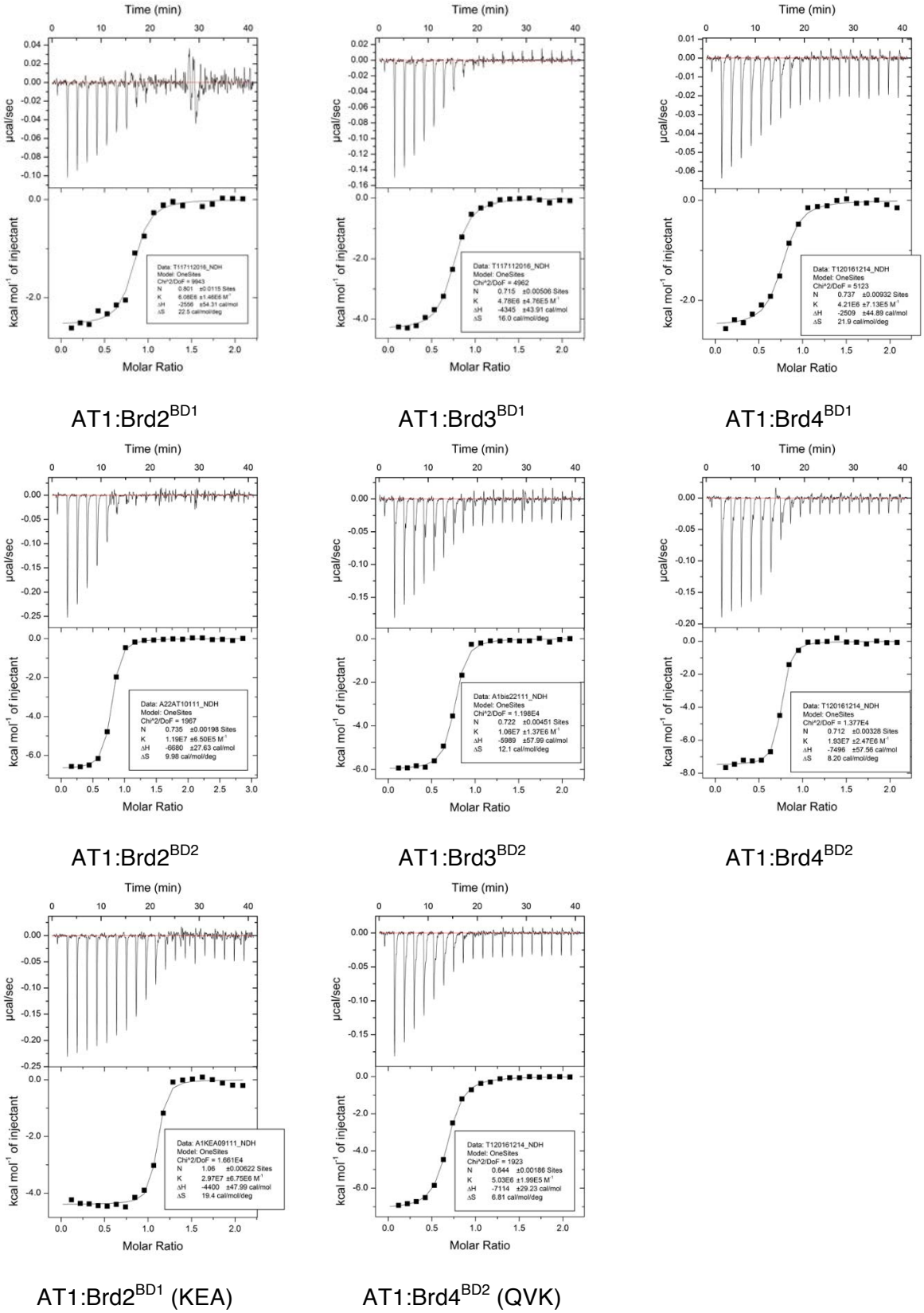


Brd4^{BD2} (QVK)

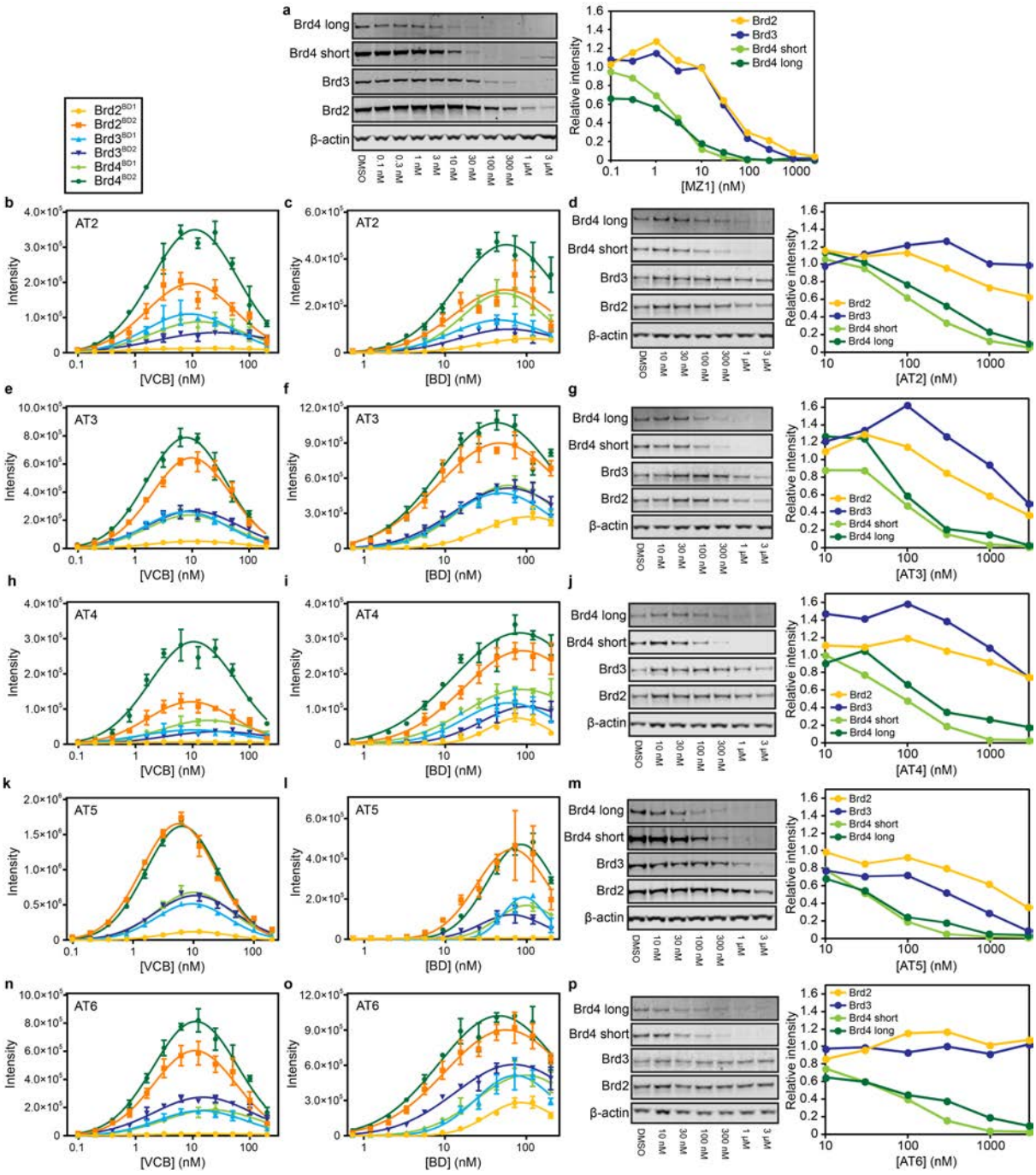


VCB

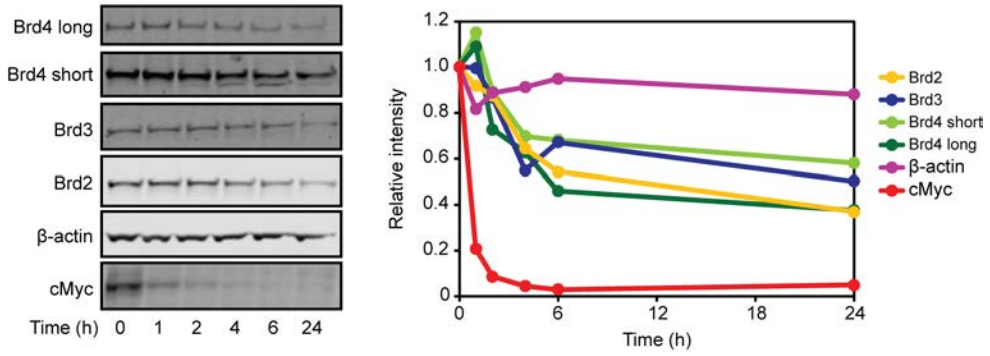
ITC titrations for ternary complex VCB:AT1:BD



Supplementary Figure 8. Representative ITC titrations to form binary and ternary complexes. Binary complexes: AT1 (20 µM in the cell), and BET BD (200 µM in the syringe). For VCB titration, AT1 (16.8 µM in the cell), VCB (168 µM) in the syringe. Ternary complexes: AT1:BD (16.8 µM) in the cell, VCB (168 µM) in the syringe. All titrations were performed at 25 °C.

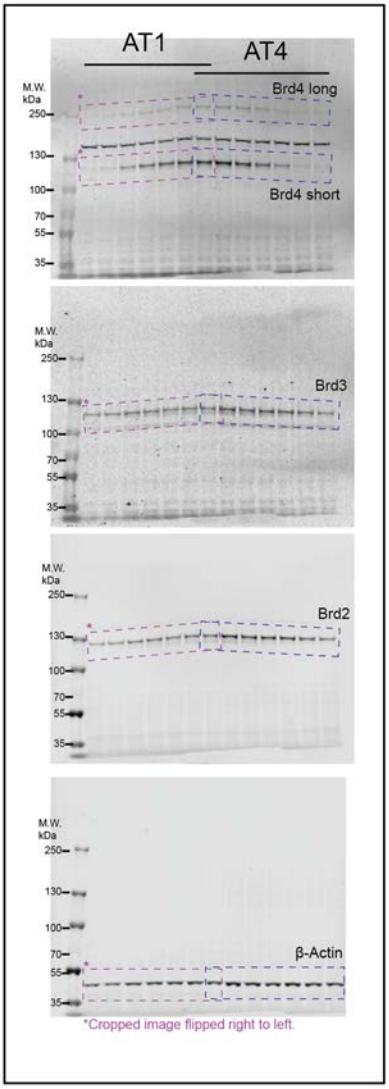


Supplementary Figure 9. Analysis of ternary complex recruitment and cellular BET protein degradation by AT1-6. a, **d**, **g**, **j**, **m**, **p**, Selective intracellular degradation of Brd4 by MZ1 (Ref. [7]) (**a**), AT2 (**d**), AT3 (**g**), AT4 (**j**), AT5 (**m**) and AT6 (**p**) in HeLa cells treated with indicated concentrations of corresponding compound. Protein levels are shown from one representative of three biological replicates, visualised by immunoblot (left) and quantified relative to DMSO control (right). **b**, **e**, **h**, **k**, **n**, AlphaLISA intensity values titrating VCB against BET-BDs with AT2 (**b**), AT3 (**e**), AT4 (**h**), AT5 (**k**) and AT6 (**n**). **c**, **f**, **i**, **l**, **o**, AlphaLISA intensity values titrating each BET-BD against VCB with, AT2 (**c**), AT3 (**f**), AT4 (**i**), AT5 (**l**) and AT6 (**o**). Western blot intensity values were quantified as described in Online Methods. AlphaLISA intensities are the mean (± 1 s.d.) of intensity values from four technical replicates.

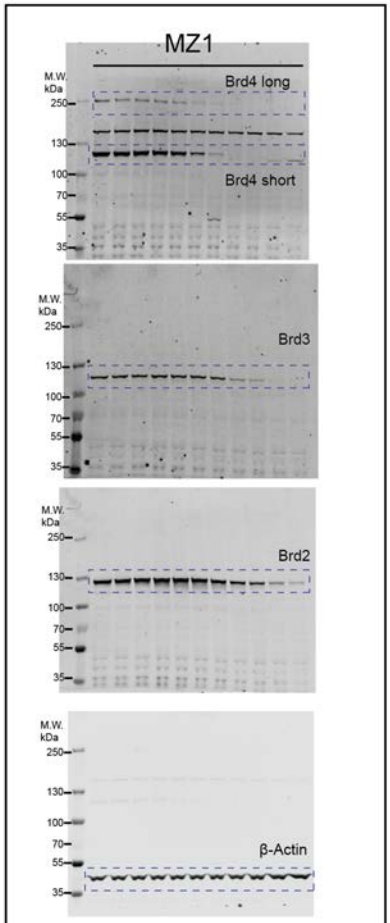


Supplementary Figure 10. Cycloheximide (CHX) chase assay to compare BET protein turnover in cells. HeLa cell protein synthesis was blocked with 100 µg/mL CHX for the given period of time from zero to 24 hours. After the treatment, 30 µg protein extracts from cells were analysed by Western blot (left) probing for BET proteins, cMyc or β-actin to examine depletion rate of these proteins in cells and quantified relative to DMSO control (right). The CHX treatment was effective as levels of cMyc, which is known to have short half-life, were depleted within two hours. All BET proteins were depleted in cells at slower rates compared to cMyc with no distinct differences between Brd2, Brd3 or Brd4. Such observation does not match with the preferential depletion of Brd4 induced by MZ1 and other PROTAC molecules, suggesting that such preference is not due to differences in native protein synthesis or turnover rate. Intensity values were quantified as described in Online Methods and are calculated from one biological replicate.

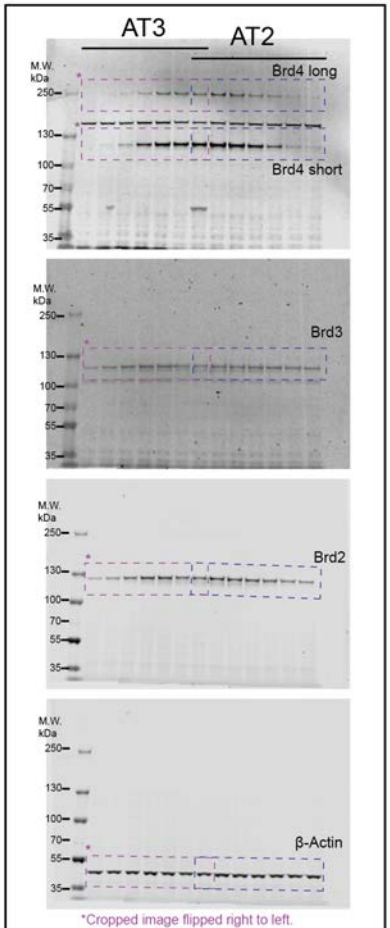
Figure 4e & Supplementary Figure 9j

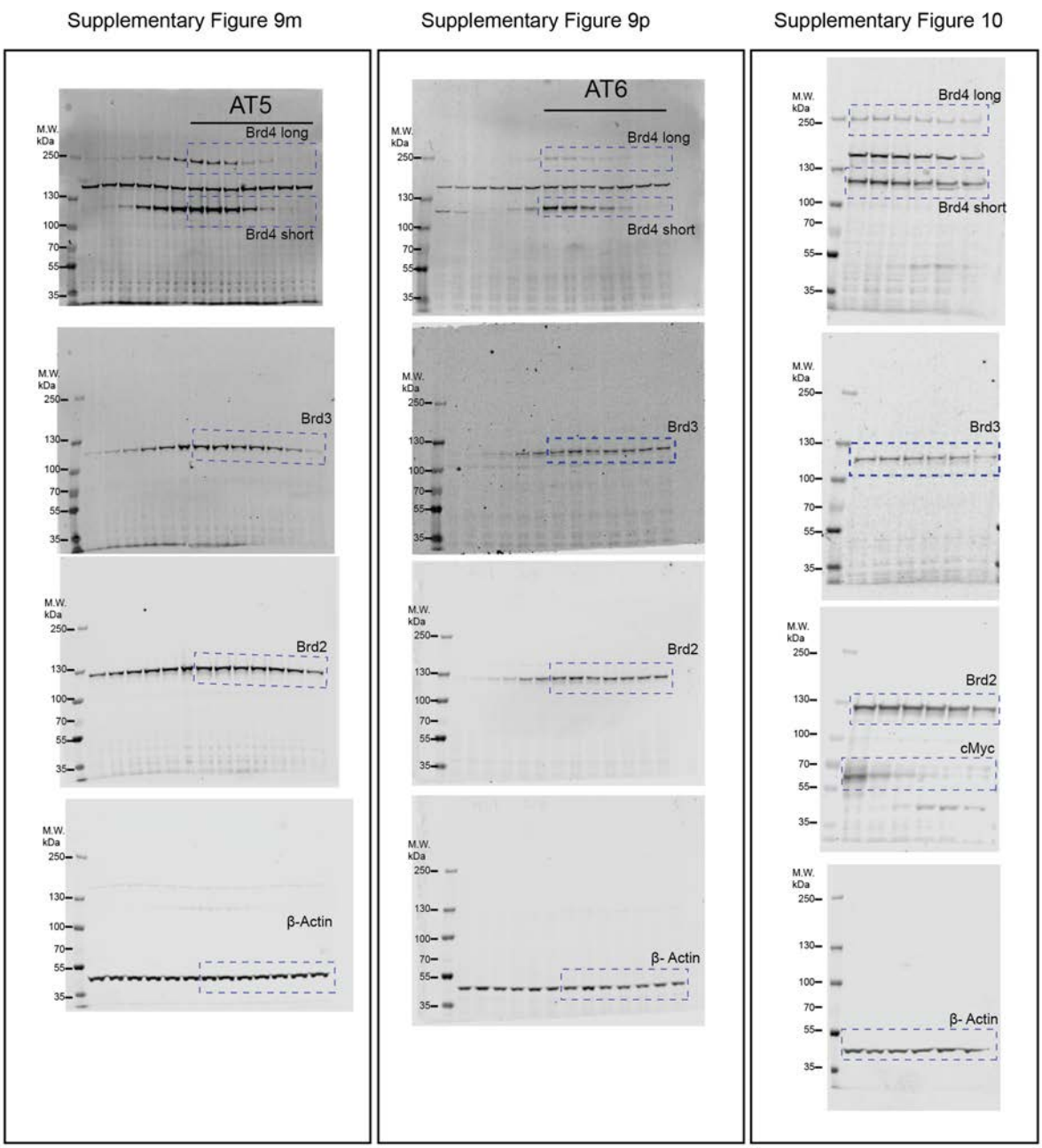


Supplementary Figure 9a

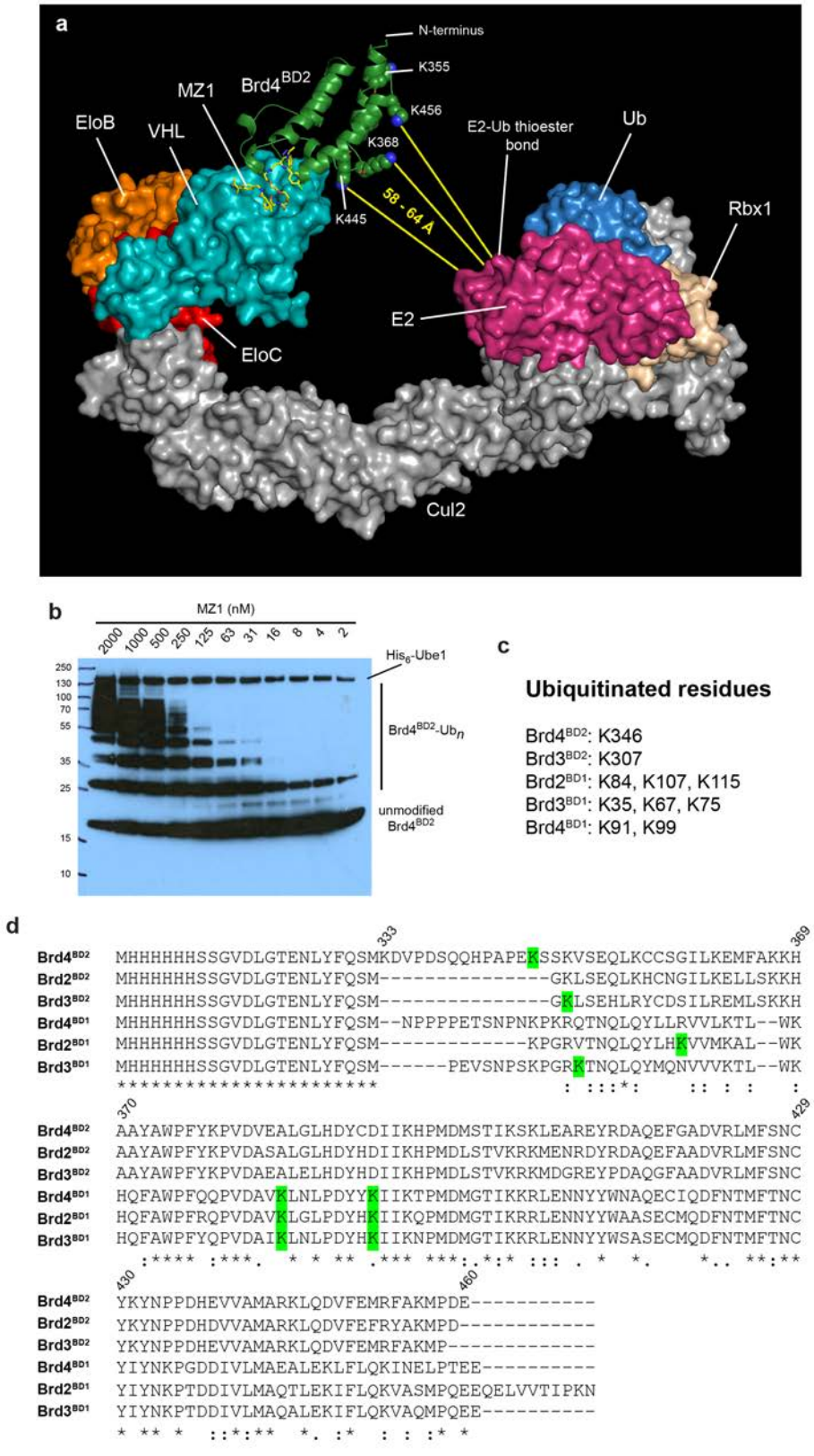


Supplementary Figure 9g,d





Supplementary Figure 11. Full uncut gel images of western blots data.



Supplementary Figure 12. Model of CRL2^{VHL}–MZ1–Brd4^{BD2}, *in vitro* ubiquitination and identification of target lysines. a, Model of CRL2^{VHL} bound to MZ1–Brd4^{BD2}. Solvent-exposed lysines of Brd4^{BD2} are highlighted as spheres and labelled. **b**, *In vitro* ubiquitination of Brd4^{BD2} by CRL2^{VHL} in the presence of increasing concentrations of MZ1. Western blot probing for His_6 purification tag was used to visualize target protein levels. **c,d**, Ubiquitination sites identified *in vitro*

by mass spectrometry, highlighted in green in the sequence alignment. Residue K346 is located at the N-terminal tail of $Brd4^{BD2}$ (as indicated in **a**) and is not visible in our co-crystal structure.

Supplementary Data Set 1:

Proteomic analysis of relative protein abundance in HeLa cells. Results are graphically represented in **Figure 4f,g**. See file "Supplementary Data Set 1.xls"

Supplementary Table 1. Data collection and refinement statistics (molecular replacement)

	$Brd4^{BD2}$:MZ1:pVHL-ElonginC-ElonginB (5T35)
Data collection	
Space group	$P3_2$
Cell dimensions	
a, *b*, *c* (Å)	102.3, 102.3, 144.3
α, β, γ (°)	90, 90, 120
Resolution (Å)	48.2–2.7 (2.79–2.70)*
R_{merge}	7.9 (66.0)
$I/\sigma(I)$	14.7 (2.1)
$CC_{1/2}$	99.8 (57.1)
Completeness (%)	99.7 (99.3)
Redundancy	4.2 (3.7)
Refinement	
Resolution (Å)	48.2–2.7
No. reflections	44157 (3322)
R_{work}	20.6 (32.3)
R_{free}	23.1 (37.4)
No. atoms	
Protein	7191
Ligand	138
Water	86
B factors	
Protein	63.7
Ligand	42.5
Water	39.7
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.14

* Highest-resolution shell is shown in parentheses.

Supplementary Table 2. Breakdown of buried surface areas in the ternary complexes ligase:ligand:target.

Buried surface areas in ternary complexes (Å^2)	VHL: MZ1: Brd4^{BD2} [a]	CRBN: lenalidomide: CK1α [b]	CRBN: CC-885: GSPT1 [c]
Ligase:target PPIs	688	1164	1263
Ligase:ligand	957	468	666
Target:ligand	976	198	461
Total	2621	1830	2390

Buried surface areas were calculated with PISA (see ref. [52]). The more extensive protein:ligand buried surface areas for our PROTAC ternary structure are consistent with the larger bifunctional nature of the compound MZ1 compared to the phthalimide-based ligands. These are in part compensated by the larger PPI contact areas in the CRBN examples, resulting in all cases in extensive total buried surface areas of > 1800 Å^2, as expected for a productive protein-protein interaction.

[a] crystal structure from this study.
[b] crystal structure from ref. [2] (PDB entry 5FQD).
[c] crystal structure from ref. [3] (PDB code 5HXB).

Supplementary Table 3. Thermodynamic parameters of formation of binary and ternary complexes between AT1, VCB and BET bromodomains measured by isothermal titration calorimetry.

Protein in syringe	Species in cell	K_d (nM)	ΔG (kcal × mol^{-1})	ΔH (kcal × mol^{-1})	$-T\Delta S$ (kcal × mol^{-1})	N	α	ΔpK_d
Brd2^{BD1}	AT1	111 ± 14	–9.49 ± 0.07	–16.2 ± 0.6	6.7 ± 0.7	0.858 ± 0.004		
Brd2^{BD2}		94 ± 9	–9.59 ± 0.06	–8.2 ± 0.1	–1.43 ± 0.07	0.81 ± 0.03		
Brd3^{BD1}		35 ± 3	–10.18 ± 0.05	–18.4 ± 0.1	8.2 ± 0.2	0.68 ± 0.02		
Brd3^{BD2}		39 ± 8	–10.1 ± 0.1	–12.73 ± 0.04	2.61 ± 0.09	0.701 ± 0.006		
Brd4^{BD1}		75 ± 23	–9.8 ± 0.2	–18.4 ± 0.7	8.7 ± 0.5	0.73 ± 0.03		
Brd4^{BD2}		44 ± 8	–10.2 ± 0.1	–9.8 ± 0.9	0 ± 1	0.69 ± 0.02		
Brd2^{BD1} KEA		35 ± 4	–10.18 ± 0.06	–15 ± 1	5 ± 1	0.9 ± 0.2		
Brd4^{BD2} QVK		38.8 ± 0.5	–10.11 ± 0.01	–10.9 ± 0.9	0.8 ± 0.9	0.69 ± 0.01		
VCB[a]	AT1[a]	335 ± 30	–8.85 ± 0.06	–6.7 ± 0.4	–2.2 ± 0.5	0.81 ± 0.04		
VCB	AT1:Brd2^{BD1}	280 ± 120	–9.0 ± 0.3	–2.61 ± 0.05	–6.4 ± 0.3	0.88 ± 0.07	1.4	0.1 ± 0.2
	AT1:Brd2^{BD2}	78 ± 6	–9.70 ± 0.05	–7.2 ± 0.6	–2.5 ± 0.5	0.732 ± 0.003	4.1	0.61 ± 0.05
	AT1:Brd3^{BD1}	207 ± 2	–9.12 ± 0.01	–4.38 ± 0.03	–4.75 ± 0.02	0.709 ± 0.006	1.5	0.19 ± 0.04
	AT1:Brd3^{BD2}	79 ± 21	–9.7 ± 0.2	–4.7 ± 0.3	–4.0 ± 0.4	0.723 ± 0.001	4.3	0.6 ± 0.1
	AT1:Brd4^{BD1}	390 ± 150	–8.8 ± 0.2	–2.2 ± 0.3	–6.57 ± 0.04	0.68 ± 0.05	1.0	–0.1 ± 0.2
	AT1:Brd4^{BD2}	46 ± 6	–11.01 ± 0.07	–6.9 ± 0.6	–3.1 ± 0.7	0.68 ± 0.03	7.0	0.84 ± 0.07
	AT1:Brd2^{BD1} KEA	52 ± 18	–10.0 ± 0.2	–5.2 ± 0.2	–5 ± 1	0.9 ± 0.2	7.0	0.8 ± 0.2
	AT1:Brd4^{BD2} QVK	160 ± 40	–9.3 ± 0.1	–6.2 ± 0.9	–3 ± 1	0.70 ± 0.06	2.0	0.3 ± 0.1

All ITC titrations were performed at 25 °C. Values reported are the mean ± s.e.m. from two independent measurements, except for VCB titration into AT1 (line [a]) for which values reported are the mean ± s.e.m. from seven independent measurements.

Supplementary Note

CHEMISTRY

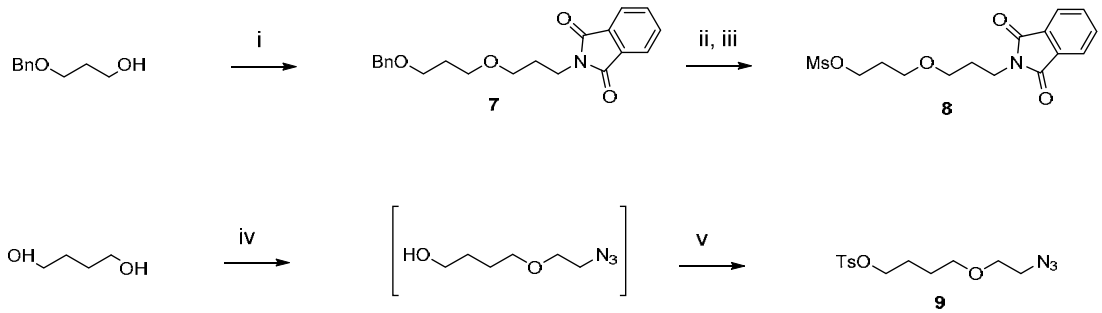
General information

All chemicals, unless otherwise stated, were commercially available and used without further purification. Enantiopure (+)-JQ1 was purchased from Medchemexpress LLC, Princeton, USA and deprotected to the free acid (+)-JQ1-COOH as previously described.[1,2] MZ1, MZ2 and MZ3 were synthesized as previously described.[1] All reactions were carried out in oven- or flame-dried glassware under nitrogen atmosphere. Normal phase TLC was carried out on pre-coated silica plates (Kieselgel 60 F254, BDH) with visualization via UV light (UV 254/365 nm) and/or basic potassium permanganate solution. All commercially available reagents were used as received. Reactions were magnetically stirred; commercially available anhydrous solvents were used. Flash column chromatography (FCC) was performed using a Teledyne Isco Combiflash Rf or Rf200i, prepacked columns RediSep Rf Normal Phase Disposable Columns were used. NMR spectra were recorded on a Bruker Ascend 400. Chemical shifts are quoted in ppm and referenced to the residual solvent signals: ^{1}H δ = 7.26 ($CDCl_3$), ^{13}C δ = 77.16 ($CDCl_3$); ^{1}H δ = 3.31 (MeOD), ^{13}C δ = 49.15 (MeOD); signal splitting patterns are described as singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m), broad (br), exchangeable hydrogen with deuterium (exch) or a combination thereof. Coupling constants ($J_{H\text{-}H}$) are measured in Hz. High Resolution Mass Spectra (HRMS) were recorded on a Bruker microTOF. Low resolution MS and analytical HPLC traces were recorded on an Agilent Technologies 1200 series HPLC connected to an Agilent Technologies 6130 quadrupole LC/MS, connected to an Agilent diode array detector. The column used was a Waters XBridge column (50 mm × 2.1 mm, 3.5 µm particle size) and the compounds were eluted with a gradient of 5−95% acetonitrile/water + 0.1% ammonia (referred in the text as "basic method") or a gradient of 5−95% acetonitrile/water + 0.1 formic acid ("acidic method").

Preparative HPLC was performed on a Gilson Preparative HPLC System with a Waters X-Bridge C18 column (100 mm x 19 mm; 5 µm particle size) and a gradient of 5 % to 95 % acetonitrile in water over 10 minutes, flow 25 mL/min, with 0.1 % ammonia in the aqueous phase.

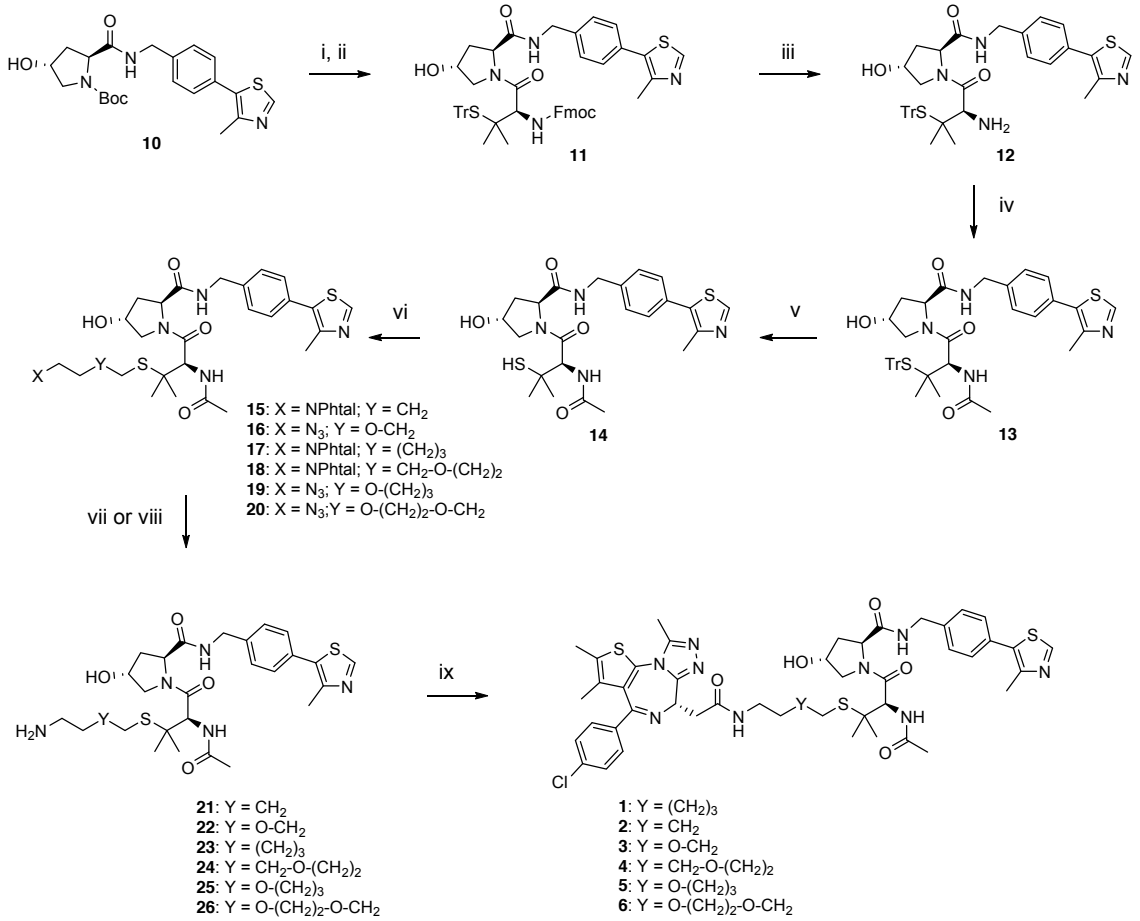
Abbreviations used: DCM for dichloromethane, EtOAc for ethyl acetate, Et_2O for diethyl ether, DMSO for dimethyl sulfoxide, DIPEA for *N,N*-diisopropylethylamine, MeOH for methanol, TEA for triethylamine, DMF for *N,N*-dimethylformamide, HATU for 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate, HOAT for 1-hydroxy-7-azabenzotriazole, DBU for 1,8-diazabicyclo[5.4.0]undec-7-ene, TIPS for triisopropylsilane, TFA for trifluoroacetic acid.

Synthesis of linkers



Scheme 1. Reagents and conditions: i) *N*-(3-Bromopropyl)phthalimide , NaH in DMF 0 °C to r.t., 75%; ii) H_2 (balloon), Pd/C 10% dry, in MeOH, 99%; iii) mesylchloride, triethylamine in DCM, 0 °C, 95%; iv) 2-azidoethyl 4-methylbenzenesulfonate, triethylamine in DCM, 0 °C to r.t.; v) tosylchloride, triethylamine in DCM, 0 °C to rt, 25% over two steps.

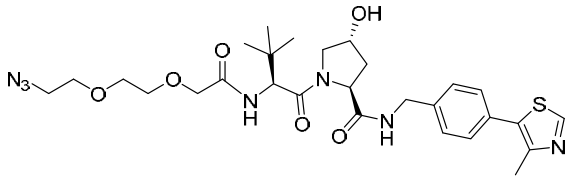
Synthesis of PROTACs AT1–6



Scheme 2. Reagents and conditions: i) HCl 2M in dioxane/DCM 1:1, r.t., 99%; Fmoc-S-trityl-L-penicillamine, HATU, HOAT, DIPEA in DMF, r.t.; iii) Piperidine 20% in DCM, r.t., 75% over two steps; iv) acetic anhydride, triethylamine in DCM 0 °C to r.t., 98%; v) TFA- TIPS 5% in DCM, r.t., 79%; vi) Linker-OMs or Linker-OTs or Linker-Br , DBU, in DMF 0 °C to r.t.,70-82%; vii) if X = N_3: H_2 (balloon), Pd/C 10%, in MeOH, 99%; viii) if X = NPhtal: Hydrazine hydrate in ethanol, 70 °C, 60-68%; (+)-JQ1-COOH, HATU, HOAT, DIPEA in DMF, r.t. 33-70%.

Synthetic methods for the preparation of PROTAC MZ4

(*2S,4R*)-1-((*S*)-2-(2-(2-(2-azidoethoxy)ethoxy)acetamido)-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (27)

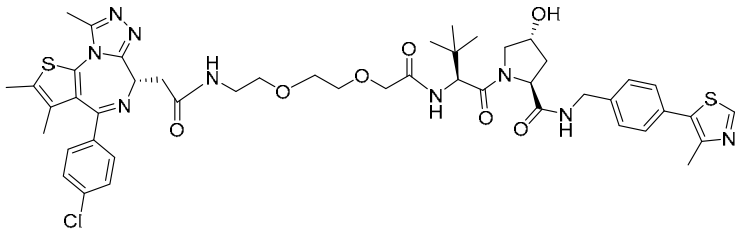


Prepared as previously described.[1] Obtained 41.1 mg, 68% yield. MS analysis: $C_{28}H_{39}N_7O_6S$ expected 601.3, found 602.3 [M+H$^+$].

^{1}H-NMR ($CDCl_3$, 400 MHz) δ: 8.68 (s, 1H), 7.38-7.33 (m, 5H), 7.24 (d, J = 8.5 Hz, 1H), 4.75 (t, J = 7.9 Hz, 1H), 4.59-4.54 (m, 2H), 4.48 (d, J = 8.6 Hz, 1H), 4.33 (dd, J = 14.9 Hz, J = 5.2 Hz, 1H) 4.12-4.09 (m, 1H), 4.06-3.96 (m, 2H), 3.70-3.66 (m, 6H), 3.61 (dd, J = 11.4 Hz, J = 3.7 Hz, 1H), 3.38-3.41 (m, 2H), 2.89 (br s, 1H), 2.63-2.58 (m, 1H), 2.52 (s, 3H), 2.08-2.13 (m, 1H), 0.95 (s, 9H).

^{13}C-NMR ($CDCl_3$, 101 MHz) δ: 171.6, 170.7, 170.6, 150.7, 138.4, 130.4, 129.7, 128.6, 128.4, 127.6, 71.2, 70.5, 70.3, 70.2, 67.2, 58.5, 57.3, 56.8, 50.7, 43.4, 35.8, 34.9, 26.5, 16.0

(*2S,4R*)-1-((*S*)-2-(*tert*-butyl)-14-((S)-4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)-4,13-dioxo-6,9-dioxa-3,12-diazatetradecanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (MZ4)



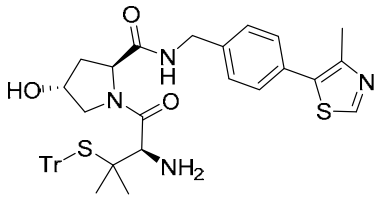
Prepared as previously described from azide **27** (ref. [1]). Yield: 22.2 mg (66 %); HRMS analysis: $C_{46}H_{57}N_9ClO_7S_S$ expected 957.3505, found 958.3498 [M+H$^+$].

^{1}H-NMR ($CDCl_3$, 400 MHz) δ: 8.65 (s, 1H), 9.30-8.325 (m, 2H), 7.63 (d, J = 10.0 Hz, 1H), 7.36 (d, J = 8.7 Hz, 2H), 7.24 (d, J = 8.3 Hz, 2H), 7.11 (d, J = 8.0 Hz, 2H), 7.02 (d, J = 8.0 Hz, 2H), 4.94-4.86 (m, 2H), 4.63-4.57 (m, 2H), 4.21 (d, J = 5.8 Hz, 2H), 4.07-3.81 (m, 5H), 3.72-3.52 (m, 7H), 3.44 (dd, J = 15.9 Hz, J = 3.0 Hz, 1H), 3.13-3.08 (m, 1H), 2.55 (s, 3H), 2.45 (s, 3H), 2.39-2.35 (m, 4H), 2.27-2.20 (m, 1H) 1.66 (s, 3H), 1.09 (s, 9 H).

^{13}C-NMR ($CDCl_3$, 101 MHz) δ: 172.1, 171.1, 170.6, 170.4, 163.3, 156.2, 150.2, 149.9, 148.4, 138.5, 136.9, 136.7, 131.9, 131.6, 131.4, 131.3, 131.2, 130.2, 130.1, 129, 128.8, 127.7, 71.5, 70.4, 70.3, 69.8, 59.3, 57.5, 56.5, 53.9, 42.7, 39.7, 38.3, 37.2, 36.3, 26.6, 16.2, 14.5, 13.3, 11.8.

Synthetic methods for the preparation of PROTACs 1–6

(*2S,4R*)-1-((*R*)-2-amino-3-methyl-3-(tritylthio)butanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (12)



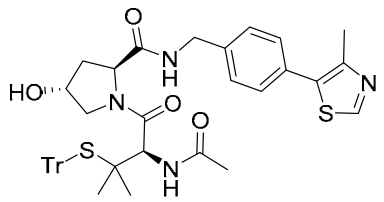
To a solution of **10** (ref. [3]) (418 mg, 1 mmol) in DCM/MeOH 9:1 (2.5 mL), 4N HCl in dioxane (2.5 mL) was added and mixture was stirred for 4 hours at room temperature. Volatiles were removed under vacuum, the residue was dissolved in water and freeze-dried to afford (*2S,4R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-ium chloride (350 mg, 99%) as a light yellow powder which was used without any further purification.

To a solution of Fmoc-S-trityl-L-penicillamine (500 mg, 0.80 mmol) in DMF (2 mL), HATU (304 mg, 0.80 mmol) and HOAT (108 mg, 0.80 mmol) were added, followed by DIPEA (250 µL, 1.2 mmol). The bright yellow solution was then added to a mixture of (*2S,4R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-ium chloride (281 mg, 0.80 mmol) and DIPEA (250 µL, 1.2 mmol) in DMF (3 mL). After 2 hours complete conversion of the starting materials was observed by HPLC-MS (basic method), water was added (10 mL) and the mixture was extracted with AcOEt (3X 25 mL). The organic phase was washed with brine (10 mL) and dried over anhydrous $MgSO_4$. Solvents were removed under vacuum to afford compound **11** which was dissolved in DCM (4 mL). Piperidine (800 µL ~ 8 mmol) was added and the reaction mixture was stirred for 1 hour. Volatiles were removed under vacuum and the crude was purified by FCC (from 5 to 15 % of 0.7 M NH_3 in MeOH in DCM) to afford the title compound **12** as a white solid (414 mg, 75% yield). MS analysis: $C_{40}H_{42}N_4O_5S_2$ expected 690.3, found 691.2 [$M+H^+$].

^{1}H-NMR (400 MHz, CD_3OD, 25 °C) δ: 8.90 (s, 1H), 7.63-7.60 (m, 6H), 7.40-7.33 (m, 4H), 7.31-7.29 (m, 6H), 7.24-7.19 (m, 3H), 4.46 (t, J = 8.2 Hz, 1H), 4.37 (br s, 1H), 4.31 (m, 2H), 3.35 (s, 1H), 3.24 (dd, J = 11.1 Hz, J = 4.1 Hz, 1H), 3.07-3.04 (m, 1H), 2.70 (s, 1H), 2.47 (s, 3H), 2.16-2.11 (m, 1H), 1.99-1.92 (m, 1H), 1.26 (s, 3H), 1.19 (s, 3H).

^{13}C-NMR (101 MHz, CD_3OD , 25 °C) δ: 172.9, 171.4, 151.5, 147.7, 144.9, 138.7, 132.0, 129.7, 129.0, 127.5, 127.4, 126.5, 69.3, 67.7, 59.4, 57.6, 57.3, 56.7, 42.1, 37.4, 24.6, 24.1, 14.4.

(*2S,4R*)-1-((*R*)-2-acetamido-3-methyl-3-(tritylthio)butanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (13)

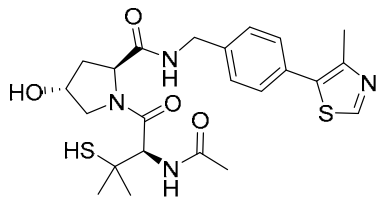


To a solution of **12** (214 mg, 0.31 mmol) in DCM (2 mL) at 0 °C, TEA (50 µL, 0.35 mmol) and acetic anhydride (30 µL, 0.32 mmol) were added. The mixture was let to react at room temperature for 2 hours. The mixture was diluted with DCM (10 mL), washed with water (2 mL) and brine (2 mL), dried over $MgSO_4$ and the solvent was removed under reduced pressure to afford the title compound **13** (222 mg, 98% yield) which was used without further purification. MS analysis: $C_{42}H_{44}N_4O_4S_2$ expected 732.3, found 733.3 [$M+H^+$].

^{1}H-NMR (400 MHz, $CDCl_3$, 25 °C) δ: 8.71 (s, 1H), 7.52-7.49 (m, 6H), 7.39-7.31 (m, 3H), 7.25-7.20 (m, 11H), 6.25 (d, J = 5.2 Hz, 1H), 4.64 (t, J = 8.1 Hz, 1H), 4.37 (br s, 1H), 4.31-4.18 (m, 2H), 3.61 (d, J = 5.3 Hz, 1H), 3.52-3.49 (m, 1H), 3.30-3.29 (m, 1H), 3.22 (dd, J = 11.5Hz, J = 3.6 Hz, 1H), 2.52 (s, 3H), 2.36-2.30 (m, 1H), 2.16-2.11 (m, 1H), 1.95 (s, 3H), 1.76 (br s, 1H), 1.18 (s, 3H), 0.97 (s, 3H).

^{13}C-NMR (101 MHz, $CDCl_3$, 25 °C) δ: 170.7, 170.5, 170.3, 150.3, 148.5, 144.2, 138.2, 131.7, 130.8, 129.7, 129.6, 129.4, 127.9, 127.0, 77.2, 70.1, 68.5, 58.5, 57.1, 56.6, 53.6, 42.8, 36.4, 26.1, 25.4, 22.9, 16.2.

(*2S,4R*)-1-((*R*)-2-acetamido-3-mercapto-3-methylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (14)

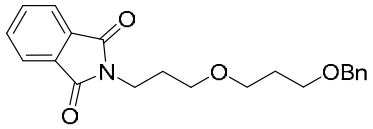


Compound **13** (114 mg, 0.16 mmol) was dissolved in 3.8 mL of DCM. TIPS (0.2 mL) and TFA (0.2 mL) were added, and the yellow mixture was let to react at room temperature for 2 h. HPLC analysis (acidic method) showed complete conversion of the starting material. Volatiles were removed and the crude was dissolved in MeOH, filtered and purified by preparative HPLC and freeze-dried to give pure compound **14** as white solid (62 mg, 79% yield). MS analysis: $C_{23}H_{30}N_4O_4S_2$ expected 490.2, found 491.1 [M+H$^+$].

^{1}H-NMR (400 MHz, $CDCl_3$, 25 °C) δ: 8.68 (s, 1H), 7.39-7.33 (m, 4H), 7.20-7.17 (m, 1H), 6.55 (d, 1H), 4.68 (t, J = 8.0 Hz, 1H), 4.59-4.52 (m, 3H), 4.35 (dd, J = 14.9 Hz, J = 5.2 Hz, 1H), 4.19-4.16 (m, 1H), 3.70 (dd, J = 11.2 Hz, J = 3.6 Hz, 1H), 3.15 (br s, 1H), 2.29 (br s, 1H), 2.52 (s, 3H), 2.48-2.42 (m, 1H), 2.20-2.15 (m, 1H), 2.01 (s, 3H), 1.36 (s, 3H), 1.31 (s, 3H).

^{13}C-NMR (101 MHz, $CDCl_3$, 25 °C) δ: 170.9, 170.7, 170.6, 150.4, 148.6, 137.9, 131.5, 131.1, 129.6, 128.1, 70.1, 58.9, 57.5, 56.6, 46.1, 43.3, 36.5, 30.7, 28.7, 23.0, 16.1.

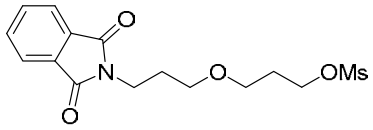
2-(3-(3-(benzyloxy)propoxy)propyl)isoindoline-1,3-dione (7)



To a suspension of NaH (60% in mineral oil, 288 mg, 7.2 mmol) in DMF (10 mL) 3-(benzyloxy)-1-propanol (1.00 g, 6 mmol) was added dropwise at 0 °C. After 45 min a solution of *N*-(4-bromopropyl)phthalimide (2.41 g, 9 mmol) in DMF (5 mL) was added dropwise. The mixture was stirred overnight, treated with $KHSO_4$ 5% to acidic pH and poured in water (50mL). The mixture was extracted with Et_2O (3 X 30 mL), the organic phase was washed with brine (10 mL) and dried over $MgSO_4$. Solvents were removed under reduced pressure and the crude was purified by FCC (10% to 30% of EtOAc in heptane) to give the title compound **7** as a viscous oil (1.59 g, 75%). MS analysis: $C_{21}H_{23}NO_4$ expected 353.2, found 354.2 [M+H$^+$].

^{1}H-NMR (400 MHz, $CDCl_3$, 25 °C) δ: 7.84-7.82 (m, 2H), 7.71-7.69 (m, 2H), 7.33-7.27 (m, 5H), 4.47 (s, 2H), 3.78 (t, J = 6.9 Hz, 2H), 3.52-3.45 (m, 6H), 1.94 (q, J = 6.5 Hz, 2H), 1.79 (q, J = 6.3 Hz, 2H).

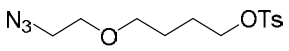
3-(3-(1,3-dioxoisoindolin-2-yl)propoxy)propyl methanesulfonate (8)



Compound **7** (1.00 g, 2.83 mmol) was dissolved in MeOH (50 mL) and a catalytic amount of Pd/C (10%) was added. The mixture was stirred under hydrogen atmosphere and TLC analysis (Heptane/EtOAc 7:3) showed complete conversion of the starting material after 3 h. The mixture was filtered through a celite pad to remove the catalyst and evaporated to dryness to give the debenzylated product as a transparent oil (745 mg, quantitative yield). A portion of the crude compound (122 mg, 0.463 mmol) was dissolved in DCM (2 mL) and treated at 0 °C with TEA (78 µL, 0.555 mmol) and mesylchloride (40 µL, 0.500 mmol). The mixture was left to react for 4 h in an ice bath, then quenched with water (1 mL) and extracted with DCM (3 X 5 mL). The organic phase was dried over $MgSO_4$. Solvents were removed under reduced pressure and the crude mesylate **8** (155 mg, considered quantitative yield) was used without further purification.

^{1}H-NMR (400 MHz, $CDCl_3$, 25 °C) δ: 7.85-7.83 (m, 2H), 7.72-7.70 (m, 2H), 4.29 (t, J = 6.3 Hz, 2H), 3.79 (t, J = 6.8 Hz, 2H), 3.49-3.46 (m, 4H), 3.01 (s, 3H), 1.97-1.89 (m, 4H).

4-(2-azidoethoxy)butyl 4-methylbenzenesulfonate (9)



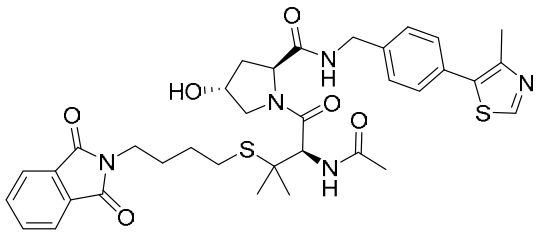
To a solution of 1,4-butanediol (220 µL, 2.48 mmol) in DMF (1 mL) NaH (60% in mineral oil, 33 mg, 0.829 mmol) was added at 0 °C. After 45 minutes, 2-azidoethyl 4-methylbenzenesulfonate[4] (200 mg, 0.829 mmol) was added dropwise at 0 °C. The reaction was stirred at room temperature overnight, TLC analysis (heptane/EtOAc 1:1) showed complete conversion of the starting material. The reaction was quenched with water (0.5 mL) and extracted with EtOAc (3 X 3 mL). The organic phase was dried over $MgSO_4$. Solvents were removed under reduced pressure and the crude was dissolved in DCM (2 mL), cooled to 0 °C and treated with TEA (115 µL, 0.829 mmol) and tosylchloride (158 mg, 0.829 mmol). The mixture was stirred at room temperature for 4 h, then water (2 mL) was added and the mixture was extracted with DCM (3 x 5 mL). The organic phase was dried over $MgSO_4$. Solvents were removed under reduced pressure and the crude was purified by FCC (from 10% to 40% of EtOAc in heptane) to give the pure title compound **9** as a transparent oil (65 mg, 25% yield over two steps). MS analysis: $C_{13}H_{19}N_3O_4S$ expected 313.1, found 314.0 [$M+H^+$].

^{1}H-NMR (400 MHz, $CDCl_3$, 25 °C) δ: 7.79 (d, J = 8.2 Hz, 2H), 7.34 (d, J = 9.0 Hz, 2H), 4.06 (t, J = 6.3 Hz, 2H), 3.57 (t, J = 3.5 Hz, 2H), 3.44 (t, J = 6.0 Hz, 2H), 3.32 (t, J = 4.9 Hz, 2H), 2.45 (s, 3H), 1.78-1.73 (m, 2H), 1.64-1.61 (m, 2H).

General procedure for the *S*-alkylation of 14

Under nitrogen and at 0 °C, a solution of compound **14** (10 mg, 0.020 mmol) in DMF (0.5 mL) was treated with DBU (3.3 µL, 0.022 mmol) followed by the alkylating reagent (0.022 mmol). The reaction mixture was let to react at room temperature until complete conversion of the starting material was observed by HPLC (acidic method, 1-3 h). The reaction was cooled to 0 °C and treated with few drops of $KHSO_4$ (5 %) to pH = 3-4. The solvent was removed under vacuum and the crude was dissolved in MeOH, filtered and purified by preparative HPLC to isolate the expected product.

(*2S,4R*)-1-((*R*)-2-acetamido-3-((4-(1,3-dioxoisoindolin-2-yl)butyl)thio)-3-methylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (15)

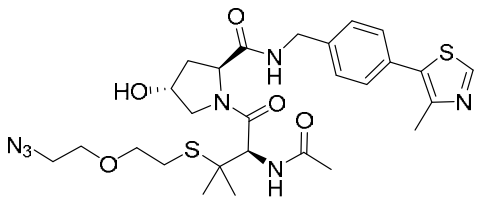


Prepared following the general procedure for the *S*-alkylation from compound **14** (10.0 mg, 0.020 mmol) and *N*-(4-Bromobutyl)phthalimide (6.2 mg, 0.022 mmol). Obtained 11.3 mg, 82% yield. MS analysis: $C_{35}H_{41}N_5O_6S_2$ expected 691.2, found 692.1 [$M+H^+$].

^{1}H NMR (400 MHz, $CDCl_3$, 25 °C) δ: 8.67 (s, 1H), 7.85-7.83 (m, 2H), 7.74-7.72 (m, 2H), 7.46-7.43 (m, 1H), 7.37-7.31 (m, 4H), 6.51 (d, J = 8.0 Hz, 1H), 4.77-4.72 (m, 2H), 4.52-4.38 (m, 3H), 4.08-4.05 (m, 1H), 3.74 (dd, J = 11.1 Hz, J = 3.8 Hz, 1H), 3.66 (t, J = 7.1 Hz, 2H) 2.60-2.48 (m, 5H), 2.40-2.34 (m, 1H), 2.30-2.25 (m, 1H), 2.00 (s, 3H), 1.77-1.70 (m, 2H), 1.55-1.47 (m, 2H), 1.32 (s, 3H), 1.30 (s, 3H).

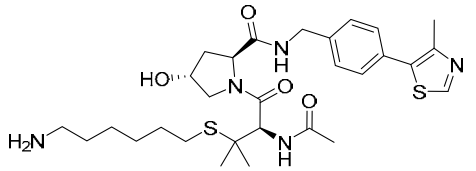
^{13}C-NMR (101 MHz, $CDCl_3$, 25 °C) δ: 171.2, 170.8, 170.5, 168.5, 150.4, 148.6, 138.2, 134.2, 132.1, 131.7,131.0, 129.5, 128.2, 123.4, 70.1, 59.3, 56.9, 56.3, 48.1, 43.1, 37.4, 37.1, 28.0, 27.6, 26.6, 25.8, 25.6, 23.1, 16.2.

(*2S,4R*)-1-((*R*)-2-acetamido-3-((2-(2-azidoethoxy)ethyl)thio)-3-methylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (16)



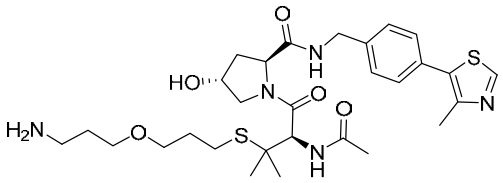
Prepared following the general procedure for the *S*-alkylation from compound **14** (10.0 mg, 0.020 mmol) and 2-(2-azidoethoxy)ethyl methanesulfonate[5] (4.6 mg, 0.022 mmol). Obtained 9.4 mg, 78% yield. MS analysis: $C_{27}H_{37}N_7O_5S_2$ expected 603.2, found 604.1 [M+H$^+$].
^{1}H NMR (400 MHz, $CDCl_3$, 25 °C) δ: 8.68 (s, 1H), 7.39-7.31 (m, 5H), 6.39 (d, J= 8.0 Hz, 1H), 4.76-4.71 (m, 2H), 4.52 (br s, 1H), 4.45 (d, J= 5.9 Hz, 1H), 4.09-4.07 (m, 1H), 3.73 (dd, J = 11.2 Hz, J = 3.8 Hz, 1H), 3.58-3.50 (m, 4H), 3.34-3.31 (m, 2H), 2.76-2.68 (m, 2H), 2.51 (s, 3H), 2.43-2.38 (m, 1H), 2.27-2.23 (m, 1H), 2.00 (s, 3H), 1.32 (s, 3H), 1.29 (s, 3H).
^{13}C-NMR (101 MHz, $CDCl_3$, 25 °C) δ: 170.8, 170.7, 170.4, 150.3, 148.5, 138.1, 131.6, 131.0, 129.5, 128.1, 70.4, 70.1, 69.8, 59.0, 56.7, 56.3, 56.2, 50.6, 48.5, 47.8, 43.1, 42.9, 36.8, 29.6, 28.2, 25.7, 25.6, 23.0, 16.1.

(*2S,4R*)-1-((*R*)-2-acetamido-3-((6-aminohexyl)thio)-3-methylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (23)



Prepared following the general procedure for the *S*-alkylation from **14** (10.0 mg, 0.020 mmol) and *N*-(4-Bromohexyl)phthalimide (6.6 mg, 0.022 mmol), obtaining 11.5 mg (80% yield) of **17**. MS analysis: $C_{37}H_{45}N_5O_6S_2$ expected 719.3, found 720.3 [M+H$^+$]. The alkylated product **17** was then dissolved in ethanol (0.5 mL) and treated with hydrazine monohydrate (24 µL, 0.32 mmol) at 60 °C for two hours. The reaction mixture was cooled to room temperature, filtered and purified by preparative HPLC to give the expected amine **23** (6.4 mg, 68% yield). MS analysis: $C_{29}H_{43}N_5O_4S_2$ expected 589.3, found 590.2 [M+H$^+$].
^{1}H NMR (400 MHz, CD_3OD, 25 °C) δ: 8.88 (s, 1H), 7.46-7.41 (m, 4H), 4.92 (s, 1H), 4.58 (t, J = 8.3 Hz 1H), 4.52-4.38 (m, 3H), 3.93-3.84 (m, 1H), 3.85 (dd, J = 10.8 Hz, J = 4.0 Hz, 1H), 2.64 (t, J = 7.3 Hz, 2H), 2.56 (t, J = 7.4 Hz, 2H) 2.48 (s, 3H), 226-2.23 (m, 1H), 2.14-2.10 (m, 1H), 2.00 (s, 3H), 1.49-1.28 (m, 16 H).
^{13}C-NMR (101 MHz, CD_3OD, 25 °C) δ: 174.4, 173.2, 171.6, 153.0, 149.2, 140.3, 133.5, 131.8, 130.6, 129.2, 71.1, 61.1, 58.0, 57.3, 43.7, 42.3, 39.2, 32.8, 30.7, 30.1, 29.3, 27.6, 26.3, 25.7, 22.5, 16.0.

(*2S,4R*)-1-((*R*)-2-acetamido-3-((3-(3-aminopropoxy)propyl)thio)-3-methylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (24)

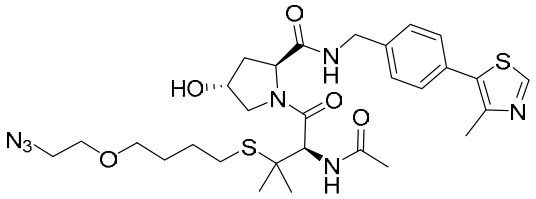


Prepared following the general procedure for *S*-alkylation from **14** (10.0 mg, 0.020 mmol) and **8** (7.5 mg, 0.022 mmol) to obtained compound **18** (11.0 mg, 75% yield). MS analysis: $C_{37}H_{45}N_5O_7S_2$ expected 735.3, found 736.3 [M+H$^+$]. Deprotection of the amino group was performed as described above, with hydrazine (24 µL, 0.32 mmol) in ethanol (0.5 mL) at 60 °C to give 5.4 mg of **24** (60% yield). MS analysis: $C_{29}H_{43}N_5O_5S_2$ expected 605.3, found 606.2 [M+H$^+$].

^{1}H NMR (400 MHz, CD_3OD, 25 °C) δ: 8.88 (s, 1H), 8.51 (br. s, 1H exch), 7.47-7.41 (m, 4H), 4.94 (s, 1H), 4.58 (t, J = 8.3, 1H), 4.51-4.37 (m, 3H), 3.91-3.89 (m, 1H), 3.85 (dd, J = 10.8 Hz, J = 4.0 Hz, 1H), 3.52 (t, J = 5.7 Hz, 2H), 3.46 (t, J = 6.2 Hz, 2H), 3.04 (t, J = 7.0 Hz, 2H), 2.67-2.61 (m, 2H), 2.48 (s, 3H), 226-2.24 (m, 1H), 2.14-2.08 (m, 1H), 2.00 (s, 3H),1.92-1.87 (m, 2H), 1.77-1.72 (m, 2H), 1.43 (s, 3H), 1.37 (s, 3H).

^{13}C-NMR (101 MHz, CD_3OD, 25 °C) δ: 174.4, 173.1, 171.4, 169.9, 153.0, 149.2, 140.3, 133.5, 131.8, 130.6, 129.2, 71.1, 69.5, 61.1, 57.9, 57.1, 43.8, 39.5, 39.2, 30.8, 28.7, 26.5, 26.2, 25.4, 22.5, 16.0.

(*2S,4R*)-1-((*R*)-2-acetamido-3-((4-(2-azidoethoxy)butyl)thio)-3-methylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (19)

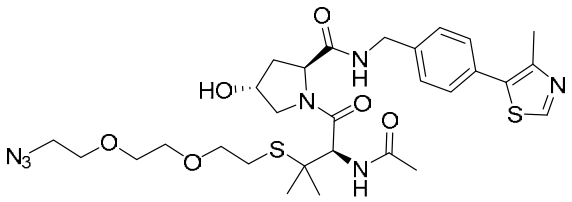


Prepared following the general procedure for the *S*-alkylation from **14** (10.0 mg, 0.020 mmol) and **9** (6.9 mg, 0.022 mmol). Obtained 9.2 mg, 73% yield. MS analysis: $C_{29}H_{41}N_7O_5S_2$ expected 631.3, found 632.2 [M+H$^+$].

^{1}H NMR (400 MHz, CD_3OD, 25 °C) δ: 8.88 (s, 1H), 8.40 (t, J = 6.2 Hz, 1H exch), 8.14 (d, J = 8.9 Hz, 1H exch), 7.46-7.41 (m, 4H), 4.94-4.92 (m, 1H), 4.58 (t, J = 8.3 Hz, 1H), 4.52-4.39 (m, 3H), 3.94-3.92 (m, 1H), 3.86 (dd, J = 10.9 Hz, J = 4.0 Hz, 1H), 3.57 (t, J = 5.0 Hz, 1H), 3.44 (t, J = 6.0 Hz, 1H), 3.33-3.30 (m, 3H), 2.60-2.57 (m, 2H), 2.48 (s, 3H), 2.29-2.24 (m, 1H), 2.14-2.08 (m, 1H), 2.00 (s, 3H), 1.65-1.53 (m, 4H), 1.40 (s, 3H), 1.36 (s, 3H).

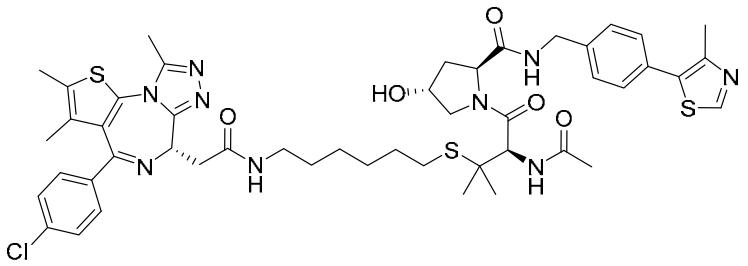
^{13}C-NMR (101 MHz, CD_3OD, 25 °C) δ: 174.5, 173.3, 173.2, 171.6, 153.0, 149.2, 140.2, 133.5, 131.8, 130.6, 129.3, 71.7, 71.0, 70.9, 61.2, 58.1, 57.4, 51.9, 43.8, 39.2, 30.3, 29.2, 27.4, 26.1, 25.9, 22.4, 16.0.

(*2S,4R*)-1-((*R*)-2-acetamido-3-((2-(2-(2-azidoethoxy)ethoxy)ethyl)thio)-3-methylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (20)



Prepared following the general procedure for *S*-alkylation from **14** (10.0 mg, 0.020 mmol) and 2-(2-(2-azidoethoxy)ethoxy)ethyl methanesulfonate[6] (5.6 mg, 0.022 mmol). Obtained 9.3 mg, 72% yield. MS analysis: $C_{29}H_{41}N_7O_6S_2$ expected 647.3, found 648.2 [M+H$^+$].
^{1}H-NMR (400 MHz, $CDCl_3$, 25 °C) δ: 8.67 (s, 1H), 7.39-7.32 (m, 5H), 6.45 (d, J = 8.0 Hz, 1H), 4.75-4.71 (m, 2H), 4.50 (br s, 1H), 4.45-4.43 (m, 2H), 4.06-4.03 (m, 1H), 3.74 (dd, J = 11.1 Hz, J = 3.8 Hz, 1H), 3.34-3.50 (m, 8H), 3.36 (t, J = 5.0 Hz, 2H), 2.78-2.66 (m, 2H), 2.50 (s, 3H), 2.39-2.33 (m, 1H), 2.28-2.22 (m, 1H), 1.97 (s, 3H), 1.88 (br s, 1H), 1.31 (s, 3H), 1.29 (s, 3H).
^{13}C-NMR (101 MHz, $CDCl_3$, 25 °C) δ: 171.0, 170.7, 170.4, 150.3, 148.5, 138.2, 131.6, 130.9, 129.5, 128.1, 70.6, 70.5, 70.4, 70.0, 59.1, 56.7, 56.2, 50.7, 47.8, 43.0, 37.0, 28.2, 25.6, 25.5, 23.0, 16.1.

(*2S,4R*)-1-((*R*)-2-acetamido-3-((6-(2-((*S*)-4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)acetamido)hexyl)thio)-3-methylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (1)

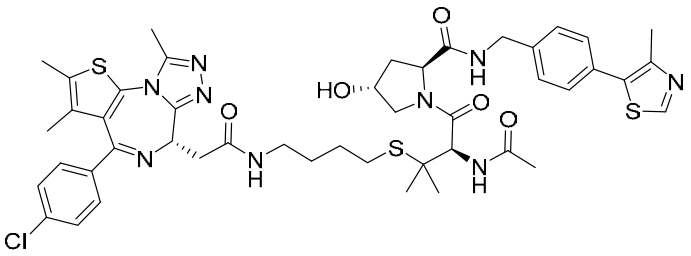


Compound **23** (6.4 mg, 0.0108 mmol) was dissolved in DMF (0.25 mL) and added to a solution of (*S*)-2-(4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)acetic acid (+)-JQ1-COOH (4.32 mg, 0.0108 mmol), HATU (4.1 mg, 0.0108 mmol), HOAT (1.5 mg, 0.0108mmol) and DIPEA (5.5 µl, 0.0324 mmol) in DMF (0.5 mL). After stirring at room temperature for 1 h, the reaction was quenched with water (0.1 mL) and the mixture of water/DMF was removed under high vacuum at room temperature (overnight). The crude mixture was dissolved in MeOH, filtered and purified by preparative HPLC to give the title compound. Obtained 7.3 mg, 70% yield. MS analysis: $C_{48}H_{58}N_9ClO_5S_3$ expected 971.3, found 972.3 [M+H$^+$].
^{1}H-NMR (400 MHz, CD_3OD, 25 °C) δ: 8.99 (s, 1H), 8. 43 (t, J = 5.9 Hz, 1 H exch.), 8.12 (d, J = 9.4 Hz, 1H exch.), 8.07 (s, 1H exch.), 7.47-7.40 (m, 8H), 4.93-4.91 (m, 1H), 4.69-4.65 (m, 1H), 4.58 (t, J = 8.3 Hz, 1H), 4.52-4.38 (m, 3H), 3.93-3.91 (m, 1H), 3.85 (dd, J = 10.8 Hz, J = 3.96 Hz,

1H), 3.45-3.39 (m, 1H), 3.27-3.16 (m, 3H), 2.72 (s, 3H), 2.26 (t, J = 7.0 Hz, 2H), 2.49 (s, 3H), 2.45 (s, 3H), 2.25-2.22 (m, 1H), 2.14-2.07 (m, 1H), 2.00 (s, 3H), 1.70 (s, 3H), 1.54-1.36 (m, 14H).

^{13}C-NMR (101 MHz, CD_3OD, 25 °C) δ: 174.4, 173.2, 172.6, 171.6, 166.7, 157.1, 153.4, 152.6, 148.5, 140.5, 138.4, 137.9, 133.9, 133.6, 132.3, 132.2, 131.6, 131.4, 130.6, 130.0, 129.7, 129.3, 71.0, 61.2, 58.1, 57.4, 55.2, 43.7, 40.5, 39.2, 38.7, 30.7, 30.5, 30.0, 29.3, 27.7, 26.2, 25.9, 22.5, 15.7, 14.6, 13.1, 11.7.

(*2S,4R*)-1-((*R*)-2-acetamido-3-((4-(2-((*S*)-4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)acetamido)butyl)thio)-3-methylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (2)

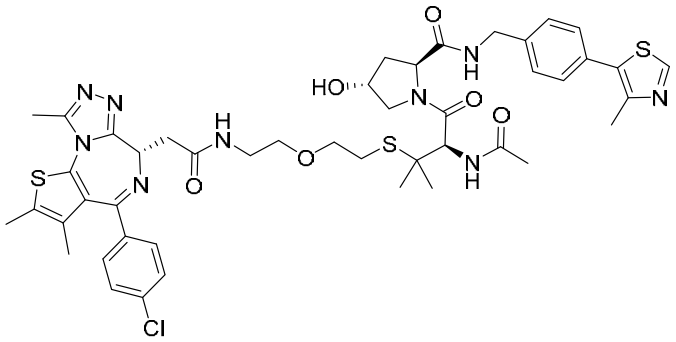


Compound **15** (11.3 mg, 0.0163 mmol) was dissolved in ethanol (0.5 mL) and treated with hydrazine monohydrate (24 µL, 0.32 mmol) at 60 °C for 2 h. The reaction mixture was cooled to room temperature, filtered and dried under vacuum to obtain compound **21**. The crude was made to react with (S)-2-(4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)acetic acid (+)-JQ1-COOH (6.1 mg, 0.015 mmol), HATU (6.0 mg, 0.015 mmol), HOAT (2.0 mg, 0.015 mmol) and DIPEA (8.0 µl, 0.045 mmol). Obtained 5.0 mg, 33% yield after preparative HPLC purification. MS analysis: $C_{46}H_{54}N_9ClO_5S_3$ expected 943.3, found 944.3 [$M+H^+$].

^{1}H-NMR (400 MHz, CD_3OD, 25 °C) δ: 9.00 (s, 1H), 7.47-7.41 (m, 8H), 4.29 (s, 1H), 4.68-4.65 (m, 1H), 4.57 (t, J = 8.3 Hz, 1H), 4.51-4.39 (m, 3H), 3.93-3.90 (m, 1H), 3.51 (dd, J = 10.9 Hz, J = 4.0 Hz, 1H), 3.44-3.39 (m, 1H), 3.28-3.16 (m, 3H), 2.73 (s, 3H), 2.62-2.59 (m, 2H), 2.48 (s, 3H), 2.45 (s, 3H), 2.27-2.22 (m, 1H), 2.14-2.07 (s, 1H), 2.00 (s, 3H), 1.70 (s, 3H), 1.65-1.55 (m, 4H)1.14 (s, 3H), 1.36 (s, 3H).

^{13}C-NMR (101 MHz, CD_3OD, 25 °C) 174.4, 173.2, 172.6, 171.5, 166.7, 157.0, 153.5,152.6, 148.4, 140.6, 138.4, 137.8, 134.0, 133.9, 133.5, 132.4, 132.2, 131.6, 131.3, 130.7, 130.6, 130.0, 129.7, 129.3, 71.1, 61.2, 58.1, 57.3, 55.1, 43.7, 40.1, 39.2, 38.7, 30.0, 29.0, 28.0, 26.2, 25.8, 22.5, 15.6, 14.6, 13.1.

(*2S,4R*)-1-((*R*)-2-acetamido-3-((2-(2-(2-((*S*)-4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)acetamido)ethoxy)ethyl)thio)-3-methylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (3)

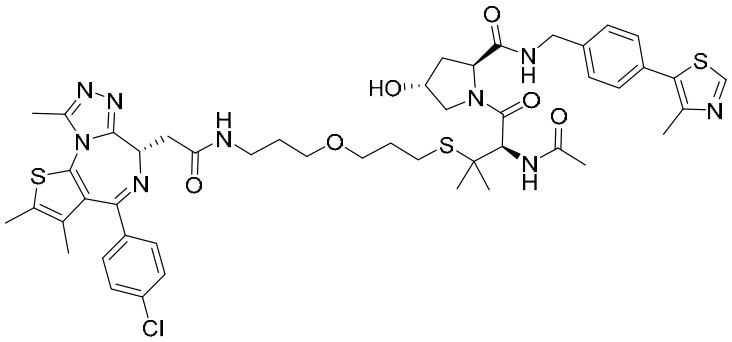


Compound **16** (9.0 mg, 0.015 mmol) was dissolved in methanol (1 ml). Catalytic amount of palladium on charcoal (10% w/w) was added and the reaction mixture stirred under an atmosphere of hydrogen for 3 h at room temperature. The reaction mixture was filtered through a syringe filter and the resulting solution evaporated to dryness to obtain the desired amine **22** which was dissolved in DMF (0.5 mL) and added to a solution of (S)-2-(4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)acetic acid (+)-JQ1-COOH (6.1 mg, 0.015 mmol), HATU (6.0 mg, 0.015 mmol), HOAT (2.0 mg, 0.015 mmol) and DIPEA (8.0 µl, 0.045 mmol) in DMF (0.5 mL). After stirring at room temperature for 1 h, the reaction was quenched with water (0.1 mL) and the mixture of water/DMF was removed under high vacuum at room temperature (overnight). The crude mixture was dissolved in MeOH, filtered and purified by preparative HPLC to give the title compound. Obtained 10.1 mg, 70% yield. MS analysis: $C_{46}H_{54}ClN_9O_6S_3$ expected 959.3, found 960.3 [M+H$^+$].

^{1}H-NMR (400 MHz, CD_3OD, 25 °C) δ: 8.98 (s, 1H), 8.49 (t, J = 6.0 Hz, 1H), 8.15 (d, J = 9.0 Hz, 1H), 7.47-7.39 (m, 9H), 4.94-4.92 (m, 1H), 4.68-4.64 (m, 1H), 4.58 (t, J = 8.4 Hz, 1H), 4.47-4.38 (m, 3H), 3.89-3.88 (m, 2H), 3.61-3.53 (m, 4H), 3.47-3.38 (m, 3H), 2.81-2.78 (m, 2H), 2.72 (s, 3H), 2.47 (s, 3H), 2.45 (s, 3H), 2.27-2.22 (m, 1H), 2.14-2.07 (m, 1H), 2.00 (s, 3H), 1.70 (s, 3H), 1.40 (s, 3H), 1.35 (s, 3H).

^{13}C-NMR (101 MHz, CD_3OD, 25 °C) δ: 174.4, 173.2, 172.9, 171.4, 166.6, 157.0, 153.3, 152.5, 148.5, 140.5, 138.4, 137.9, 133.9, 133.8. 133.6, 132.3, 132.1, 131.6, 131.4, 130.6, 130.5, 130.0, 129.6, 129.3, 71.6, 71.1, 70.5, 61.2, 58.1, 57.4, 55.1, 43.8, 40.6, 39.2, 38.6, 29.6, 26.5, 26.0, 22.5, 15.7, 14.6, 13.1, 11.7.

(*2S,4R*)-1-((*R*)-2-acetamido-3-((3-(3-(2-((*S*)-4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)acetamido)propoxy)propyl)thio)-3-methylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (4)

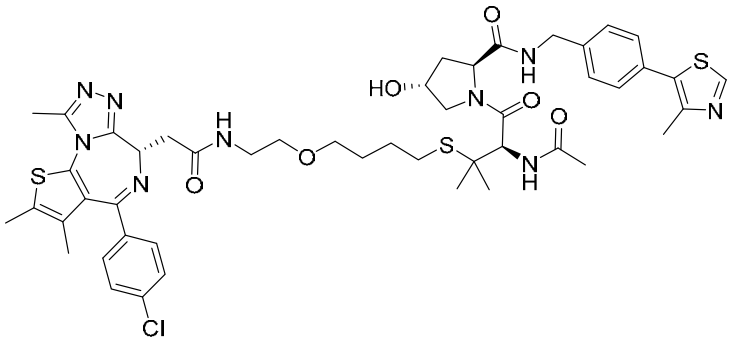


Prepared as described for **1** starting from compound **24**. Obtained 6.0 mg (68% yield) after preparative HPLC purification. MS analysis: $C_{48}H_{58}N_9ClO_6S_3$ expected 987.3, found 988.3 [M+H$^+$].

^{1}H-NMR (400 MHz, CD_3OD, 25 °C) δ: 8.27 (s, 1H), 7.45-7.39 (m, 8H), 4.91 (s, 1H), 4.63-4.54 (m, 2H), 4.52-4.38 (m, 3H), 3.93-3.91 (m, 1H), 3.86 (dd, J = 10.8 Hz, J = 3.8 Hz, 1H), 3.49-3.26 (m, 10H), 2.68 (s, 3H), 2.64-3.61 (m, 2H), 2.45 (s, 3H), 2.44 (s, 3H), 2.27-2.16 (m, 1H), 2.13-2.06 (m, 1H), 2.00 (s, 3H), 1.81-1.78 (m, 2H), 1.72-1.70 (m, 5H), 1.39 (s, 3H), 1.35 (s, 3H).

^{13}C-NMR (101 MHz, CD_3OD, 25 °C) δ: 174.4, 173.2, 172.8, 171.6, 166.4, 157.2, 153.0, 152.3, 149.2, 140.3, 138.3, 138.1, 133.7, 133.5, 133.4, 132.2, 132.1, 131.8, 131.5, 130.6, 129.9, 129.2, 71.1, 70.7, 69.5, 61.2, 58.1, 57.4, 55.4, 43.7, 39.2, 39.0, 38.0, 31.0, 30.8, 26.2, 25.8, 22.5, 16.0, 14.6, 13.1, 11.8.

(*2S,4R*)-1-((*R*)-2-acetamido-3-((4-(2-(2-((S)-4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)acetamido)ethoxy)butyl)thio)-3-methylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (5)

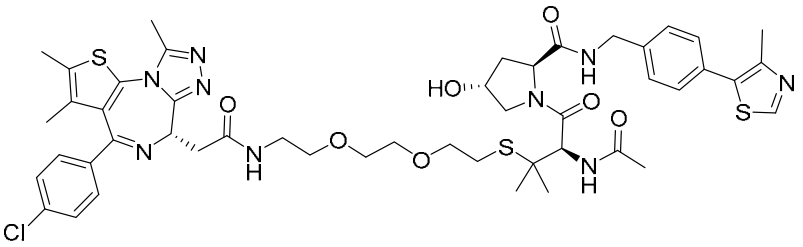


Prepared as described for **3** from azide **19**. Obtained 9.2 mg (62% yield) after preparative HPLC purification. MS analysis: $C_{48}H_{58}N_9ClO_6S_3$ expected 987.3, found 988.3 [$M+H^+$].

^{1}H-NMR (400 MHz, CD_3OD, 25 °C) δ: 8.87 (s, 1H), 8.43-8.37 (m, 2H exch), 8.12 (d, J = 9.4 Hz, 1H exch), 7.46-7.39 (m, 8H), 4.92-4.90 (m, 1H), 4.64-4.55 (m, 2H), 4.51-4.38 (m, 3H), 3.93-3.90 (m, 1H), 3.85 (dd, J = 10.8 Hz, J = 3.9 Hz, 1H), 3.53-3.41 (m, 7H), 2.69(s, 3H), 2.57 (t, J = 7.1 Hz, 2H), 2.47 (s, 3H), 2.44 (s, 3H), 2.87-2.22 (m, 1H), 2.14-2.07 (m, 1H), 2.00 (s, 3H), 1.69 (s, 3H), 1.62-1.52 (m, 4H), 1.38 (s, 3H), 1.35 (s, 3H).

^{13}C-NMR (101 MHz, CD_3OD, 25 °C) δ: 174.4, 173.2, 173.1, 171.6, 166.3, 157.2, 153.0, 152.3, 149.2, 140.3, 138.3, 138.1, 133.7, 133.5, 133.4, 132.2, 132.1, 131.8, 131.5, 130.6, 129.9, 129.2, 71.7, 71.0, 70.4, 61.2, 58.1, 57.4, 55.3, 43.7, 40.7, 39.2, 38.9, 30.3, 29.2, 27.5, 26.2, 25.9, 22.5, 16.0, 14.6, 13.1, 11.8.

(*2S,4R*)-1-((*R*)-14-acetamido-1-((*S*)-4-(4-chlorophenyl)-2,3,9-trimethyl-6H-thieno[3,2-f][1,2,4]triazolo[4,3-a][1,4]diazepin-6-yl)-13,13-dimethyl-2-oxo-6,9-dioxa-12-thia-3-azapentadecan-15-oyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (6)



Prepared as described for **3** starting from azide **20**. Obtained 8.4 mg (60% yield) after preparative HPLC purification. MS analysis: $C_{48}H_{58}N_9ClO_7S_3$ expected 1,003.3, found 1,004.3 [M+H$^+$].

^{1}H-NMR (400 MHz, CD_3OD, 25 °C) δ: 8.86 (s, 1H), 7.46-7.39 (m, 8H), 4.91 (s, 1H), 4.64-4.60 (m, 1H), 4.57 (t, J = 8.3 Hz, 1H), 4.49-4.43 (m, 3H), 3.90-3.89 (m, 2H), 3.61-3.41 (m, 12H), 3.33-3.22 (m, 2H), 2.78 (t, J = 6.5 Hz, 2H), 2.69(s, 3H), 2.46 (s, 3H), 2.44 (s, 3H), 2.28-2.22 (m, 1H), 2.13-2.06 (m, 1H), 1.99 (s, 3H), 1.69 (s, 3H), 1.38 (s, 3H), 1.35 (s, 3H).

^{13}C-NMR (101 MHz, CD_3OD, 25 °C) δ:174.4, 173.2, 173.0, 171.5, 166.3, 157.2, 153.0, 152.3, 149.2, 140.3, 138.3, 138.1, 133.7, 133.5, 133.3, 132.2, 132.1, 131.7, 131.5, 130.5, 129.9, 129.3, 71.8, 71.5, 71.1, 70.8, 61.2, 58.1, 57.4, 55.3, 43.7, 40.7, 39.2, 38.9, 29.5, 26.3, 25.9, 22.5, 16.0, 14.6, 13.1, 11.8.

References

1. Zengerle, M.; Chan, K. H.; Ciulli, A. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. *ACS Chem. Biol.* **2015***, 10*, 1770-1777.

2. Anders, L.; Guenther, M. G.; Qi, J.; Fan, Z. P.; Marineau, J. J.; Rahl, P. B.; Loven, J.; Sigova, A. A.; Smith, W. B.; Lee, T. I.; Bradner, J. E.; Young, R. A. Genome-wide localization of small molecules. *Nat. Biotechnol.* **2014***, 32*, 92-96.

3. Galdeano, C.; Gadd, M. S.; Soares, P.; Scaffidi, S.; Van Molle, I.; Birced, I.; Hewitt, S.; Dias, D. M.; Ciulli, A. 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. *J. Med. Chem.* **2014***, 57*, 8657-8663.

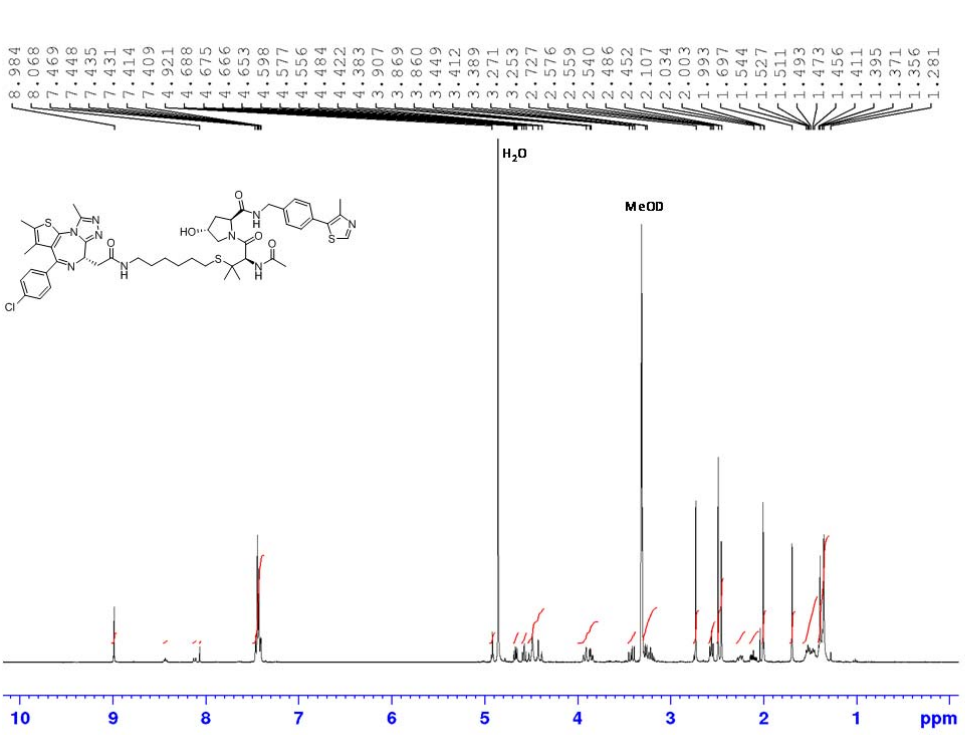
4. Demko, Z.; Sharpless, K. B. An Intramolecular [2 + 3] Cycloaddition Route to Fused 5-Heterosubstituted Tetrazoles. *Org. Lett.* **2001**, *3*, 4091- 4094.

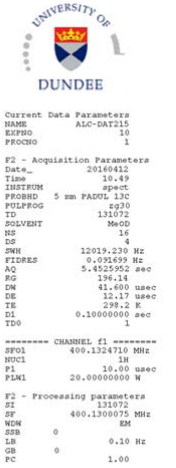
5. Yu, X.; Eymur, S.; Singh, V.; Yang, B.; Tonga, M.; Bheemaraju, A.; Cooke, G.; Subramani, C.; Venkataraman, D.; Stanley, R. J.; Rotello, V. M. Flavin as a photo-active acceptor for efficient energy and charge transfer in a model donor-acceptor system. *Phys. Chem.* **2012**, *14*, 6749- 6754.

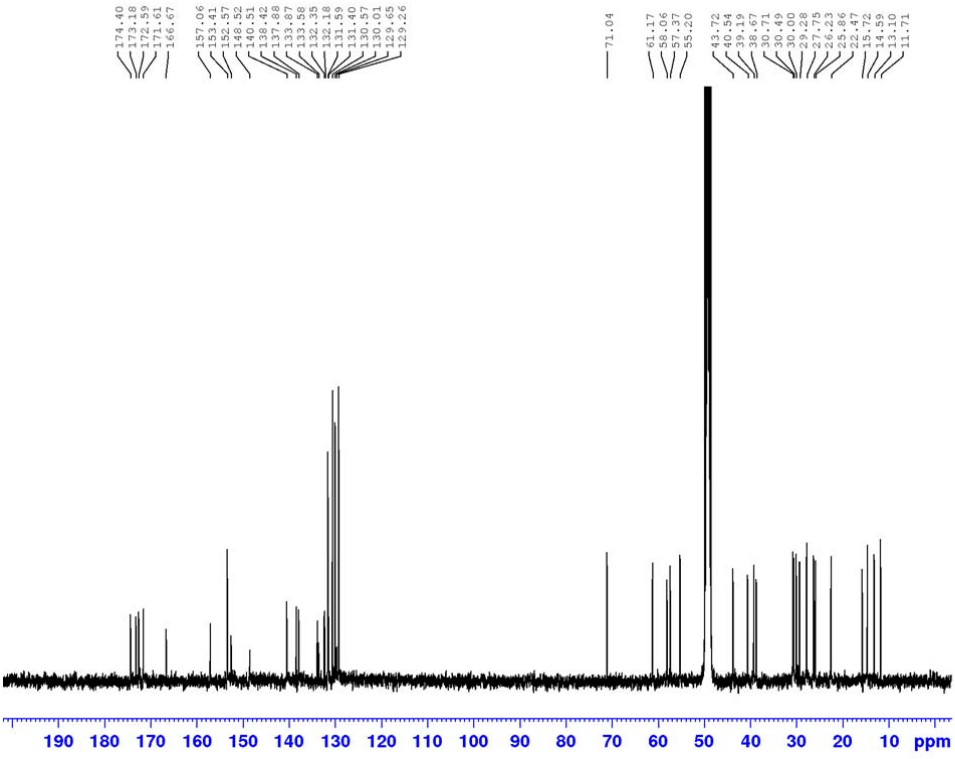
6. Sakamoto, J.; Takita, C.; Koyama, T.; Hatano, K.; Terunuma, D.; Matsuoka, K. Use of a recycle-type SEC method as a powerful tool for purification of thiosialoside derivatives. *Carbohydr. Res.* **2008***, 343*, 2735-2739.

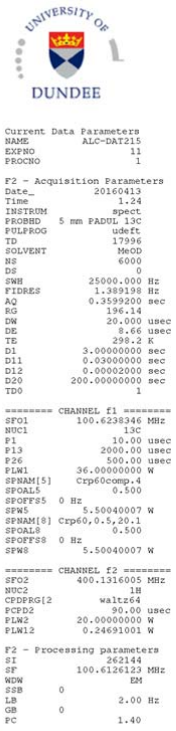
NMR SPECTRA OF NEW-GENERATION PROTACS 1-6

1, ^{1}H-NMR and ^{13}C-NMR in CD_3OD

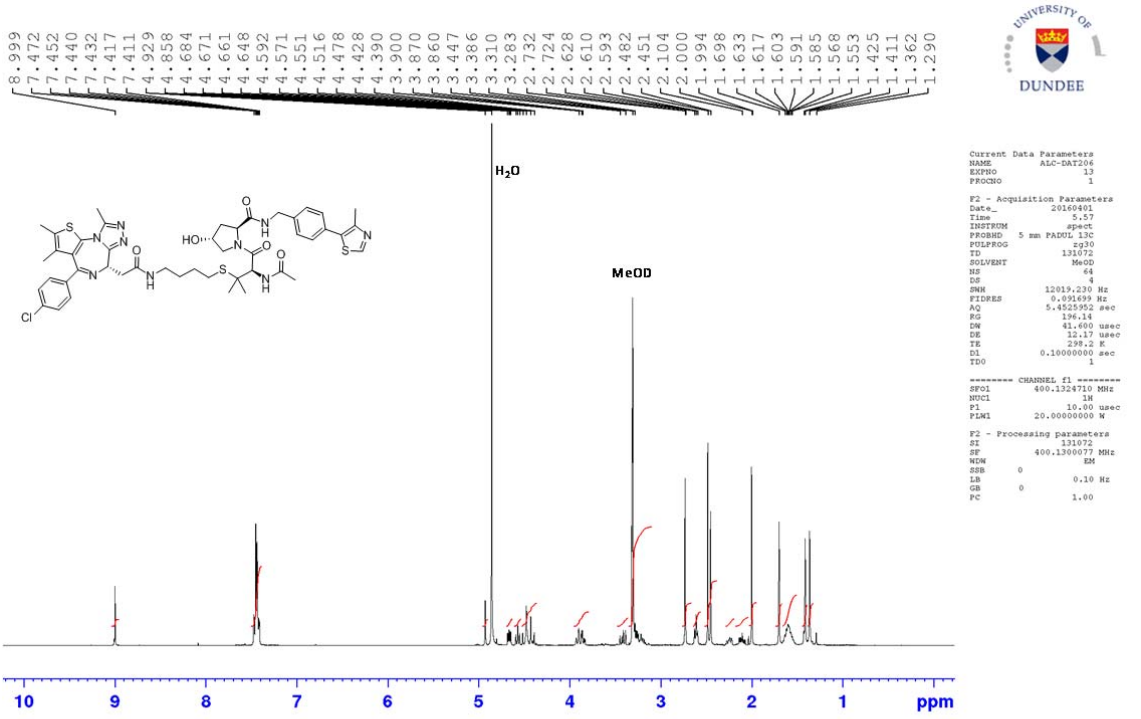




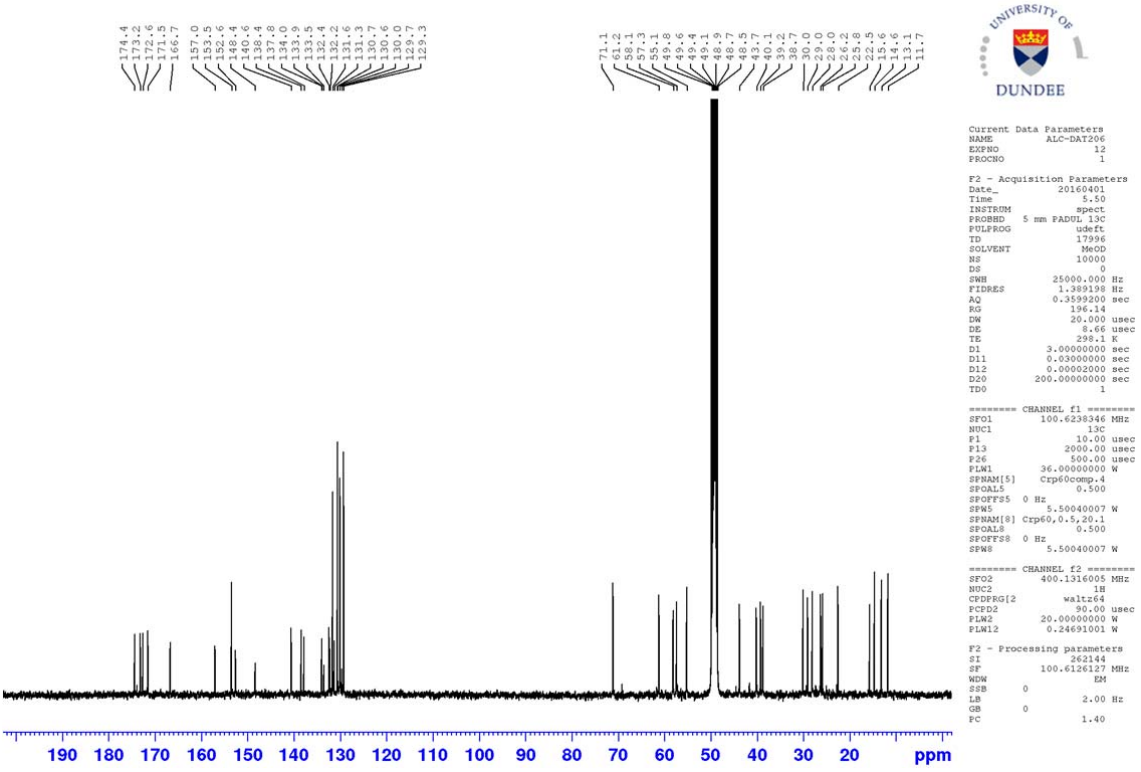




503 **2**, ^{1}H-NMR and ^{13}C-NMR in CD_3OD



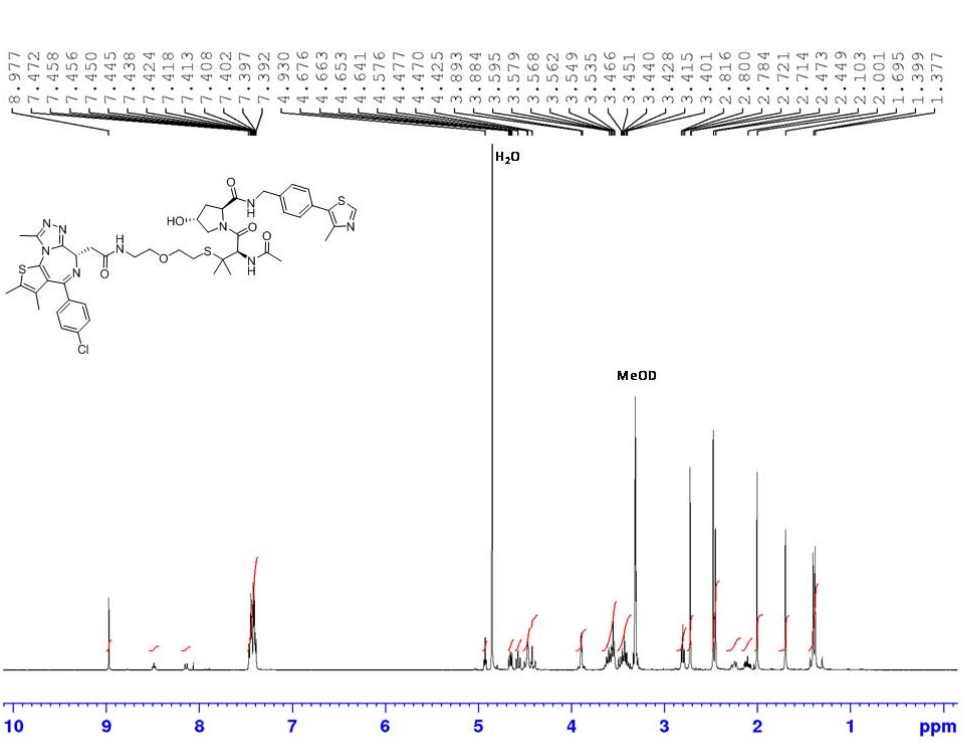
504

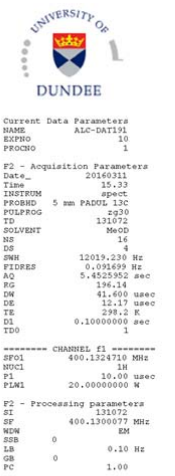


505

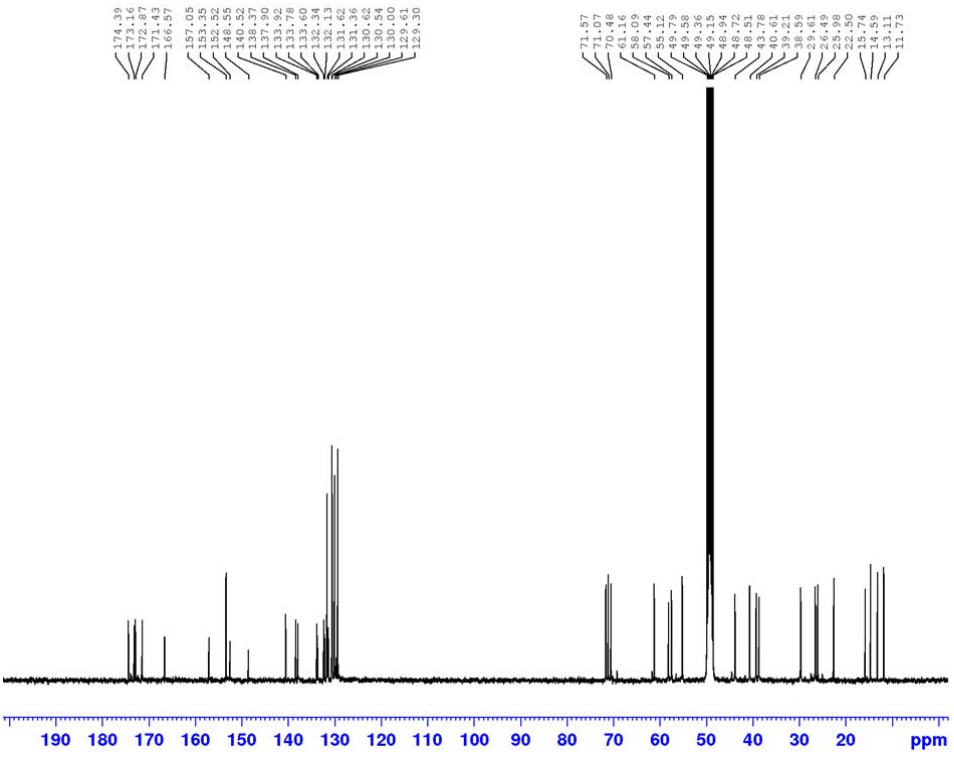
506

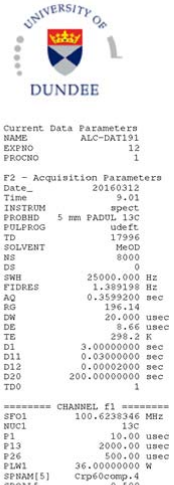
507 **3**, ^{1}H-NMR and ^{13}C-NMR in CD_3OD





508

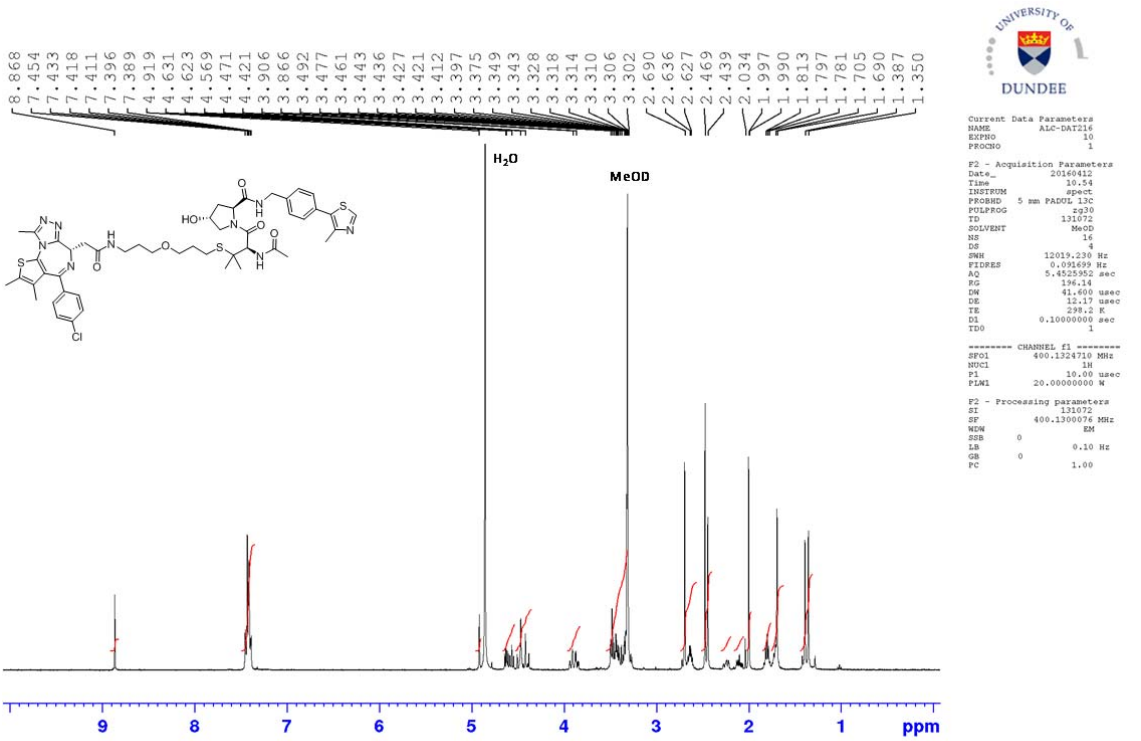




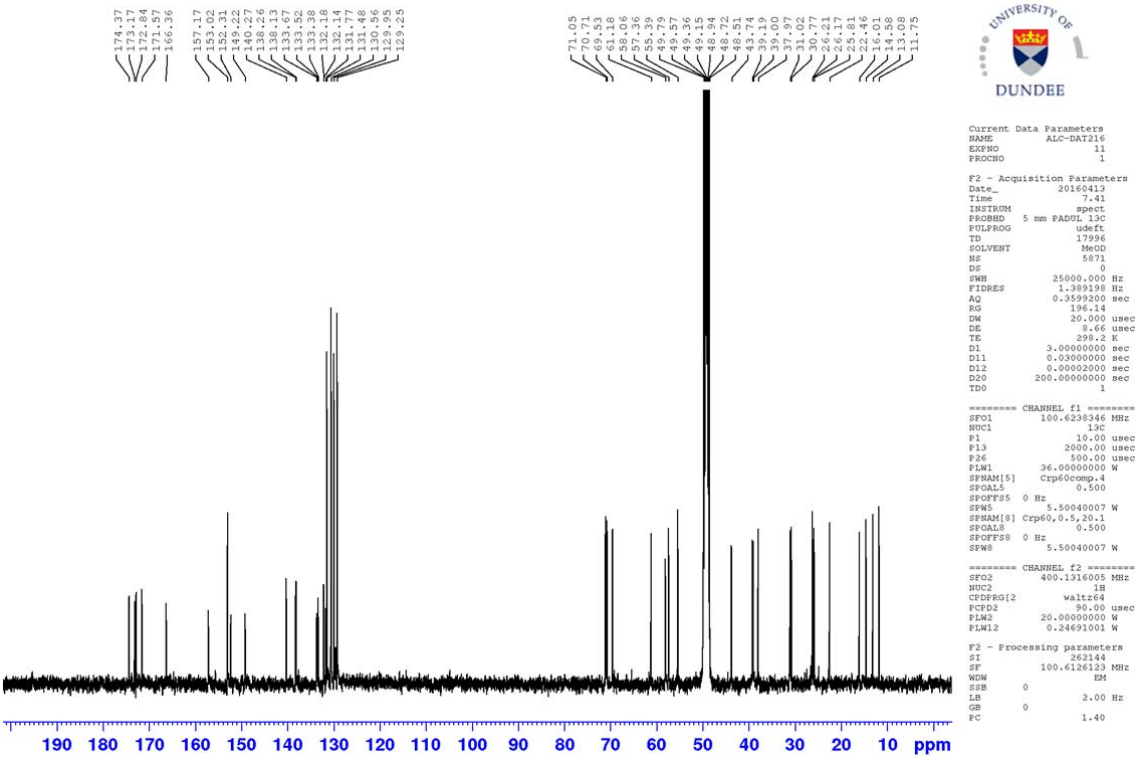
509

510

511 **4**, ^{1}H-NMR and ^{13}C-NMR in CD_3OD



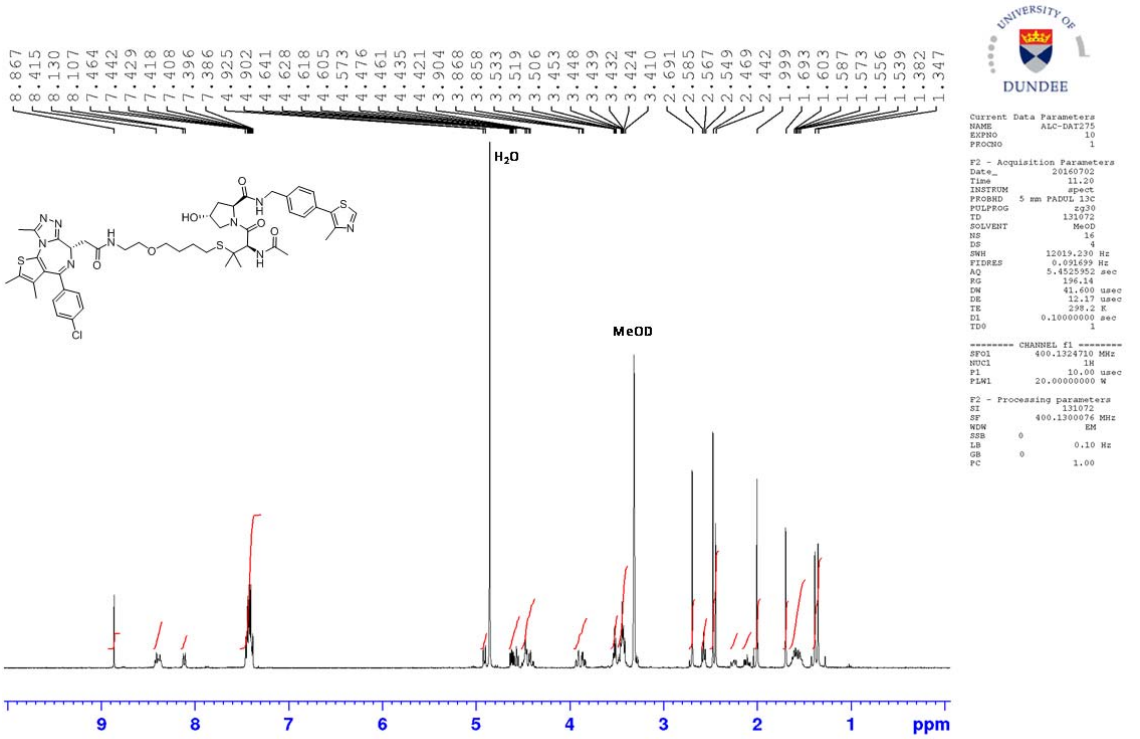
512



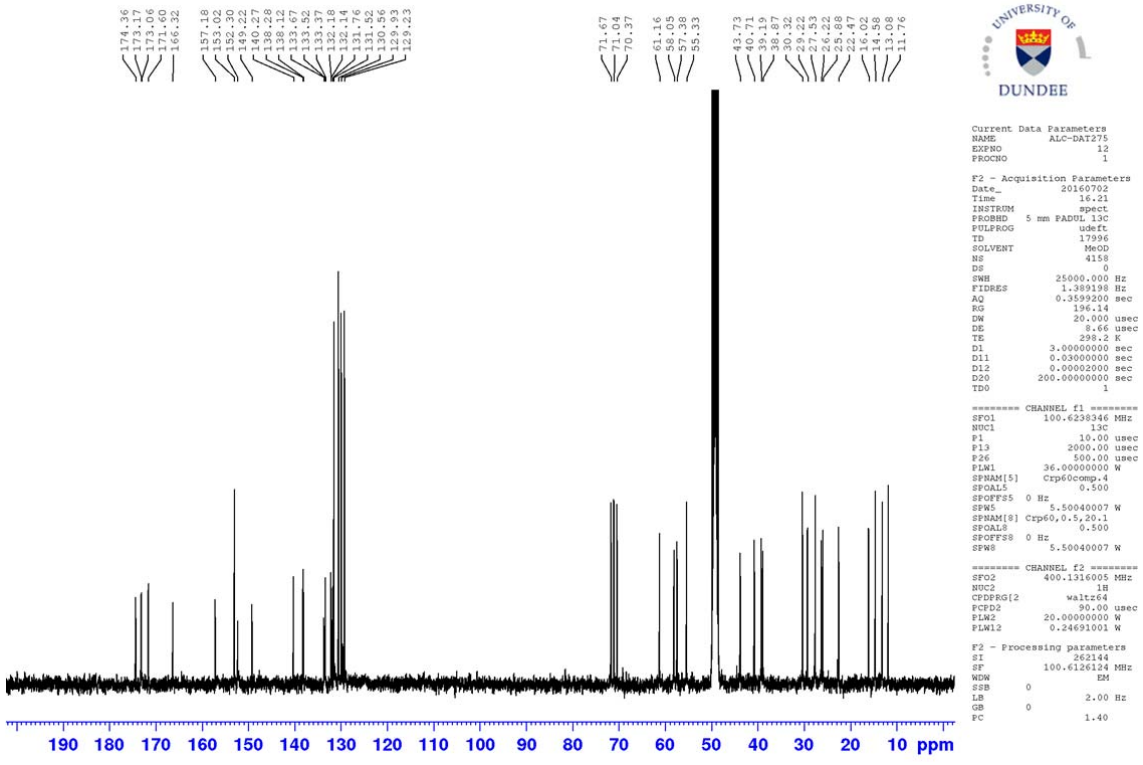
513

514

515 **5**, ^{1}H-NMR and ^{13}C-NMR in CD_3OD



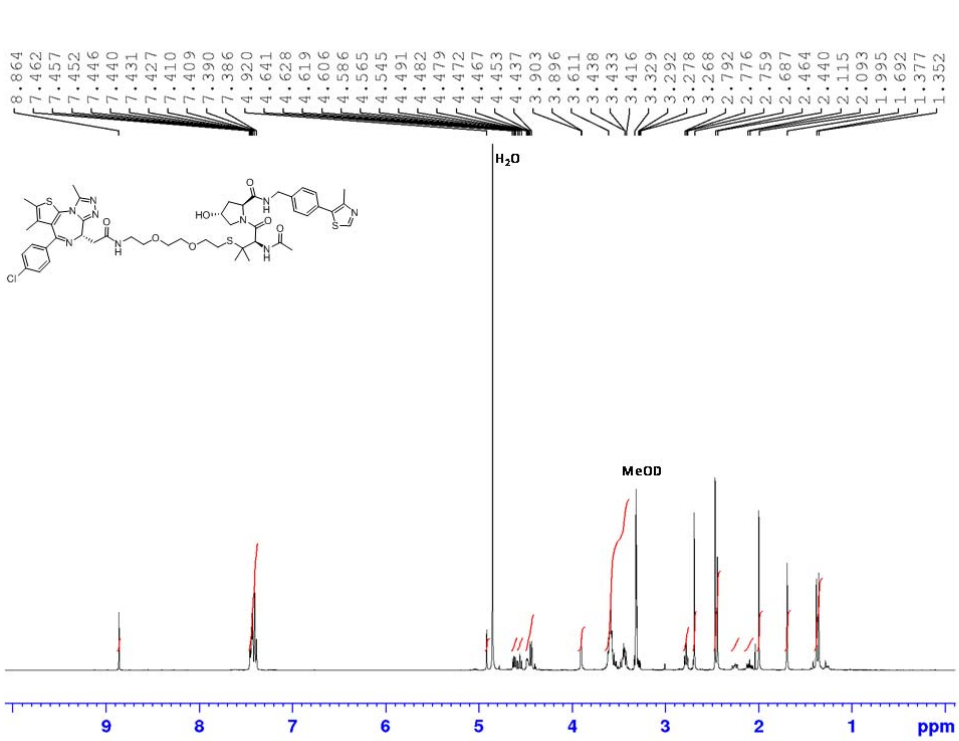
516

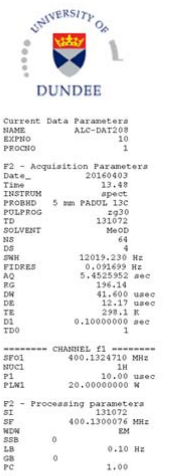


517

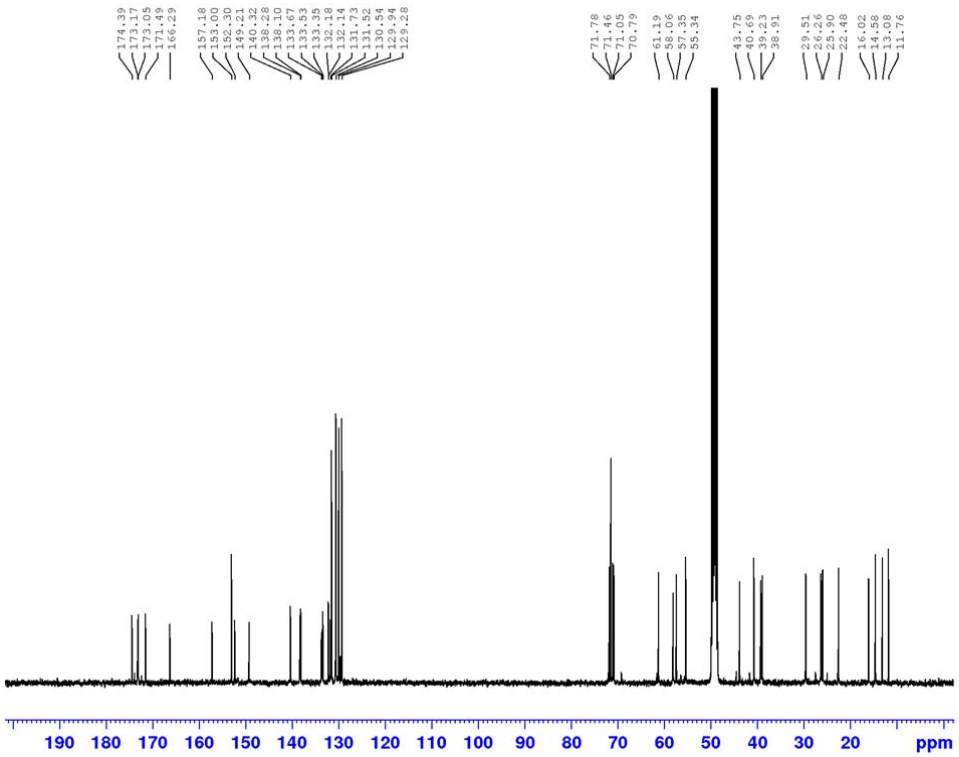
518

519 **6**, ^{1}H-NMR and ^{13}C-NMR in CD_3OD





520



Current Data Parameters
NAME ALC-DAI208
EXPNO 12
PROCNO 1

F2 - Acquisition Parameters
Date_ 20160403
Time 20.10
INSTRUM spect
PROBHD 5 mm PADUL 13C
PULPROG udeft
TD 17996
SOLVENT MeOD
NS 5780
DS 0
SWH 25000.000 Hz
FIDRES 1.389198 Hz
AQ 0.3599200 sec
RG 196.14
DW 20.000 usec
DE 8.66 usec
TE 298.2 K
D1 3.00000000 sec
D11 0.03000000 sec
D12 0.00002000 sec
D20 200.00000000 sec
TD0 1

======== CHANNEL f1 ========
SFO1 100.6238346 MHz
NUC1 13C
P1 10.00 usec
P13 2000.00 usec
P26 500.00 usec
PLW1 36.00000000 W
SPNAM[5] Crp60comp.4
SPOAL5 0.500
SPOFFS5 0 Hz
SPW5 5.50040007 W
SPNAM[8] Crp60,0.5,20.1
SPOAL8 0.500
SPOFFS8 0 Hz
SPW8 5.50040007 W

======== CHANNEL f2 ========
SFO2 400.1316005 MHz
NUC2 1H
CPDPRG[2 waltz64
PCPD2 90.00 usec
PLW2 20.00000000 W
PLW12 0.24691001 W

F2 - Processing parameters
SI 262144
SF 100.6126127 MHz
WDW EM
SSB 0
LB 2.00 Hz
GB 0
PC 1.40

521