Supplementary Information

High accuracy prediction of PROTAC complex structures

Mikhail Ignatov[a,b,+] ,Akhil Jindal[a,b,+], Sergei Kotelnikov[a,b,+], Dmitri Beglov[c,d], Ganna Posternak[e,f], Xiaojing Tang[e], Pierre Maisonneuve[e], Gennady Poda[g,h], Robert A. Batey[f], Frank Sicheri[e,i,j], Adrian Whitty[k*], Peter J. Tonge[l*], Sandor Vajda[c,k,*], and Dima Kozakov[a,b,*]

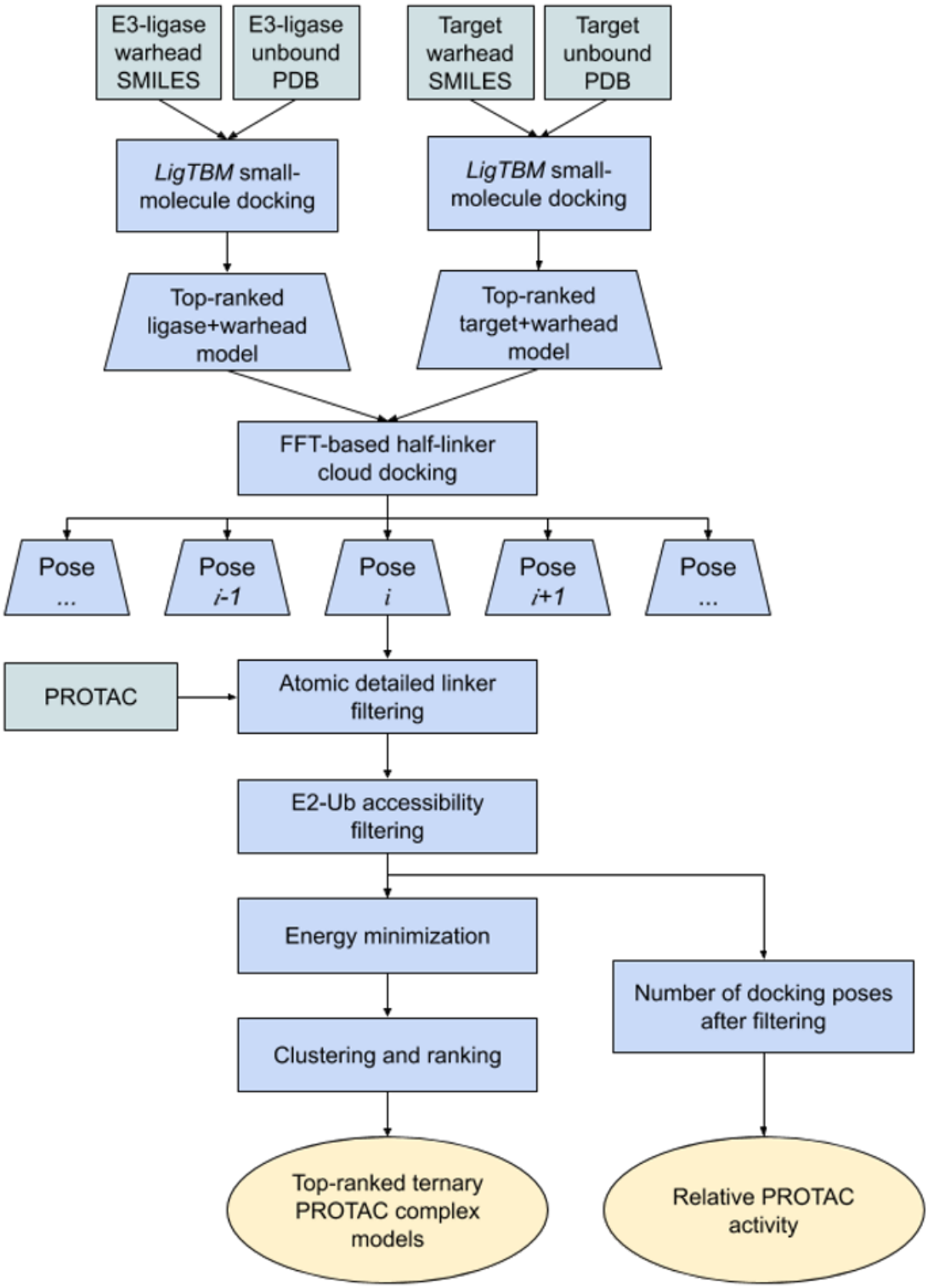
Mikhail Ignatov[a,b,+] , Akhil Jindal[a,b,+], Sergei Kotelnikov[a,b,+], Dmitri Beglov[c,d], Ganna Posternak[e,f], Xiaojing Tang[e], Pierre Maisonneuve[e], Gennady Poda[g,h], Robert A. Batey[f], Frank Sicheri[e,i,j], Adrian Whitty[k*], Peter J. Tonge[l*], Sandor Vajda[c,k,*], and Dima Kozakov[a,b,*]

[a]Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, New York 11794, USA.
[b]Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, New York 11794, USA.
[c]Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215 USA.
[d]Acpharis Inc., Holliston, Massachusetts 01746, USA.
[e]Center for Molecular, Cell and Systems Biology, Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, Ontario L4K-M9W, Canada.
[f]Department of Chemistry, University of Toronto, Toronto, Ontario L4K-M9W, Canada.
[g]Drug Discovery Program, Ontario Institute for Cancer Research, Toronto, Ontario L4K-M9W, Canada.
[h]Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Ontario L4K-M9W, Canada.
[i]Department of Molecular Genetics, University of Toronto, Toronto, Ontario L4K-M9W, Canada.
[j]Department of Biochemistry, University of Toronto, Toronto, Ontario L4K-M9W, Canada.
[k]Department of Chemistry, Boston University, Boston, Massachusetts 02215, USA.
[l]Department of Chemistry, Stony Brook University, Stony Brook, New York 11794, USA.

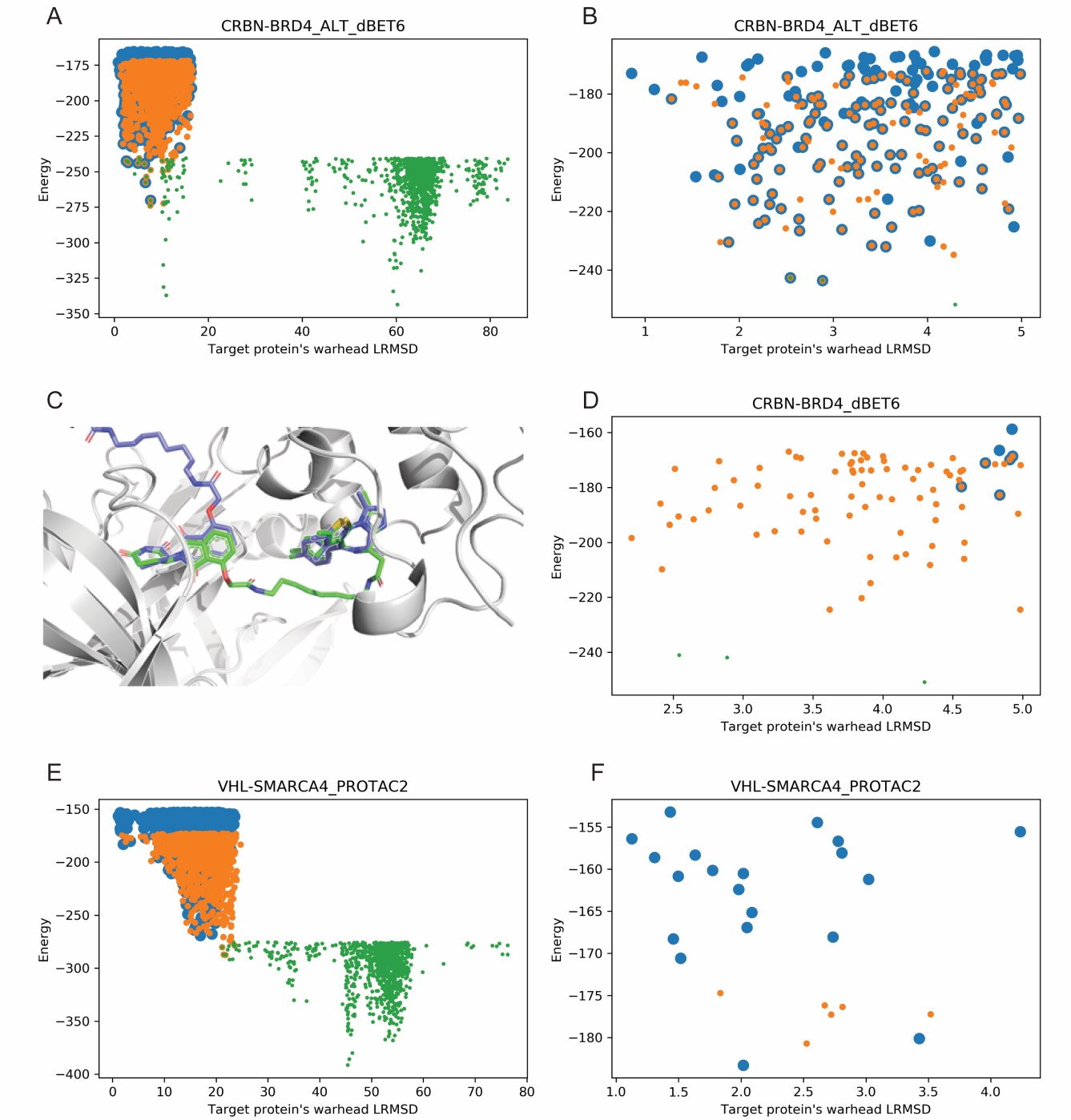
Supplementary Figures 1 – 5

Supplementary Tables 1 – 4

Supplementary Notes

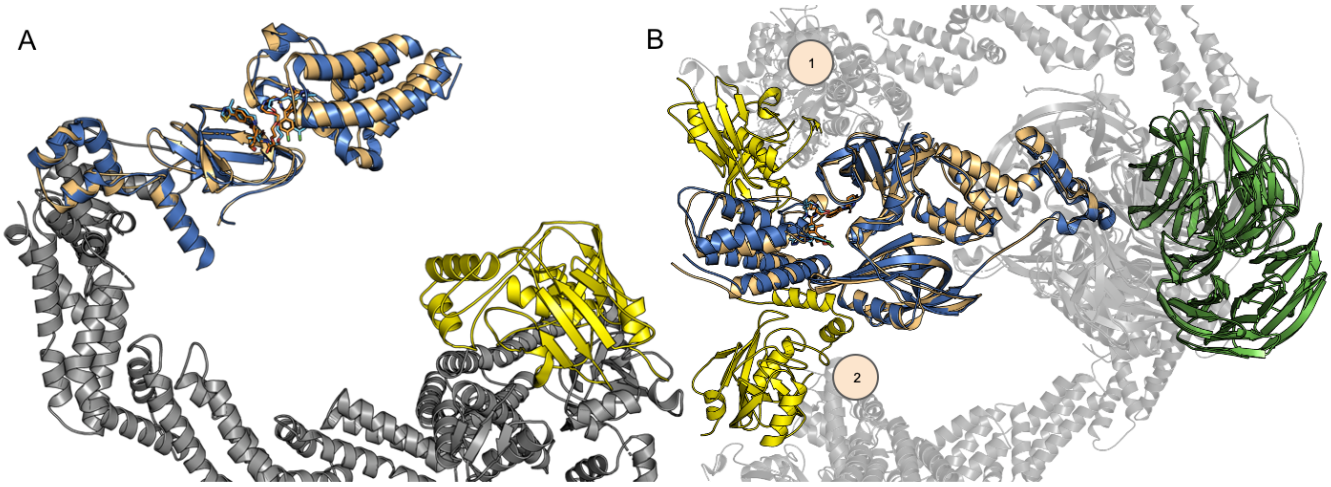


Supplementary Figure 1 | Block diagram of the proposed method.

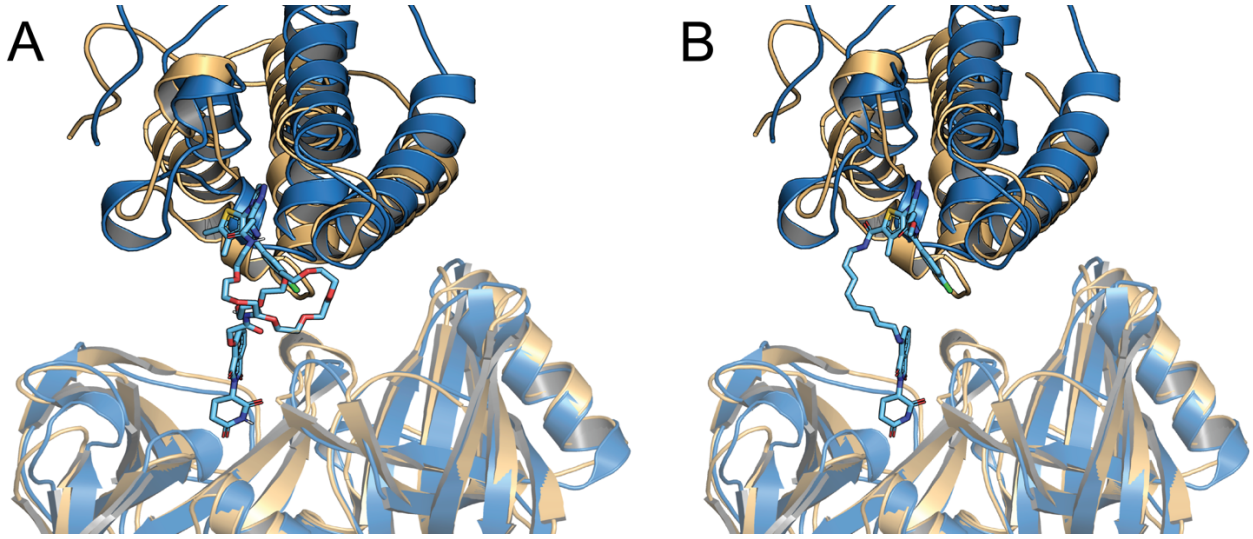


Supplementary Figure 2. Sampling by FFT with and without accounting for the "half-linker clouds" applied to the CRBN–dBET6–BRD4 complex. In these calculations the CRBN protein with the bound thalidomide ligand was fixed and we sampled poses of the BRD4 protein. The RMSD of the BRD4-bound JQ1 ligand from its native conformation in the ternary complex, denoted as target protein warhead ligand RMSD (LRMSD), provides a measure of how closely each resulting predicted geometry

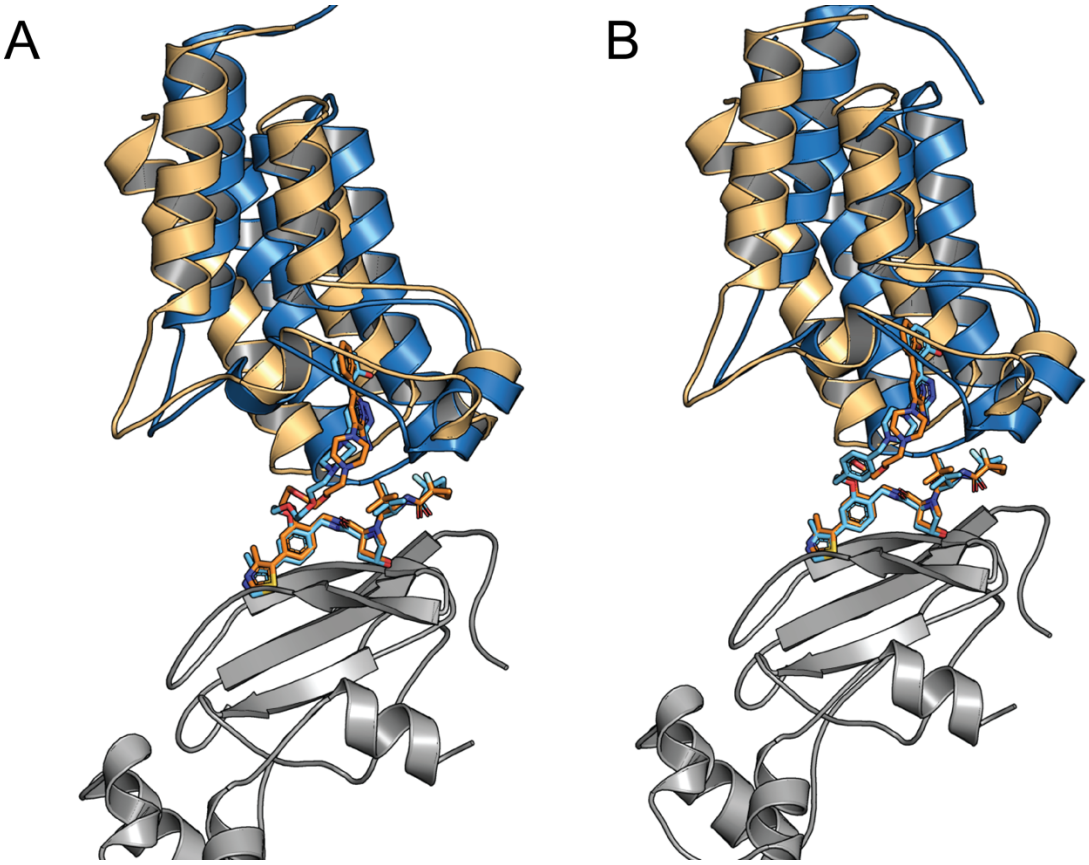
for the complex reproduces the native state. (*A*) Energy values of the CRBN-dBET6-BRD4 complex, generated by global sampling without the linker and without any restraints (green points), using simple distance restraints (orange points), and using the half-linker cloud as restraint (blue points). As shown by the green points, the protein-protein interaction energy function has a deep minimum around 10 Å LRMSD from the experimentally observed structure. As shown by the orange and blue points, the use of restraints substantially enhances sampling in the near-native region. (*B*) Same as Figure 2A, but restricted to the 5 Å neighborhood of the native state. Only our approach of explicitly generating linker conformations yields any structure within 1 Å LRMSD. (*C*) The CRBN ligand thalidomide has two rotamers, and only one (shown in green) yields the LRMSD values shown in Supplementary Figure 2B. (*D*) Energy values obtained when sampling the alternative thalidomide rotamer shown in grey in Figure 2C. (*E*) Energy values obtained by the three sampling schedules for the VHL–PROTAC2–SMARCA4 complex. (*F*) Same as Figure 2E, but restricted to the 5 Å neighborhood of the native state.



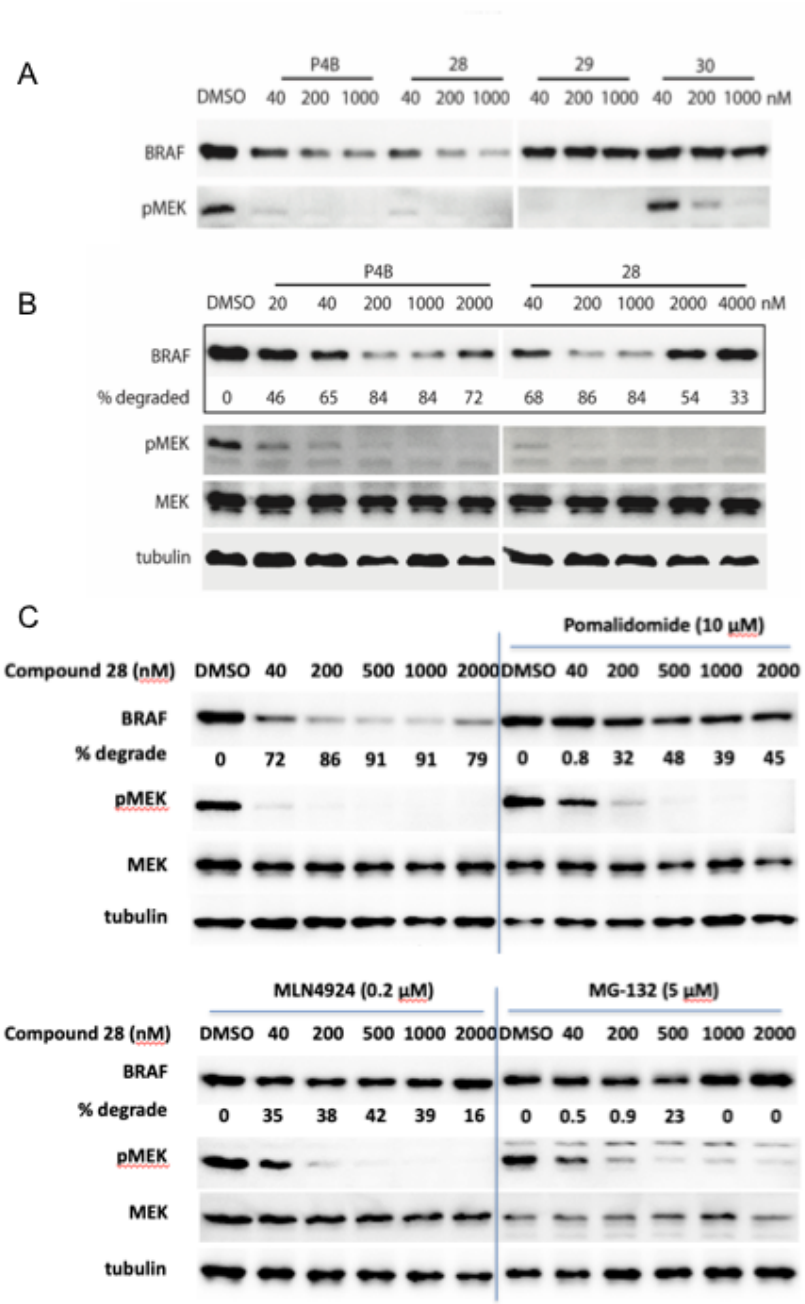
Supplementary Figure 3. Testing ubiquitin accessibility. (A**)** The native structure (PDB ID: 6BN7) and top ranked model of CRBN - dBET23 - BRD4 in orange and blue, respectively, bound to the Cul4 portion of the CRL4CRBN multi-subunit complex in grey, and the E2-Rbx1 proteins in yellow. (B) The native structure (PDB ID: 5T35) and the rank 1 model of the VHL - MZ1 - BRD4 complex (shown in blue and orange, respectively), bound to the Cul2 portion of the CRL2VHL multi-subunit complex and the E2-Rbx1 proteins, shown in grey and yellow, respectively. One important distinction between CRL4CRBN and CRL2VHL shown in Figures A and B, respectively is the DDB1 structure shown in green. This DDB1 structure reveals an intertwined three-propeller cluster, which contains two tightly coupled beta propellers with a large pocket between and a third beta propeller flexibly attached on the side. Distinct conformations of the beta propeller cluster offer disparate Cul4 arrangements, and can play a critical role in understanding the structural framework for targeted protein degradation with CRBN E3 ligase.



Supplementary Figure 4. Consensus models. (A) Superimposing the consensus prediction of CRBN-dBET6-BRD4BD1 (blue) and its X-ray structure (orange, PDB ID: 6BOY) (B) Superimposing the consensus prediction (model 1) of CRBN-dBET23-BRD4BD1 (blue) and its X-ray structure (light orange, PDB ID: 6BN7).



Supplementary Figure 5. **Further consensus models.** (A) Superimposing the consensus prediction of VHL- PROTAC2-SMARCA2 and its X-ray structure (PDB ID: 6HAX). (B) Superimposing the consensus prediction of VHL-PROTAC2-SMARCA4 and its X-ray structure (PDB ID: 6HR2). Note that the two X-ray structures in 6HAX and 6HR2 are very similar, the overall RMSD is 0.49 Å.



Supplementary Figure 6. Testing the BRAF (V600E) degrading activity of new compounds 28, 29, and 30 in A375 cells and comparison to P4B. A. 3-point dose response analyses. A375 cells were treated with the indicated compounds at the indicated concentrations for 24 h prior to immunoblot analysis of whole cell lysates. B. Broader dose response analyses of P4B and compound 28. C. compound 28 activity to degrade BRAF is suppressed by pomalidomide, MLN 4924 or MG132 treatment, demonstrating that BRAF degradation is through ubiquitin-proteasome mechanism. All data shown in are representative of at least two independent experiments.

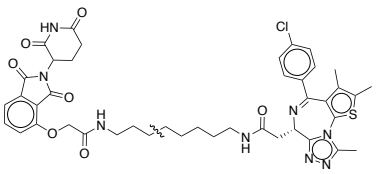
Supplementary Table 1. PROTAC benchmark set of known X-ray structures

PROTAC ternary complex		Unbound E3 Ligase				Unbound Target			
E3:PROTAC:Target [PDB Entry ID]	Res. (Å)	PDB Entry ID	Res. (Å)	Similarity (%)		PDB Entry ID	Res. (Å)	Similarity (%)	
				Sequence	MCS			Sequence	MCS
CRBN: dBET6: BRD4 BD1 [6BOY]	3.33	5FQD	2.5	95	71	3MXF	1.6	100	75
CRBN: dBET23: BRD4 BD1 [6BN7]	3.50								
CRBN: dBET55: BRD4 BD1 [6BN8]*	3.99								
CRBN: dBET57: BRD4 BD1 [6BNB]*	6.34								
CRBN: dBET70: BRD4 BD1 [6BN9]*	4.38								
VHL: MZ1: BRD4 BD2 [5T35]	2.70	4W9H	2.1	99	91	6DUV	1.8	100	89
VHL: PROTAC1: SMARCA2 [6HAY]	2.24	5NVX	2.2	100	100	6HAZ	1.3	100	100
VHL: PROTAC2: SMARCA2 [6HAX]	2.35	5NVX	2.2	100	100	6HAZ	1.3	100	100
VHL: PROTAC1: SMARCA4 [6HR2]	1.76	5NVX	2.2	100	100	3UVD	1.8	100	51
VHL: PROTAC9: BRD4 BD1 [7KHH]**	2.28	5NVX	2.2	100	100	6CKS	1.7	100	100
VHL: PROTAC: FAK [7PI4]	2.24	4W9H	2.1	100	100	6YXV	2.3	100	100
VHL: MS33: WDR5 [7JTP]	2.12	5NVX	2.2	100	100	3SMR	1.8	100	100
VHL: PROTAC6: Bcl-xL [6XHC]	1.60	4W9H	2.1	99	91	1MAZ	2.2	100	100

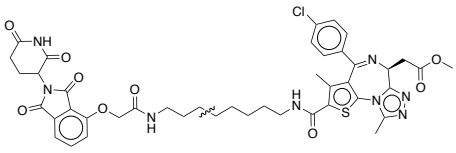
*PROTAC is not visible in the structure, warhead positions are inferred from protein orientations;

** VHL warhead placement with LigTBM failed, copied bound conformation of the warhead in the unbound VHL model

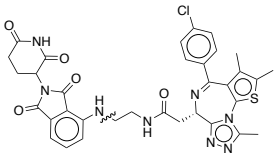
Supplementary Table 2 | PROTAC structures with half linkers.



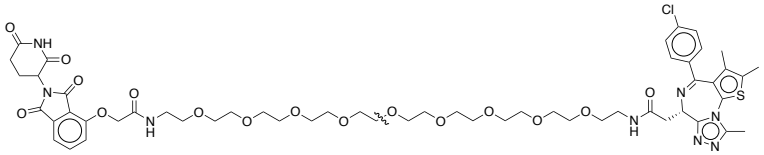
CRBN: dBET6: BRD4 BD1



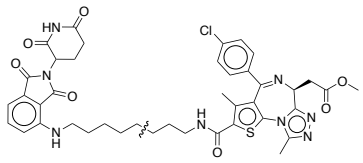
CRBN: dBET23: BRD4 BD1



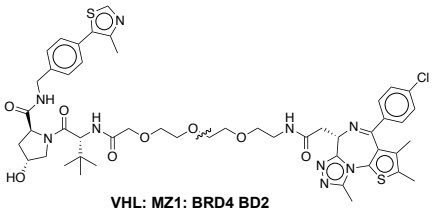
CRBN: dBET57: BRD4 BD1



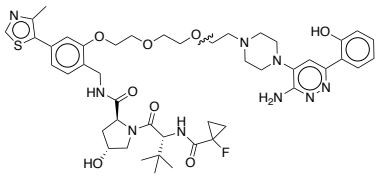
CRBN: dBET55: BRD4 BD1



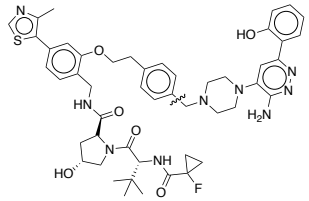
CRBN: dBET70: BRD4 BD1



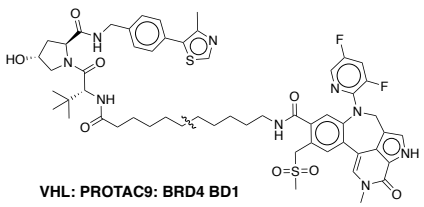
VHL: MZ1: BRD4 BD2



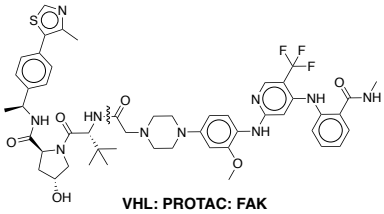
VHL: PROTAC1: SMARCA2



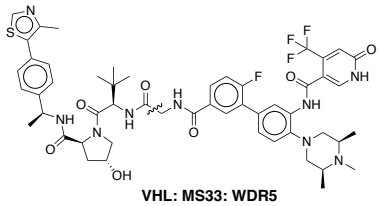
VHL: PROTAC2: SMARCA2



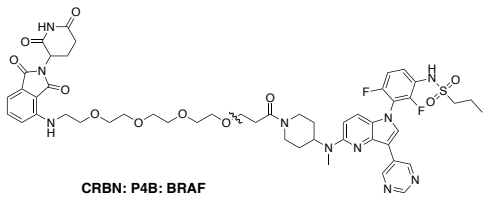
VHL: PROTAC9: BRD4 BD1



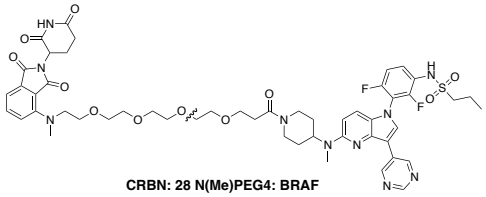
VHL: PROTAC: FAK



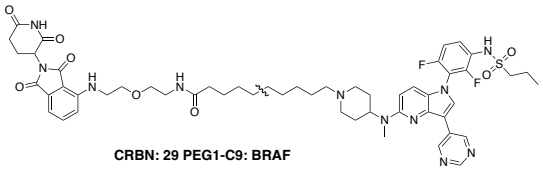
VHL: MS33: WDR5



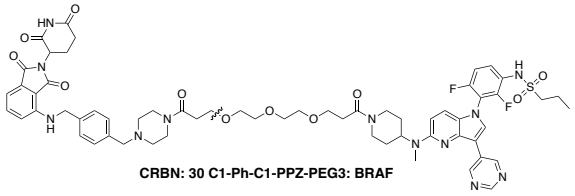
CRBN: P4B: BRAF



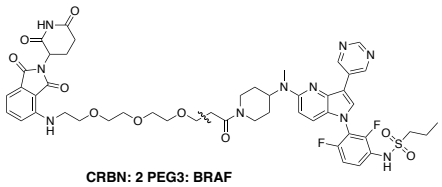
CRBN: 28 N(Me)PEG4: BRAF



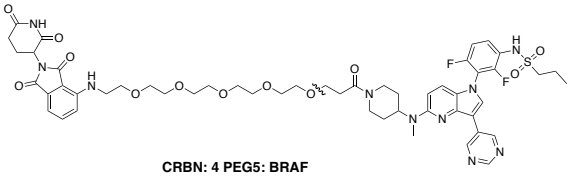
CRBN: 29 PEG1-C9: BRAF



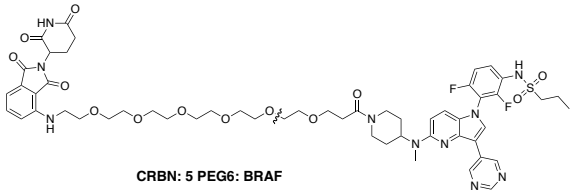
CRBN: 30 C1-Ph-C1-PPZ-PEG3: BRAF



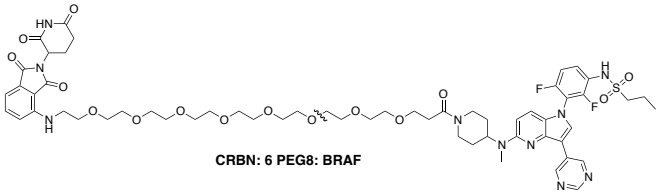
CRBN: 2 PEG3: BRAF



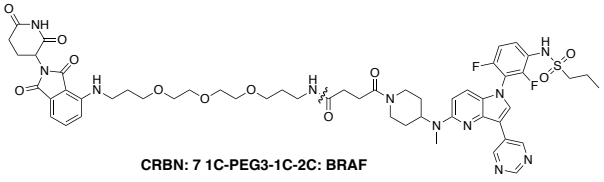
CRBN: 4 PEG5: BRAF



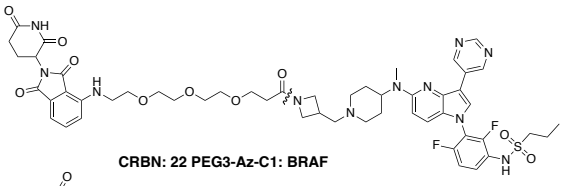
CRBN: 5 PEG6: BRAF



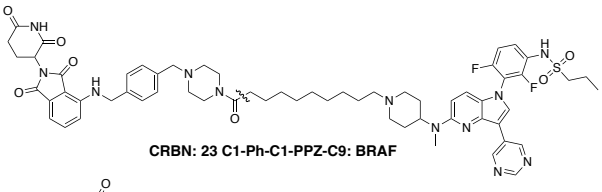
CRBN: 6 PEG8: BRAF



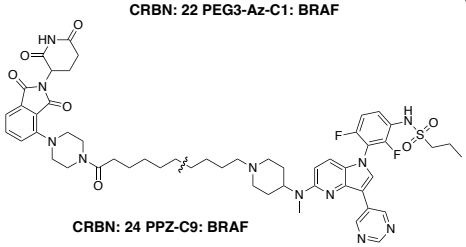
CRBN: 7 1C-PEG3-1C-2C: BRAF



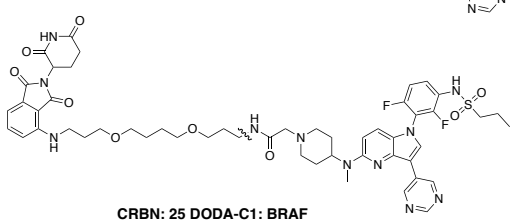
CRBN: 22 PEG3-Az-C1: BRAF



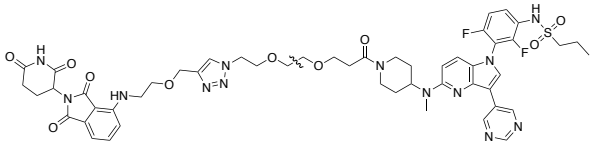
CRBN: 23 C1-Ph-C1-PPZ-C9: BRAF



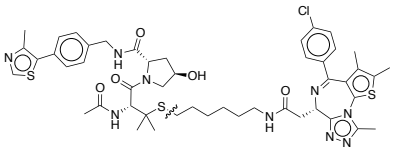
CRBN: 24 PPZ-C9: BRAF



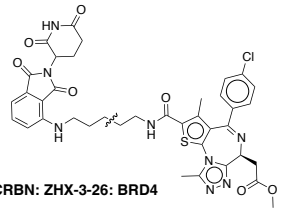
CRBN: 25 DODA-C1: BRAF



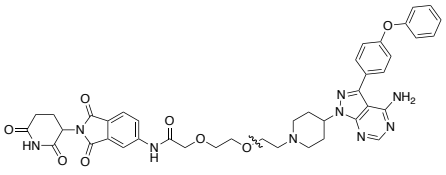
CRBN: 26 PEG1-Trz-PEG2: BRAF



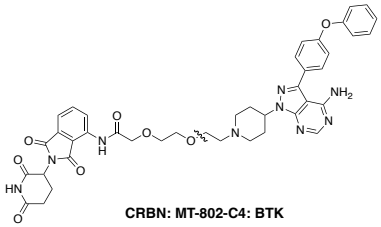
VHL: AT1: BRD4



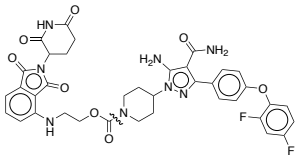
CRBN: ZHX-3-26: BRD4



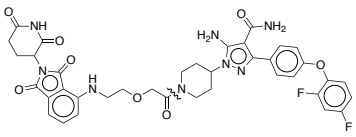
CRBN: MT-802: BTK



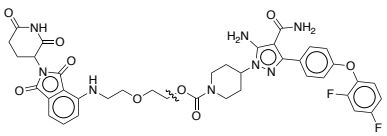
CRBN: MT-802-C4: BTK



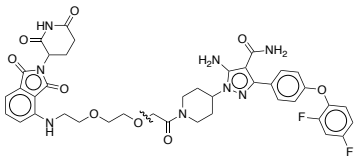
CRBN: PROTAC1: BTK



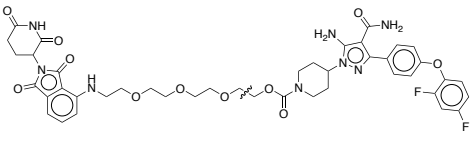
CRBN: PROTAC2: BTK



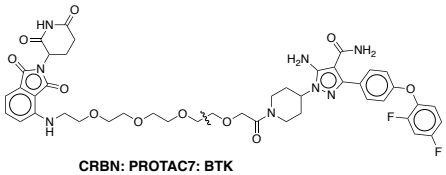
CRBN: PROTAC3: BTK



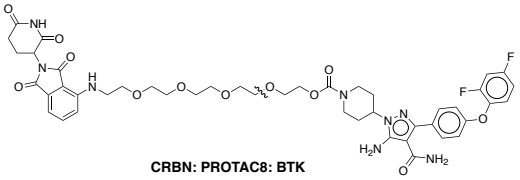
CRBN: PROTAC4: BTK



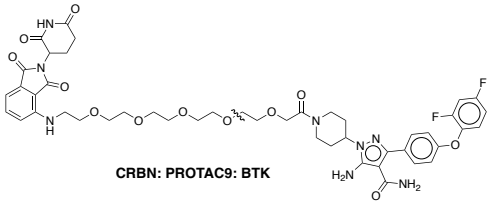
CRBN: PROTAC6: BTK



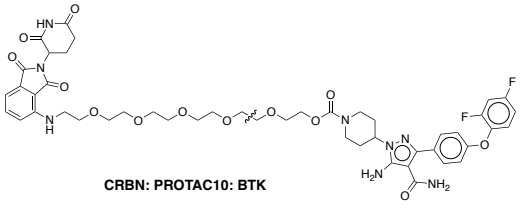
CRBN: PROTAC7: BTK



CRBN: PROTAC8: BTK



CRBN: PROTAC9: BTK



CRBN: PROTAC10: BTK

Supplementary Table 3 | SmRMSD[a] values (Å) for the top 10 models of PROTACs in the benchmark set.

Complex	PROTAC	Rank									
		1	2	3	4	5	6	7	8	9	10
CRBN - BRD4 BD1	dBET6	4.65	6.82	**2.54**	3.12	**2.59**	4.76	5.02	8.07	6.58	5.83
CRBN - BRD4 BD1	dBET23	2.83	4.9	**2.19**	4.69	**2.74**	11.41	7.21	4.63	**1.86**	**2.71**
CRBN - BRD4 BD1	dBET55	**2.21**	3.36	**2.52**	8.69	7.66	5.56	5.60	18.21	8.76	8.43
CRBN- BRD4 BD1	dBET57	3.24	8.12	3.21	3.36	**1.63**	7.18	3.79	**2.73**	**2.57**	3.12
CRBN- BRD4 BD1	dBET70	**1.35**	**2.59**	2.70	5.30	7.25	3.78	6.05	5.01	3.26	**2.33**
CRBN - BRD4 BD1	MZ1	**2.01**	5.63	4.32	9.83	5.80	5.48	7.68	10.98	8.34	8.01
VHL - SMARCA2	PROTAC1	8.92	6.69	4.23	6.34	**1.29**	5.52	9.90	7.31	7.83	11.53
VHL - SMARCA2	PROTAC2	4.45	**1.56**	2.87	4.09	**1.91**	9.64	8.78	7.18	10.14	10.98
VHL - SMARCA4	PROTAC2	11.27	10.17	9.58	8.10	5.72	**1.52**	10.64	10.38	3.34	9.19
VHL - BRD4 BD1	PROTAC9	6.33	4.38	8.37	**2.41**	15.55	11.59	19.00	16.70	18.99	18.67
VHL - FAK	PROTAC	7.56	10.17	11.15	**1.97**	6.88	3.88	8.17	4.85	5.11	5.85
VHL - WDR5	MS67	8.32	6.29	6.37	3.45	4.33	**1.86**	3.55	2.19	2.37	2.53

[a]smRMSD (small molecular RMSD) is defined as the root mean square deviation between the experimental and predicted structures of the protein-bound ligands.

Supplementary Table 4 | Models and their smRMSD values in the largest consensus clusters.

Complex	PROTAC	Rank	smRMSD, Å	In largest cluster	Consensus smRMSD, Å [c]
CRBN-BRD4	dBET6	3	2.54	X[a]	1.89
CRBN-BRD4	dBET6	4	3.12	X	
CRBN-BRD4	dBET6	5	2.59	X	
CRBN-BRD4	dBET23	1	2.83	X	1.86
CRBN-BRD4	dBET23	3	2.19	X	
CRBN-BRD4	dBET23	5	2.74	X	
CRBN-BRD4	dBET23	9	1.86	X	
CRBN-BRD4	dBET23	10	2.71	X	
CRBN-BRD4	dBET55	1	2.21	X	1.92
CRBN-BRD4	dBET55	3	2.52	X	
CRBN-BRD4	dBET70	1	1.35	X	1.58
CRBN-BRD4	dBET70	3	2.70	X	
CRBN-BRD4	dBET70	10	2.33	X	
CRBN-BRD4	dBET57	5	1.63	No	---
CRBN-BRD4	dBET57	8	2.73	No	
CRBN-BRD4	dBET57	9	2.57	No	
VHL-BRD4	MZ1	1	2.01	---	---
VHL-SMARCA2	PROTAC1	5	1.29	X[b]	1.56
VHL-SMARCA2	PROTAC2	2	1.56	X	1.56
VHL-SMARCA2	PROTAC2	3	2.87	X	
VHL-SMARCA2	PROTAC2	5	1.91	X	
VHL-SMARCA4	PROTAC2	6	1.52	X	1.56
VHL-SMARCA4	PROTAC2	9	3.34	X	

[a]All 50 models of CRBN-PROTAC6/23/55/70/57-BRD4 are superimposed. X indicates the models in the largest consensus cluster, which does not include any of the CRBN-PROTAC57-BRD4 models.

[b]All 30 models of VHL-PROTAC1/2-SMARCA2/4 are superimposed. X indicates the models that are members of the largest consensus cluster.

[c]smRMSD of the model closest to the center of the largest consensus cluster.

Supplementary Table 5 | Dissociation constant and cooperativity compared to the weighted sum of acceptable poses for the VHL - PROTAC - bromodomain ternary complexes

		ITC[a]		SPR[b]		Weghted sum of poses
PROTAC	Target	K_d	α	K_d	α	
MZ1	BRD2 BD1	24	2.9	23	1.3	10.5
	BRD2 BD2	28	2.3	0.9	32	10.9
	BRD3 BD1	19	3.5	12	2.4	16.3
	BRD3 BD2	7	10.7	8	3.6	23.6
	BRD4 BD1	*28*	*2.3*	*30*	*0.9*	*6.20*
	BRD4 BD2	**3.7**	**17.6**	**1**	**22**	**31.1**
	BRD2 BD1 KEA	12	7.9			18.7
	BRD4 BD2 QVK	14.9	4.2			2.50
AT1	BRD2 BD1	280	1.4	402	0.3	0.89
	BRD2 BD2	78	4.1	27	4.1	0.54
	BRD3 BD1	207	1.5	133	0.8	1.43
	BRD3 BD2	79	4.3	163	0.7	0.89
	BRD4 BD1	390	1.0	578	0.2	0.37
	BRD4 BD2	**46**	**7.0**	**24**	**4.7**	**2.54**
	BRD2 BD1 KEA	52	7.0			1.24
	BRD4 BD2 QVK	160	2.0			0.10

[a] Determined by isothermal calorimetry (Gadd et al. Structural basis of PROTAC cooperative recognition for selective protein degradation. Nat. Chem. Biol. 13, 514-521 (2017)).

[b] Determined by surface plasmon resonance (Roy et al. SPR-measured dissociation kinetics of PROTAC ternary complexes influence target degradation rate. ACS Chem Biol 14, 361-368 (2019)).

Supplementary Table 6 | Characterization of PROTACs directed against BRAF mutant V600E						
Compound ID	Name	Linker Length (Å)	IC_{50} (nM)	D_{max} (%)	DC_{50} (nM)	Computed activity
2	Pmd-PEG3-BI	15.6	65	65	100	68.29
3	Pmd-PEG4-BI(P4B)	19.1	12	82	15	108.14
4	Pmd-PEG5-BI(P5B)	22.7	25	82	70	113.28
5	Pmd-PEG6-BI	26.1	83	84	90	155.04
6	Pmd-PEG8-BI	33.2	85	65	45	193.19
7	Pmd-1C-PEG3-1C-2C-BI	22.9	82	51	100	21.32
8	VH-PEG4-BI	19.2	67	40	100	
22	Pmd-PEG3-Az-C1-BI	16.2	2	0	-	122.38
23	Pmd-C1-Ph-C1-PPZ-C9-BI	21.5	33	0	-	11.52
24	Pmd-PPZ-C9-BI	17.0	3	0	-	12.51
25	Pmd-DODA-C1-BI	18.5	25	71	25	25.68
26	Pmd-PEG1-Trz-PEG2-BI	18.5	15	72	80	50.49
27	Len(C)≡PEG4-BI	17.1	90	74	50	
28	N(Me)-PEG4	19.1	23	86	15	119.38
29	PEG1-C9	21.0	7	0	-	38.03
30	C1-Ph-C1-PPZ-PEG3	25.6	90	0	-	66.23

Abbreviations: P4B is a pomalidomide- and BI 882370-based PROTAC, bearing a linker with 4 PEG moieties; PEG4 is a linker, containing 4 PEG subunits connected together in a linear fashion; N(Me) is the N-methylated version of pomalidomide; PEG1-C9 is a linker where the PEG1 moiety is connected to a C9-aliphatic carbon chain (nonane) using an amide bond; C1-Ph-C1 is a phenyl group surrounded by methyl (C1) groups on each side; PPZ is a piperazine ring; PEG3 is a linker, containing 3 PEG subunits connected together; linker length was estimated based on its 4 different extended conformations and averaged afterwards (see Methods); IC_{50} (nM) are half-inhibitory concentrations in the Eurofins KinaseProfiler radiometric protein kinase assay for BRAF V600E and BRAF wild type (see Methods); D_{max} (%) is the maximal level of degradation observed for BRAF V600E and wild type (in percent); DC_{50} (nM) is the PROTAC concentration at which 50% of the maximal degradation was observed for BRAF V600E and wild type (in nM).

Supplementary Notes

Chemistry: General Methods

All solvents and commercially available reagents were used as obtained. Oxygen and/or moisture sensitive reactions were carried out under a nitrogen atmosphere in oven- or flame-dried glassware. Chromatography purification was performed using a Teledyne ISCO Combiflash system, Isolera™ One system and prepacked RediSep Rf Normal-phase Silica (60 Å mesh) Flash Cartridges or RediSep Rf C18 Reverse-phase (60 Å mesh) Flash Cartridges.

Nuclear magnetic resonance (NMR) analysis was performed on a Bruker 500 MHz NMR, 400 MHz Bruker Avance III NMR and 500 MHz Agilent DD2 NMR Spectrometers. Spectra were measured at 298K, unless indicated otherwise and were referenced relative to the solvent chemical shift. NMR chemical shifts are expressed in ppm and coupling constants (J) are expressed in Hz. Data were reported as follows: (s) singlet, (d) doublet, (t) triplet, (q) quartet, (m) multiplet, (br) broad.

Target compounds and/or intermediates were characterized by liquid chromatography/mass spectrometry (LCMS) using a Waters Acquity separation module. General conditions are as follows. Mass spectra were acquired on LC/MS systems using electrospray ionization methods from Waters ACQUITY UPLC system with a SQ (single quadrupole) MS or Waters ACQUITY UPLC H-Class system with a 3100 (single quadrupole) MS. *[M+H]* refers to the protonated molecular ion of the chemical species. The mobile phase was 0.1% formic acid in water (solvent A) and 0.1% formic acid in ACN (solvent B). The gradient that was used is presented in the table below.

Time (min)	Flow (mL/min)	%A	%B
Initial	0.4	90	10
1.8	0.4	5	95
2.3	0.4	5	95
2.5	0.4	90	10
3	0.4	90	10
5	0	90	10

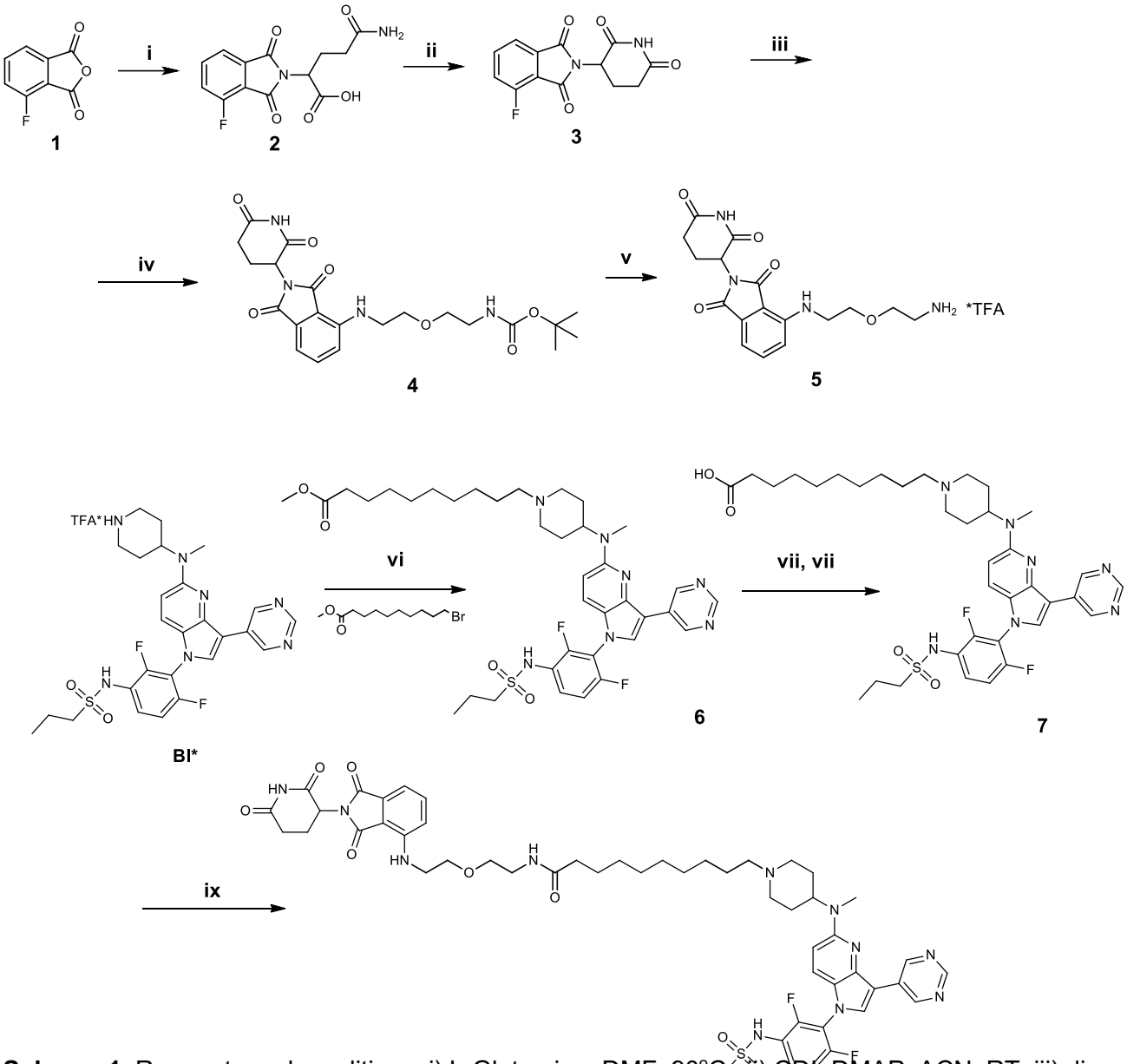
Column 1: Acquity UPLC CSH C18 (2.1 x 50 mm, 130 Å, 1.7 µm. Part No. 186005296) or Column 2: Acquity UPLC BEH C8 (2.1 x 50 mm, 130 Å, 1.7 µm. Part No. 186002877). The columns were used with column temperature maintained at 25 °C. The sample solution injection volume was 1 µL. MassLynx 4.1 was used for data analysis. High resolution mass spectrometry (HRMS) values were measured and calculated with Agilent 6538 UHD Q-TOF MS. All of the values provided correspond to the ionic species of interest and the given ionic formula includes the charging agent.

Abbreviations used: DCM for dichloromethane, EtOAc for ethyl acetate, DMSO for dimethyl sulfoxide, DIPEA for *N,N*-diisopropylethylamine, MeOH for methanol, DMF for *N,N*-dimethylformamide, HATU for 1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluoro-phosphate, TFA for

trifluoroacetic acid, CDI for 1,1'-carbonyldiimidazole, DMAP for 4-dimethylaminopyridine, ACN for acetonitrile, RT for room temperature, MPLC for medium-pressure liquid chromatography.

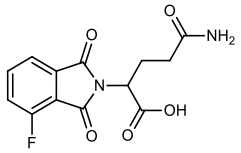
Synthetic methods for the preparation of BI-based PROTACs

Synthesis of 10-(4-((1-(2,6-difluoro-3-(propylsulfonamido)phenyl)-3-(pyrimidin-5-yl)-1H-pyrrolo[3,2-b]pyridin-5-yl)(methyl)amino)piperidin-1-yl)-N-(2-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)ethoxy)ethyl)decanamide (Pmd-PEG1-C9-BI)



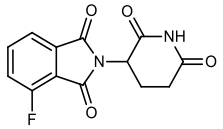
Scheme 1. Reagents and conditions: i) L-Glutamine, DMF, 90 °C; ii) CDI, DMAP, ACN, RT; iii) dioxane, DIPEA, 100-110 °C: iv) dioxane, DIPEA, 100-110 °C: v) TFA, DCM, RT; vi) ACN, K_2CO_3, 50 °C; vii) THF/Water, LiOH, RT; viii) 1M HCl; ix) **5**, HATU, DIPEA, DMF, RT.

5-Amino-2-(4-fluoro-1,3-dioxoisoindolin-2-yl)-5-oxopentanoic acid (2)



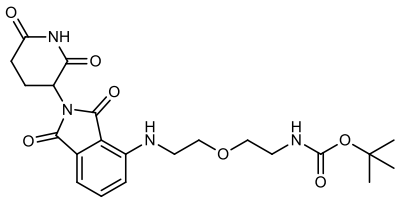
Using a literature procedure[1] starting from 3-fluorophthalic anhydride **1** (8.81 g, 53.0 mmol) and L-glutamine (10 g, 68.4 mmol), derivative **2** (11.4 g, 70%) was obtained. ^{1}H NMR (500 MHz, DMSO-d_6): δ 7.97-7.91 (m, 1H), 7.80-7.90 (m, 2H), 7.20 (br. s., 1H), 6.72 (br. s., 1H), 4.75 (dd, *J*=4.52, 10.88 Hz, 1H), 2.39-2.21 (m, 2H), 2.17-2.07 (m, 2H). ^{19}F NMR (471 MHz, DMSOd_6): δ -114.82 (s, 1F). LCMS (ESI); *m/z*: $[M+H]^+$ calcd. for $C_{13}H_{12}FN_2O_5$, 295.07. Found 295.34.

2-(2,6-Dioxopiperidin-3-yl)-4-fluoroisoindoline-1,3-dione (3)



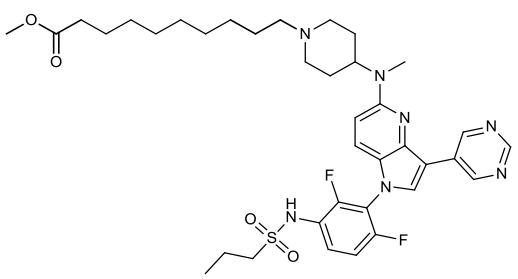
A literature procedure[1] using 5-amino-2-(4-fluoro-1,3-dioxoisoindolin-2-yl)-5-oxopentanoic acid **2** (5.7 g, 19.37 mmol) afforded intermediate **3** (1.1 g, 20%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.11 (br. s., 1H), 7.81-7.71 (m, 2H), 7.44 (t, *J* = 8.4 Hz, 1H), 5.00 (dd, *J* = 12.5, 5.4 Hz, 1H), 2.96-2.73 (m, 3H), 2.21 - 2.14 (m, 1H). ^{19}F NMR (471 MHz, $CDCl_3$): δ -111.69 (s, 1F). LCMS (ESI); *m/z*: $[M+H]^+$ calcd. for $C_{13}H_{10}FN_2O_4$, 277.06. Found 277.30.

tert-Butyl (2-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)ethoxy)ethyl)-carbamate (4)



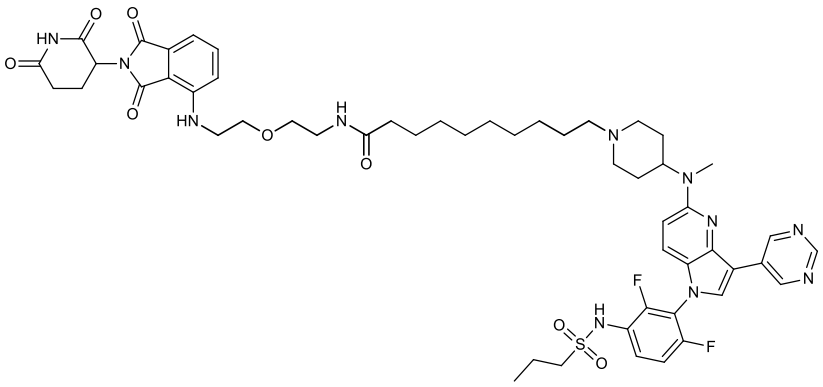
A mixture of *tert*-butyl (2-(2-aminoethoxy)ethyl)carbamate (103 mg, 0.504 mmol) and 2-(2,6-dioxopiperidin-3-yl)-4-fluoroisoindoline-1,3-dione (100 mg, 0.362 mmol) in 1,4-dioxane (5 mL) was stirred at 100 °C overnight in a sealed vial. The reaction mixture was concentrated onto Celite and purified by silica gel chromatography, eluting with DCM/ACN. The relevant fractions were collected and concentrated with a rotary evaporator to afford intermediate **4** (204 mg, 88%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.20 (br s, 1H), 7.42 (dd, *J* = 8.4, 7.1 Hz, 1H), 7.06 (d, *J* = 7.1 Hz, 1H), 6.87 (d, *J* = 8.4 Hz, 1H), 6.46 (br s, 1H), 4.92 (br s, 1H), 4.85 (dd, *J* = 12.2, 5.3 Hz, 1H), 3.61 (t, *J* = 5.3 Hz, 2H), 3.48 (t, *J* = 5.1 Hz, 2H), 3.40 - 3.36 (m, 2H), 3.27 - 3.23 (m, 2H), 2.83 - 2.59 (m, 3H), 2.06 - 2.03 (m, 1H), 1.35 (s, 9H). ^{13}C NMR (126 MHz, $CDCl_3$): δ 171.24, 169.37, 168.47, 167.57, 155.98, 146.81, 136.06, 132.45, 116.76, 111.74, 110.33, 79.24, 70.24, 48.88, 42.21, 40.36, 31.40, 28.37, 22.72. LCMS (ESI); *m/z*: $[M+Na]^+$ calcd. for $C_{22}H_{28}N_4NaO_7$ 483.18. Found 483.19.

Methyl 10-(4-((1-(2,6-difluoro-3-(propylsulfonamido)phenyl)-3-(pyrimidin-5-yl)-1H-pyrrolo[3,2-b]pyridin-5-yl)(methyl)amino)piperidin-1-yl)decanoate (6)



Potassium carbonate (41 mg, 0.3 mmol) was added to the mixture of (*N*-(2,4-difluoro-3-(5-(methyl(piperidin-4-yl)amino)-3-(pyrimidin-5-yl)-1H-pyrrolo[3,2-b]pyridin-1-yl)phenyl)propane-1-sulfonamide trifluoroacetate (**BI***) (67 mg, 0.101 mmol, synthesized using a known procedure[2]) and methyl 10-bromodecanoate (80 mg, 0.3 mmol) in ACN. The reaction mixture was stirred at 50 °C for 8 h. The reaction was diluted with semi-concentrated $NaHCO_3$ solution and extracted with DCM. Organic layer was concentrated and purified by reverse phase chromatography (Water/ACN) to give the title compound **6** (54 mg, 50%) as an off-white semi-solid. ^{1}H NMR (400 MHz, $CDCl_3$): δ 9.53 (s, 2H), 9.12 (s, 1H), 8.54 (s, 2H), 7.76 – 7.59 (m, 2H), 7.35 (dt, *J* = 9.2, 1.9 Hz, 1H), 7.20 (td, *J* = 9.2, 1.9 Hz, 1H), 6.58 (d, *J* = 9.2 Hz, 1H), 5.05 (t, *J* = 12.3 Hz, 1H), 3.73, (s, 1H), 3.70 (s, 3H), 3.26 – 3.12 (m, 2H), 3.00 – 2.91 (m, 6H), 2.83 (t, *J* = 12.1 Hz, 2H), 2.42 (d, *J* = 12.5 Hz, 1H), 2.33 (t, *J* = 7.5 Hz, 2H), 2.02 – 1.91 (m, 4H), 1.64 (q, *J* = 7.3 Hz, 3H), 1.33 (t, *J* = 5.8 Hz, 12H), 1.10 (t, *J* = 7.4 Hz, 3H). ^{19}F NMR (377 MHz, $CDCl_3$): δ -121.94, -129.78. ^{13}C NMR (101 MHz, CDCl3): δ 174.31, 156.04, 155.67, 153.92, 141.94, 128.60, 127.12, 124.47, 123.65, 122.68, 121.01, 116.19, 113.22, 110.90, 104.15, 77.37, 77.06, 76.74, 54.78, 52.69, 51.46, 43.12, 34.08, 32.81, 30.68, 29.37, 29.34, 29.20, 29.10, 29.06, 27.00, 26.51, 25.73, 24.91, 24.63, 17.32, 12.96. LCMS (ES^+): *m/z* 726.4 $[M+1]^+$.

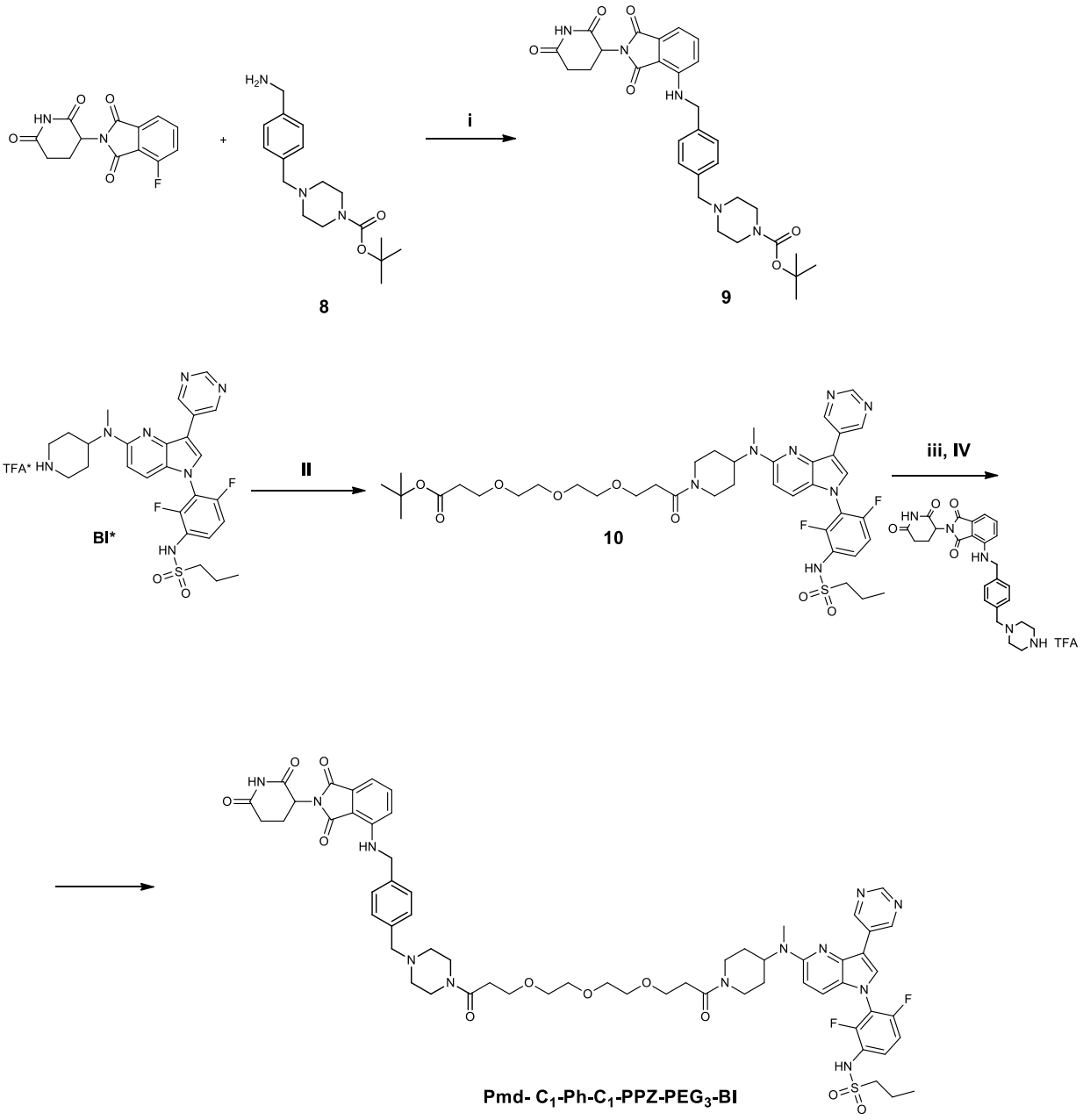
10-(4-((1-(2,6-Difluoro-3-(propylsulfonamido)phenyl)-3-(pyrimidin-5-yl)-1H-pyrrolo[3,2-b]pyridin-5-yl)(methyl)amino)piperidin-1-yl)-N-(2-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)ethoxy)ethyl)decanamide (Pmd-PEG1-C9-BI)



A solution of methyl 10-(4-((1-(2,6-difluoro-3-(propylsulfonamido)phenyl)-3-(pyrimidin-5-yl)-1H-pyrrolo[3,2-b]pyridin-5-yl)(methyl)amino)piperidin-1-yl)decanoate **6** (15 mg, 0.02 mmol) and LiOH (1 mg, 0.02 mmol) in 1 mL THF/water (1:1) was stirred at RT overnight, then the reaction was acidified until it reached pH 2 with 1M HCl, extracted with EtOAc and the organic layer was separated and dried to afford 10-(4-((1-(2,6-difluoro-3-

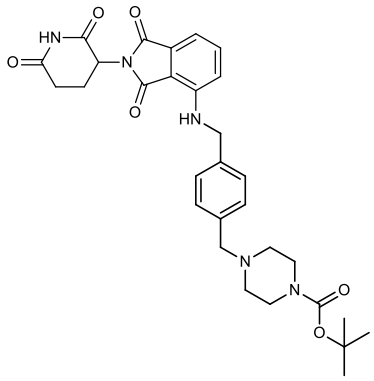
(propylsulfonamido)phenyl)-3-(pyrimidin-5-yl)-1H-pyrrolo[3,2-b]pyridin-5-yl)(methyl)amino)-piperidin-1-yl)decanoic acid **(7)** (11 mg, 100%). The crude product was dissolved in DMF (1 mL), then HATU (9 mg, 0.023 mmol) was added at RT. The mixture was stirring during 5 min before *N,N*-diisopropylethylamine (DIPEA) (14 mg, 0.1 mmol) was added and the mixture was stirred for another 10 min. 4-((2-(2-Aminoethoxy)ethyl)amino)-2-(2,6-dioxopiperidin-3-yl)isoindoline-1,3-dione trifluoroacetate **5** (21 mg, 0.04 mmol) was dissolved in 1 mL of DMF and added to the reaction mixture. After stirring overnight at RT, the reaction mixture was evaporated onto Celite and purified by reverse phase chromatography (Water/ACN). The desired fractions were collected, concentrated and dried under vacuum to afford the desired compound as a dark yellow solid (1.5 mg, 7%). HRMS (*m/z*): $[M]^+$ calcd. for $C_{53}H_{66}F_2N_{11}O_8S$, 1054.4779; found, 1054.4789.

Synthesis of N-(3-(5-((1-(3-(2-(2-(3-(4-(4-(((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)methyl)benzyl)piperazin-1-yl)-3-oxopropoxy)ethoxy)ethoxy)propanoyl)piperidin-4-yl)(methyl)amino)-3-(pyrimidin-5-yl)-1H-pyrrolo[3,2-b]pyridin-1-yl)-2,4-difluorophenyl)propane-1-sulfonamide (Pmd- C1-Ph-C1-PPZ-PEG3-BI)



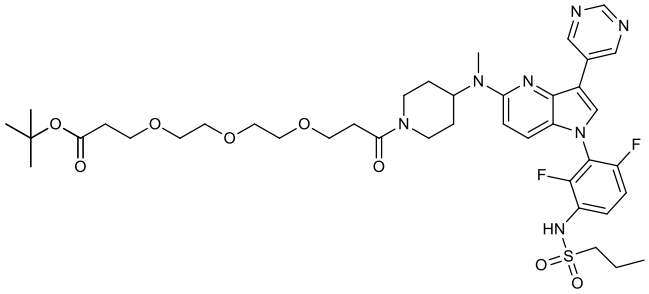
Scheme 2. Reagents and conditions: i) dioxane, DIPEA, 100 ºC; ii) Acid-PEG3-t-butyl ester, HATU, DIPEA, DMF, RT; iii) DCM, TFA, RT; iv) HATU, DIPEA, DMF, RT.

tert-Butyl 4-(4-(((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)methyl)-benzyl)piperazine-1-carboxylate (9)



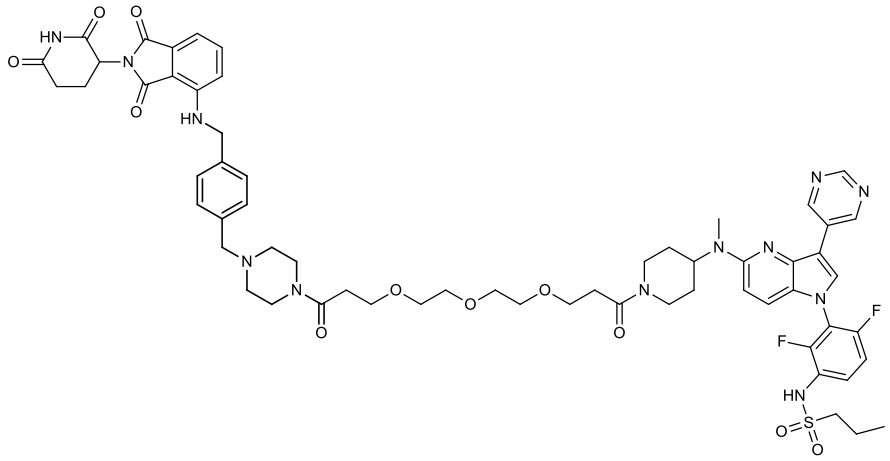
The procedure described above for intermediate **4** using *tert*-butyl 4-(4-(aminomethyl)benzyl)piperazine-1-carboxylate (110 mg, 0.360 mmol) and 2-(2,6-dioxopiperidin-3-yl)-4-fluoroisoindoline-1,3-dione **3** (100 mg, 0.361 mmol) afforded the crude product, which was purified by normal phase chromatography eluting with the mixture MeOH/ACN in DCM. Fractions with pure compound were collected, concentrated with a rotary evaporator to afford intermediate **9** as a bright yellow powder (240 mg, 43%). ^{1}H NMR (400 MHz, CDCl3): δ 8.17 (s, 1H), 7.38 (dd, *J* = 8.5, 7.2 Hz, 1H), 7.24 (d, *J* = 1.9 Hz, 4H), 7.05 (dd, *J* = 7.2, 0.6 Hz, 1H), 6.78 (d, *J* = 8.5 Hz, 1H), 6.61 (t, *J* = 5.9 Hz, 1H), 4.85 (dd, *J* = 12.1, 5.3 Hz, 1H), 4.42 (d, *J* = 5.8 Hz, 2H), 3.46 (s, 2H), 3.37 (d, *J* = 5.3 Hz, 4H), 2.88 – 2.62 (m, 3H), 2.35 (t, *J* = 4.9 Hz, 4H), 2.11 – 2.04 (m, 1H), 1.38 (s, 9H). 13C NMR (126 MHz, CDCl3): δ 171.13, 169.46, 168.42, 167.54, 153.56, 146.63, 136.15, 132.43,129.82, 127.05, 117.07, 111.99, 110.44, 79.71, 62.36, 52.74, 48.90, 46.50, 43.33, 42.81, 31.40, 28.40, 22.78. LCMS (ES^{+}): *m/z* 562.2 $[M+1]^{+}$.

tert-Butyl 3-(2-(2-(3-(4-((1-(2,6-difluoro-3-(propylsulfonamido)phenyl)-3-(pyrimidin-5-yl)-1H-pyrrolo[3,2-b]pyridin-5-yl)(methyl)amino)piperidin-1-yl)-3-oxopropoxy)ethoxy)-ethoxy)propanoate (10)



Acid-PEG3-t-butyl ester (56 mg, 0.183 mmol) was dissolved in DMF (2 mL), and then HATU (64 mg, 0.177 mmol) was added at RT. The mixture was stirring during 10 min before N,N-diisopropylethylamine (DIPEA) (100 mg, 0.77 mmol) was added and the mixture was stirred for another 5 min. N-(3-(5-(2-aminopyrimidin-4-yl)-2-(piperidin-4-yl)thiazol-4-yl)-2-fluorophenyl)-2,6-difluorobenzene-sulfonamide trifluoroacetic acid salt (**BI***) (84 mg, 0.129 mmol) was dissolved in 1 mL of DMF and added to the reaction mixture. After stirring overnight at RT, water (1 mL) was added to the reaction mixture and extracted with EtOAc. The organic extract was evaporated onto Celite and purified by reverse phase chromatography, eluting with Water/ACN. The desired fractions were collected, concentrated and dried under vacuum to afford the desired compound as an off-white solid **10** (54 mg, 51%). ^{1}H NMR (500 MHz, DMSO-d_6) δ = 10.03 - 9.83 (m, 1H), 9.68 - 9.64 (m, 2H), 9.03 (s, 1H), 8.41 (s, 1H), 7.58 (dt, *J* = 6.0, 8.7 Hz, 1H), 7.49 - 7.37 (m, 2H), 6.82 (d, *J* = 9.3 Hz, 1H), 4.66 - 4.57 (m, 2H), 4.13 - 4.03 (m, 1H), 3.61 - 3.43 (m, 12H), 3.27 - 3.07 (m, 4H), 3.01 - 2.89 (m, 3H), 2.73 - 2.55 (m, 4H), 2.47 - 2.36 (m, 2H), 1.81 - 1.69 (m, 5H), 1.46 - 1.20 (m, 9H), 0.99 (t, *J* = 7.5 Hz, 3H). ^{19}F NMR (471 MHz, $CDCl_3$): δ -130.94 (s, 1F) -121.56 (s, 1F). ^{13}C NMR (126MHz, $CDCl_3$): δ 170.95, 169.36, 156.07, 155.86, 153.94, 153.75, 150.88, 148.91, 146.91, 142.15, 128.43, 126.89, 124.27, 123.63 - 123.55 (m), 122.94 – 122.85 (m), 120.82, 116.54 - 116.28 (m), 112.94 - 112.61 (m), 111.20, 104.03, 80.52, 70.54, 70.52, 70.47, 70.42, 70.32, 67.50, 66.89, 54.73, 53.57, 45.72, 41.78, 36.27, 33.60, 30.57, 29.55, 28.94, 28.10, 17.31, 12.92 .LCMS (ES$^+$): *m/z* 830.4 [M+1]$^+$.

N-(3-(5-((1-(3-(2-(2-(3-(4-(4-(((2-(2,6-Dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)-amino)methyl)benzyl)piperazin-1-yl)-3-oxopropoxy)ethoxy)ethoxy)propanoyl)piperidin-4-yl)(methyl)amino)-3-(pyrimidin-5-yl)-1H-pyrrolo[3,2-b]pyridin-1-yl)-2,4-difluorophenyl)-propane-1-sulfonamide (Pmd-C1-Ph-C1-PPZ-PEG3-BI)

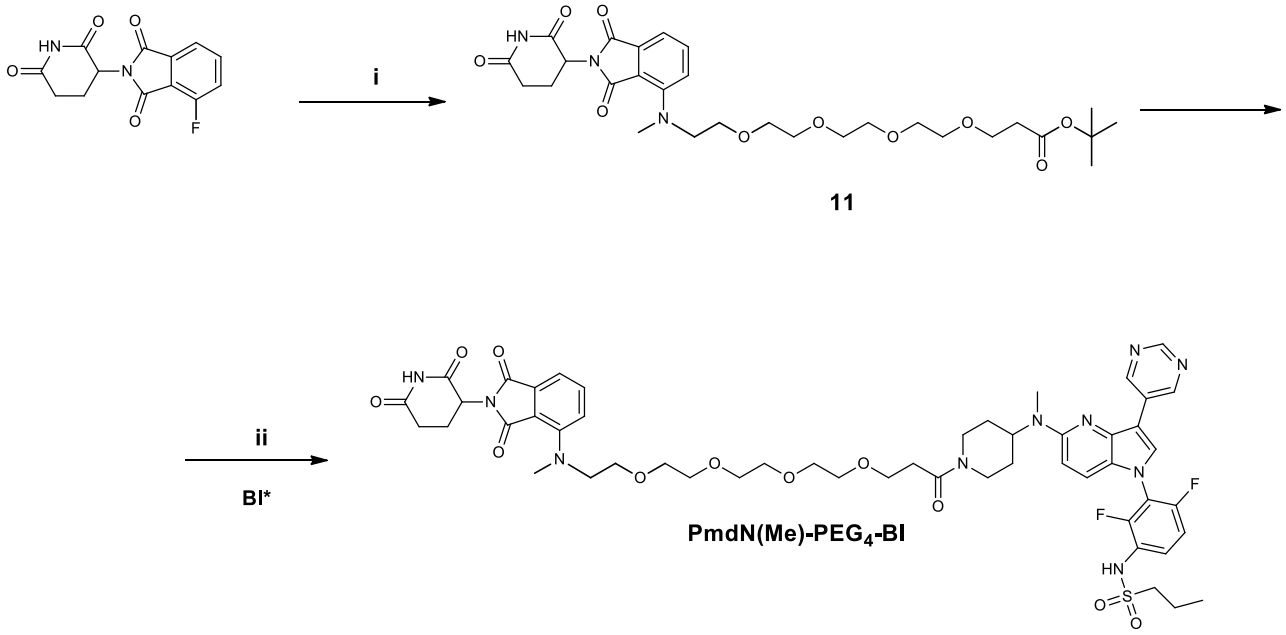


Tert-butyl 3-(2-(2-(3-(4-((1-(2,6-difluoro-3-(propylsulfonamido)phenyl)-3-(pyrimidin-5-yl)-1H-pyrrolo[3,2-b]pyridin-5-yl)(methyl)amino)piperidin-1-yl)-3-oxopropoxy)ethoxy)-ethoxy)propanoate **(10)** (42 mg, 0.05 mmol) was dissolved in a mixture of TFA (1 mL) and DCM (1 mL). The reaction mixture was stirred for 5 h at RT. After the reaction was complete, the solvent was evaporated and crude product was dried under high vacuum overnight to afford 3-(2-(2-(3-(4-((1-(2,6-difluoro-3-(propylsulfonamido)phenyl)-3-(pyrimidin-5-yl)-1H-pyrrolo[3,2-b]pyridin-5-yl)(methyl)amino)piperidin-1-yl)-3-oxopropoxy)ethoxy)ethoxy)propanoic acid in a quantitative yield (40 mg, 0.055 mmol) which was used in the next step without any further purification. This

material was dissolved in DMF (3 mL), and then HATU (36 mg, 0.1 mmol) was added at RT. The mixture was stirring during 10 min before DIPEA (52 mg, 0.4 mmol) was added and the mixture was stirred for another 5 min. 2-(2,6-dioxopiperidin-3-yl)-4-((4-(piperazin-1-ylmethyl)benzyl)amino)-isoindoline-1,3-dione trifluoroacetate (27 mg, 0.048 mmol) was dissolved in 1 mL of DMF and added to the reaction mixture. After stirring overnight at RT, the reaction mixture was evaporated onto Celite and purified by reverse phase chromatography, eluting with Water/ACN (0% to 100% of ACN) with 0.1% formic acid. The desired fractions were collected, concentrated and dried under vacuum to afford the desired compound as a dark yellow semi-solid (10 mg, 18%). HRMS (*m/z*): $[M+Na]^+$ calcd. for $C_{61}H_{70}F_2N_{12}NaO_{11}S$, 1239.4868; found, 1239.4863.

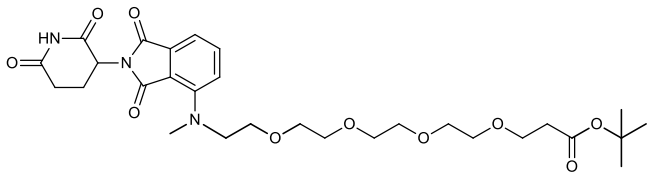
Synthesis of N-(3-(5-((1-(2-(2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)-5,8,11,14-tetraoxa-2-azaheptadecan-17-oyl)piperidin-4-yl)(methyl)amino)-3-(pyrimidin-5-yl)-1H-pyrrolo[3,2-b]pyridin-1-yl)-2,4-difluorophenyl)propane-1-sulfonamide

(PmdN(Me)-PEG4-Bl)



Scheme 15. Reagents and conditions: i) dioxane, methylamino-PEG4-t-butyl ester, DIPEA, 100 °C; ii) DCM, TFA, RT; iii) DMF, DIPEA, HATU, RT.

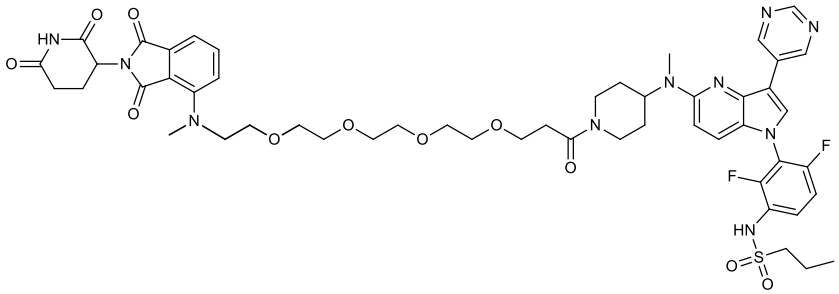
tert-Butyl 2-(2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)-5,8,11,14-tetraoxa-2-azaheptadecan-17-oate (11)



The title compound was prepared according to procedure similar to that described above for intermediate **4** from *tert*-butyl 5,8,11,14-tetraoxa-2-azaheptadecan-17-oate and 2-(2,6-dioxopiperidin-3-yl)-4-fluoroisoindoline-

1,3-dione **(3)**. The title compound was obtained in a partially pure form as a dark yellow oil, which was purified by normal phase chromatography eluting with hexane/isopropanol to give the product as a bright yellow powder (50 mg, 30%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.33 (s, 1H), 7.51 (dd, *J* = 8.5, 7.1 Hz, 1H), 7.12 (d, *J* = 7.1 Hz, 1H), 6.95 (d, *J* = 8.5 Hz, 1H), 4.98 – 4.88 (m, 1H), 3.81 – 3.60 (m, 16H), 3.51 (t, *J* = 5.5 Hz, 2H), 3.23 (s, 3H), 3.04 – 2.92 (m, 1H), 2.80 – 2.70 (m, 2H), 2.64 (t, *J* = 6.0 Hz, 2H), 2.11(m, 1H), 1.45 (9H). LCMS (ES$^+$): *m/z* 592.2 $[M+1]^+$.

N-(3-(5-((1-(2-(2-(2,6-Dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)-5,8,11,14-tetraoxa-2-azaheptadecan-17-oyl)piperidin-4-yl)(methyl)amino)-3-(pyrimidin-5-yl)-1H-pyrrolo[3,2-b]pyridin-1-yl)-2,4-difluorophenyl)propane-1-sulfonamide (PmdN(Me)-PEG4-BI)



tert-Butyl 2-(2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)-5,8,11,14-tetraoxa-2-azaheptadecan-17-oate **(11)** (49 mg, 0.082 mmol) was dissolved in a mixture of TFA (1 mL) and DCM (1 mL). The reaction mixture was stirred for 3 h at RT. After the reaction, the solvent was evaporated and crude product was dried under high vacuum overnight. 2-(2-(2,6-Dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)-5,8,11,14-tetraoxa-2-azaheptadecan-17-oic acid was obtained in a quantitative yield (45 mg, 100 %) and used in the next step without any further purification. This material was then dissolved in DMF (2 mL), then HATU (33 mg, 0.091 mmol) was added at RT. The mixture was stirred for 10 min before DIPEA (30 mg, 0.23 mmol) was added. The mixture was stirred for another 10 min, then *N*-(3-(5-(2-aminopyrimidin-4-yl)-2-(piperidin-4-yl)thiazol-4-yl)-2-fluorophenyl)-2,6-difluorobenzene-sulfonamide trifluoroacetic acid salt (**BI***) (33 mg, 0.05 mmol) was dissolved in 1 mL of DMF and added to the reaction mixture. After stirring overnight at RT, the mixture was evaporated onto Celite and purified by reverse phase chromatography, eluting with Water/ACN. The desired fractions were collected, concentrated and dried to afford the residue, which was repurified by normal phase chromatography eluting with MTBE/MeOH to obtain the desired compound as a yellow powder (5.3 mg, 11%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 9.58 (s, 2H), 9.13 (s, 1H), 8.59 (s, 1H), 7.74 – 7.65 (m, 2H), 7.51 (dd, *J* = 8.6, 7.0 Hz, 1H), 7.34 (dt, *J* = 7.7, 1.6 Hz, 2H), 7.24 – 7.13 (m, 2H), 6.64 (d, *J* = 9.2 Hz, 1H), 5.02 – 4.79 (m, 2H), 4.08 (d, *J* = 13.6 Hz, 1H), 3.85 - 3.59 (m, 19H), 3.28 (t, *J* = 12.5 Hz, 1H), 3.21 – 3.15 (m, 2H), 3.13 (s, 3H), 2.97 (s, 3H), 2.92 – 2.67 (m, 6H), 2.16 – 2.09 (m, 1H), 2.02 – 1.81 (m, 4H), 1.72 (tt, *J* = 13.5, 7.1 Hz, 2H), 1.11 (t, *J* = 7.4 Hz, 3H). HRMS (*m/z*): $[M]^+$ calcd. for $C_{51}H_{61}F_2N_{10}O_{11}S$, 1059.4205; found, 1059.4207.

REFERENCES

1. Lu, J.; Qian, Y; Altieri, M; Dong, H; Wang, J; Raina, K; Hines, J; Winker, J. D.; Crew, A. P., Coleman K.; Crews C. M. Hijacking the E3 Ubiquitin Ligase Cereblon to Efficiently Target BRD4. *Chemistry & Biology* **22***,* 755-763 (2015).

2. Stadtmueller, H., New azaindolylphenyl sulfonamides as serine/threonine kinase inhibitors. WO2012/104388 A1 (2012).