SUPPLEMENTARY INFORMATION

A selective USP1-UAF1 inhibitor links deubiquitination to DNA damage responses

Qin Liang[1], Thomas S Dexheimer[2], Ping Zhang[1], Andrew S Rosenthal[2], Mark A Villamil[1], Changjun You[3], Qiuting Zhang[4], Junjun Chen[1], Christine A Ott[1], Hongmao Sun[2], Diane K Luci[2], Bifeng Yuan[3], Anton Simeonov[2], Ajit Jadhav[2], Hui Xiao[5], Yinsheng Wang[3], David J Maloney[2*] & Zhihao Zhuang[1*]

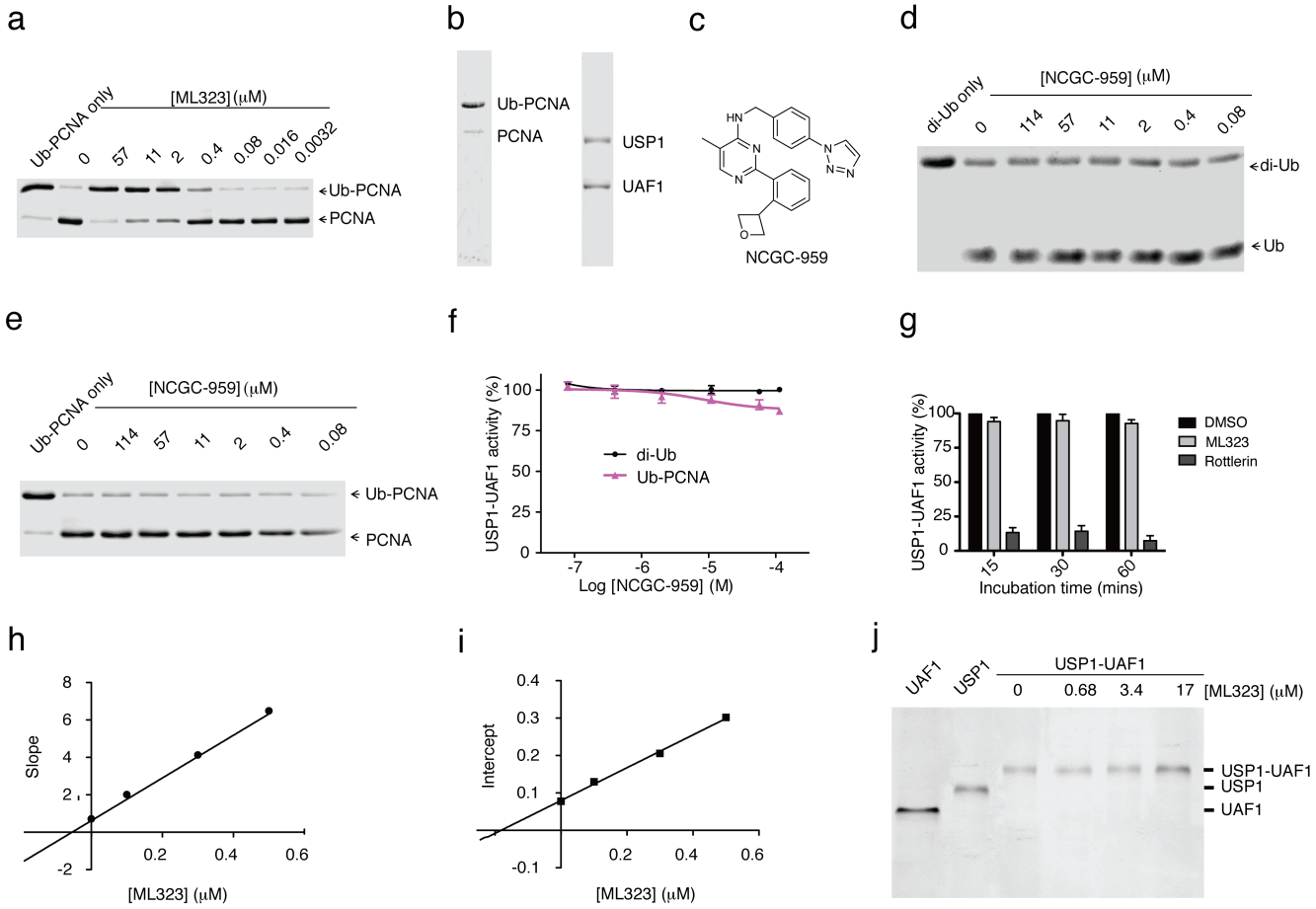
[1]Department of Chemistry and Biochemistry, University of Delaware, Newark, Delaware, USA. [2]National Center for Advancing Translational Sciences, National Institutes of Health, Bethesda, Maryland, USA. [3]Department of Chemistry, University of California-Riverside, Riverside, California, USA. [4]State Key Laboratory of Food Science and Technology, Nanchang University, Nanchang, Jiangxi, China. [5]Laboratory of Macromolecular Analysis and Proteomics, Albert Einstein College of Medicine, Bronx, New York, USA.

[*] To whom correspondence should be addressed

Email: zzhuang@udel.edu

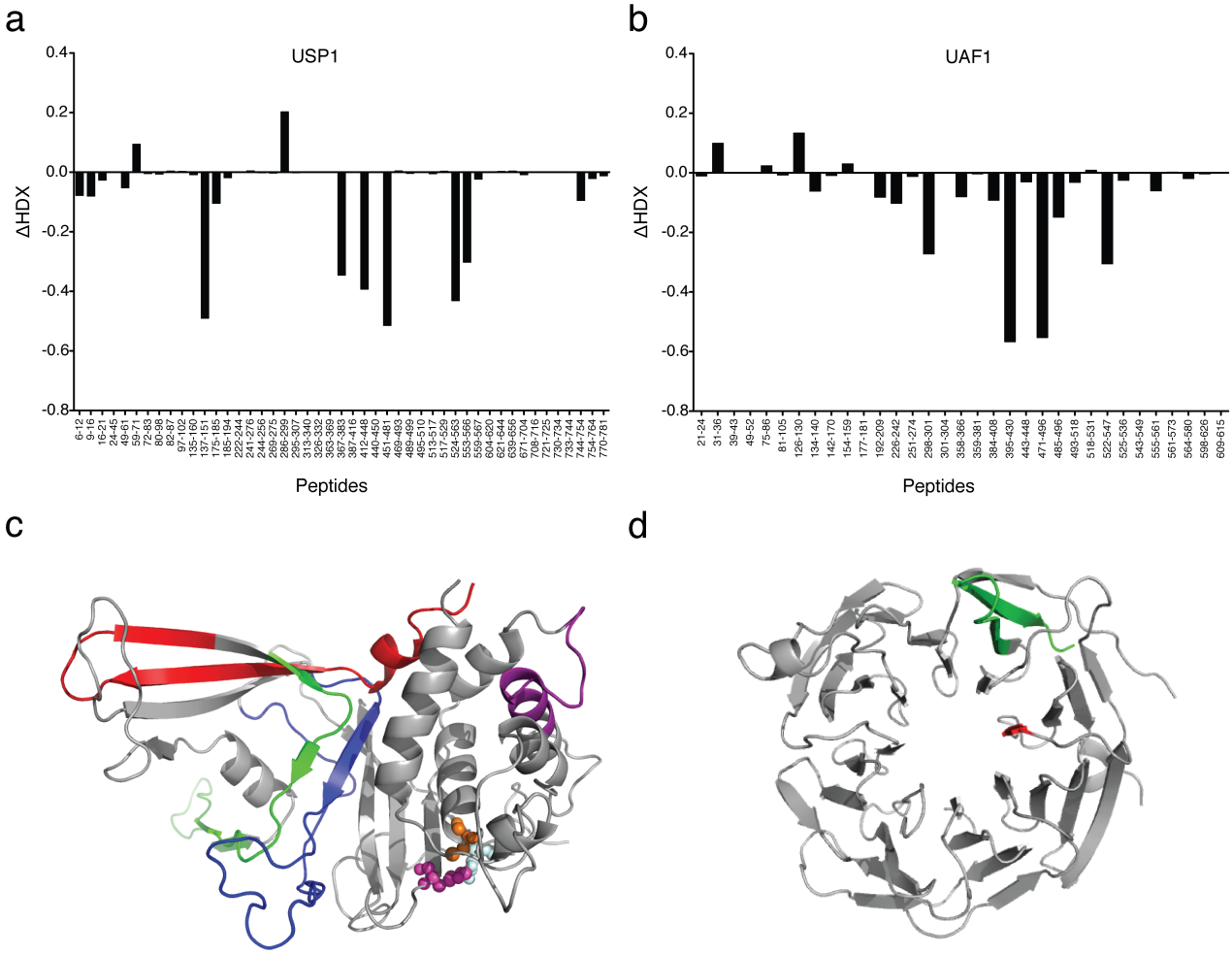
maloneyd@mail.nih.gov

SUPPLEMENTARY RESULTS

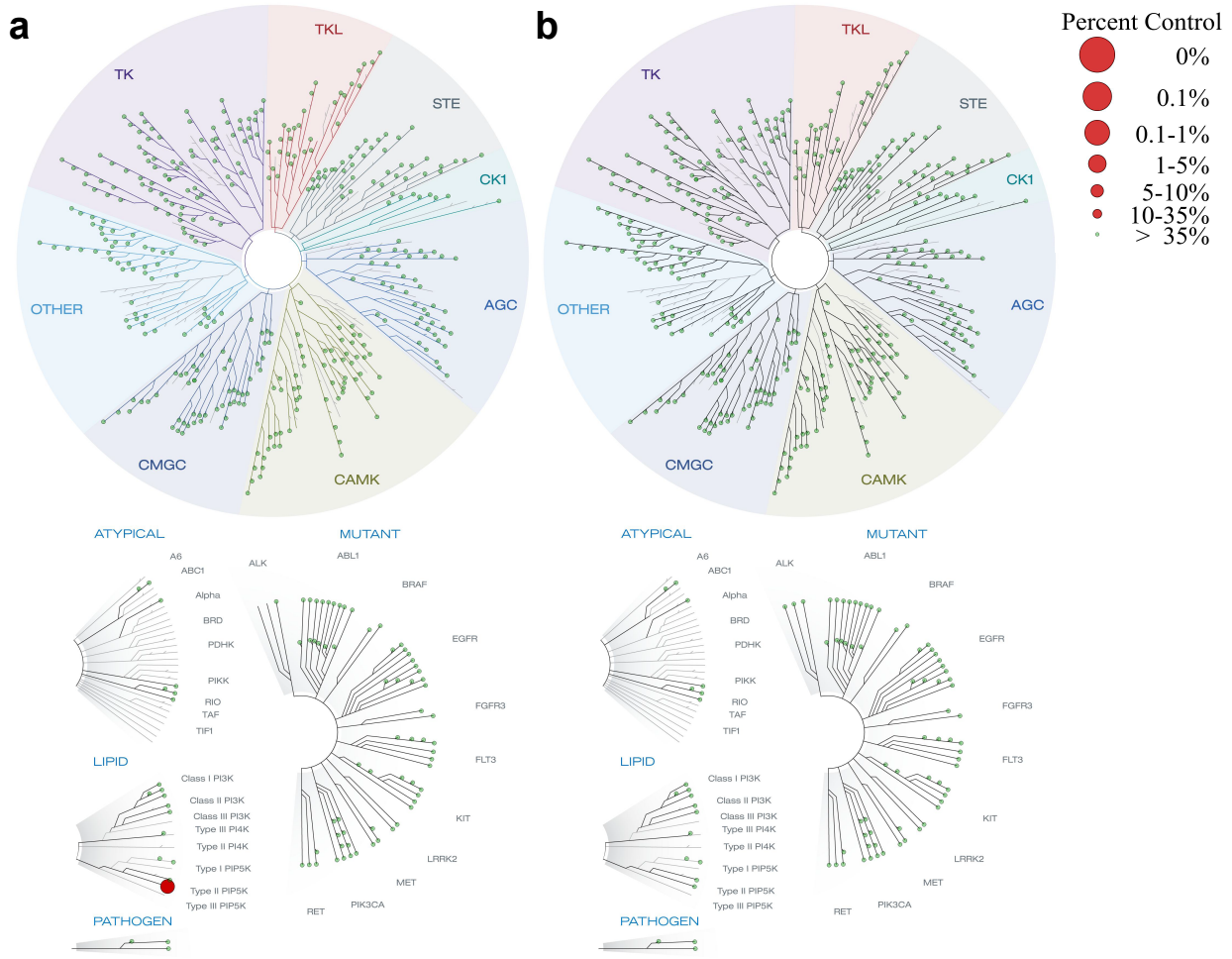


Supplementary Figure 1. **Biochemical characterizations of ML323 as a USP1-UAF1 inhibitor**. (**a**) SDS-PAGE analysis of the cleavage of Ub-PCNA by USP1-UAF1 in the presence of different concentrations of ML323. (**b**) Purified native Ub-PCNA and USP1-UAF1 complex. (**c**) Structure of the negative control compound NCGC-959. (**d**) SDS-PAGE analysis of the cleavage of K63-linked diubiquitin (di-Ub) by USP1-UAF1 in the presence of different concentrations of NCGC-959. (**e**) SDS-PAGE analysis of the cleavage of Ub-PCNA by USP1-UAF1 in the presence of different concentrations of NCGC-959. (**f**) Quantification of the USP1-UAF1-catalyzed hydrolysis of K63-linked di-Ub and Ub-PCNA in the presence of NCGC-959. NCGC-959 had little to no inhibition against USP1-UAF1. (**g**) The reversibility test of ML323 in Ub-AMC assay. USP1-UAF1 was pre-incubated with DMSO, ML323, rottlerin for 15, 30 and 60 minutes respectively and then rapidly diluted into assay buffer containing Ub-AMC. Rottlerin, a known irreversible inhibitor of USP1-UAF1, served as a control. (**h**) The plot of slopes in Lineweaver-Burk analysis against ML323 concentrations to determine the inhibition constants K_i. (**i**) The plot of intercepts in Lineweaver-Burk analysis against ML323 concentrations to determine the inhibition constants K'_i. (**j**) Native gel analysis of USP1-UAF1 treated with increasing concentrations of ML323. No dissociation of the USP1-UAF1 complex was

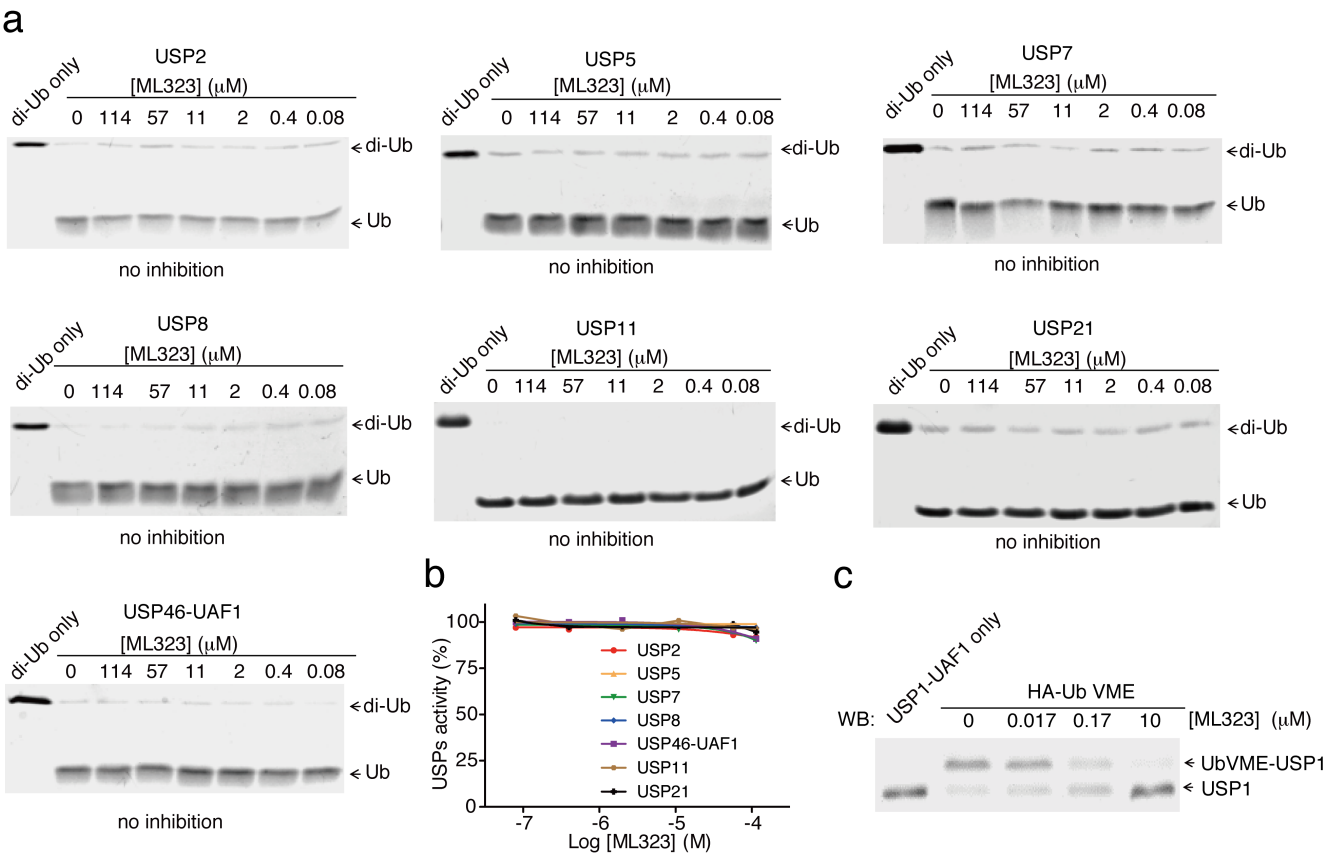
observed when incubated with ML323 up to 100 fold of its IC_{50} (174 nM in the di-Ub assay). USP1 and UAF1 alone were loaded as controls. For (**f**) and (**g**), data are presented as mean ± s.d.; for (**f**), N = 3 experiments/group; for (**g**), N = 3-5 experiments/group. The full SDS-PAGE gels for this figure are shown in Supplementary Fig. 14.



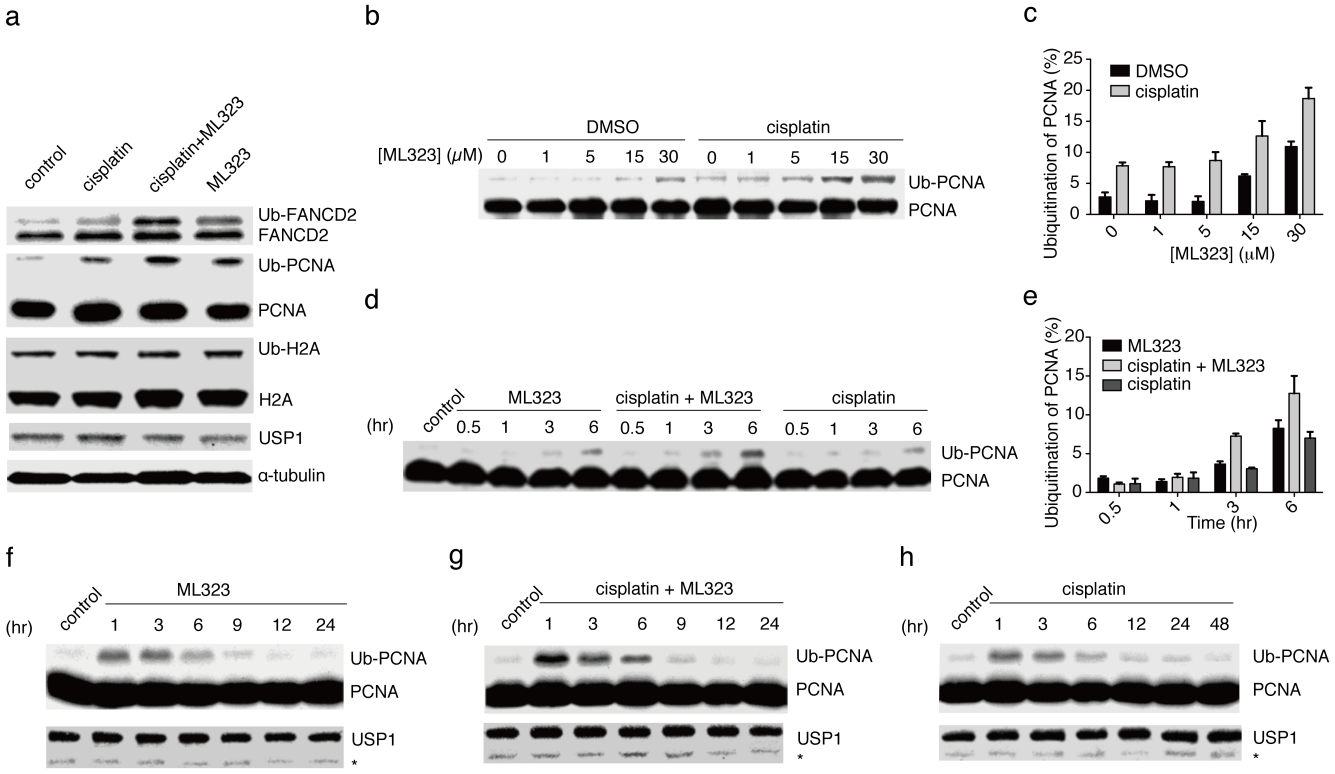
Supplementary Figure 2. ML323-induced alterations in deuteration of USP1-UAF1 complex. **(a)** The differential deuterium incorporation in mass difference (ΔHDX) of the USP1 subunit. ΔHDX of each peptide was obtained by comparing the centroid mass of the USP1 peptides in USP1-UAF1 complexed with 20 μM ML323 to that of USP1-UAF1 in the absence of ML323. Peptides with larger than 0.5 ΔHDX are considered to be strongly protected upon the binding of ML323. (**b**) The differential deuterium incorporation in mass difference (ΔHDX) of the UAF1 subunit was obtained in the same way as described for USP1. (**c**) Mapping the local HDX alterations on the modeled USP1 catalytic core structure described previously[1]. The catalytic residues consisting of Cys90 (orange), His593 (magenta) and Asp751 (light blue) are shown. The peptides that are perturbed by the binding of ML323 are color coded as follows: a. a. 137-151 (purple), 412-448 (red), 451-481 (green), 524-563 (553-566) (blue). The peptide 367-383 located in an inserted domain that was previously shown to interact with UAF1 is not included in the model[2]. (**d**) Mapping the local HDX alterations on the modeled UAF1 structure using the WD40 domain of Sif2 as a template. The peptides perturbed by the binding of ML323 are color coded as follows: a. a. 395-430 (green) and 298-301 (red). Peptides a. a. 471-496 and 522-547 are not included in the model.



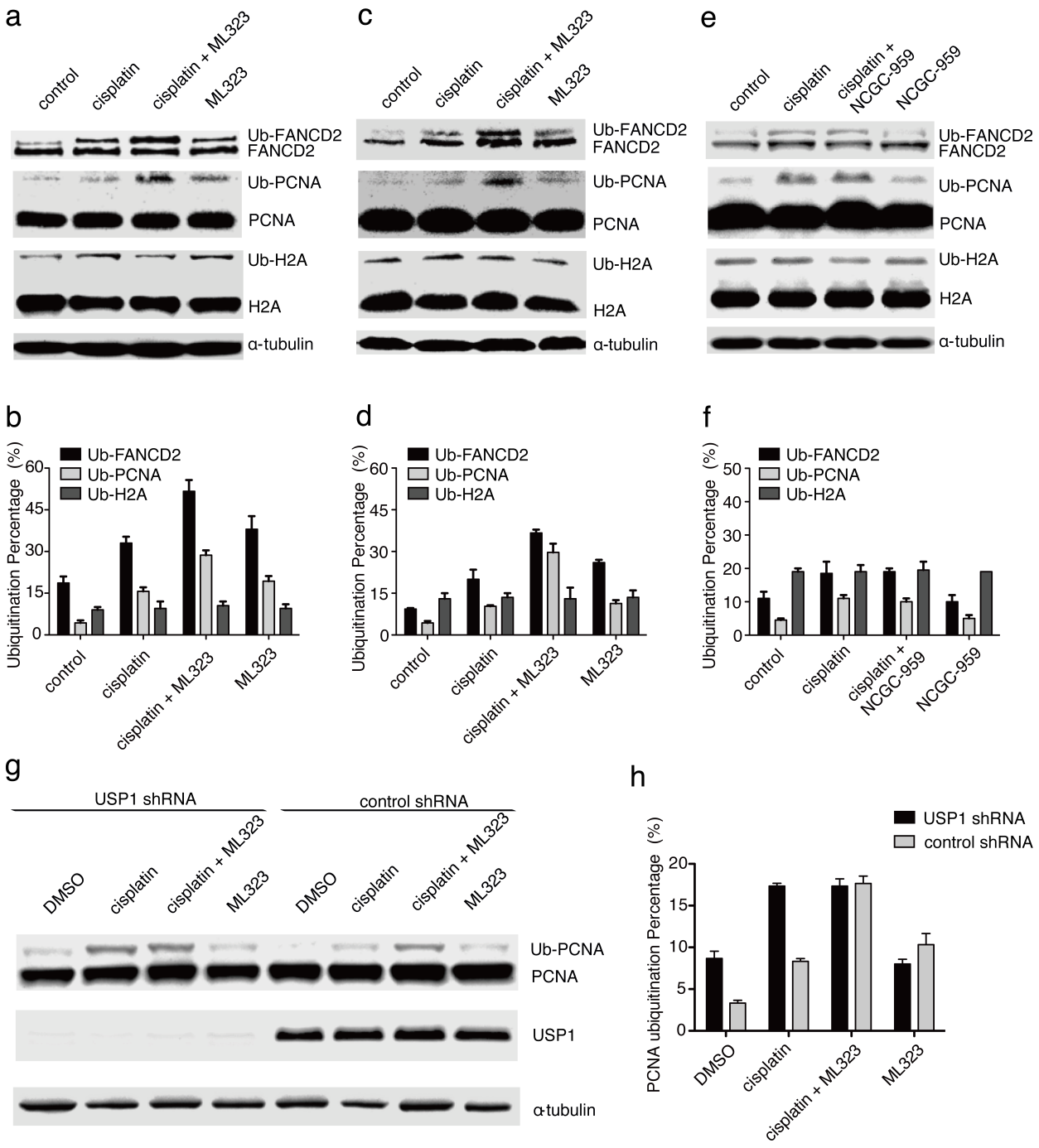
Supplementary Figure 3. Dendrograms of hit compound "1" (a) and ML323 (b) in KINOMEscan profiling by DiscoveRx. The compounds were screened at 10 μM against 451 kinases. The size of the red circle is proportional to the degree of inhibition relative to control as depicted in the figure legend. See www.discoverx.com for details on profiling assays.



Supplementary Figure 4. **Selectivity studies of ML323 against other USPs in the K63-linked diubiquitin assay and inhibition of the labeling of USP1-UAF1 by HA-Ub-VME in an *in vitro* assay**. (**a**) SDS-PAGE analysis of the cleavage of K63-linked diubiquitin (di-Ub) by USP2, USP5, USP7, USP8, USP11, USP21 and USP46-UAF1 at the different concentrations of ML323. (**b**) Dose dependence of the DUB activity of the tested USPs in the presence of increasing concentrations of ML323. (**c**) Western blotting analysis of USP1-UAF1 incubated with increasing concentrations of ML323 (0, 0.017, 0.17 and 10 μM) and 5 μM HA-Ub-VME using anti-USP1 antibody. The full SDS-PAGE and immunoblotting gels for this figure are shown in Supplementary Fig. 15.

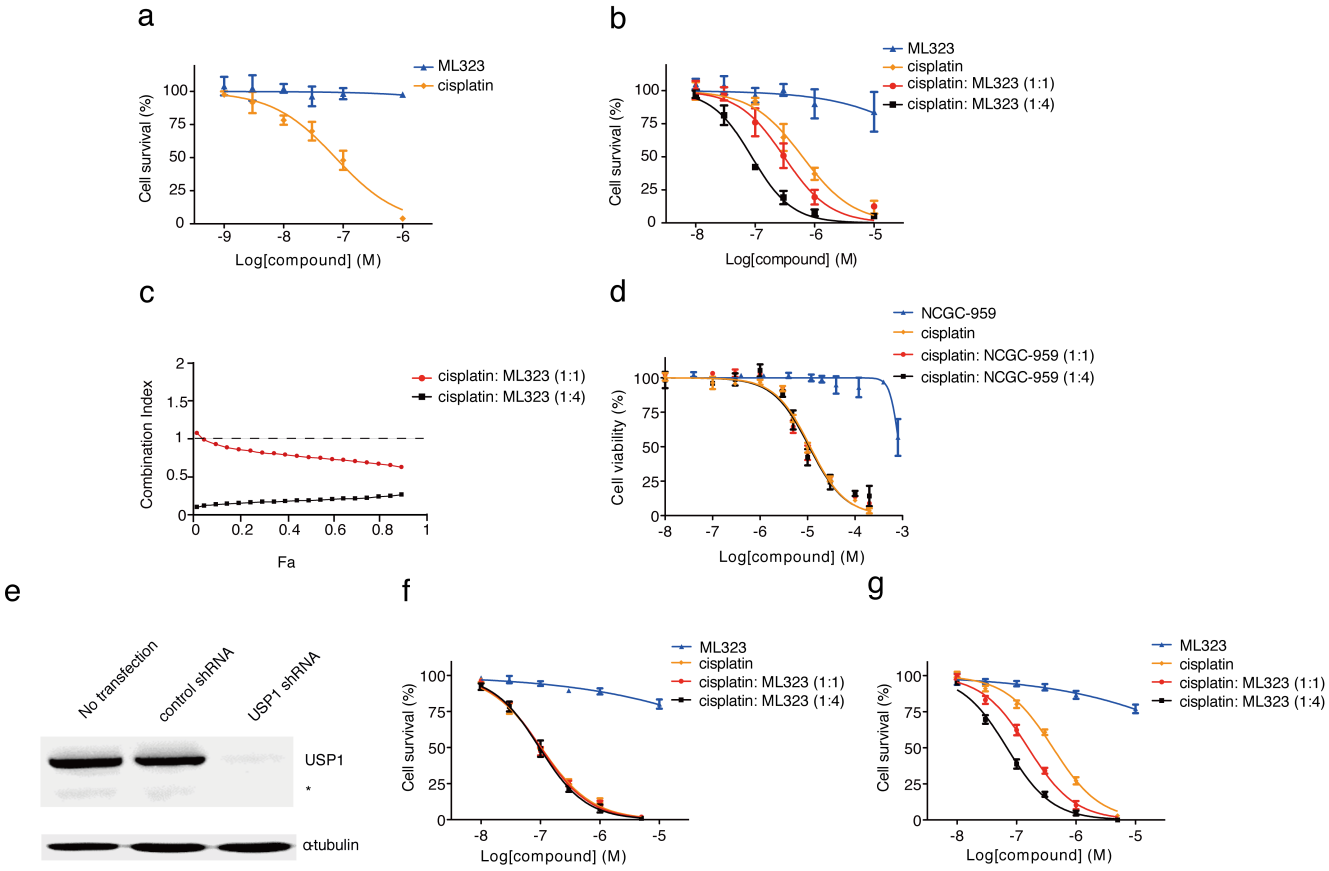


Supplementary Figure 5. The effect of ML323 on the Ub-PCNA level in cells. (**a**) H596 cells were treated with cisplatin (100 μM), ML323 (30 μM), or a combination of cisplatin (100 μM) and ML323 (30 μM). The cells treated with the vehicle (DMSO) were used as a control. Cell lysate was analyzed by Western blotting using the anti-FANCD2, anti-PCNA, anti-H2A, anti-USP1 and anti-α-tubulin antibodies. (**b, c**) Dose-dependent analysis of the inhibitory effect of ML323 in the absence or presence of cisplatin on PCNA ubiquitination in H596 cells. In **b**, cells were treated with 0, 1, 5, 15 or 30 μM ML323 in the absence or presence of 100 μM cisplatin for 6 hours and analyzed by immunoblotting with anti-PCNA antibody. In **c**, the percentages of monoubiquitinated PCNA as determined in **b** under the indicated treatments. (**d, e**) Time profile analysis of the effect of ML323, cisplatin, or a combination of ML323 and cisplatin on PCNA ubiquitination in H596 cells. In **d**, cells were treated with 30 μM ML323, 100 μM cisplatin, or a combination of 30 μM ML323 and 100 μM cisplatin at indicated time, and analyzed by immunoblotting with anti-PCNA antibody. In **e**, the percentages of monoubiquitinated PCNA as determined in **d** under the indicated treatments. A time-dependent increase in the level of Ub-PCNA was observed when cells were treated with ML323 or a combination of ML323 and cisplatin. (**f, g, h**) The inhibitory effect of ML323 in cells can be reversed. H596 cells were treated with 30 μM ML323 (**f**), with a combination of ML323 (30 μM) and cisplatin (100 μM) (**g**), or with 100 μM cisplatin (**h**) for 6 hours, washed with fresh media to remove the compound(s), incubated with fresh growth medium for the indicated time, and analyzed by immunoblotting with anti-PCNA and anti-USP1 antibodies. Three to six hours following the removal of the compound, the Ub-PCNA level began to decrease in the treated cells; while the USP1 level remained largely unchanged. Asterisk indicates truncated USP1 band. For (**c**) and (**e**), data are presented as mean ± s.d.; for (**c**), N = 3 experiments/group; for (**e**), N = 2 experiments/group. The full immunoblots for this figure are shown in Supplementary Fig. 16, 17.

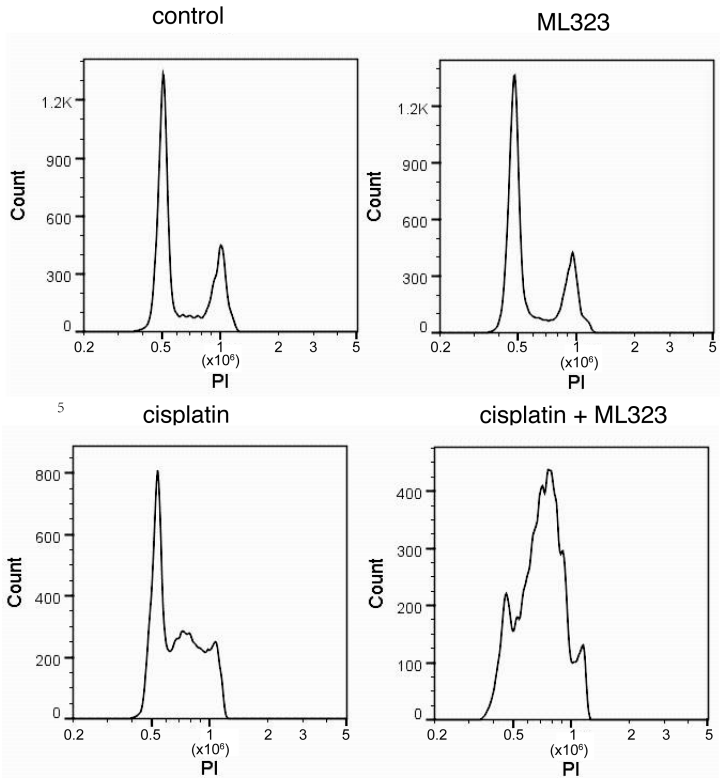


Supplementary Figure 6. ML323, but not NCGC-959, inhibits the cellular activity of USP1-UAF1. Western blotting analysis and quantified percentages of FANCD2, PCNA and H2A monoubiquitination in U2OS (**a, b**) and HEK293T cells (**c, d**) treated with 100 μM cisplatin, 30 μM ML323, or a combination of 100 μM cisplatin and 30 μM ML323. The cells treated with the vehicle (DMSO) were used as a control. The treatment with ML323 increased the percentage of monoubiquitinated PCNA and FANCD2 in cells. A combination of cisplatin and ML323 further increased the percentage of monoubiquitinated PCNA and FANCD2. ML323 had little effect on histone H2A monoubiquitination. (**e, f**) Western blotting analysis and quantified percentages of

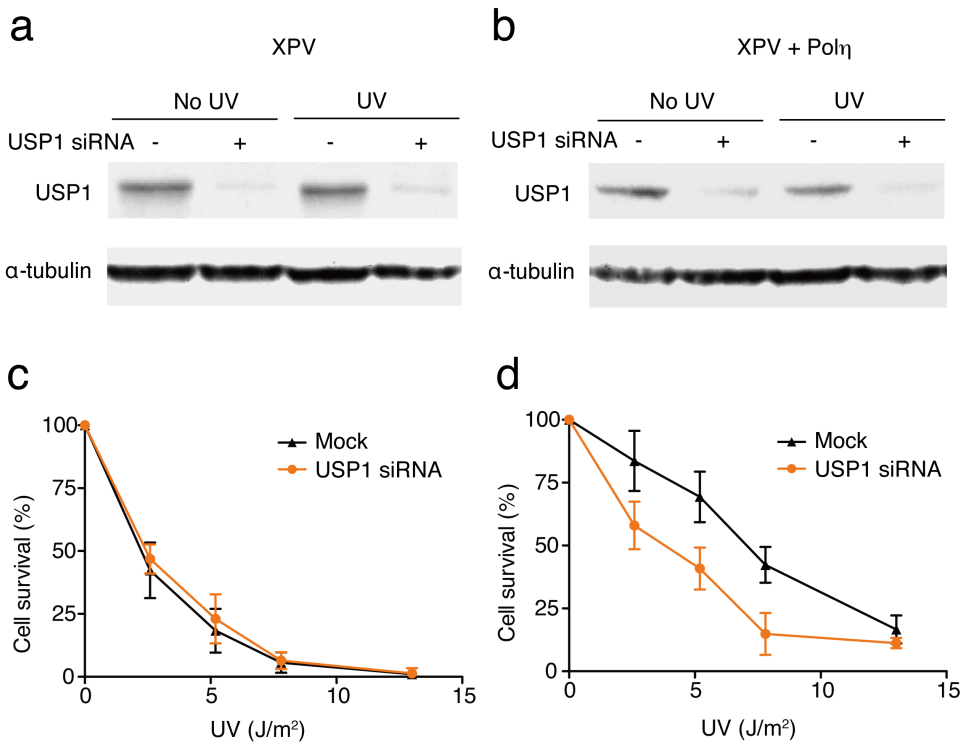
FANCD2, PCNA and H2A monoubiquitination in H596 cells treated with cisplatin (100 μM), NCGC-959 (30 μM), or a combination of cisplatin (100 μM) and NCGC-959 (30 μM). The cells treated with the vehicle (DMSO) were used as a control. α-tubulin was used as a loading control. The treatment with NCGC-959 had no effect on the percentage of monoubiquitinated PCNA, FANCD2 and H2A in H596 cells. A combination of cisplatin and NCGC-959 did not further increase the percentage of monoubiquitinated PCNA and FANCD2 compared to cisplatin alone. (**g, h)** Western blotting analysis and quantified percentages of PCNA monoubiquitination in USP1-shRNA knockdown and control shRNA knockdown H596 cells treated with 100 μM cisplatin, 30 μM ML323, or a combination of 100 μM cisplatin and 30 μM ML323, respectively. The cells treated with DMSO were used as control. For (**b**), (**d**), (**f**) and (**h**), data are presented as mean ± s.d.; for (**b**), (**d**) and (**h**), N = 3 experiments/group; for (**f**), N = 2 experiments/group. The full immunoblots for this figure are shown in Supplementary Fig. 18, 19.



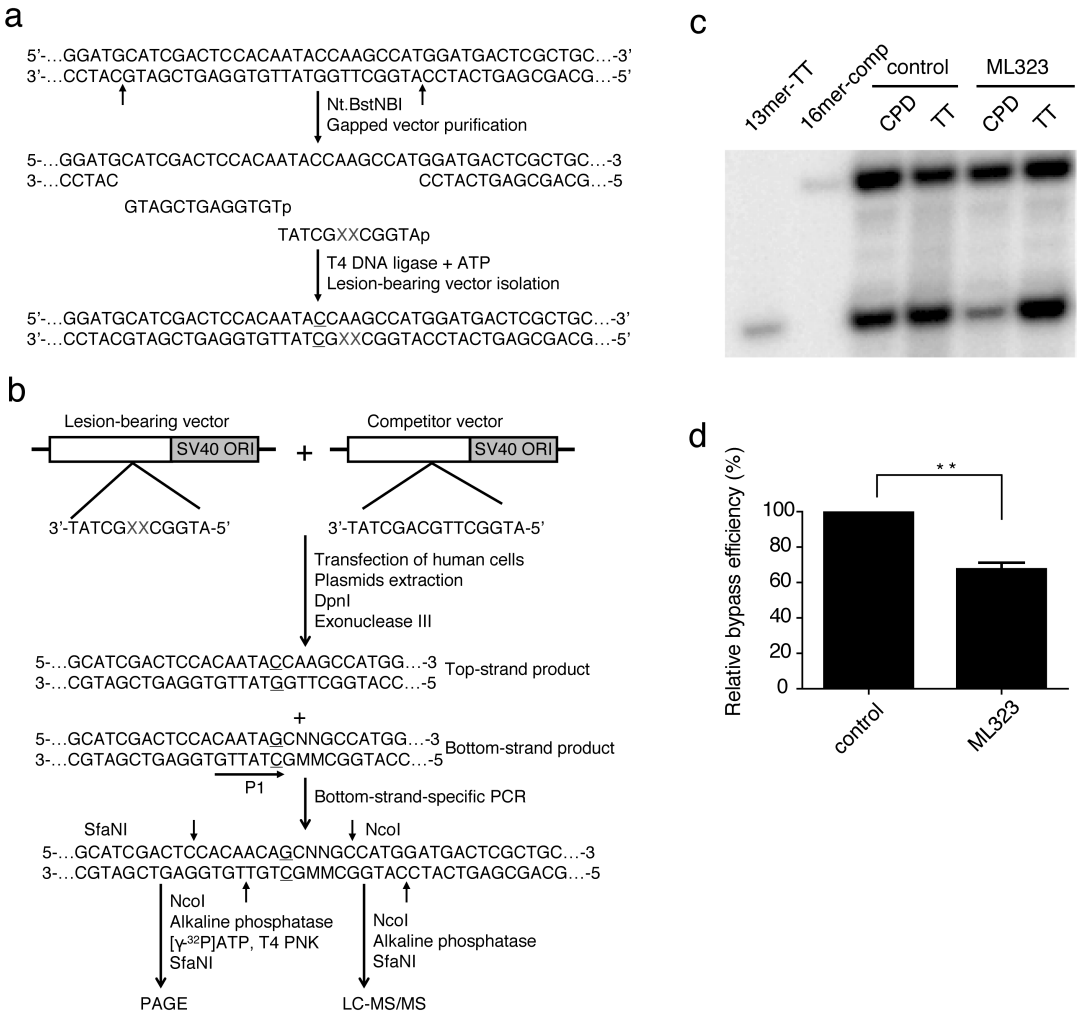
Supplementary Figure 7. Cytotoxicity of ML323 in combination with cisplatin in cancer cells. (**a**) Cytotoxicity of cisplatin and ML323 in H460 cells in colony-forming assay. For cisplatin alone, the EC_{50} is 0.074 μM; for ML323 alone, the EC_{50} is higher than 10 μM. (**b**) Cytotoxicity of cisplatin alone, ML323, or a combination of cisplatin and ML323 in U2OS cells at a ratio of 1:1, 1:4 in colony-forming assay. The number of colonies obtained under the treatment with vehicle was treated as 100%. For cisplatin alone, the EC_{50} is 0.629 μM; for ML323 alone, the EC_{50} is higher than 10 μM; for a combination of cisplatin and ML323**,** the EC_{50} is 0.308 μM (for 1:1 ratio of cisplatin: ML323) and 0.088 μM (for 1:4 ratio of cisplatin: ML323). (**c**) Synergistic interaction of cisplatin and ML323 in colony-forming assay in U2OS cells at the ratio of 1:1 and 1:4. The dashed horizontal line in the figure represents a combination index (CI) of 1. (**d**) Cytotoxicity of cisplatin, NCGC-959, or a combination of cisplatin and NCGC-959 at ratios of 1:1, 1:4 in H596 cells in CCK assay. Cell viability determined with equal volume of vehicle was treated as 100%. For cisplatin alone, the EC_{50} is 11.4 μM; for NCGC-959 alone, the EC_{50} is higher than 400 μM; for a combination of cisplatin and NCGC-959, the EC_{50} at the 1:1 ratio is 11.1 μM and at the 1:4 ratio is 10.9 μM. (**e**) Western blots showing the USP1-shRNA knockdown in H596 cells. (**f & g**) Colony-forming assay using USP1-shRNA knockdown (**f**) and control-shRNA knockdown (**g**) H596 cells treated with cisplatin, ML323, or a combination of cisplatin and ML323 at ratios of 1:1 and 1:4. For (**a**), (**b**), (**d**), (**f**) and (**g**), data are presented as mean ± s.d.; for (**a**), (**b**), (**f**) and (**g**), N = 3 experiments/group; for (**d**), N = 2 experiments/group. The full immunoblots for this figure are shown in Supplementary Fig. 19.



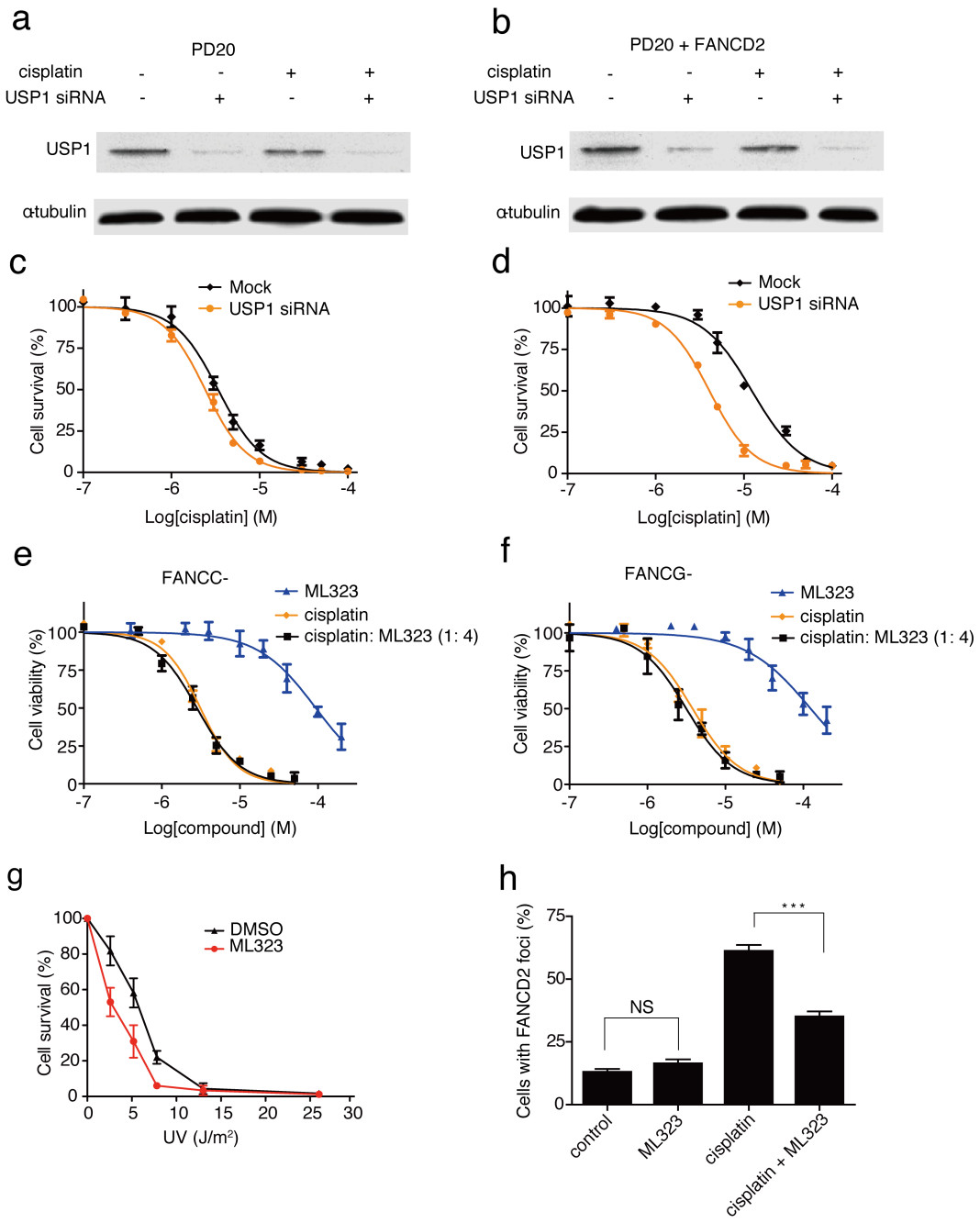
Supplementary Figure 8. Cell cycle analysis of H596 cells treated with ML323 and cisplatin. H596 cells were treated with DMSO, ML323 alone (20 μM), cisplatin alone (5 μM), or a combination of ML323 (20 μM) and cisplatin (5 μM), and analyzed by flow cytometry.



Supplementary Figure 9. **The effect of USP1 knockdown in XPV and XPV + Polη cells.** Knockdown of USP1 by anti-USP1 siRNA in XPV (XP115LO) cells (**a**) and XPV + Polη cells (**b**) in the absence or presence of UV irradiation. Cell lysate was analyzed by Western blotting with anti-human USP1 antibody. α-tubulin was the loading control. Colony-forming assay of XPV cells (**c**) and XPV + Polη cells (**d**) irradiated with UV or a combination of UV and anti-USP1 siRNA. In XPV cells, for UV alone, LD_{50} is 2.6 J m^{-2}; for a combination of UV irradiation and USP1 siRNA, the LD_{50} is 2.9 J m^{-2}. In XPV + Polη cells, for UV alone, LD_{50} is 7.4 J m^{-2}; for a combination of UV irradiation and USP1 siRNA, the LD_{50} is 4.2 J m^{-2}. The numbers of colonies determined with no UV irradiation were treated as 100%. For (**c**) and (**d**), data are presented as mean ± s.d.; N = 3 experiments/group. The full immunoblots for this figure are shown in Supplementary Fig. 20.

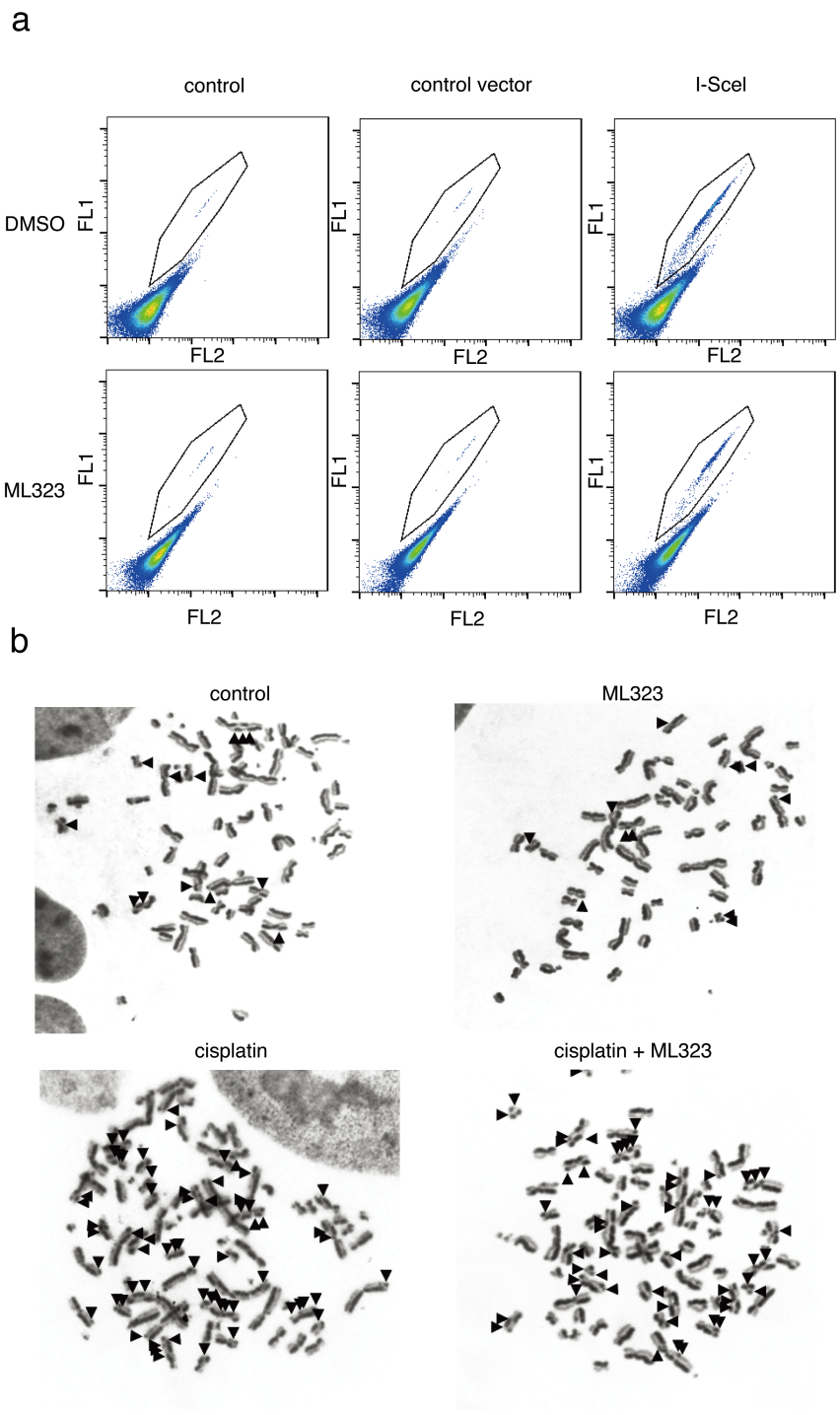


Supplementary Figure 10. The strand-specific PCR-based competitive replication and adduct bypass (SSPCR-CRAB) assay. Schematic diagrams showing the procedures for the preparation of the CPD-bearing plasmid (**a**) and the SSPCR-CRAB assay (**b**). Only the construction of CPD-bearing vector is shown. "SV40 ORI" and "XX" indicate SV40 replication origin and CPD, respectively. C/C mismatch site is underlined. The cleavage sites of Nt.BstNBI, NcoI and SfaNI are designated with arrows. P1 represents one of the primers for PCR, i.e., d(GCTAGCGGATGCATCGACTCCTTCACAG). P1 contains a G as the terminal 3'-nucleotide corresponding to the C/C mismatch site of the lesion-bearing genome, and it also contains a C/A mismatch three bases away from the 3'-end for improving the specificity of PCR. (**c**) A representative gel image of cellular replication studies of CPD using HEK293T cells that are treated with DMSO alone or treated with ML323. TT stands for undamaged DNA containing two neighboring thymines. The restriction fragments arising from the competitor vector, i.e., d(CATGGCTTGCAGCTGT), is designated with '16mer-comp'; d(CATGGCTTGCTGT) is designated with standard ODN '13mer-TT'. (**d**) The percentage of relative bypass efficiency of the CPD lesion in HEK293T cells that were treated with ML323 in the SSPCR-CRAB assay. The relative bypass efficiency of CPD lesion in cells treated with DMSO alone was designated as 100%. The data are presented as mean ± s.d.; N = 6 experiments/group. (**, $P < 0.01$ by Student's *t*-test). The full gel image for panel **c** is shown in Supplementary Fig. 20.

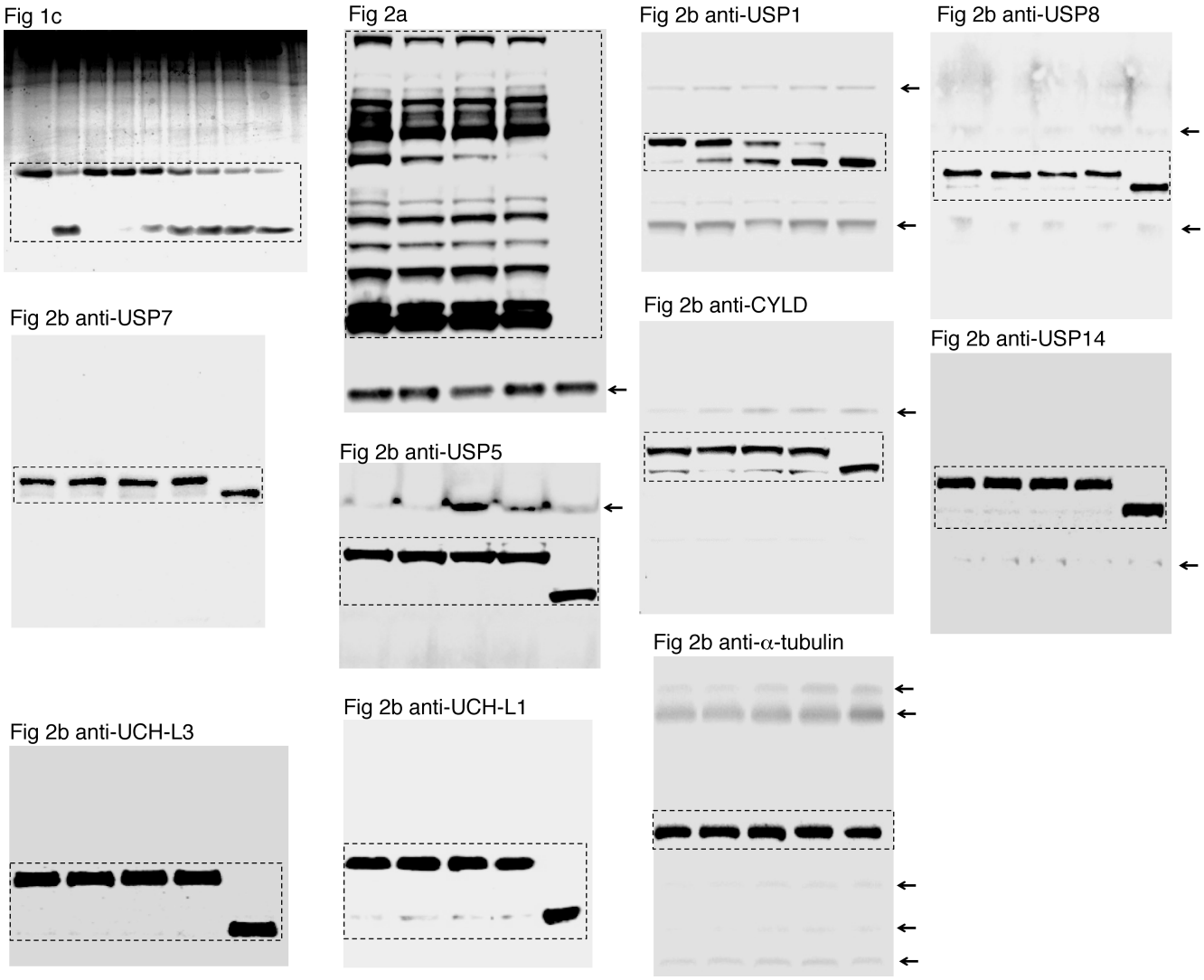


Supplementary Figure 11. Cytotoxicity of ML323 in combination with cisplatin in USP1-knockdown cells, and cells deficient in FANCC and FANCG. (a, b) Knockdown of USP1 expression by anti-USP1 siRNA in PD20 cells (**a**) and PD20 + D2 cells (**b**) in the absence or presence of cisplatin. Cell lysate was analyzed by Western blotting with anti-human USP1 antibody. α-tubulin was used as a loading control. (**c, d**) Cytotoxicity of PD20 cells (**c**) and PD20 + D2 cells (**d**) treated with cisplatin alone or a combination of cisplatin and anti-USP1 siRNA in CCK assay. In PD20 cells, for cisplatin alone, the EC_{50} is 3.2 μM; for a combination of cisplatin and siRNA, the EC_{50} is 2.5 μM. In PD20 + D2 cells, for cisplatin alone, the EC_{50} is 12.3 μM; for a combination of cisplatin and siRNA, the EC_{50} is 4.3 μM. (**e, f**) Cytotoxicity in FANCC-deficient cells (**e**) and FANCG-deficient cells (**f**) treated with cisplatin, ML323, or a combination of cisplatin and ML323 at the ratio of 1:4 in CCK assay. In FANCC-deficient cells, for cisplatin alone, the EC_{50} is 3.1 μM; for ML323 alone, the EC_{50} is

94 μM; for the combination of ML323 and cisplatin, the EC_{50} is 2.7 μM. In FANCG-deficient cells, for cisplatin alone, the EC_{50} is 3.8 μM; for ML323 alone, the EC_{50} is higher than 100 μM; for the combination of ML323 and cisplatin, the EC_{50} is 3.1 μM. (**g**) Colony-forming assay of PD20 + D2 cells irradiated with UV alone or a combination of UV and 20 μM ML323. For UV alone, LD_{50} is 5.7 J m^{-2}; for a combination of UV irradiation and ML323, the LD_{50} is 3.4 J m^{-2}. (**h**) Quantification of the FANCD2 foci formation in U2OS cells treated with ML323, cisplatin, and a combination of ML323 and cisplatin. The FANCD2 foci percentage was calculated as such, the number of cells bearing 5 and more FANCD2 foci was divided by total number of the cells. Each value was based on at least 200 nuclei in five independent experiments, which is presented as mean ± s.d. (NS, $P > 0.05$; ***, $P < 0.001$ by Student's *t*-test). For (**c**), (**d**), (**e**), (**f**) and (**g**), data are presented as mean ± s.d.; for (**c**), (**d**) and (**g**) N = 3 experiments/group; for (**e**) and (**f**), N = 2 experiments/group. The full immunoblots for this figure are shown in Supplementary Fig. 20.

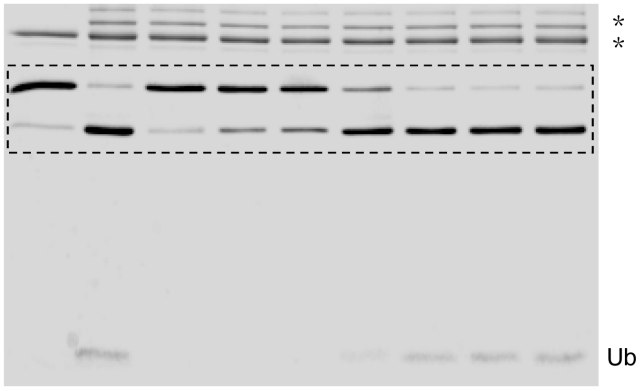


Supplementary Figure 12. HR and SCE assays of U2OS cells treated with ML323. (a) The DR-U2OS cells were treated with DMSO or ML323, and transfected with the control vector or I-SecI. Cells treated with only the transfection reagent was used as control. Cells were analyzed by flow cytometry. The green fluorescent cells reporting successful homologous recombination were indicated. (**b**) Representative figures of sister chromatid exchange in U2OS cells treated with the vehicle, ML323, cisplatin, a combination of ML323 and cisplatin. SCE events are indicated with arrow heads.

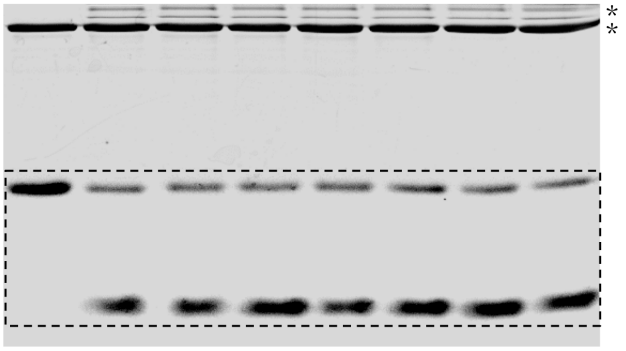


Supplementary Figure 13. Full gels for **Fig 1c** and **Fig 2a, b**. Arrows indicate non-specific bands in immunoblotting.

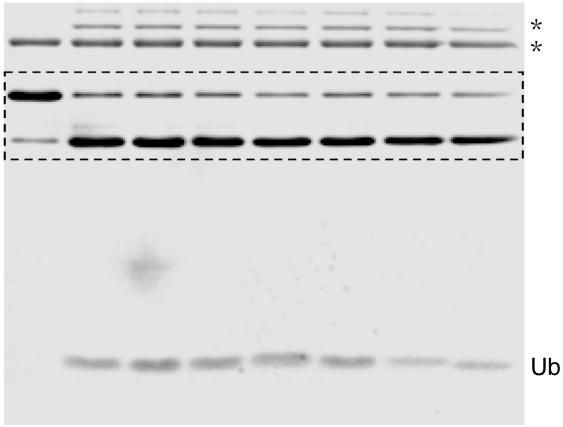
Sup Fig 1a



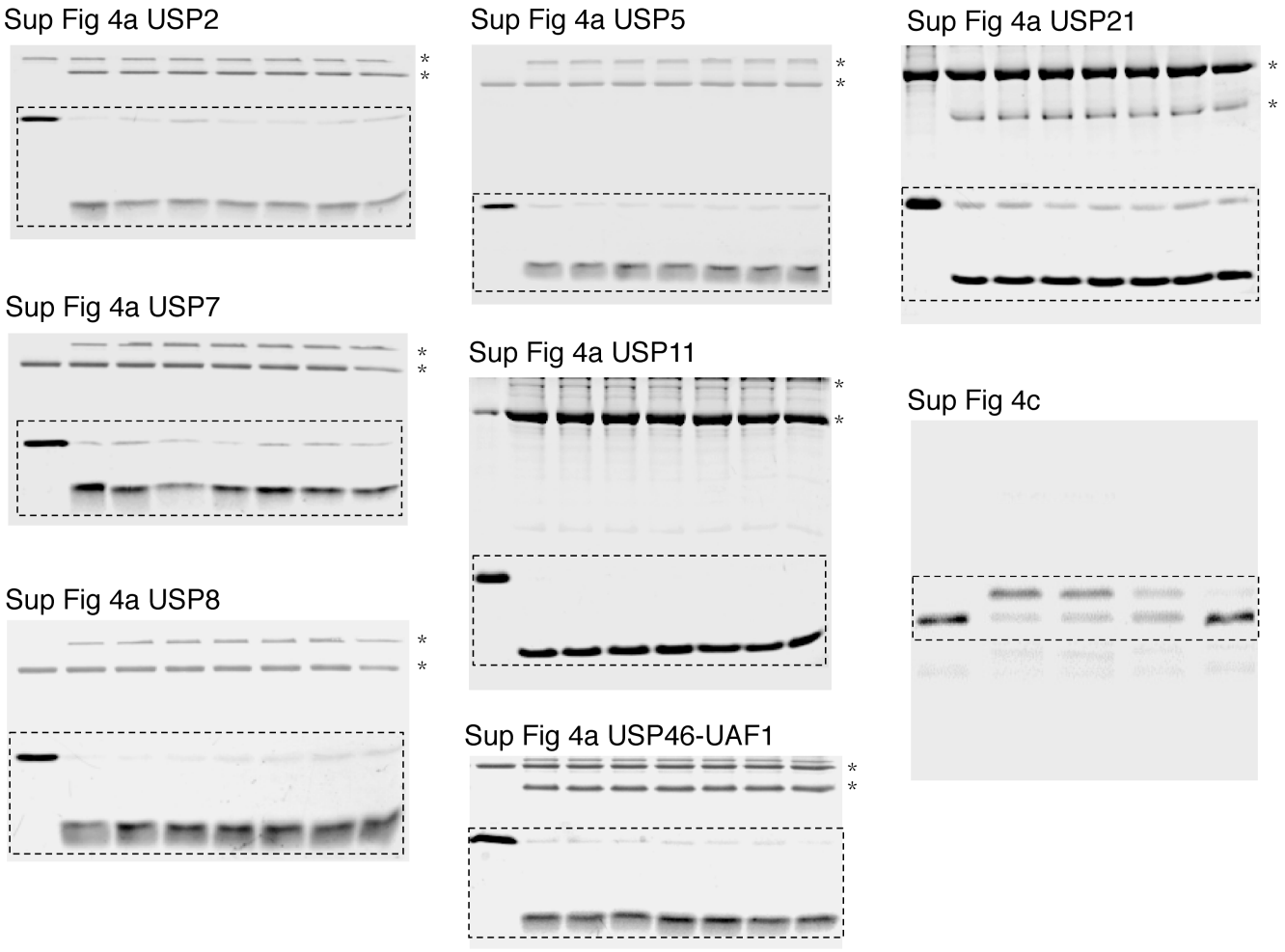
Sup Fig 1d



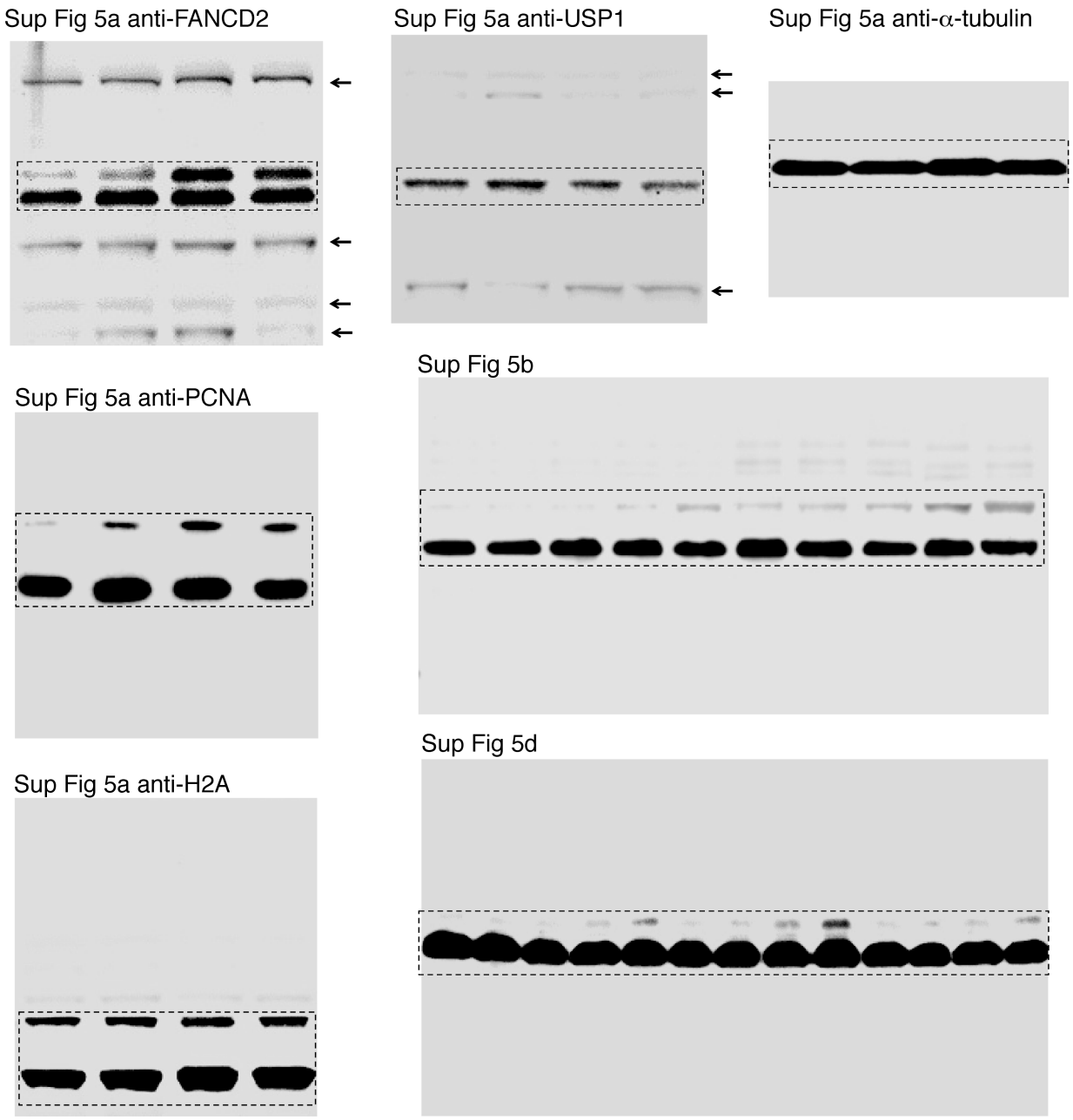
Sup Fig 1e



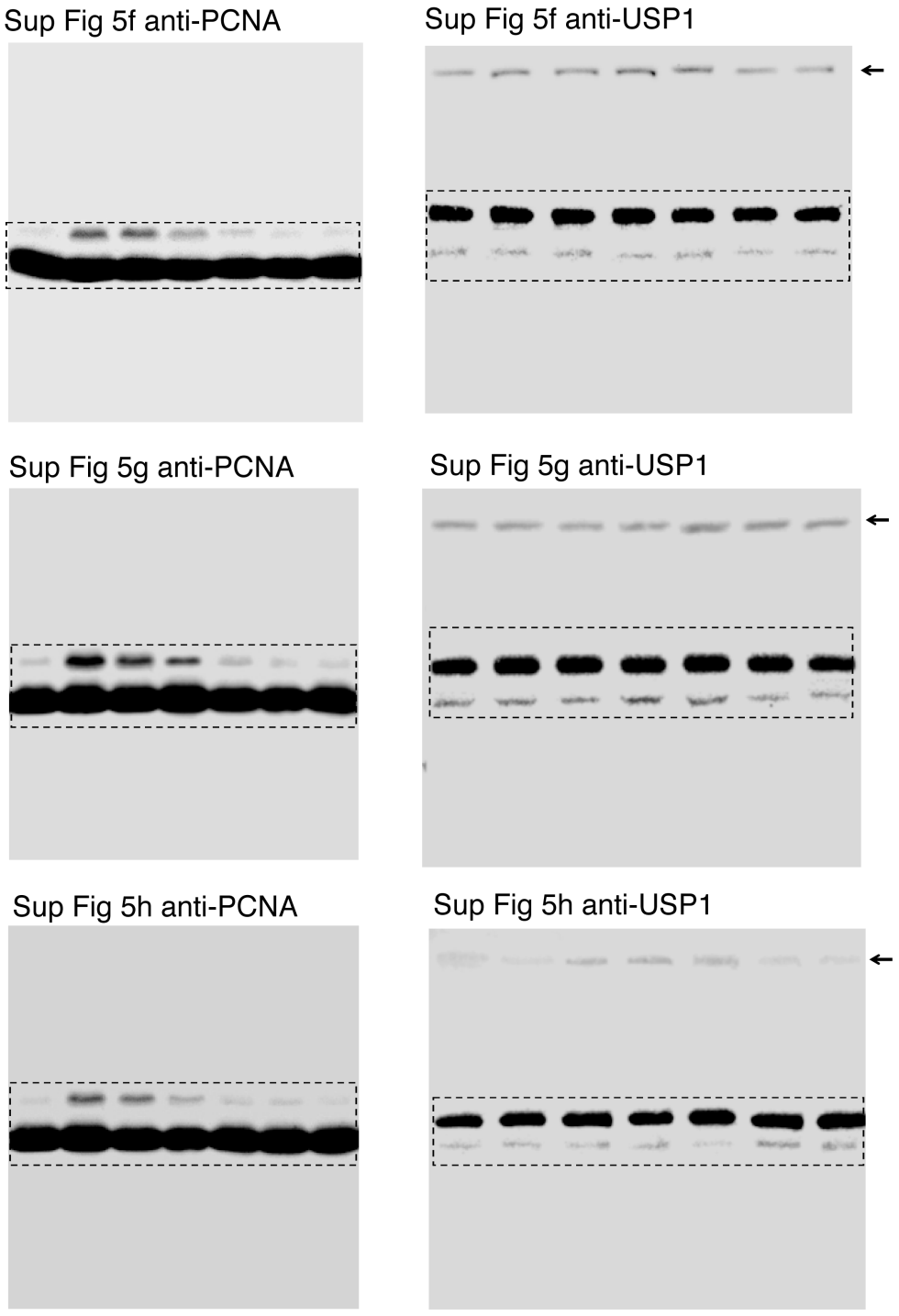
Supplementary Figure 14. Full gels for **Supplementary Fig 1a, d, e**. * indicates the enzyme and BSA bands.



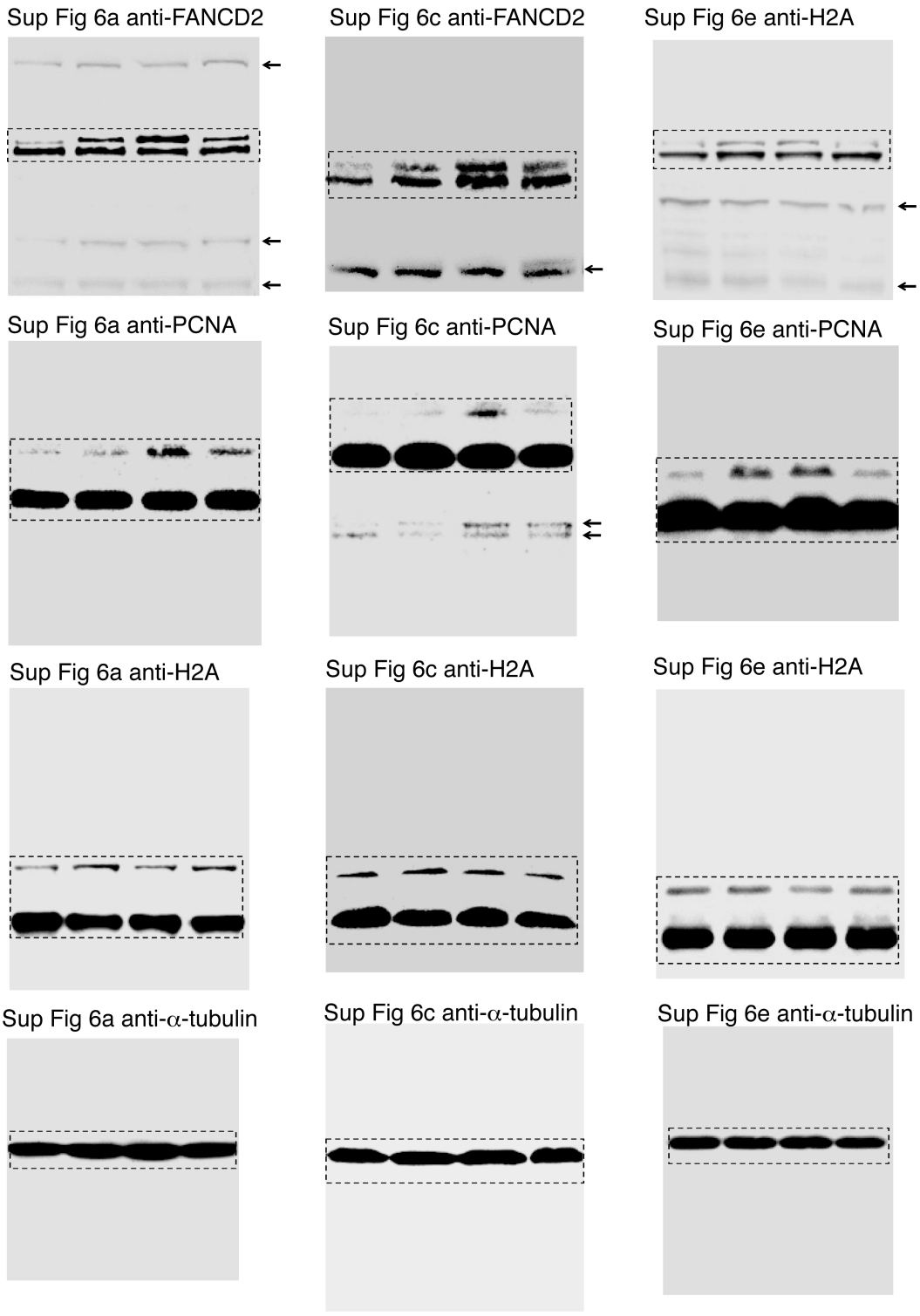
Supplementary Figure 15. Full gels for **Supplementary Fig 4a, c**. * indicates the enzyme and BSA bands.



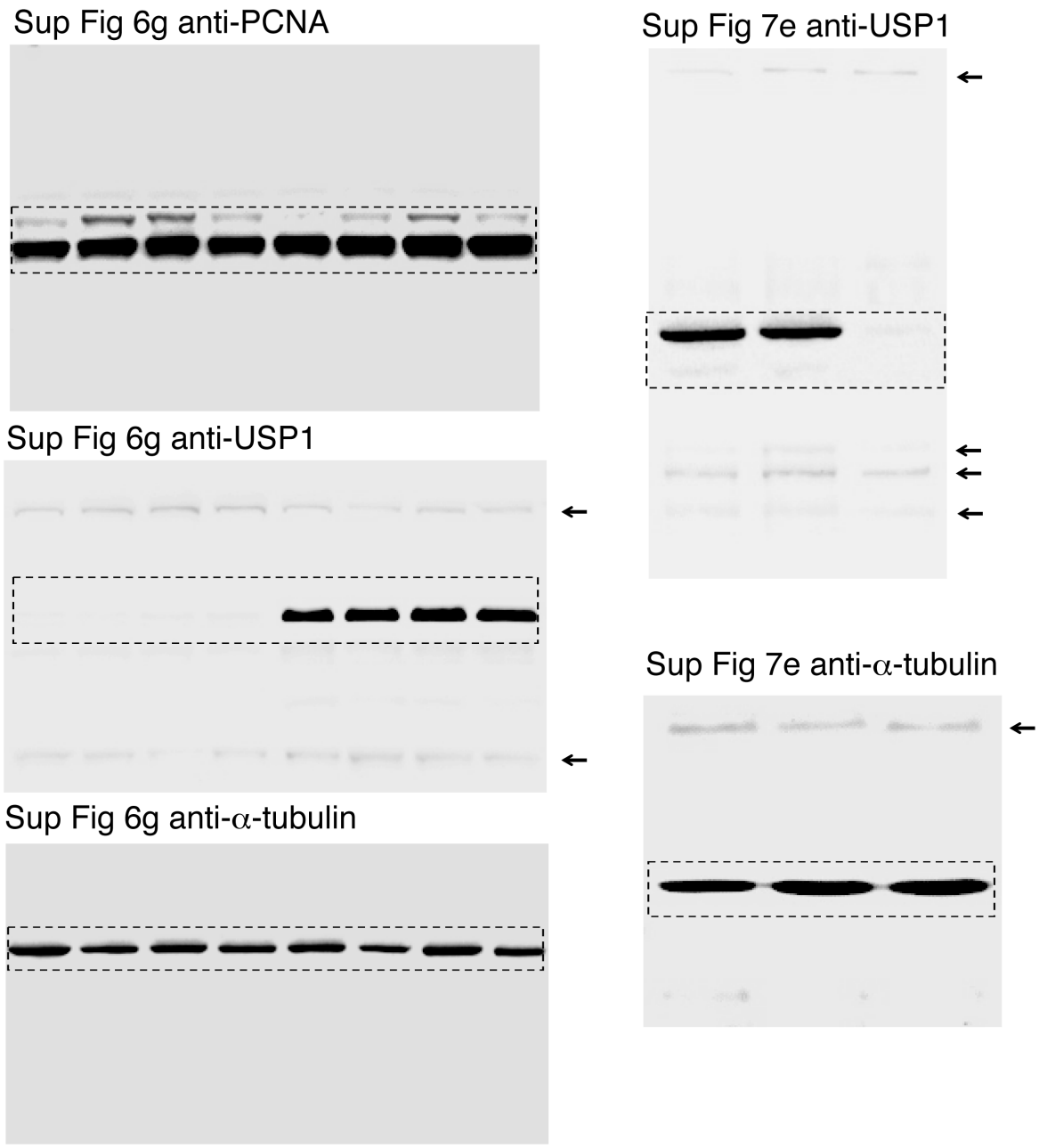
Supplementary Figure 16. Full gels for **Supplementary Fig 5a, b, d.** Arrows indicate non-specific bands in immunoblotting.



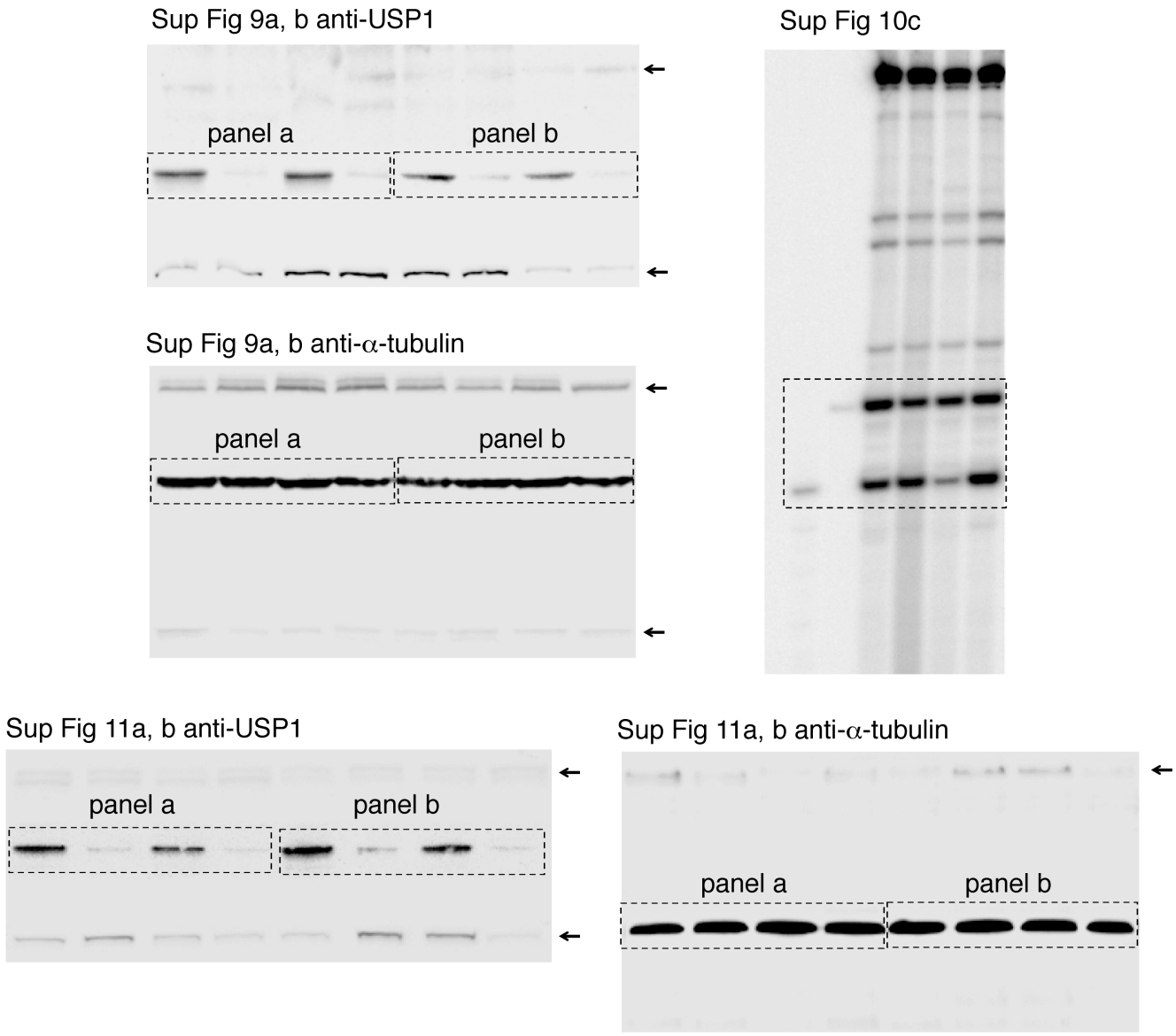
Supplementary Figure 17. Full gels for **Supplementary Fig 5f, g, h.** Arrows indicate non-specific bands in immunoblotting.



Supplementary Figure 18. Full gels for **Supplementary Fig 6a, c, e.** Arrows indicate non-specific bands in immunoblotting.



Supplementary Figure 19. Full gels for **Supplementary Fig 6g, 7e.** Arrows indicate non-specific bands in immunoblotting.



Supplementary Figure 20. Full gels for **Supplementary Fig 9a, b, 10c, 11a, b.** Arrows indicate non-specific bands in immunoblotting.

Supplementary Table 1. USP1-UAF1 HTS summary table.

Category	Parameter	Description
Assay	Type of assay	qHTS fluorometric assay using ubiquitin-rhodamine110 as a substrate
	Target	*Ubiquitin-specific protease 1 (USP1)* / USP1-associated factor 1 (UAF1) complex
	Primary measurement	Fluorescence intensity
	Key reagents	Ubiquitin-rhodamine110 (LifeSensors, Inc.) and recombinant USP1-UAF1 complex, preparation detailed in reference 3.
	Assay protocol	See methods, described in reference 3.
Library	Library size	402,701
	Library composition	Approved and investigational drugs; bioactive collections
	Source	MicroSource Discovery Systems, Inc. (Gaylordsville, CT); SigmaAldrich (St. Louis, MO) LOPAC1280; Prestwick Chemical (Illkirch, France); Tocris Biosciences (Ellisville, MO); TimTec, Inc (Newark, DE); NCGC Pharmaceutical Collection (reference 4); NCGC Kinase Focused Pharmacopeia library (reference 5); NIH Small Molecule Repository (http://mli.nih.gov/mli/compound-repository/); NCGC drug-like "Sytravon" diversity library.
	Additional comments	Compounds arrayed as 5-fold interplate titration series in DMSO from top concentration of 10 mM
Screen	Format	1536-well flat bottom, solid black, medium binding plates (Greiner Bio-one)
	Concentration(s) tested	114 μM, 57 μM, 11 μM, 2.3 μM, 460 nM
	Plate controls	No enzyme, DMSO, and GW7647 titration (n=2) on each plate
	Reagent/ compound dispensing system	Fully automated Kalypsys robotic system, as detailed in reference 6.
	Detection instrument and software	Wallac ViewLux CCD detector with ViewLux Data Manager software (Perkin Elmer, Waltham, MA)
	Assay validation/QC	Detailed in reference 3.
	Correction factors	Correction factors were derived from DMSO only plates interleaved approximately every 50 plates throughout the screen to monitor any systematic trend in the assay signal
	Normalization	Normalized to inhibited and uninhibited controls.
Post-HTS analysis	Hit criteria	Dose response behavior
	Hit rate	Actives: 1,756 compounds or 0.44% hit rate (Curve Class -1.1, -1.2, -2.1, -2.2 and inhibition at top concentration greater than 50%)

	Inconclusives: 20,519 compounds or 5.09% inconclusives which were weakly active or resulted in moderate inhibition. See reference 7 for details on processing of qHTS data.
Additional assay(s)	Compounds confirmed by re-test, gel-based assay, and Western blot, see text for details.
Confirmation of hit purity and structure	Resynthesis
Additional comments	

Supplementary Table 2. Selectivity of ML323 in inhibiting USP1-UAF1.

The IC_{50} (μM) value of compound ML323. NI, no inhibition.

Enzyme	USP1/UAF1	USP2	USP5	USP7	USP8	USP11	USP21	USP46/UAF1
Substrate	di-Ub	di-Ub	di-Ub	di-Ub	di-Ub	di-Ub	di-Ub	di-Ub
IC_{50}	0.17	NI	NI	NI	NI	NI	NI	NI
Enzyme	**USP1**	**UCH-L3**	**UCH-L1**	**SENP-1**	**Caspase-1**	**Caspase-3**	**Caspase-6**	**Papain**
Substrate	Ub-Rho	Ub-Rho	Ub-Rho	SUMO1 -AMC	Ac-WEHD -AFC	Ac-DEVD -AFC	Ac-VEID -AFC	Z-FR -AMC
IC_{50}	>200	>200	>200	>200	>200	>200	>200	>200

Supplementary Table 3. *In Vitro* profiling of DUBs and proteases by ML323

ML323 has no significant inhibition against the tested DUBs at the indicated concentration of ML323 and the other proteases with 3-fold serial dilutions of ML323 starting at 20 μM.

Enzyme	Substrate	[ML323] (μM)	Enzyme	Substrate	[ML323] (μM)
A20	Ub-AMC	10	Granzyme B	Ac-IEPD-AMC	20
Ataxin-3	Ub AMC	10	HCV NS3/4A	Anaspec EnzoLyte	20
BAP1	Ub-AMC	10	Kallikrein 1	Z-GPR-AMC	20
USP10	Ub-AMC	10	Kallikrein 5	Z-VVR-AMC	20
USP14	Ub-AMC	10	Kallikrein 8	VPR-AMC	20
NEDP1	NEDD8-AMC	10	Kallikrein 12	VPR-AMC	20
Calpain 1	N-Succinyl-Leu-Tyr-AMC	20	Kallikrein 13	VPR-AMC	20
Caspase 2	Ac-LEHD-AMC	20	Kallikrein 14	VPR-AMC	20
Caspase 4	Ac-LEHD-AMC	20	MMP1	(5-FAM/QXLTM) FRET peptide	20
Caspase 5	Ac-LEHD-AMC	20	MMP2	(5-FAM/QXLTM) FRET peptide	20
Caspase 7	Ac-DEVD-AMC	20	MMP3	(5-FAM/QXLTM) FRET peptide	20
Caspase 8	Ac-LEHD-AMC	20	MMP7	(5-FAM/QXLTM) FRET peptide	20
Caspase 9	Ac-LEHD-AMC	20	MMP8	(5-FAM/QXLTM) FRET peptide	20
Caspase 10	Ac-LEHD-AMC	20	MMP9	(5-FAM/QXLTM) FRET peptide	20
Caspase 11	Ac-LEHD-AMC	20	MMP10	(5-FAM/QXLTM) FRET peptide	20
Caspase 14	Ac-LEHD-AMC	20	MMP12	(5-FAM/QXLTM) FRET peptide	20
Cathepsin B	Z-FR-AMC	20	MMP13	(5-FAM/QXLTM) FRET peptide	20
Cathepsin C	Z-FR-AMC	20	MMP14	(5-FAM/QXLTM) FRET peptide	20
Cathepsin D	MCA-GKPILFFRLK(Dnp)-D-RNH$_2$	20	Plasma Kallikrein	Z-FR-AMC	20
Cathepsin E	MCA-GKPILFFRLK(Dnp)-D-RNH$_2$	20	Plasmin	H-D-CHA -Ala-Arg-AMC.2AcOH	20
Cathepsin G	Suc-AAPF-AMC	20	Proteinase A	Z-GPR-AMC	20
Cathepsin H	R-AMC	20	Proteinase K	H-D-CHA -Ala-Arg-AMC.2AcOH	20
Cathepsin K	Z-GPR-AMC	20	TACE	MCA-PLAQAV-Dpa-RSSSR-NH$_2$	20
Cathepsin L	Z-FR-AMC	20	Thrombin a	H-D-CHA -Ala-Arg-AMC.2AcOH	20
Cathepsin S	Z-FR-AMC	20	tPA	Z-GPR-AMC	20
Cathepsin V	Z-FR-AMC	20	Trypsin	H-D-CHA -Ala-Arg-AMC.2AcOH	20
Cathepsin X/Z	MCA-RPPGFSAFK(Dnp)	20	Tryptase b2	Z-GPR-AMC	20

Chymase	Suc-AAPF-AMC	20	Tryptase g1	Z-GPR-AMC	20
Chymotrypsin	Suc-AAPF-AMC	20	Urokinase	Bz-b-Ala-Gly-Arg-AMC.AcOH	20
DPP IV	H-GP-AMC	20	ACE1	MCA-RPPGFSAFK(Dnp)	20
DPP-VIII	H-GP-AMC	20	ACE2	MCA-RPPGFSAFK(Dnp)	20
DPP-IX	H-GP-AMC	20	ADAM-9	MCA-PLAQAV-Dpa-RSSSR-NH_3	20
Elastase	AR-AMC	20	ADAM-10	MCA-PLAQAV-Dpa-RSSSR-NH_3	20
Factor VIIa	Z-VVR-AMC	20	BACE 1	MCA-SEVNLDAEFRK(Dnp)-RR-NH_2	20
Factor Xa	CH_3SO_2-D-CHA-Gly-Arg-AMC-AcOH	20	Complement Component C1s	Dabcyl-SLGRKIQI-EDANS	20
Factor XIa	(Boc-Glu(OBzl)-Ala-Arg)-MCA	20			

Supplementary Note

General methods for chemistry.

All air or moisture sensitive reactions were performed under positive pressure of nitrogen with oven-dried glassware. Chemical reagents and anhydrous solvents were obtained from commercial sources and used as-is. Preparative purification was performed on a Waters semi-preparative HPLC. The column used was a Phenomenex Luna C18 (5 micron, 30 x 75 mm) at a flow rate of 45 mL/min. The mobile phase consisted of acetonitrile and water (each containing 0.1% trifluoroacetic acid). A gradient of 10% to 50% acetonitrile over 8 minutes was used during the purification. Fraction collection was triggered by UV detection (220 nM). Analytical analysis was performed on an Agilent 1200 LC/MS (Agilent Technologies). Method 1: A 7 minute gradient of 4% to 100% Acetonitrile (containing 0.025% trifluoroacetic acid) in water (containing 0.05% trifluoroacetic acid) was used with an 8 minute run time at a flow rate of 1 mL/min. A Phenomenex Luna C18 column (3 micron, 3 x 75 mm) was used at a temperature of 50 °C. Method 2: A 3 minute gradient of 4% to 100% Acetonitrile (containing 0.025% trifluoroacetic acid) in water (containing 0.05% trifluoroacetic acid) was used with a 4 minute run time at a flow rate of 1 mL/min. All of the analogs for assay have purity greater than 95% based on both analytical methods. ^{1}H and ^{13}C NMR spectra were recorded on a Varian 400 (100) MHz spectrometer. *High resolution mass* spectrometry data was recorded on Agilent 6210 Time-of-Flight LC/MS system.

Optimization of lead compound 1

The optimization strategy for HTS "hit", compound **1**, involved systematic modification of the three regions of the molecule. First, SAR exploration of the northern thiophene moiety was pursued with a

variety of different substituents. These efforts revealed that replacement of the thiophene group with a phenyl group results in comparable activity. This provided an additional function handle by which further modifications could be explored. Ultimately, through extensive SAR we found that the phenyl triazole moiety (found on ML323) had optimal activity. Optimization of the pendant CF_3-phenyl moiety revealed that substitution at the 2-position was required for activity. A potency breakthrough was achieved by changing the CF_3 group to an isopropyl group (as found on ML323). Finally, in the interest of reducing the molecular weight of the parent molecule, we explored numerous variations to the core quinazoline scaffold. In doing so, we found that the quinazoline could be replaced with the 5-methyl-pyrimidine core which led to improved potency and reduced molecule weight and thus increased ligand efficiency (LE). These efforts were conducted in an iterative fashion using parallel synthesis techniques and while taking into account *in vitro* ADME parameters such as aqueous solubility, microsomal stability and PAMPA permeability. Importantly, all final compounds were extensively purified and characterized for their identity and purity (>95%).

Synthesis and characterization of compounds.

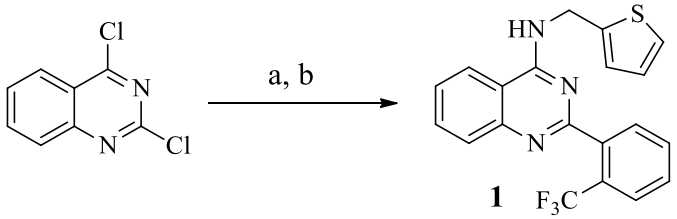


Figure 1. Synthesis of hit compound **1**. (a) thiophen-2-ylmethanamine (1.2 equiv), NEt_3 (3 equiv), CH_2Cl_2, 16 h (b) 2-(trifluoromethyl)phenylboronic acid (2 equiv), 0.4 M Na_2CO_3 (1.5 equiv), $Pd(PPh_3)_4$ (15 mol%), MeCN, MW, 150 ^oC, 10 min.

***N*-(thiophen-2-ylmethyl)-2-(2-(trifluoromethyl)phenyl)quinazolin-4-amine**

2,4-Dichloroquinazoline (0.50 g, 2.51 mmol), thiophen-2-ylmethanamine (0.34 g, 3.0 mmol), triethylamine (1 mL, 7.54 mmol) was stirred overnight in CH_2Cl_2 (6 mL) at room temperature. The reaction mixture was poured in to water and extracted (3 X) with CH_2Cl_2, the organic layers were combined washed (1 X) with brine, dried over Na_2SO_4, filtered, concentrated, to provide 2-chloro-N-(thiophen-2-ylmethyl)quinazolin-4-amine which was used without further purification in the next reaction. LC-MS Retention Time (Method 2: 3 min) = 3.334 min.

A 5 mL microwave reaction vessel was charged with a mixture of the above-mentioned 2-chloro-N-(thiophen-2-ylmethyl)quinazolin-4-amine (0.09 g, 0.31 mmol), 2-(trifluoromethyl)phenylboronic acid (0.18 g, 0.96 mmol), $Pd(PPh_3)_4$ (0.06 g, 0.05 mmol) , sodium carbonate (0.40 M in water, 1.2 mL, 0.48 mmol), and acetonitrile (1.2 mL). The vessel was sealed and heated, with stirring, at 150 ºC for 10 minutes via microwave irradiation. The organic portion was concentrated under reduced pressure and the residue was purified by silica gel chromatography using (80 percent hexanes, 20 percent ethyl acetate) to give the desired product. LC-MS Retention Time (Method 1: 7 min) = 4.633 min and (Method 2: 3 min) = 2.895 min; ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.47−8.37 (m, 1H), 8.03−7.75 (m, 7H), 7.71 (ddd, J = 1.25, 7.10, 8.34 Hz, 1H), 7.38 (dd, J = 1.27, 5.12 Hz, 1H), 7.07 (dd, J = 1.22, 3.43 Hz, 1H), 6.95 (dd, J = 3.46, 5.09 Hz, 1H) and 5.03 (d, J = 5.70 Hz, 2H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 160.09, 159.72, 158.66, 158.33, 140.73, 135.43, 132.43, 131.94, 130.96, 128.01, 127.70, 127.40, 126.24, 125.67, 123.93, 122.95, 118.24, 115.30, 112.89 and 40.60. LC-MS Retention Time (Method 1: 7 min) = 4.808 min and (Method 2: 3 min) = 3.128 min. HRMS: *m/z* (M+H) = (Calculated for $C_{20}H_{15}F_3N_3S$ 386.0933) found, 386.0942.

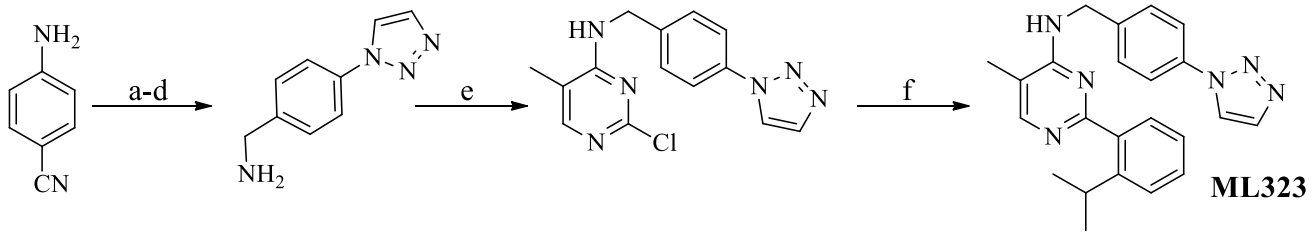


Figure 2. Synthesis of ML323. (a) TFA (1.0 equiv), rt, 5 min, *t*-butyl nitrite (1.5 equiv), azidotrimethylsilane (1.4 equiv), 0 ºC, 30 min. (98%) (b) ethynyltrimethylsilane (6.0 equiv), sodium ascorbate (0.8 equiv), Cu(II)SO_4 (0.07 equiv), DMSO/H_2O, 80 ºC 24 h (c) TFA (1 equiv), acetonitrile, reflux, 2 h, (57%). (d) H-Cube Pro®, 70 mm 10% Pd/C Catcart, 50 ºC, 40 bar, TFA, MeOH/DMF (10/1), (98%) (e) 2,4-dichloro-5-methylpyrimidine NEt_3 (3.0 equiv), DMF, 100 ºC, 18 h. (f) (2-isopropylphenyl)boronic acid (3.0 equiv), 2 M Na_2CO3 (4.0 equiv), DPP-Pd silica bound Silicycle® 0.26 mmol/g (19 mol%), DME, MW, 150 ºC, 30 min, 35-50% yield.

***N*-(4-(1H-1,2,3-triazol-1-yl)benzyl)-2-chloro-5-methylpyrimidin-4-amine**

4-aminobenzonitrile (1.0 g, 8.5 mmol), trifluoroacetic acid (TFA) (0.65 mL, 8.46 mmol) in acetonitrile (70 mL) was stirred at room temperature for 5 min. The reaction mixture was cooled to 0 ºC in a salt ice bath before the dropwise addition of *tert*-butyl nitrite (1.51 mL, 12.70 mmol) followed by azidotrimethylsilane (1.35 mL, 10.16 mmol). This reaction mixture was stirred for 30 min at 0 ºC, allowed to warm to room temperature (rt) before pouring into ethyl acetate (50 mL) and water (75 mL). The water layer was extracted (2 X) with ethyl acetate, the organic layers were combine and washed (1 X) with brine. The organic layer was dried over Na_2SO_4, filtered, and concentrated under reduced pressure to give 1.22 g of the product as a reddish brown solid in a 98% yield. The compound was used as is in the next reaction: LC-MS Retention Time (Method 2: 3 min) = 3.331 min.

4-azidobenzonitrile (1.22 g, 8.33 mmol), sodium ascorbate (1.32 g, 6.66 mmol), copper (II) sulfate (93 mg, 0.0.58 mmol), and ethynyltrimethylsilane (6.20 mL, 50.0 mmol), was heated in DMSO/water (80

mL/40 mL) to 80 ºC in a sealed tube for 24 h. The reaction was allowed to cool to rt and poured into ethyl acetate, and washed (3X) with 100 mL water. The organic layer was dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The residue was taken up in acetonitrile (16 mL), and TFA (0.64 mL, 8.33 mmol) and heated to reflux for 1.5 h. After this time, the reaction was cooled to rt and poured into ethyl acetate (30 mL), washed (2X) with saturated sodium bicarbonate, dried over Na_2SO_4, filtered, and concentrated. The residue was placed on a reversed-phase flash system for purification (gradient 20-100% acetonitrile w/ 0.1% TFA in water w/ 0.1% TFA). LC-MS Retention Time (Method 2: 3 min) = 2.671 min; ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.98 (d, J = 1.26 Hz, 1H), and 8.22 – 8.00 (m, 5H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 139.81, 135.13, 133.94, 133.14, 121.64, 120.71, 120.69, 117.68, and 112.40. *m/z* $(M+H)^+$ = 171.1. The above-mentioned 4-(1H-1,2,3-triazol-1-yl)benzonitrile (1.2 g, 7.05 mmol), TFA (0.60 mL, 7.8 mmol), was dissolved in methanol (100 mL)/DMF (10 mL) and passed through a H-Cube Pro® flow reactor using a 10% Pd/C 70 mm Catcart, at 40 bar and 50 ºC. After completion of the reaction the MeOH was concentrated and the crude used in the next reaction sequence. LC-MS Retention Time (Method 2: 3 min) = 1.386 min (*m/z* $(M+H)^+$ = 174.2).

_N_-(4-(1H-1,2,3-triazol-1-yl)benzyl)-2-chloro-5-methylpyrimidin-4-amine:

(4-(1H-1,2,3-triazol-1-yl)phenyl)methanamine, TFA (7.05 mmol), 2,4-dichloro-5-methylpyrimidine (1.16 g, 7.05 mmol), triethylamine (3.0 mL, 21.3 mmol), was heated overnight to 100 ºC in DMF (25 mL). The completed reaction was poured into water (30 mL) and extracted with ethyl acetate. The ethyl acetate layer was washed (2X) with water (1X) with saturated sodium bicarbonate, dried over Na_2SO_4, filtered, and concentrated. The residue was purified on a reverse phase flash system

(gradient 10-100% acetonitrile w/ 0.1% TFA in water w/ 0.1% TFA) to give 0.19 g of desired product. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.76 (d, J = 1.17 Hz, 1H), 7.97 – 7.90 (m, 2H), 7.86 –7.80 (m, 3H), 7.54 – 7.48 (m, 2H), 4.63 (d, J = 5.98 Hz, 2H), and 2.18 – 1.75 (m, 3H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 162.60, 157.80, 154.97, 140.20, 135.90, 134.78, 129.01, 123.61, 120.61, 113.64, 43.54, and 13.50; LC-MS Retention Time (Method 2: 3 min) = 2.770 min; *m/z* $(M+H)^+$ = 301.1.

***N*-(4-(1H-1,2,3-triazol-1-yl)benzyl)-2-(2-isopropylphenyl)-5-methylpyrimidin-4-amine (ML323):** *N*-(4-(1H-1,2,3-triazol-1-yl)benzyl)-2-chloro-5-methylpyrimidin-4-amine (0.19 g, 0.63 mmol) was combined with 2-isopropylphenylboronic acid (0.31 g, 1.90 mmol), sodium carbonate (2.0 M in water, 1.23 mL, 2.53 mmol), and DPP-Pd Silicycle® 0.26 mmol/g (0.30 g) in DMF (4.50 mL). The reaction was heated at 150 ºC for 30 min in a Biotage Initiator® microwave reactor. The resulting mixture was filtered over celite and purified by HPLC (gradient 20-100% acetonitrile w/ 0.1% TFA in water w/ 0.1% TFA) to yield, after lyophilization, *N*-(4-(1H-1,2,3-triazol-1-yl)benzyl)-5-methyl-2-(2-isopropylphenyl)-pyrimidin-4-amine as a TFA salt (0.01 g, 0.02 mmol, 30%): ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.82 (dd, J = 7.9, 3.4 Hz, 1H), 8.02 – 7.93 (m, 2H), 7.76 (d, J = 1.4 Hz, 1H), 7.64 – 7.56 (m, 2H), 7.51 – 7.33 (m, 5H), 7.25 – 7.17 (m, 1H), 4.88 (d, J = 5.9 Hz, 2H), 3.19 (d, J = 6.8 Hz, 1H), 2.22 (s, 3H), 1.10 (d, J = 1.2 Hz, 3H) and 1.09 – 1.07 (m, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 162.02, 160.18, 148.00, 140.63, 138.03, 136.14, 134.35, 134.28, 131.89, 130.34, 129.25, 129.00, 126.61, 126.04, 122.04, 121.99, 120.74, 113.81, 44.62, 29.43, 23.95, 13.71 and 13.67; LC-MS Retention Time (Method 1: 7 min) = 2.938 min and (Method 2: 3 min) = 1.756 min; HRMS (ESI) *m/z* $(M+H)^+$ calcd. for $C_{23}H_{25}N_6$, 385.2135; found 385.2146.

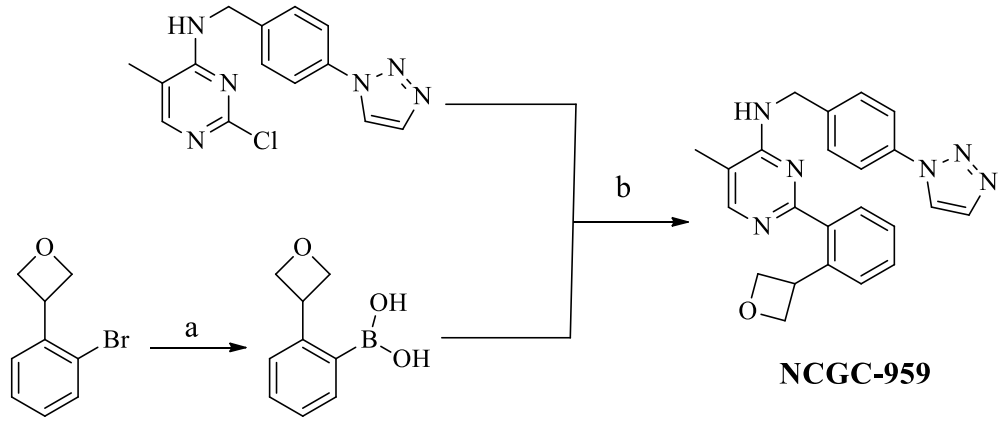


Figure 3. Synthesis of NCGC-959. (a) *N*-butyl-lithium (2 equiv), tripropyl borate (2 equiv) THF, -78 ^{o}C, 18 h; (b), 2 M Na_2CO3 (4.0 equiv), DPP-Pd silica bound Silicycle® 0.26 mmol/g (19 mol%), DME, MW, 150 ºC, 30 min, 35-50% yield.

N-(4-(1H-1,2,3-triazol-1-yl)benzyl)-2-(2-(oxetan-3-yl)phenyl)-5-methylpyrimidin-4-amine TFA (NCGC-959): 3-(2-bromophenyl)oxetane was synthesized using the method described by Jacobsen et. al,[8] except diethyl 2-(2-bromophenyl)malonate[9] was used as the starting material instead of diethyl 2-(2-(benzyloxy)phenyl)malonate. This synthetic sequence was carried out without further purification/characterization of the intermediates.

In an oven-dried 50 mL round bottom flask, the above-mentioned 3-(2-bromophenyl)oxetane (0.14 g, 0.66 mmol) and THF (4 mL) was cooled to -78 ºC. The dropwise addition of *N*-butyllithium (0.85 mL, 1.36 mmol), by syringe occurred over a 15 min period. After this time, the reaction mixture was allowed to stir at -78 ºC for an additional 2 h. The contents were transferred via cannula to a second 50 mL oven-dried round bottom flask containing a stirred solution of tripropyl borate (0.29 mL, 1.32 mmol), and THF (2.0 mL) cooled to -78 ºC. Once transferred, the mixture was allowed to stir at -78 ºC for 10 min after which time the ice bath was removed and the reaction mixture stirred overnight. 1 N HCl was used to adjust the pH to ~1, and the acidic mixture was stirred for 45 min, poured into

water, and extracted with diethyl ether (3 X). The organic layers were combined, dried over $MgSO_4$, and concentrated and once the solvent was evaporated, a ~1 M solution of the crude (2-(oxetan-3-yl)phenyl)boronic acid in 1,2-DME was made and used in next step with no further purification. In a sealed tube *N*-(4-(1H-1,2,3-triazol-1-yl)benzyl)-2-chloro-5-methylpyrimidin-4-amine (0.10 g, 0.33 mmol) was combined with 2-oxetainphenylboronic acid (1 M solution in DME, 0.67 mL, 0.67 mmol) sodium carbonate (2.0 M in water, 0.33 mL, 0.67 mmol), DPP-Pd Silicycle® 0.26 mmol/g (0.20 g) in DME (2.0 mL), and heated at 150 ºC for 30 min in a Biotage Initiator® microwave reactor. The resulting mixture was filtered over celite and purified by reversed-phase HPLC (gradient 20-100% acetonitrile w/ 0.1% TFA in water w/ 0.1% TFA) to give the desired product (NCGC-959): ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 9.36−9.47 (m, 1 H), 8.80 (s, 1 H), 8.35 (s, 1 H), 8.29 (d, *J* = 7.83 Hz, 1 H), 7.97 (s, 1 H), 7.90 (d, *J* = 8.22 Hz, 2 H), 7.68 (d, *J* = 8.61 Hz, 3 H), 7.52 (s, 2 H), 5.12 (brs, 1 H), 4.93 − 5.00 (m, 2 H), 4.41 − 4.50 (m, 2 H), 3.61 (dd, *J* =10.56, 4.30 Hz, 2 H), and 2.22 (s, 3 H); ^{13}C NMR (151 MHz, CDCl3) δ 161.40, 153.48, 144.27, 138.36, 137.52, 136.20, 134.57, 134.28, 129.62, 128.71, 128.52, 128.14, 126.14, 121.96, 120.88, 115.47, 109.98, 62.49, 51.61, 44.94, 39.85, and 14.15; LC/MS: Method 2 (short), retention time: 2.526 min. HRMS: *m/z* (M+H) = (Calculated for $C_{23}H_{23}N_6O$ 399.1928) found, 399.1924.

REFERENCE

1 Villamil, M. A., Chen, J., Liang, Q. & Zhuang, Z. A Noncanonical Cysteine Protease USP1 Is Activated through Active Site Modulation by USP1-Associated Factor 1. *Biochemistry* **51**, 2829-2839, (2012).

2 Villamil, M. A. *et al.* Serine phosphorylation is critical for the activation of ubiquitin-specific protease 1 and its interaction with WD40-repeat protein UAF1. *Biochemistry* **51**, 9112-9123, (2012).

3 Chen, J. *et al.* Selective and Cell-Active Inhibitors of the USP1/ UAF1 Deubiquitinase Complex Reverse Cisplatin Resistance in Non-small Cell Lung Cancer Cells. *Chemistry & biology* **18**, 1390-1400, (2011).

4 Huang, R. *et al.* The NCGC pharmaceutical collection: a comprehensive resource of clinically approved drugs enabling repurposing and chemical genomics. *Sci Transl Med* **3**, 80ps16, (2011).

5 Yasgar, A. *et al.* Compound Management for Quantitative High-Throughput Screening. *JALA Charlottesv Va* **13**, 79-89, (2008).

6 Michael, S. *et al.* A robotic platform for quantitative high-throughput screening. *Assay and drug development technologies* **6**, 637-657, (2008).

7 Inglese, J. *et al.* Quantitative high-throughput screening: a titration-based approach that efficiently identifies biological activities in large chemical libraries. *Proc Natl Acad Sci U S A* **103**, 11473-11478, (2006).

8 Loy, R. N. & Jacobsen, E. N. Enantioselective intramolecular openings of oxetanes catalyzed by (salen)Co(III) complexes: access to enantioenriched tetrahydrofurans. *J Am Chem Soc* **131**, 2786-2787, (2009).

9 Zhang, D., Yum, E. K., Liu, Z. & Larock, R. C. Synthesis of indenes by the palladium-catalyzed carboannulation of internal alkynes. *Org Lett* **7**, 4963-4966, (2005).