Supporting Information

Discovery of new binders for DCAF1, an emerging ligase target in the Targeted Protein Degradation field

Anna Vulpetti,[*[a]] Philipp Holzer, [a] Niko Schmiedeberg, [a] Patricia Imbach-Weese, [a] Carole Pissot-Soldermann, [a] Gregory J. Hollingworth, [a] Thomas Radimerski, [b, d] Claudio Thoma, [b, d] Therese-Marie Stachyra, [b] Matthias Wojtynek, [c, e] Magdalena Maschlej, [c] Suzanne Chau, [c] Ansgar Schuffenhauer, [c] César Fernández, [c] Martin Schröder, [c] Martin Renatus [c, d]

Corresponding author:
Dr. Anna Vulpetti
Global Discovery Chemistry, Novartis Institutes for BioMedical Research, Basel 4002 (Switzerland)
E-mail: anna.vulpetti@novartis.com

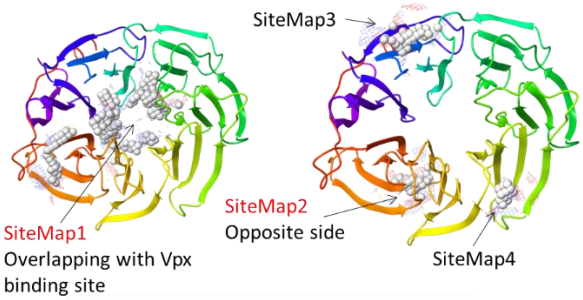
[a] Global Discovery Chemistry, Novartis Institutes for BioMedical Research, Basel 4002 (Switzerland);
[b] Oncology Drug Discovery, Novartis Institutes for BioMedical Research, Basel 4002, (Switzerland);
[c] Chemical Biology & Therapeutics, Novartis Institutes for BioMedical Research, Basel 4002 (Switzerland);
[d] Ridgeline Discovery GmbH, Aeschenvorstadt 36, Basel 4051 (Switzerland);
[e] Institute of Biochemistry, ETH, Otto-Stern-Weg 3, Zurich 8093 (Switzerland).

CONTENTS

SUPPLEMENTARY DATA

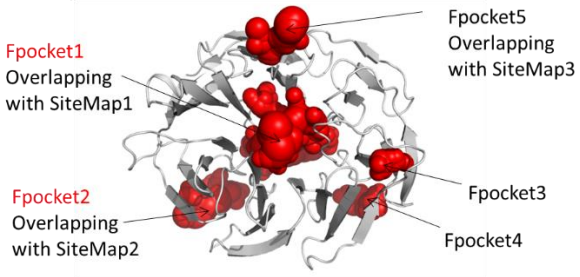
EXPERIMENTAL SECTION

a)



SiteMap	SiteScore	DScore	Size	Volume (Å³)
SiteMap1	1.055	1.085	272	1331.87
SiteMap2	0.751	0.578	45	136.52
SiteMap3	0.591	0.522	29	62.08
SiteMap4	0.488	0.406	17	51.45

b)



Fpocket	Score	Druggability Score	Total SASA	Volume (Å³)
Fpocket1	52.834	0.654	1297.795	3047.063
Fpocket2	8.714	0.082	409.026	989.772
Fpocket3	7.653	0.017	34.614	123.459
Fpocket4	2.860	0.004	193.488	328.318
Fpocket5	1.561	0.008	248.733	592.836

Figure S1. Crystal structure of human DCAF1 (amino acid residues 1058-1396) bound to human SAMHD1 (amino acid residues 582-626) and Vpx isolated from sooty mangabey (PDB ID code: 4CC9)[1]. Both in silico pocket detection approaches were performed on DCAF1 after removal of the other two proteins: **a)** Sitemap and **b)** fpocket. Both methods identified a central region as the major druggable pocket. This corresponds to the site called SiteMap1 and Fpocket1.

Additional cavities, with much lower druggability scores, are also reported. The second ranked binding pocket by both SiteMap and fpocket (i.e. SiteMap2 and Fpocket2) is also in agreement.

Two scores, SiteScore and Dscore, are calculated by SiteMap[2] using different weights for three descriptors: n = the number of site points (related to the size of the pocket) (capped at 100); e = enclosure (related to how open the site is to solvent); p = hydrophilic score (subtracted as penalty) are defined as: SiteScore = 0.0733 sqrt(n) + 0.6688 e - 0.20 p and Dscore = 0.094 sqrt(n) + 0.60 e - 0.324 p. The SiteMap DScore < 0.83 is an indication of undruggable target; 0.83 < DScore < 0.98 of a difficult target; DScore > 0.98 of a druggable target.

The fpocket Druggability Score[3] < 0.5 is an indication of undruggable target and > 0.5 of a druggable target. SASA stands for solvent accessible surface area.

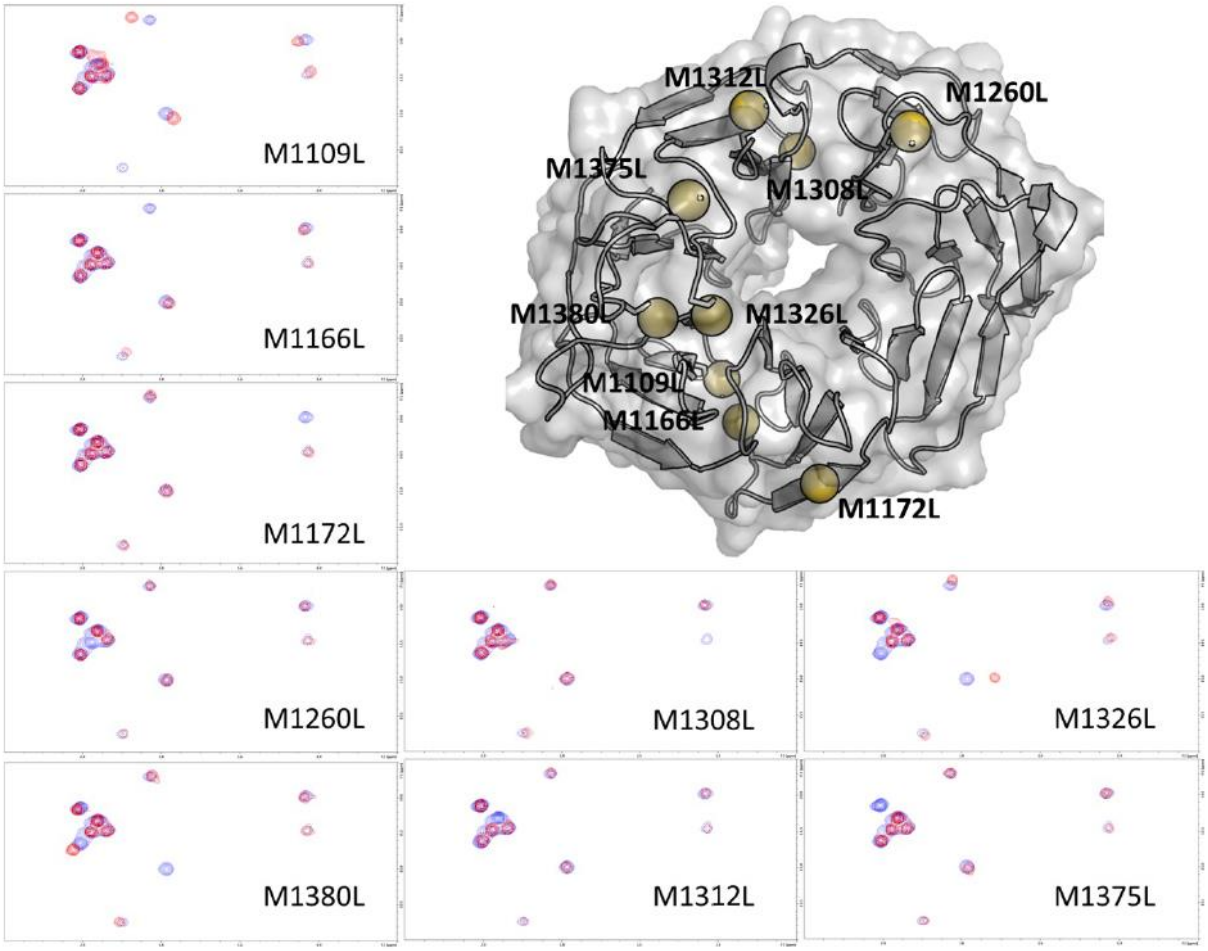


Figure S2. Assignment of methionine methyl resonances for DCAF1. Superposition of 2D [^{13}C, ^{1}H]-HMQC spectra for the methionine to leucine mutants (red) with wt-DCAF1 (blue), all of them selectively ^{13}C-labeled at the methionine C$^{\varepsilon}$ positions. Nine individual M→L mutants were expressed, isotopically labeled and purified similarly as the wildtype protein. Analysis of the corresponding spectra allowed us to unambiguously assign all methionine methyl signals, manifested by lack of the peak for the mutated residues and by chemical shift perturbation for methionine residues located nearby to the mutated one. On the top right, the DCAF1 structure is shown in grey with the sulfur of the methionine residues shown as yellow spheres.

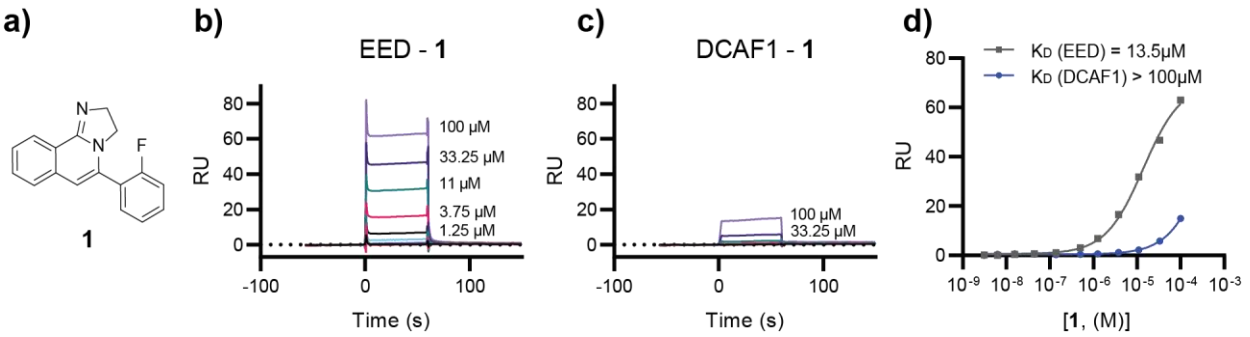


Figure S3. **a)** Compound **1** structure. **b)** SPR sensorgrams of **1** binding to EED at a 3-fold concentration series ranging from 1.25 µM to 100 µM. **c)** SPR sensorgrams of **1** binding to DCAF1 at a 3-fold concentration series ranging from 1.25 µM to 100 µM. **d)** Affinity for the two proteins was calculated using equilibrium 1:1 binding model.

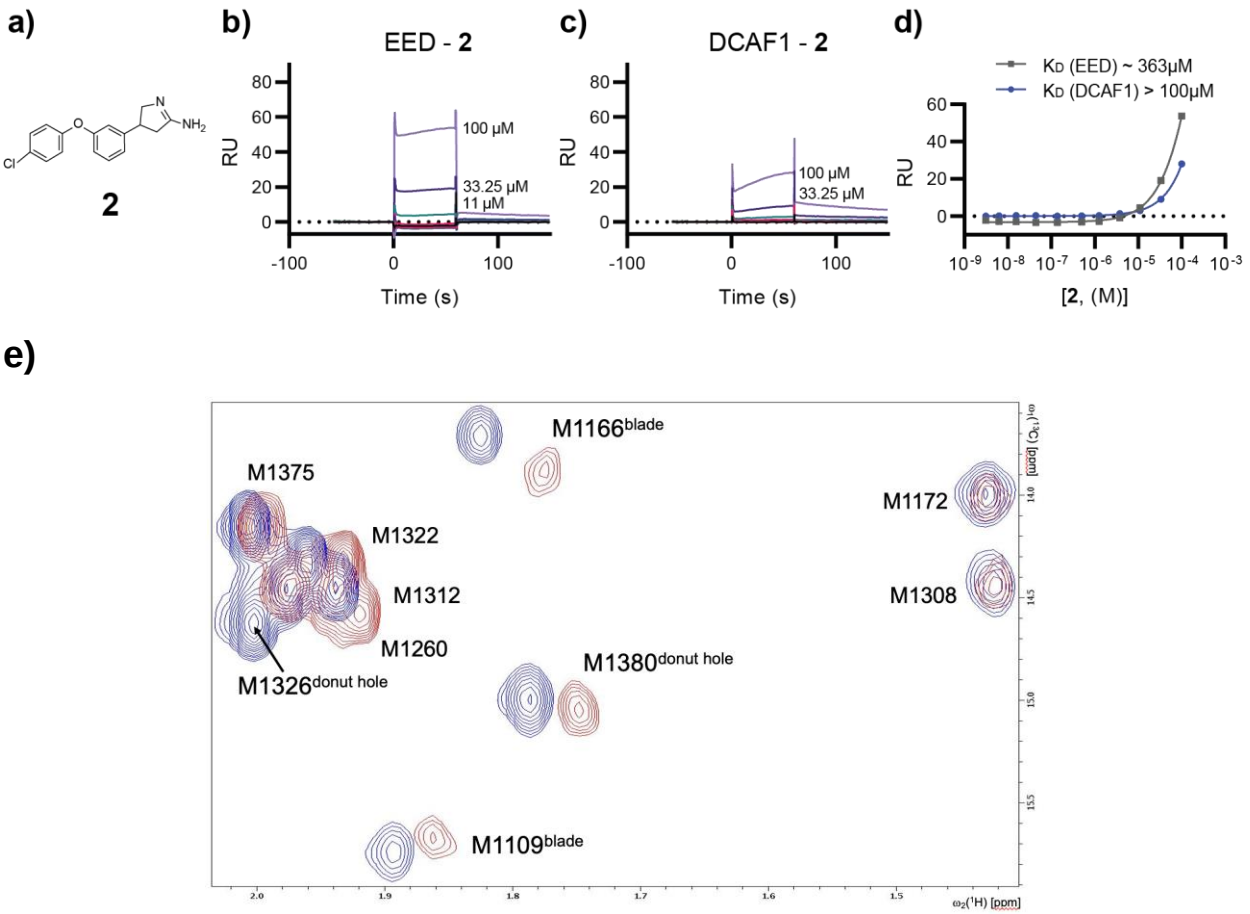


Figure S4. a) Compound **2** structure. **b)** SPR sensorgrams of **2** binding to EED at a 3-fold concentration series ranging from 1.25 µM to 100 µM. **c)** SPR sensorgrams of **2** binding to DCAF1 at a 3-fold concentration series ranging from 1.25 µM to 100 µM. **d)** Affinity for the two proteins was calculated using equilibrium 1:1 binding model. **e)** Overlay of the methyl region of 2D [^{13}C,^{1}H]-HMQC spectra of selectively ^{13}C$^{\varepsilon}$-methionine labeled DCAF1 in the absence (blue) and in the presence of the compound **2** (red). The concentrations of the protein and the compound were 15 µM and 480 µM, respectively.

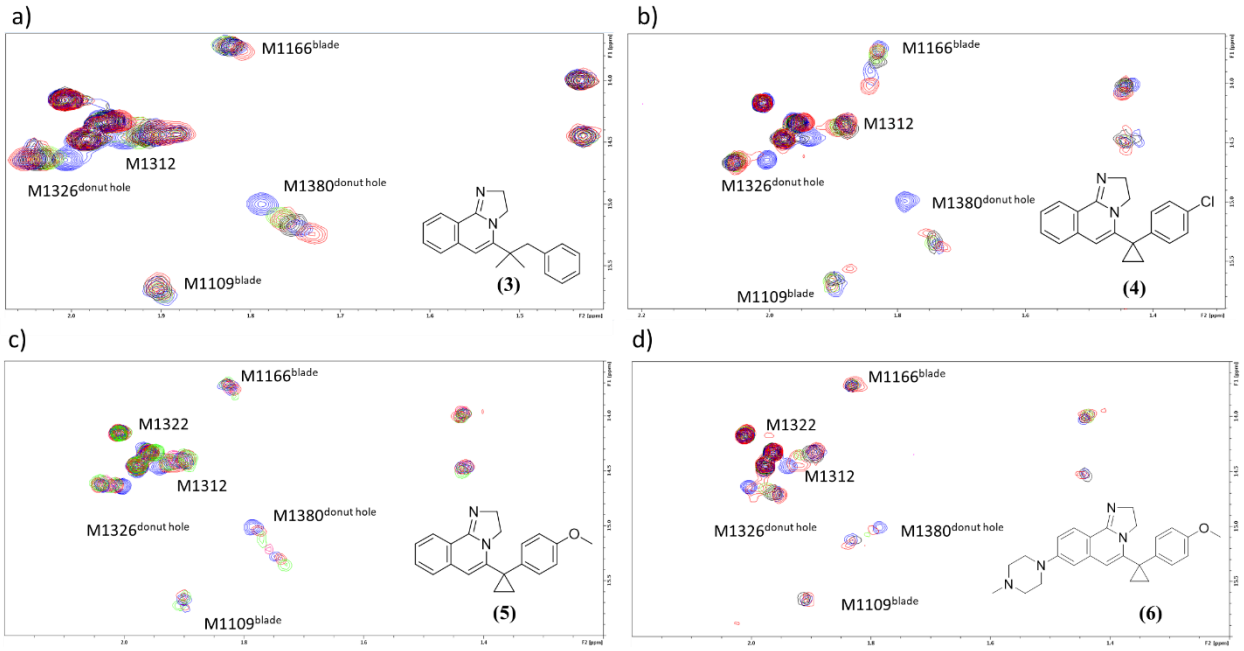


Figure S5. 2D [^{13}C,^{1}H]-HMQC spectra of selectively $^{13}C^{\varepsilon}$-methionine labeled DCAF1 in the absence and in the presence of different concentrations of compounds **3-6**. The chlorine group of compound **4** appears to mediate stronger interactions than the methoxy group of **5** in the blade pocket (see M1166 cross peak). The results for compound **6** indicate that the extension out of the donut hole is tolerated and helpful to abolish binding to the blade region. **a)** The concentrations of the protein and the compound were 30 µM and 0, 30, 60, 120, 240, 480 µM; **b)** The concentrations of the protein and the compound were 15 µM and 0, 30, 60, 120, 240, 480 µM; **c)** The concentrations of the protein and the compound were 15 µM and 0, 15, 30, 60, 120, 240, 480 µM; **d)** The concentrations of the protein and the compound were 15 µM and 0, 7.5, 15, 30, 60, 120 µM. The compound concentrations were chosen according to their binding affinities and the compound solubility in NMR buffer, which has been previously determined by quantitative 1D proton NMR.

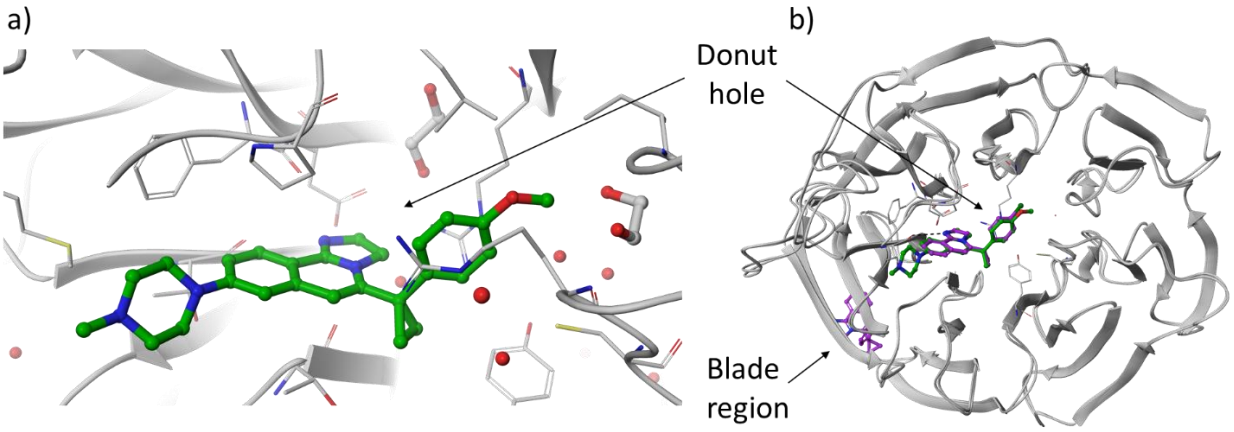


Figure S6. a) Close up view of the DCAF1 crystal structure in complex with compound **6** (8OGA PDB id, green). **b)** DCAF1 crystal structure in complex with compound **6** (ligand bound to donut hole shown in green) overlaid with DCAF1 crystal structure in complex with compound **4** (ligands bound to donut hole and blade region shown in purple). DCAF1 protein is shown in grey.

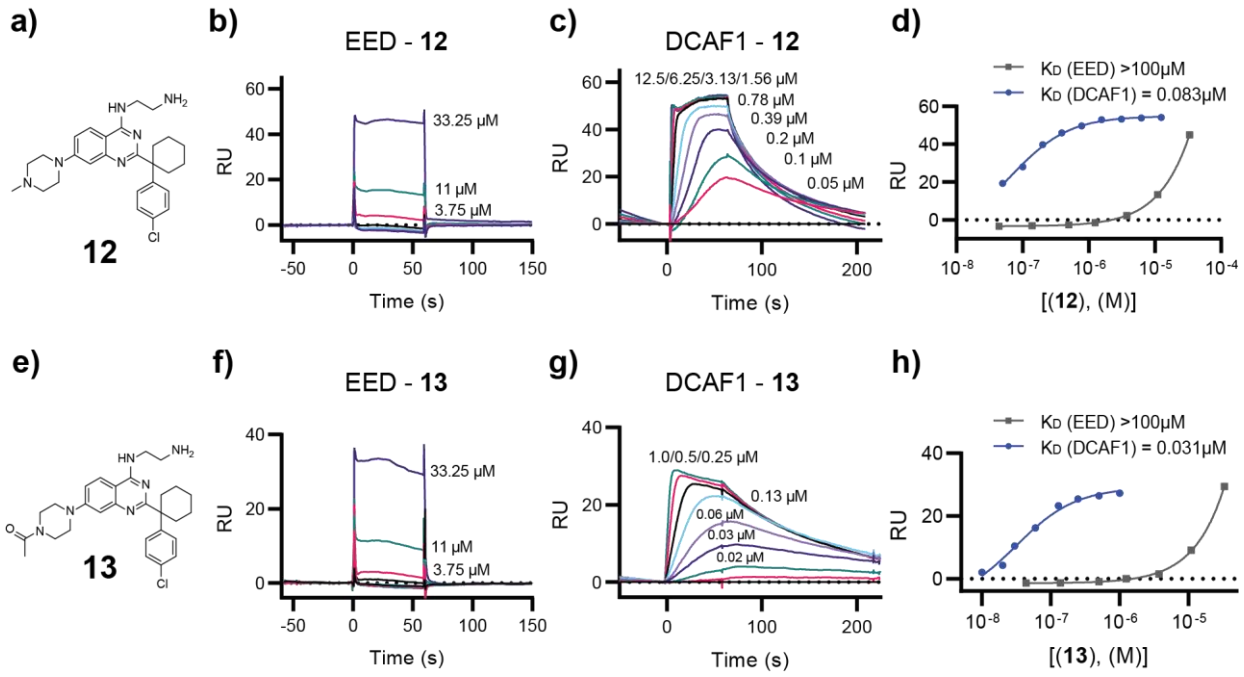


Figure S7. a) Compound **12** structure. **b)** SPR sensorgrams of **12** binding to EED. **c)** SPR sensorgrams of **12** binding to DCAF1. **d)** Affinity of compound **12** for the two proteins was calculated using equilibrium 1:1 binding model. **e)** Compound **13** structure. **f)** SPR sensorgrams of **13** binding to EED. **g)** SPR sensorgrams of **12** binding to DCAF1. **h)** Affinity of compound **13** for the two proteins was calculated using equilibrium 1:1 binding model.

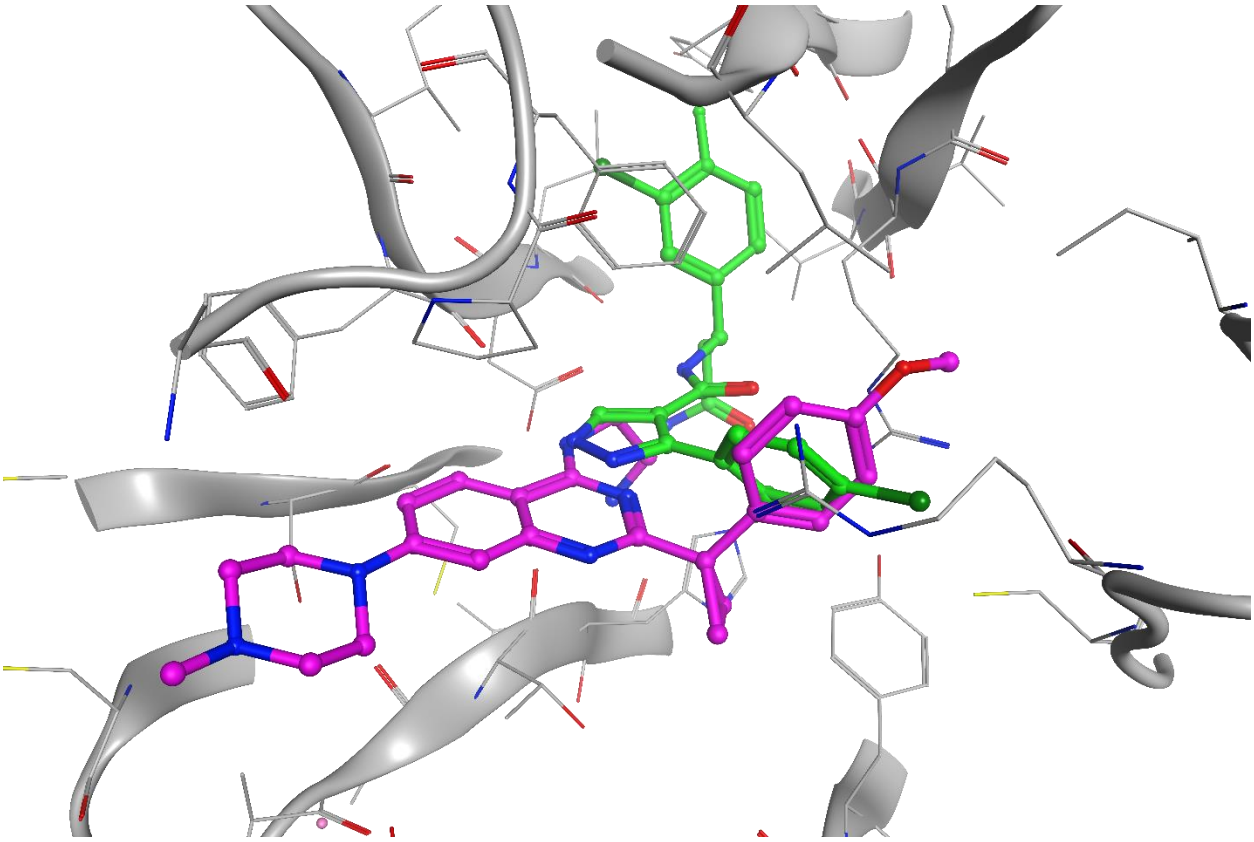


Figure S8. Overlay of DCAF1 crystal structure in complex with compound **8** (8OGB PDB id, purple), of which only the ligand is shown, with the DCAF1 crystal structure in complex with compound **26e** of reference[13] (8F8E PDB id, green).

Table S1. Data collection and refinement statistics.

Complex	**DCAF1-1**	**DCAF1-1**	**DCAF1-2**
PDB accession code	**8OG6**	**8OG5**	**8OG7**
Beamline	**SLS PXII**	**SLS PXII**	**SLS PXII**
Protein	DCAF1(1039-1401)Q1250L Trypsin digested	DCAF1(1039-1401)Q1250L Trypsin digested	DCAF1(1039-1401)Q1250L Trypsin digested
Co-crystallization/soak	Soak	Soak	Soak
Data Collection			
Spacegroup	*$P6_122$*	*$P6_122$*	*$P6_122$*
Cell dimensions	*a = b = 82.2* Å, *c = 237.8* *α = γ* = 90.0˚; *β* = 120˚	*a* = b = 81.9 Å, *c* = 236.8 Å *α = γ* = 90.0˚; *β* = 120˚	*a* = b = 81.8 Å, *c* = 235.6 Å *α = γ* = 90.0˚; *β* = 120˚
Resolution (Å) [a]	47.56-2.24 (2.32-2.24)	52.76-2.20 (2.27-2.20)	52.59-2.64 (2.77-2.64)
No. unique reflections [a]	23,772 (2,126)	24,916 (2,097)	13,467 (1,764)
Completeness[a] (%)	100.0 (99.9)	100.0 (100.0)	94.4 (95.9)
I/σI[a]	23.3 (4.3)	17.2 (1.9)	11.1 (2.4)
R_{merge}[a]	0.071 (0.590)	0.106 (1.640)	0.202 (1.402)
CC (1/2) [a]	0.999 (0.932)	0.999 (0.717)	0.994 (0.752)
Redundancy[a]	12.7 (12.8)	12.6 (13.3)	13.3 (13.5)
Refinement			
No. atoms in refinement (P/L/W)[b]	2,416/44/197	2,382/68/207	2,412/28/162
B factor (P/L/W)[b] (Å^2)	51/71/55	52/75/58	50/89/50
R_{fact} (%)	20.3	19.7	15.8
R_{free} (%)	23.1	23.5	20.6
Rmsd [c] (Å)	0.020	0.016	0.013
Rmsd [c] angle (°)	2.15	1.96	1.92
Molprobity Ramachandran			
Favoured (%)	95.62	95.22	96.62
Outliers (%)	0.34	0.34	0.34
Crystallization conditions	0.1M Tris pH 7.0, 2.13M Lithium acetate	0.1M Tris pH 7.0, 2.13M Lithium acetate	0.1M MES pH6.5, 2.13M Lithium acetate

[a] Values in brackets show the statistics for the highest resolution shell.

[b] P/L/W indicate protein, ligand molecules, and water molecules, respectively.

[c] Rmsd indicates root-mean-square deviation

Table S1 (cont.). Data collection and refinement statistics.

Complex	**DCAF1-3**	**DCAF1-4**	**DCAF1-6**
PDB accession code	**8OG8**	**8OG9**	**8OGA**
Beamline	**SLS PXII**	**SLS PXII**	**SLS PXII**
Protein	DCAF1(1039-1401)Q1250L Trypsin digested	DCAF1(1039-1401)Q1250L Trypsin digested	DCAF1(1079-1393)Q1250L
Co-crystallization/soak	Soak	Soak	Soak
Data Collection			
Spacegroup	*$P6_122$*	*$P6_122$*	*$P6_122$*
Cell dimensions	*a = b = 81.1* Å, *c = 233.0* $\alpha = \gamma = 90.0^{\circ}$; $\beta = 120^{\circ}$	*a* = b = 83.0 Å, *c* = 237.6 Å $\alpha = \gamma = 90.0^{\circ}$; $\beta = 120^{\circ}$	*a* = b = 81.6 Å, *c* = 234.8 Å $\alpha = \gamma = 90.0^{\circ}$; $\beta = 120^{\circ}$
Resolution (Å) [a]	46.59-2.11 (2.17-2.11)	53.22-2.94 (3.12-2.94)	46.96-2.20 (2.27-2.20)
No. unique reflections [a]	27,598 (2,186)	11,003 (1,680)	24,491 (2,058)
Completeness[a] (%)	100.0 (100.0)	99.6 (97.8)	100. (100.0)
I/σI[a]	14.4 (2.6)	7.2 (1.5)	16.0 (0.8)
R_{merge}[a]	0.190 (1.507)	0.387 (2.308)	0.364 (3.438)
CC (1/2) [a]	0.998 (0.718)	0.994 (0.654)	0.995 (0.703)
Redundancy[a]	14.3 (15.0)	18.5 (18.7)	29.5 (29.2)
Refinement			
No. atoms in refinement (P/L/W)[b]	2,409/77/173	2,393/46/11	2,412/51/57
B factor (P/L/W)[b] (Å^2)	32/48/37	62/69/34	63/86/63
R_{fact} (%)	17.4	19.1	18.5
R_{free} (%)	21.3	24.0	20.9
Rmsd [c] (Å)	0.018	0.007	0.016
Rmsd [c] angle (°)	2.15	1.48	1.988
Molprobity Ramachandran			
Favoured (%)	96.97	94.95	96.62
Outliers (%)	0	0.67	0.34
Crystallization conditions	0.1M MES pH6.5, 2.13M Lithium acetate	0.1M MES pH6.5, 2.30M Lithium acetate	0.1M Tris pH 7.5, 1.66M Lithium acetate

[a] Values in brackets show the statistics for the highest resolution shell.

[b] P/L/W indicate protein, ligand molecules, and water molecules, respectively.

[c] Rmsd indicates root-mean-square deviation

Table S1 (cont.). Data collection and refinement statistics.

Complex	**DCAF1-8**	**DCAF1-11**
PDB accession code	**8OGB**	**8OGC**
Beamline	**SLS PXII**	**SLS PXII**
Protein	DCAF1(1039-1401)Q1250L Trypsin digested	DCAF1(1039-1401)Q1250L Trypsin digested
Co-crystallization/soak	Soak	Soak
Data Collection		
Spacegroup	*$P6_122$*	*$P6_122$*
Cell dimensions	*a* = *b* = *85.0* Å, *c* = *219.4* *α* = *γ* = 90.0˚; *β* = 120˚	*a* = b = 81.7 Å, *c* = 234.1 Å *α* = *γ* = 90.0˚; *β* = 120˚
Resolution (Å) [a]	219.36-2.27 (2.34-2.27)	234.1-2.09 (2.15-2.09)
No. unique reflections [a]	22,640 (2,016)	28,474 (2.159)
Completeness[a] (%)	100.0 (100.0)	100.0 (100.0)
I/σI[a]	14.8 (2.6)	14.3 (2.2)
R_{merge}[a]	0.157 (1.347)	0.183 (1.700)
CC (1/2) [a]	0.998 (0.784)	0.999 (0.746)
Redundancy[a]	18.8 (18.0)	19.0 (19.2)
Refinement		
No. atoms in refinement (P/L/W)[b]	2,411/44/233	2,426/87/193
B factor (P/L/W)[b] (Å^2)	42/54/45	36/48/43
R_{fact} (%)	16.8	17.1
R_{free} (%)	21.2	20.2
Rmsd [c] (Å)	0.018	0.018
Rmsd [c] angle (°)	2.05	2.21
Molprobity Ramachandran		
Favoured (%)	97.97	96.28
Outliers (%)	0.68	0
Crystallization conditions	0.1M Tris pH 7.5, 2.13M Lithium acetate	0.1M Tris pH 7.5, 2.18M Lithium acetate

[a] Values in brackets show the statistics for the highest resolution shell.

[b] P/L/W indicate protein, ligand molecules, and water molecules, respectively.

[c] Rmsd indicates root-mean-square deviation

EXPERIMENTAL SECTION

Preparation and characterization of compounds

General conditions. Unless otherwise mentioned, all reagents and solvents were obtained from commercial sources and used without purification. In some cases, intermediates were characterized by LC-MS to confirm that the mass matched the structure and carried on to the next step without further purification. Air sensitive procedures were performed under an atmosphere of nitrogen or argon. Purification of the final compounds to > 95% purity was carried out either using prepacked silica gel cartridge (Analogix. Biotage or ISCO) or reverse phase C18 column. ^{1}H and ^{13}C NMR spectra were recorded on Bruker 400 and 600 Avance spectrometer. NMR chemical shift (δ) are quoted in parts per million (ppm) referenced to the residual solvent peak [DMSO-d6] set at 2.49 ppm or [$CDCl_3$] set at 7.26 ppm. Purity of all compounds was determined to be >95% by analytical HPLC.

LC-MS method A: Column ACQUITY UPLC HSS T3 1.8 μm, 2.1x50 mm, eluent A: water + 0.05% formic acid + 3.75 mM acetic acid, eluent B: acetonitrile + 0.04% formic acid, column temperature 60 °C, flow 1 mL / min, gradient: from 5 to 98% B in 1.4 min.

HR-MS: Acquity UPLC/Xevo G2-S QTof MS (ESI), Column Type:CORTECS™ C18+ 2.7 μm, Gradient: from 5 to 60 % eluent B (Isopropanol + 0.05 % TFA) in 4.0 min; 60 to 98 % B in 0.5 min.

Synthesis/characterization of

-(2-fluorophenyl)-2,3-dihydroimidazo[2,1-a]isoquinoline (1)

Characterized as HCl salt

HR-MS: retention time r.t. (1.38 min), $[M+H]^+$=265.13, purity >98%, MW=264.30, Formula=$C_{17}H_{13}FN_2$

^{1}H NMR (600 MHz, DMSO) δ 11.33 (s, 1H), 8.59 (d, J = 8.2 Hz, 1H), 8.04 – 7.97 (m, 2H), 7.83 (ddd, J = 8.2, 6.6, 1.7 Hz, 1H), 7.73 – 7.63 (m, 2H), 7.50 (t, J = 9.3 Hz, 1H), 7.48 – 7.43 (m, 1H), 7.29 (s, 1H), 4.40 (dd, J = 11.6, 8.8 Hz, 2H), 4.09 (dd, J = 11.4, 9.0 Hz, 2H).

^{13}C NMR (151 MHz, DMSO) δ 158.96 (d, J=247 Hz), 156.38, 136.78, 135.23, 134.27, 132,96 (d, J=8.28 Hz), 131.71 (d, J = 1.6 Hz), 129.12, 127.67, 126.51, 125.17 (d, J = 3.4 Hz), 120.11 (d, J=1.2 Hz), 116.27 (d, J=21.10 Hz), 115.06, 112.91, 49.73 (d, J = 2.7 Hz), 43.40.

3-(3-(4-chlorophenoxy)phenyl)-3,4-dihydro-2H-pyrrol-5-amine (2)

Characterized as HCl salt

HR-MS: r.t (2.16 min), $[M+H]^+$=287.10, purity >98%, MW=286.76, Formula=$C_{16}H_{15}ClN_2O$

^{1}H NMR (600 MHz, DMSO) δ 9.62 (s, 1H), 9.23 (s, 1H), 8.96 (s, 1H), 7.46 – 7.42 (m, 2H), 7.38 (t, J = 7.9 Hz, 1H), 7.17 – 7.13 (m, 1H), 7.08 (t, J = 2.1 Hz, 1H), 7.06 – 7.01 (m, 2H), 6.95 – 6.90 (m, 1H), 3.93 (dd, J = 10.5, 8.2 Hz, 1H), 3.78 (p, J = 8.5 Hz, 1H), 3.51 (dd, J = 10.5, 8.0 Hz, 1H), 3.17 (dd, J = 17.2, 8.7 Hz, 1H), 2.95 (dd, J = 17.1, 9.2 Hz, 1H).

^{13}C NMR (151 MHz, DMSO) δ 169.81, 156.39, 155.51, 142.89, 130.39, 129.87, 127.18 (d, J = 0.7 Hz),122.67, 120.15, 117.83, 117.37, 53.27, 37.31.

5-(2-methyl-1-phenylpropan-2-yl)-2,3-dihydroimidazo[2,1-a]isoquinoline (3)

The compound was prepared according to the scheme for compound **6**, (intermediate B) using methyl 2,2-dimethyl-3-phenylpropanoate in the first step, isolated as HCl salt.

HR-MS: r.t. (1.85 min), $[M+H]^+$=303.20, purity >98%, MW=302.40, Formula=$C_{21}H_{22}N_2$

^{1}H NMR (600 MHz, DMSO) δ 11.17 (s, 1H), 8.54 (d, J = 8.3 Hz, 1H), 7.93 – 7.89 (m, 1H), 7.86 (d, J = 8.0 Hz, 1H), 7.73 (ddd, J = 8.2, 6.9, 1.2 Hz, 1H), 7.17 (dt, J = 10.0, 6.7 Hz, 3H), 7.03 – 7.00 (m, 2H), 6.90 (s, 1H), 5.10 (dd, J = 10.8, 8.8 Hz, 2H), 4.17 (t, J = 9.8 Hz, 2H), 3.09 (s, 2H), 1.44 (s, 6H).

^{13}C NMR (151 MHz, DMSO) δ 157.87, 146.63, 137.07, 136.76, 135.07, 130.14, 128.52, 127.76, 127.50, 126.42, 125.97, 114.17, 110.44, 52.60, 44.92, 43.32, 39.71, 27.63.

5-(1-(4-chlorophenyl)cyclopropyl)-2,3-dihydroimidazo[2,1-a]isoquinoline (4)

The compound was prepared according to the scheme for compound **6**, (intermediate B), isolated as HCl salt.

HR-MS: r.t. (1.95 min), $[M+H]^+$=321.13, purity >98%, MW=320.80, Formula=$C_{20}H_{17}ClN_2$

^{1}H NMR (600 MHz, DMSO) δ 11.11 (s, 1H), 8.52 (d, J = 8.2 Hz, 1H), 8.02 – 7.95 (m, 2H), 7.77 (ddd, J = 8.3, 6.4, 1.9 Hz, 1H), 7.40 (s, 1H), 7.38 – 7.34 (m, 2H), 7.23 – 7.19 (m, 2H), 4.38 (t, J = 10.2 Hz, 2H), 4.02 (t, J = 10.2 Hz, 2H), 1.67 – 1.62 (m, 2H), 1.51 (d, J = 5.1 Hz, 2H).

^{13}C NMR (151 MHz, DMSO) δ 156.73, 141.76, 139.76, 137.28, 135.00, 131.11, 128.69, 128.61, 127.62, 127.46, 126.28, 114.68, 112.32, 48.75, 43.18, 26.02, 16.84.

5-(1-(4-methoxyphenyl)cyclopropyl)-2,3-dihydroimidazo[2,1-a]isoquinoline (5)

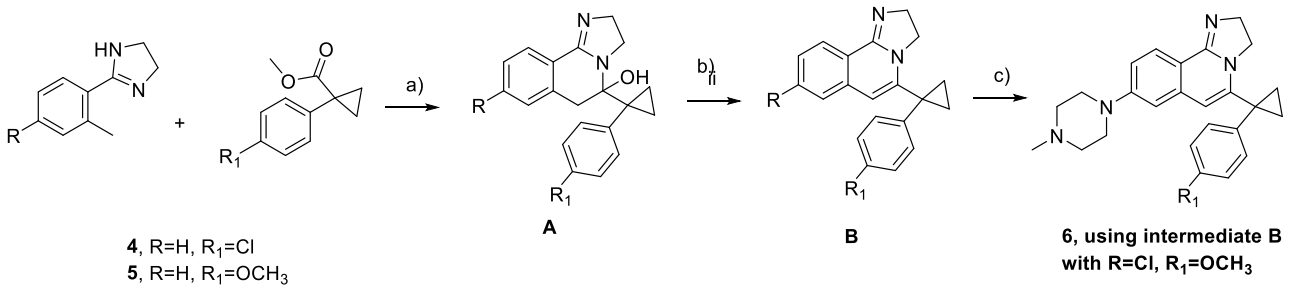
The compound was prepared according to the scheme for compound **6**, (intermediate B).

LC-MS method A: r.t.(0.81 min), $[M+H]^+$=317.2, purity >98 %, MW=316.40, Formula=$C_{21}H_{20}N_2O$

^{1}H NMR (600 MHz, DMSO-d6) δ 7.89 (d, J = 7.9 Hz, 1H), 7.49 (t, J = 7.5 Hz, 1H), 7.40 (d, J = 7.9 Hz,1H), 7.27 (t, J = 7.5 Hz, 1H), 7.14 (d, J = 8.5 Hz, 2H), 6.87 (d, J = 8.6 Hz, 2H), 6.11 (s, 1H), 3.77 (q, J =9.9, 8.5 Hz, 4H), 3.72 (s, 3H), 1.35 (s, 2H), 1.25 (d, J = 17.1 Hz, 2H).

5-(1-(4-methoxyphenyl)cyclopropyl)-8-(4-methylpiperazin-1-yl)-2,3-dihydroimidazo[2,1-a]isoquinoline (6)

Synthesis was performed according to the following procedure:



Reagents and conditions: a) n-BuLi (1.6M in THF), THF ; b) AcOH, 3h, 125°C; c), sodium tert-butoxide, Pd(OAc)2, 2-dicyclohexylphosphino-2', 6'-diisopropoxybiphenyl (RuPhos), N-methyl piperazine, 110°C.

LC-MS method A: r.t.(0.60 min), $[M+H]^+$=415.4, purity 97%, MW=414.55, Formula=$C_{26}H_{30}N_4O$

^{1}H NMR (600 MHz, DMSO) δ 7.78 (d, J = 8.9 Hz, 1H), 7.11 (d, J = 8.7 Hz, 2H), 7.03 (d, J = 8.9 Hz, 1H), 6.86 (d, J = 8.4 Hz, 3H), 6.18 (d, J = 4.2 Hz, 1H), 3.82 (t, J = 10.1 Hz, 2H), 3.78 – 3.73 (m, 2H), 3.71 (s, 3H), 3.28 (t, J = 5.0 Hz, 4H), 2.44 (t, J = 5.0 Hz, 4H), 2.22 (s, 3H), 1.33 (s, 2H), 1.24 (s, 2H).

^{13}C NMR (151 MHz, DMSO) δ 157.63, 156.28, 153.02, 143.76, 137.92, 133.48, 126.94, 126.61, 114.64, 114.01, 111.35, 108.43, 103.62, 55.03, 54.37, 50.62, 47.05, 47.03, 45.71, 26.58, 15.48.

N1-(2-(1-(4-methoxyphenyl)cyclopropyl)quinazolin-4-yl)ethane-1,2-diamine (7)

The compound was prepared according to the scheme for compound **10**.

LC-MS method A: r.t.(50 min), $[M+H]^+$=335.2, purity >98 %, MW=334.42, Formula=$C_{20}H_{22}N_4O$

^{1}H NMR(400MHz, DMSO-d6) δ 8.18 – 8.01 (m, 2H), 7.63 (t, J = 7.6Hz, 1H), 7.44 (d, J = 8.2Hz, 1H), 7.33 (dd, J = 28.0, 8.3Hz, 3H), 6.86 (d, J = 8.7Hz, 2H), 3.76 (s, 3H), 2.69 (t, J = 6.4Hz, 2H), 1.63 (q, J = 3.4Hz, 2H), 1.16 (q, J = 3.5Hz, 2H).

N1-(2-(1-(4-methoxyphenyl)cyclopropyl)-7-(4-methylpiperazin-1-yl)quinazolin-4-yl)ethane-1,2-diamine (8)

The compound was prepared according to the scheme for compound **10**.

LC-MS method A: r.t.(0.40 min), $[M+H]^+$=433.3, purity >98 %, MW=432.57, Formula=$C_{25}H_{32}N_6O$

^{1}H NMR(400MHz, DMSO-d6) δ 7.97 – 7.86 (m, 1H), 7.80 – 7.67 (m, 1H), 7.26 (d, J = 8.6Hz, 2H), 7.11 (d, J = 8.9Hz, 1H), 6.84 (d, J = 8.6Hz, 2H), 6.62 (s, 1H), 3.74 (s, 3H), 3.24 (s, 4H), 2.70 – 2.61 (m, 1H), 2.45 – 2.38 (m, 4H), 2.20 (s, 3H), 1.67 (s, 2H), 1.62 – 1.54 (m, 2H), 1.23 (t, J = 6.7Hz, 2H), 1.08 (s, 2H), 0.90 – 0.78 (m, 1H).

N1-(2-(1-(4-methoxyphenyl)cyclopropyl)-7-(4-(prop-1-en-2-yl)piperazin-1-yl)quinazolin-4-yl)ethane-1,2-diamine (9)

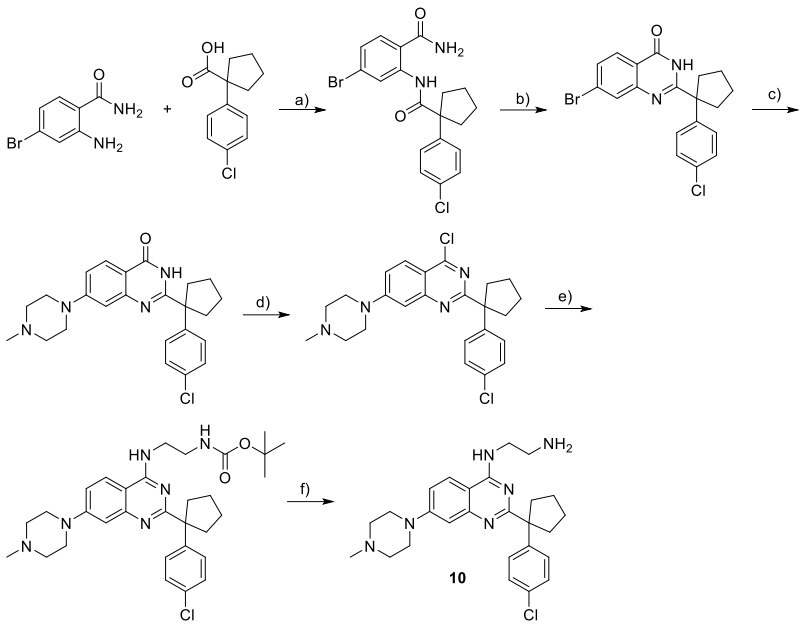
The compound was prepared according to the scheme for compound **10**.

LC-MS method A: r.t.(0.55 min), $[M+H]^+$=461.3, purity >98 %, MW=460.58, Formula=$C_{26}H_{32}N_6O_2$

^{1}H NMR(400MHz, DMSO-d6) δ 8.09 – 7.62 (m, 2H), 7.27 (d, J = 8.5Hz, 2H), 7.18 – 7.05 (m, 1H), 6.84 (d, J = 6.9Hz, 2H), 6.64 (d, J = 8.8Hz, 1H), 3.74 (s, 3H), 3.55 (s, 5H), 3.14 (d, J = 18.8Hz, 2H), 2.67 (s, 1H), 2.03 (s, 3H), 1.59 (s, 4H), 1.23 (s, 2H), 1.08 (d, J = 11.4Hz, 2H).

N1-(2-(1-(4-chlorophenyl)cyclopentyl)-7-(4-methylpiperazin-1-yl)quinazolin-4-yl)ethane-1,2-diamine (10)

Synthesis was performed according to the following procedure:



Reagents and conditions: a) 1-Chloro-N,N,2-trimethyl-1-propenylamine, THF, Et_3N, 0 °C to rt, 7h, continued crude; b) NaOH 10% THF 1:1 80 °C; 1h, 60 % after crystallization from reaction mixture; c) N-methylpiperazin, $Pd(OAc)_2$, RuPhos G3, NaOtBu, tert.-BuOH, 100 °C, 0.5h, 65 %; d) $PhNEt_2$, $POCl_3$, toluene, 110 °C, 1.5h, continued crude; e) N-Boc-ethylenediamine, ACN, 60 °C 45h, continued crude; f) TFA, DCM, 46 % over 3 steps.

LC-MS method A: r.t.(0.5 min), $[M+H]^+$=465.5, 25 %, Cl-isotope 467.5, purity >98 %, MW=465.05, Formula=$C_{26}H_{33}N_6Cl$

HR-MS: $[M+H]^+$=465.25241 (-0.83488 ppm)

^{1}H NMR (400 MHz, DMSO-d6) δ 9.88 (t, J = 5.5 Hz, 1H), 8.18 (d, J = 9.5 Hz, 1H), 7.99 (s, 3H), 7.54 – 7.49 (m, 1H), 7.47 (d, J = 8.6 Hz, 2H), 7.39 (d, J = 8.6 Hz, 2H), 7.17 (d, J = 2.1 Hz, 1H), 4.24 – 4.03 (m, 2H), 3.96 (q, J = 5.9 Hz, 2H), 3.66 – 3.46 (m, 2H), 3.32 – 3.02 (m, 6H), 2.85 (s, 3H), 2.82 – 2.73 (m, 2H), 2.29 – 2.19 (m, 2H), 1.80 – 1.60 (m, 4H).

^{13}C NMR (101 MHz, DMSO) δ 166.55, 159.83, 154.26, 142.18, 141.15, 132.29, 129.00, 128.69, 125.82, 117.13, 103.33, 99.71, 58.24, 51.99, 44.01, 42.28, 39.41, 38.19, 36.64, 22.81.

N1-(2-(1-(4-chlorophenyl)cyclopentyl)-7-(4-(prop-1-en-2-yl)piperazin-1-yl)quinazolin-4-yl)ethane-1,2-diamine (11)

The compound was prepared according to the scheme for compound **10**.

LC-MS method A: r.t.(0.62 min), $[M+H]^+$=493.4, 25 % Cl-isotope 493.4, purity >98 %, MW=493.06, Formula=$C_{27}H_{33}ClN_6O$

^{1}H NMR (400 MHz, DMSO-d6) δ 12.87 (s, 1H), 9.73 (t, J = 5.4 Hz, 1H), 8.12 (d, J = 9.5 Hz, 1H), 7.94 (s, 3H), 7.45 (d, J = 8.6 Hz, 3H), 7.40 (d, J = 8.7 Hz, 2H), 7.01 (d, J = 2.2 Hz, 1H), 3.94 (q, J

= 5.9 Hz, 2H), 3.59 (s, 4H), 3.52 (s, 2H), 3.49 – 3.41 (m, 2H), 3.16 (q, J = 5.5 Hz, 2H), 2.81 – 2.72 (m, 2H), 2.28 – 2.19 (m, 2H), 2.04 (s, 3H), 1.79 – 1.60 (m, 4H).

N1-(2-(1-(4-chlorophenyl)cyclohexyl)-7-(4-methylpiperazin-1-yl)quinazolin-4-yl)ethane-1,2-diamine (12)

The compound was prepared according to the scheme for compound **10**.

LC-MS method A:, r.t(0.57 min), $[M+H]^+$=479.3, 25 % Cl-isotope 481.3, purity >98 %, MW=479.07, Formula=$C_{27}H_{35}ClN_6$

^{1}H NMR (400 MHz, DMSO) δ 12.87 (s, 1H), 10.16 (s, 1H), 9.89 (t, *J* = 5.6 Hz, 1H), 8.19 (d, *J* = 9.4 Hz, 1H), 8.08 – 7.90 (m, 3H), 7.53 (dd, *J* = 9.5, 2.4 Hz, 1H), 7.48 (d, *J* = 8.8 Hz, 2H), 7.40 (d, *J* = 8.7 Hz, 2H), 7.26 (d, *J* = 2.4 Hz, 1H), 4.13 (d, *J* = 12.2 Hz, 2H), 3.95 (q, *J* = 6.1 Hz, 2H), 3.58 (d, *J* = 11.8 Hz, 2H), 3.26 (s, 2H), 3.16 (q, *J* = 6.3 Hz, 4H), 2.86 (s, 3H), 2.76 (d, *J* = 13.0 Hz, 2H), 2.12 (t, *J* = 11.5 Hz, 2H), 1.70 – 1.61 (m, 2H), 1.49 – 1.44 (m, 2H), 1.35-1.15 (m,2H)

1-(4-(4-((2-aminoethyl)amino)-2-(1-(4-chlorophenyl)cyclohexyl)quinazolin-7-yl)piperazin-1-yl)ethan-1-one (13)

The compound was prepared according to the scheme for compound **10**.

LC-MS method A: r.t.(0.63 min), $[M+H]^+$=507.4, 30 % Cl-isotope 509.4, purity >98 %, MW=507.08, Formula=$C_{28}H_{35}ClN_6O$

^{1}H NMR(400MHz, DMSO-d6) δ 7.95 (dd, J = 21.6, 9.2Hz, 1H), 7.82 (d, J = 5.3Hz, 1H), 7.45 (dd, J = 19.3, 8.7Hz, 2H), 7.26 (d, J = 8.6Hz, 2H), 7.23 – 7.17 (m, 1H), 6.85 (d, J = 2.2Hz, 1H), 3.57 (d, J = 5.2 Hz, 4H), 3.45 (dt, J = 11.8, 7.0Hz, 2H), 2.82 (d, J = 12.4Hz, 2H), 2.77 – 2.66 (m, 2H), 2.05 (s, 3H), 1.95 – 1.82 (m, 2H), 1.58 – 1.41 (m, 5H), 1.38 – 1.13 (m, 5H).

Protein production

NMR: His6-TEV-DCAF1(1039-1401)Q1250L was subcloned into pIEx/Bac-3 derived vectors for expression in Sf9 cells using the flashBAC™ system. Recombinant baculoviruses were generated and amplified followed by large scale expression in Sf9 cells. Transfected cells were typically grown at 27°C for 48 h in protein, serum and animal-origin free Sf900-III medium. For [^{13}C]-methyl methionine labeled protein production, Met-free Sf4 medium was supplemented with 1 g/l [^{13}C]-methyl methionine. Cells were harvested by centrifugation. Cell pellets were resuspended in lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl, 0.1% Triton X-100, 4 mM TCEP) supplemented with o'complete (EDTA free) protease inhibitor cocktail (Roche) and Benzonase

(Novagen) and lysed by sonication. After removal of cell debris by centrifugation at 4°C at 50,000xg for at least 30 minutes the cleared supernatants were purified using Ni-NTA columns (HisTrap FF crude) on AKTA™ systems (Cytiva). The cleavable affinity tags were cleaved using TEV protease. Proteins were further purified by Ion exchange chromatography (ResourceQ) followed by a final Size Exclusion Chromatography step on a Superdex S75 column (Cytiva) in buffer containing 20 mM HEPES pH 7.5, 200 mM NaCl, 2 mM TCEP. Proteins were concentrated, aliquoted and flash-frozen in liquid nitrogen.

SPR: Avi-DCAF1(1073-1399)E1398S-His was subcloned into pIEx/Bac-3 derived vectors for expression in Sf9 cells using the flashBAC™ system. Recombinant baculoviruses were generated and amplified followed by large scale expression in Sf9. Co-transfected cells with BirA for biotinylation were grown at 27°C for up to 62 h before harvesting by centrifugation. Cell pellets were resuspended in lysis buffer (50 mM Tris pH 8, 300 mM NaCl, 10% Glycerol (w/v), 5 mM TCEP, 5 mM Imidazole) supplemented with complete (EDTA free) protease inhibitor cocktail (Roche) and Benzonase (Novagen) and lysed by sonication. After removal of cell debris by centrifugation at 4°C at 50,000xg for at least 30 minutes the cleared supernatants were purified using Ni-NTA columns (HisTrap FF crude) on AKTA™ systems (Cytiva). Proteins were further purified by ion exchange chromatography (Q HP column) followed by a final Size Exclusion Chromatography step on Superdex S75 columns (Cytiva). The protein in final buffer (50 mM Tris pH 8, 300 mM NaCl, 10% Glycerol (w/v), 2 mM TCEP) was concentrated, analysed by RP-HPLC and RP-LC/MS, and found fully biotinylated with 40367 Da molecular as expected. Protein aliquots were flash-frozen in liquid nitrogen.

His-Px-Avi-EED(76-441) was co-expressed with BirA in *E. coli* BL21-DE3. Transformed and selected cells were grown to an OD=1.5 at 37°C in Terrific broth (VWR) supplemented with 0.05 mg/mL Kanamycin before reducing the temperature to 18°C and induction with 0.005 mM IPTG. Cells were grown overnight and harvesting by centrifugation. Cell pellets were resuspended in lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, 1mM TCEP) supplemented with o'complete (EDTA free) protease inhibitor cocktail (Roche) and Benzonase (Novagen) and lysed by sonication. After removal of cell debris by centrifugation at 4°C at 50,000xg for at least 30 minutes the cleared supernatants were purified using Ni-NTA columns (HisTrap FF crude) on AKTA™ systems (Cytiva). Proteins were further purified by a final Size Exclusion

Chromatography step on Superdex S200 (Cytiva) in buffer containing 30 mM HEPES pH 7.5, 300 mM NaCl, 1 mM TCEP, 5% glycerol. Proteins were concentrated, aliquoted and flash-frozen in liquid nitrogen. The biotinylation of the Avi-tag was confirmed by intact MS analysis.

X-ray crystallography: His6-TEV-DCAF1(1039-1401)Q1250L and His-TEV-DCAF1(1079-1393) Q1250L were all subcloned into pIEx/Bac-3 derived vectors for expression in Sf9 cells using the flashBAC™ system. Recombinant baculoviruses were generated and amplified followed by large scale expression in Sf9 cells. Transfected cells were typically grown at 27°C for up to 62h before harvesting by centrifugation. Cell pellets were resuspended in lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl, 0.1% Triton X-100, 4 mM TCEP) supplemented with o'complete (EDTA free) protease inhibitor cocktail (Roche) and Benzonase (Novagen) and lysed by sonication. After removal of cell debris by centrifugation at 4°C at 50,000xg for at least 30 minutes the cleared supernatants were purified using Ni-NTA columns (HisTrap FF crude) on AKTA™ systems (Cytiva). The cleavable affinity tags were cleaved using TEV protease. Proteins were further purified by Ion exchange chromatography (ResourceQ) followed by a final Size Exclusion Chromatography step on Superdex S200 or Superdex S75 columns (Cytiva) in buffer containing 20 mM HEPES pH 7.5, 200 mM NaCl, 2 mM TCEP. Proteins were concentrated, aliquoted and flash-frozen in liquid nitrogen.

Nuclear magnetic resonance and NMR screening

Solutions of unlabeled or selectively $^{13}C^{\varepsilon}$-methionine labeled DCAF1(1039-1401)Q1250L were prepared in NMR buffer consisting of 25 mM d_{11}-Tris, 100 mM NaCl, 1 mM d_{16}-TCEP, 10% D_2O, pH = 8.0. All NMR spectra were measured in 3 mm NMR tubes with a sample volume of 170 µl at 296 K. Protein observation 1D proton spectra with excitation sculpting water suppression and 2D [^{13}C,^{1}H]-HMQC spectra were recorded on Bruker Avance III HD 800 MHz or Bruker Avance III HD 600 MHz NMR spectrometer equipped with 5 mm triple ^{1}H,^{13}C,^{15}N-resonance cryogenic probes with shielded xyz-gradient coils. Ligand observation one-dimensional proton $T_{1\rho}$ relaxation experiments[4] were measured on a Bruker Avance III 600 MHz NMR spectrometer equipped with a 5 mm ^{1}H/^{19}F,^{13}C,^{15}N quadruple resonance cryogenic probe with shielded xyz-gradient coils. The data were processed and analyzed with the software Topspin 3.6 (Bruker, Switzerland).

NMR screening of the 21 *in silico* WDR focused set compounds to identify donut binders was performed by 1D proton spectroscopy (protein observation and $T_{1\rho}$ experiments) and 2D [^{13}C,^{1}H]-HMQC spectra at concentrations of 30 μM and 600 μM of DCAF1 and compound, respectively. Selected hits were titrated into 30 μM $^{13}C^{\varepsilon}$-Met labeled DCAF1 samples at different concentrations to determine the binding constants as described in reference.[5]

Surface Plasmon Resonance (SPR) assessment of ligand affinity

Ligand affinities (K_D) were determined by SPR using a BiacoreTM T200 device (GE Healthcare). Biotinylated avi-DCAF1(1073-1399)E1398S-His was immobilized on a Streptavidin-coated sensorchip (GE Healthcare, BR-1005-31) to a density of ~3000 RU. The running buffer contained 50 mM Tris-HCl pH 8, 100 mM NaCl, 1 mM TCEP, 0.05% Tween-20 and 1 or 2% DMSO.

Ligand affinities (K_D) for EED were determined by SPR using a BiacoreTM T200 device (GE Healthcare). Biotinylated His-Px-Avi-EED(76-441) was immobilized on a Streptavidin-coated sensor chip (GE Healthcare, BR-1005-31) to a density of ~5000 RU. The running buffer contained 20 mM HEPES pH 7.5, 200 mM NaCl, 1 mM TCEP, 0.05% Tween-20, 5% Glycerol and 1% DMSO.

Experiments were carried out at 22 °C (15°C for EED) using a flow-rate of 60 mL/min (30 mL/min for EED). A fixed concentration of in-house identified binder was injected between test compounds to assess the stability of the signal throughout the run. Compounds were tested in standard successive injection mode at least 8 different concentrations. Curve fitting was performed using the Biacore T200 evaluation software. The sensorgrams were fitted by applying a 1:1 binding model to calculate kinetic rate constants and equilibrium dissociation constants.

Crystal structure determination

The DCAF1 apo crystals were grown by mixing cleaved His-TEV-DCAF1(1079-1393)Q1250L at 9 mg/mL with 1.66 M $LiSO_4$ and 0.1 M Tris pH 7.5 in a 1:1 ratio (200 nL + 200 nL) in sitting drop crystallization plates at 20°C. Crystals were soaked with 1 mM (6) for 15 minutes and cryo-protected by the addition of 20% ethylene glycol and frozen in liquid nitrogen.

The apo crystals for the remain soaks were obtained by incubation of cleaved His6-TEV-DCAF1(1039-1401)Q1250L at 1.6mg/mL with a 1:200 dilution of Trypsin Gold (Promega) for 2h on ice. The cleaved protein was further purified on HiLoad 16/60 Superdex 75 equilibrated with

50 mM HEPES, 200 mM NaCl, 0.5 mM TCEP. Fractions containing a fragment of ~27kDa were pooled and concentrated to 13 mg/mL. Protein was then mixed with 2 mM compound and incubated briefly on ice before mixing with reservoir solution 1:1 (100 nL + 100 nL) in sitting drop crystallization plates at 20°C. Crystals were obtained within 1-3 days with reservoir solutions containing crystallization buffer as indicated in Table X. Crystals were soaked with the different compounds for 10-30 minutes and cryo-protected by the addition of 20% ethylene glycol and frozen in liquid nitrogen.

Datasets were collected at the PXII beamline at the Swiss Lightsource (SLS). Data was integrated by XDS6 and successively merged and scaled by AIMLESS7 in the CCP4I2 interface8. Molecular replacement was performed using Phaser[9] with the starting coordinates 4PXW. Refinement was performed in iterative cycles of modelbuilding in Coot[10] and refinement in Refmac5[11] (**Table 1**). Geometrical correctness of the model was validated by Molprobity[12]. All coordinates were deposited in the PDB with accession codes: 8OG5 and 8OG6 (compound **1**), 8OG7 (compound **2**), 8OG8 (compound **3**), 8OG9 (compound **4**), 8OGA (compound **6**), 8OGB (compound **8**), 8OGC (compound **11**).

References

1. Schwefel, D.; Groom, H. C. T.; Boucherit, V. C.; Christodoulou, E.; Walker, P. A.; Stoye, J. P.; Bishop, K. N.; Taylor, I. A. Structural basis of lentiviral subversion of a cellular protein degradation pathway. *Nature* **2014,** 505, 234-238.

2. Halgren, T. A. Identifying and Characterizing Binding Sites and Assessing Druggability. *Journal of Chemical Information and Modeling* **2009,** 49, 377-389.

3. Schmidtke, P.; Barril, X. Understanding and Predicting Druggability. A High-Throughput Method for Detection of Drug Binding Sites. *Journal of Medicinal Chemistry* **2010,** 53, 5858-5867.

4. Hajduk, P. J.; Olejniczak, E. T.; Fesik, S. W. One-Dimensional Relaxation- and Diffusion-Edited NMR Methods for Screening Compounds That Bind to Macromolecules. *Journal of the American Chemical Society* **1997,** 119, 12257-12261.

5. Fielding, L. NMR methods for the determination of protein–ligand dissociation constants. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2007,** 51, 219-242.

6. Kabsch, W. XDS. *Acta Crystallogr D Biol Crystallogr* **2010,** 66, 125-32.

7. Evans, P. R.; Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **2013,** 69, 1204-14.

8. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **2011,** 67, 235-42.

9. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J Appl Crystallogr* **2007,** 40, 658-674.

10. Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **2004,** 60, 2126-32.

11. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **1997,** 53, 240-55.

12. Williams, C. J.; Headd, J. J.; Moriarty, N. W.; Prisant, M. G.; Videau, L. L.; Deis, L. N.; Verma, V.; Keedy, D. A.; Hintze, B. J.; Chen, V. B.; Jain, S.; Lewis, S. M.; Arendall, W. B., 3rd; Snoeyink, J.; Adams, P. D.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* **2018,** 27, 293-315.

13. Li, A. S. M.; Kimani, S.; Wilson, B.; Noureldin, M.; González-Álvarez, H.; Mamai, A.; Hoffer, L.; Guilinger, J. P.; Zhang, Y.; von Rechenberg, M.; Disch, J. S.; Mulhern, C. J.; Slakman, B. L.; Cuozzo, J. W.; Dong, A.; Poda, G.; Mohammed, M.; Saraon, P.; Mittal, M.; Modh, P.; Rathod, V.; Patel, B.; Ackloo, S.; Santhakumar, V.; Szewczyk, M. M.; Barsyte-Lovejoy, D.; Arrowsmith, C. H.; Marcellus, R.; Guié, M.-A.; Keefe, A. D.; Brown, P. J.; Halabelian, L.; Al-awar, R.; Vedadi, M. Discovery of Nanomolar DCAF1 Small Molecule Ligands. *Journal of Medicinal Chemistry* **2023,** *66*, 5041-5060.