Supporting Information

Discovery and optimization of a selective ligand for the Switch/Sucrose Non-Fermenting-related bromodomains of Polybromo protein-1 by the use of virtual screening and hydration analysis.

Vassilios Myrianthopoulos[1], Nicolas Gaboriaud-Kolar[1], Cynthia Tallant[2-3], Michelle-Lynn Hall[4], Stylianos Grigoriou[1], Peter Moore Brownlee[5], Oleg Fedorov[2-3], Catherine Rogers[2-3], David Heidenreich[6], Marek Wanior[6], Nikolaos Drosos[1], Nikitia Mexia[1], Pavel Savitsky[2-3], Tina Bagratuni[7], Efstathios Kastritis[7], Evangelos Terpos[7], Panagis Filippakopoulos[2-3], Susanne Müller[2,3,6], Alexios-Leandros Skaltsounis[1], Jessica Ann Downs[5], Stefan Knapp[2,3,6] and Emmanuel Mikros[1*]*

[1]Department of Pharmacy, University of Athens, Panepistimiopolis Zografou, GR-15771, Athens, Greece.

[2]Nuffield Department of Clinical Medicine, University of Oxford, Structural Genomics Consortium, Old Road Campus Research Building, Roosevelt Drive, Oxford OX3 7DQ, UK.

[3]Nuffield Department of Clinical Medicine, University of Oxford, Target Discovery Institute (TDI), Roosevelt Drive, Oxford OX3 7BN, UK.

[4]Schrödinger Inc., 222 Third Street, Cambridge MA 02139, United States.

[5]Genome Damage and Stability Centre, School of Life Sciences, University of Sussex, Brighton BN1 9RQ, UK.

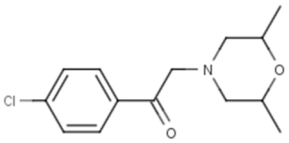
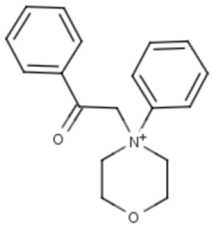
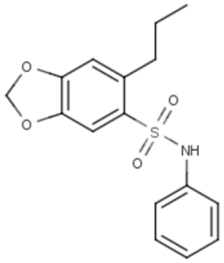
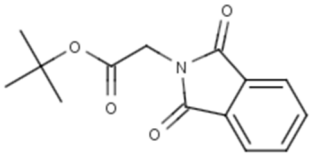
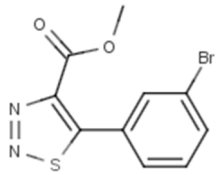
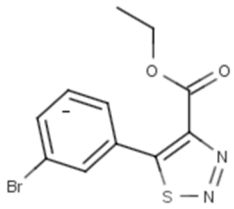
[6]Johann Wolfgang Goethe-University, Institute for Pharmaceutical Chemistry and Buchmann Institute for Life Sciences, Max-von-Laue-Str. 9, D-60438 Frankfurt am Main, Germany.

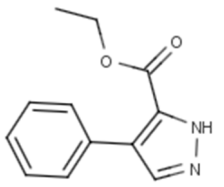
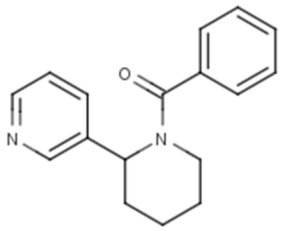
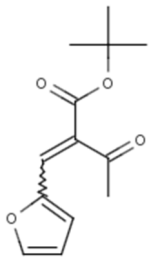
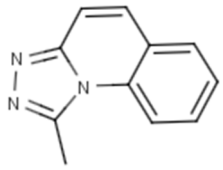
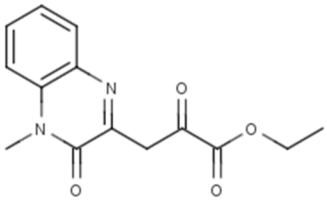
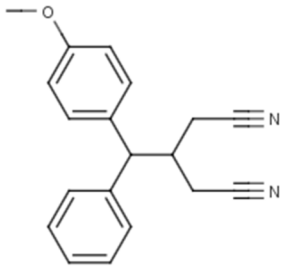
[7]Department of Clinical Therapeutics, School of Medicine, University of Athens, Mikras Asias 75, GR-11527, Athens, Greece.

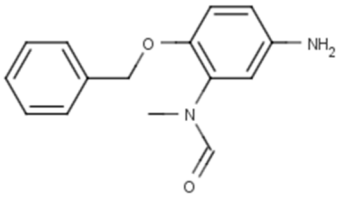
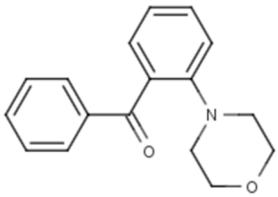
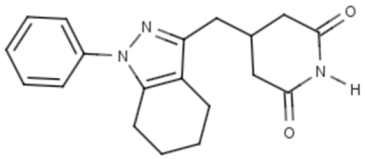
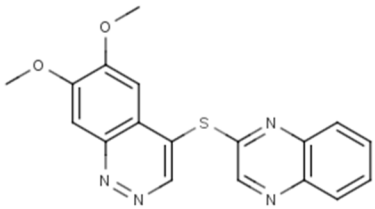
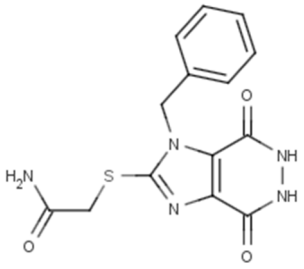
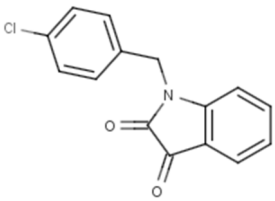
Contents:

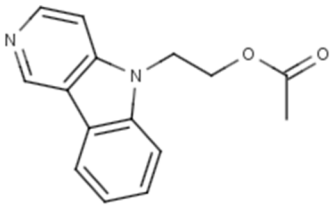
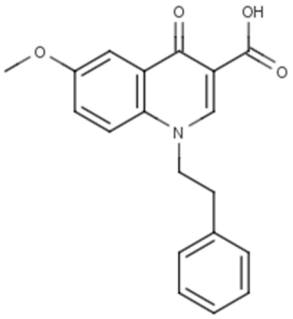
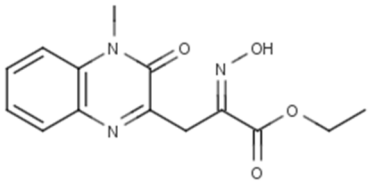
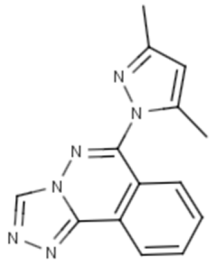
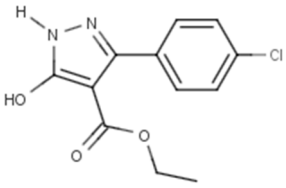
Fingerprint type	Atom typing method	Scaling
MOLPRINT2D	Carhart	no scaling
MOLPRINT2D	Daylight	no scaling
Radial	Carhart	square root feature
Radial	Carhart	square root unity
Radial	Daylight	no scaling
Radial	Daylight	square root feature
Radial	ElemR	no scaling
Radial	Hybrid	square root feature
Radial	Hybrid	square root feature
Radial	Mol2	no scaling
Radial	Mol2	square root feature

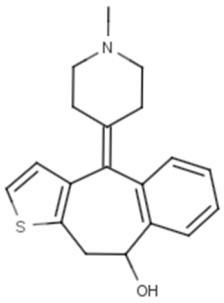
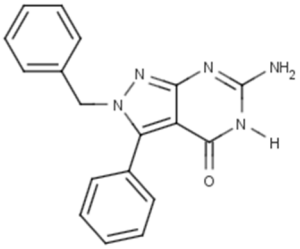
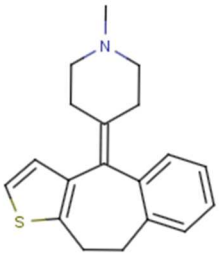
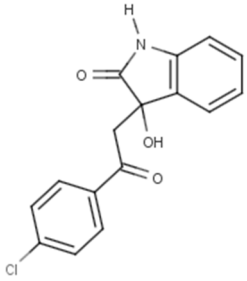
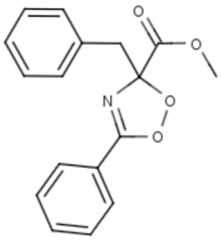
Table S1. The 11 most efficient FP combinations of Canvas used subsequently for screening. These FP screens afforded high enrichment quantified as the enrichment factor at 20% of screened library ($EF_{20\%}$). The 11 FP combinations affording high recovery of known actives showed an $EF_{20\%}$ greater than 2.25 while the remaining combinations resulted to lower $EF_{20\%}$ ranging from 0 to 1.25. The $EF_{20\%}$ ranges from 0 (no enrichment) to 5 (maximum enrichment).

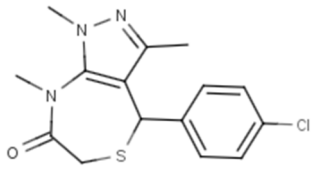
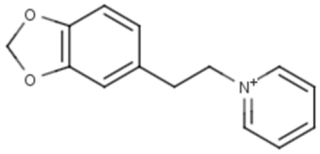
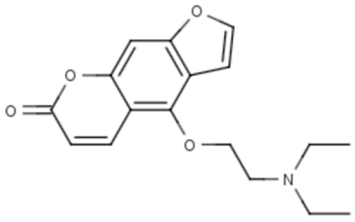
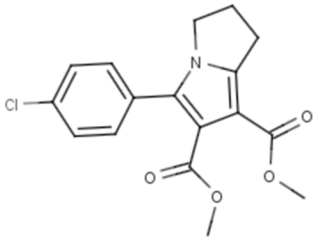
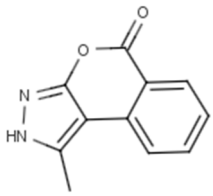
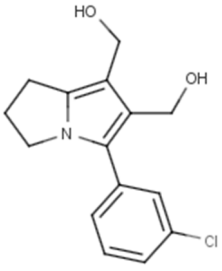
Compound	Structure	T_m (°C) PB1(5)	T_m (°C) BRD4(2)
NSC32894		-1.9	-2.0
NSC35427		-2.3	-1.8
NSC38116		0.0	0.0
NSC45837 (**3**)		0.0	0.8
NSC48742 (**4**)		-0.2	**2.0**
NSC48745		-2.4	-0.5

NSC55495 (**5**)		-0.2	**1.6**
NSC55917 (**6**)		0.6	**1.8**
NSC75086		-0.4	-0.6
NSC76484 (**8**)		0.1	**3.2**
NSC86927		-1.5	0.4
NSC97717 (**7**)		-0.4	**1.5**

NSC99768		-2.8	0.3
NSC111167		-0.4	1.1
NSC112143		-0.1	0.4
NSC114077		0.0	0.0
NSC114088		-0.3	0.1
NSC127077		-1.4	-4.4

NSC164955		-0.9	0.3
NSC211010		-0.3	0.9
NSC251944		-0.8	-0.2
NSC270404		1.0	0.5
NSC270420		0.0	-0.4

NSC282480		0.2	0.4
NSC282740		-0.3	0.3
NSC291564		-1.2	-1.4
NSC297830		-0.3	-0.9
NSC303574		0.2	-1.0

NSC313672		-1.1	-0.3
NSC340031		0.0	-0.2
NSC341960		-1.1	0.4
NSC356242		0.2	0.2
NSC356476 (**9**)		**4.1**	**1.9**
NSC358365		0.5	-0.5

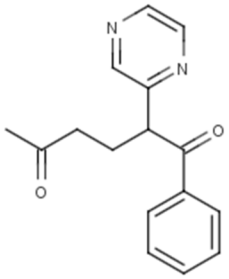
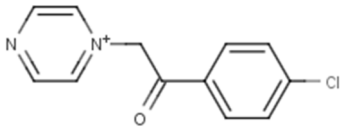
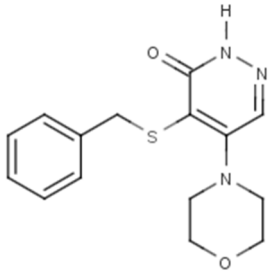
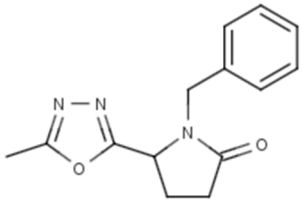
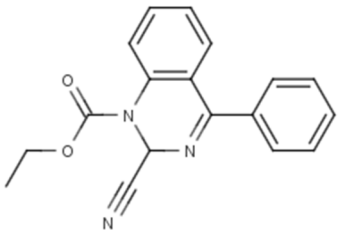
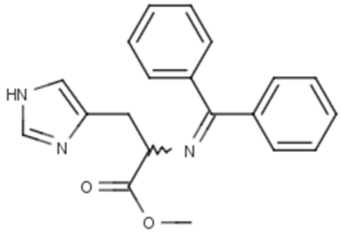
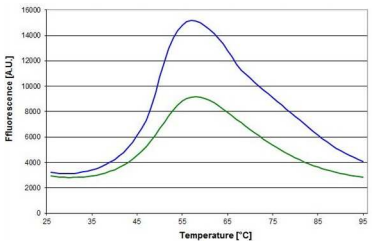
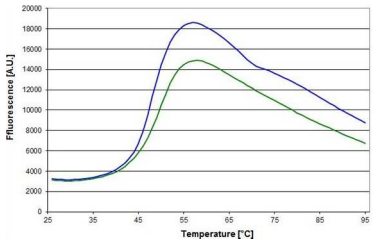
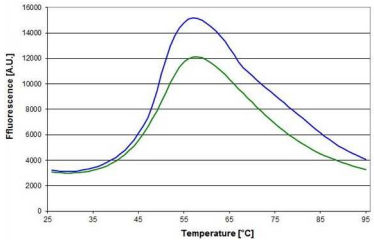
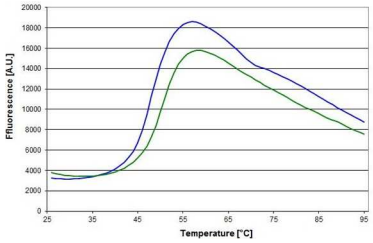
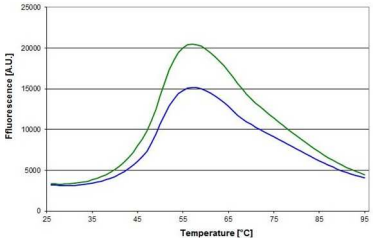
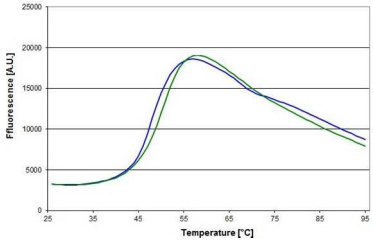
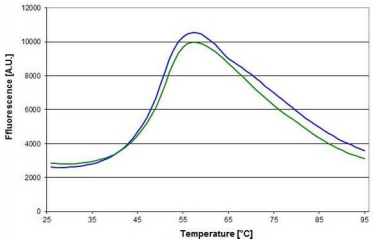
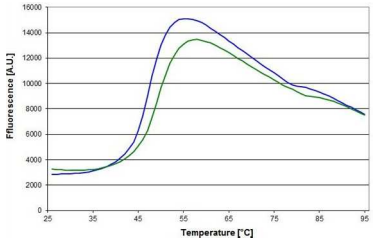
NSC400371		0.1	0.2
NSC400609		-1.4	-1.5
NSC409712		-5.1	-3.7
NSC607601		0.0	-0.2
NSC612976		-1.1	-0.6
NSC669640		0.1	-0.5

Table S2. The 40 compounds from NCI/DTP that were evaluated in the DSF screen and their T_m values towards PB1(5) and BRD4(2). The compounds were screened at 200μM while the protein was at 20μM. Shown are the compounds, their NSC code numbers and the respective T_m values. Stabilization values exceeding 1.5 °C are marked in bold. Compound **9** showed significant affinity for PB1(5) and was thus subjected to co-crystallization and X-ray crystallography to determine its binding mode to PB1(5) and suggest possible routes for optimization.

Compound	T_m PB1(5)	T_m BRD4(2)
	0 °C	0.8 °C
3 (NSC45837) C, R		
	-0.2 °C	2.0 °C
4 (NSC48742) C, R		
	0 °C	1.6°C
5 (NSC55495) G		
	0.6 °C	1.8 °C
6 (NSC55917) G		

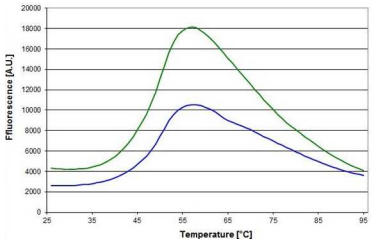
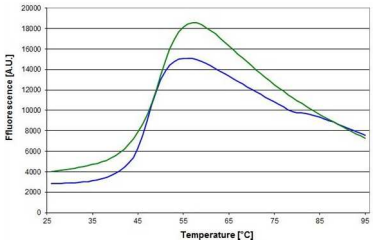
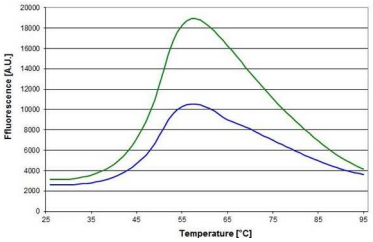
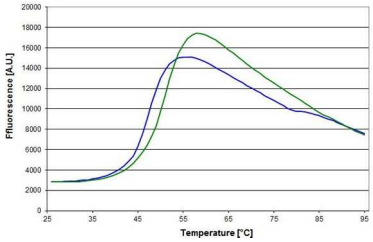
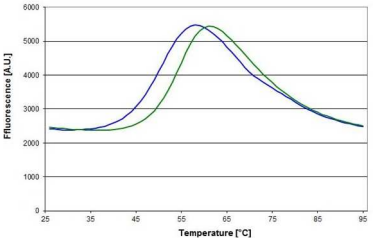
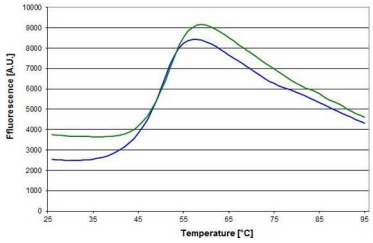
	-0.4 °C	1.5 °C
7 (NSC97717) G		
	0.1 °C	3.2 °C
8 (NSC76484) C, R		
	4.1 °C	1.9 °C
9 (NSC356476) G		

Table S3. The T_m curves of the seven compounds showing measurable binding affinity in the DSF screen. The compounds were screened at 200μM while the protein was at 20μM. Shown are the NSC number of the compounds and the T_m in °C observed for PB1(5) and BRD4(2). In the panel BRD2(2), BRD3(2), BRD4(1) and BRDT were included in addition to BRD4(2) and PB1(5) but no binding was observed for those proteins. The screening methodology affording each compound is shown below the NSC code; C: Canvas; G: Glide; R: ROCS.

PDB ID	5II1	5HRV	5HRW
Protein	PB1(5)	PB1(5)	PB1(5)
Compound	**9**	**10**	**11**
Space group	$P2_12_12_1$	C 2 2 2_1	P 2_1 2_1 2_1
Cell dimensions:			
a, b, c (Å)	41.11 57.97 106.01	57.82 140.3 41.55	41.11 59.10 105.50
α, β, γ (deg)	90.00 90.00 90.00	90.00 90.00 90.00	90.00 90.00 90.00
Resolution* (Å)	24.10 (2.02)	28.91 (1.70)	29.55 (1.80)
Unique observations*	17373 (2488)	19011 (883)	23971 (1312)
Completeness* (%)	99.8 (100.0)	99.3 (89.7)	97.4 (93.2)
Redundancy*	4.3 (4.3)	7.8 (4.2)	5.0 (4.8)
Rmerge*	0.070 (0.671)	0.043 (0.360)	0.053 (0.735)
I/σI*	12.5 (2.0)	28.2 (3.9)	18.5 (2.2)
Refinement			
Resolution (Å)	2.02	1.70	1.80
R_{work} / R_{free} (%)	18.8/26.5	17.78 / 20.76	20.21 / 23.57
Number of atoms (protein/other/water)	1826/30/113	970 / 20 / 99	1796 / 59 / 81
B-factors (A^2) (protein/other/water)	36.24/25.55/39.24	26.39/22.55/34.40	29.69/33.56/33.37
r.m.s.d bonds (Å)	0.015	0.017	0.019
r.m.s.d angles ($^{o)}$	1.381	1.554	1.931
Ramachandran			
Favoured (%)	98.62	100.00	99.53
Allowed (%)	1.38	0.00	0.00
Disallowed (%)	0.00	0.00	0.47

* Values in parentheses correspond to the highest resolution shell.

PDB ID	5HRX	5II2	5IID
Protein	PB1(5)	PB1(5)	PB1(5)
Compound	**12**	**15**	**14**
Space group	P 2_1 2_1 2_1	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions:			
a, b, c (Å)	41.18 58.42 105.80	41.52 58.51 138.23	41.50 56.40 139.48
α, β, γ (deg)	90.00 90.00 90.00	90.00 90.00 90.00	90.00 90.00 90.00
Resolution* (Å)	29.21 (1.73)	25.00 (2.21-2.10)	19.39 (2.53-2.40)
Unique observations*	26823 (1398)	127884 (16300)	62035 (6533)
Completeness* (%)	97.8 (96.3)	99.0 (95.3)	99.3 (97.0)
Redundancy*	5.6 (5.0)	6.3 (5.9)	4.6 (3.5)
Rmerge*	0.051 (0.893)	0.128 (0.368)	0.104 (0.784)
I/σI*	17.3 (1.8)	9.1 (4.0)	10.4 (1.5)
Refinement			
Resolution (Å)	1.73	2.10	2.40
R_{work} / R_{free} (%)	21.38 / 25.37	17.6/22.7	22.1/28.1
Number of atoms (protein/other/water)	1810/44/75	1878/57/82	1894/40/49
B-factors (A^2) (protein/other/water)	32.71/29.78/36.95	28.13/22.88/33.88	45.22/45.64/41.39
r.m.s.d bonds (Å)	0.019	0.016	0.014
r.m.s.d angles ($^{o)}$	1.888	1.623	1.471
Ramachandran			
Favoured (%)	99.53	98.65	99.12
Allowed (%)	0.00	1.35	0.88
Disallowed (%)	0.47	0.00	0.00

* Values in parentheses correspond to the highest resolution shell.

Table S4. Data collection and refinement statistics.

Protein	**9**	**10**	**11**	**12**	**2**	**1**
PB1(1)	0.3 ± 0.8	0.7 ± 0.4	0.4 ± 0.4	0.8 ± 0.9	-1.4 ± 0.5	0.6 ± 0.6
PB1(2)	-0.4 ± 0.4	0.2 ± 0.1	0.1 ± 0.2	0.3 ± 0.1	0.9 ± 0.1	-0.7 ± 0.1
PB1(3)	0 ± 0.2	0.6 ± 0.2	0.4 ± 0.3	1.1 ± 0.2	1.3 ± 0	1.4 ± 0.6
PB1(4)	0.1 ± 0.3	0.3 ± 0.2	0.1 ± 0.3	0.2 ± 0.1	0.1 ± 0	-0.6 ± 0.5
PB1(5)	0.7 ± 0.3	2.8 ± 0.2	1.1 ± 0.2	2.2 ± 0.3	9.1 ± 0.1	-0.5 ± 0.8
PB1(6)	-0.1 ± 0.2	0 ± 0.2	-0.2 ± 0.4	-0.1 ± 0.2	-0.1 ± 0	-1.6 ± 0.3
SMARCA 2A	0.4 ± 0.1	0.5 ± 0.3	0.1 ± 0.2	0.7 ± 0.2	6.3 ± 0.1	0.3 ± 0.2
SMARCA 2B	-0.1 ± 0.3	0.3 ± 0.2	0.1 ± 0.2	0.5 ± 0.3	4.4 ± 0.2	0.6 ± 0.2
SMARCA 4A	0.5 ± 0.4	0.3 ± 0.2	-0.2 ± 0.3	0.5 ± 0.3	5.6 ± 0.1	-0.6 ± 0.2
TAF1(1)	-0.6	0.3	0.1	0	0 ± 0	-0.15 ± 0.1
TAF1(2)	0.03	0.3	-0.3	0.1	-0.1 ± 0.3	0.25 ± 0.1
TAF1(1:2)	0.2	-0.7	-0.6	-1.5	-0.8 ± 0.1	-1.3 ± 0
TAF1L(1)	-0.7	0.3	0.3	-0.6	0 ± 0.1	-0.2 ± 0.4
TAF1L(2)	0.4	0.5	-0.1	-1	-0.3 ± 0.5	-0.1 ± 0.1
TAF1L(1:2)	0.4	-0.8	-0.8	-0.8	-1.2 ± 0.4	-2.4 ± 0.3
BRD1	0.1	-0.3	-0.9	-0.6	0.1 ± 0.1	-0.1 ± 0.1
BRD2(1)	0	-0.8	-1	-0.9	-0.7 ± 0.2	7.1 ± 0.6
BRD4(1)	0	-0.5	-0.7	-0.4	-0.6 ± 0	8.8 ± 0
BRD4(2)	-0.1	-0.3	-0.5	0.1	0.2 ± 0	10.1 ± 0.1

Table S5. DSF assay summary. ΔT_m (°C) values are shown. The compounds were screened at 100μM while the protein was at 2.5μM. Compounds **9**-**12** were measured in triplicate on PB1 and SMARCA. **1** and **2** were measured in duplicate.

Protein – Compound	K_d (μM)	n	ΔG (kcal/mol)	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)
PB1(5) – **2**	0.042	1.2	-9.7	-7.9	-1.8
PB1(5) – **9**	11.5	1.0	-6.5	-5.6	-0.9
PB1(5) – **10**	3.4	1.2	-7.2	-6.7	-0.6
PB1(5) – **11**	3.3	0.8	-7.2	-2.9	-4.3
PB1(5) – **12**	5.1	1.1	-7.0	-6.4	-0.6
PB1(3) – **12**	No binding				
SMARCA2(2A) – **12**	No binding				
BRD4(1) – **12**	No binding				

Table S6. Isothermal titration calorimetry summary. Protein (200μM) was titrated into compound solution (15μM) at 15°C.

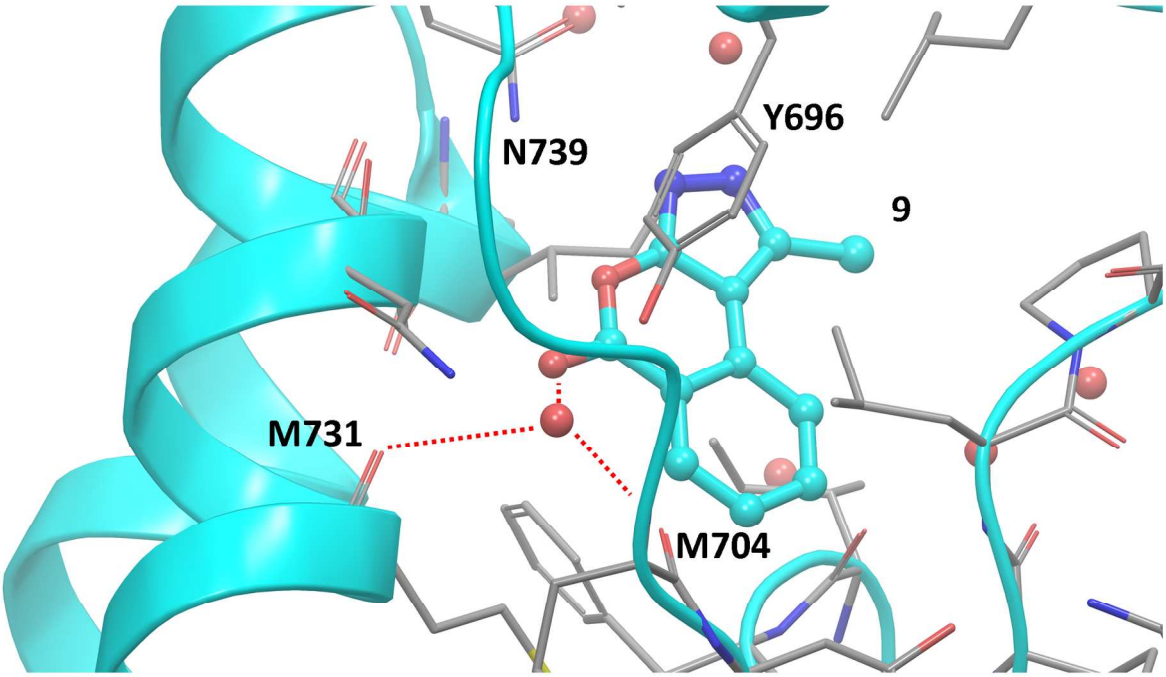


Figure S1. The buried water molecule that is not displaced upon binding of **9** in PB1(5) is stabilized by an extended network of H-bonds through surrounding residues such as the sidechain of Y696, the backbone carbonyls of M704 and M731 as well as the carbonyl of **9**.

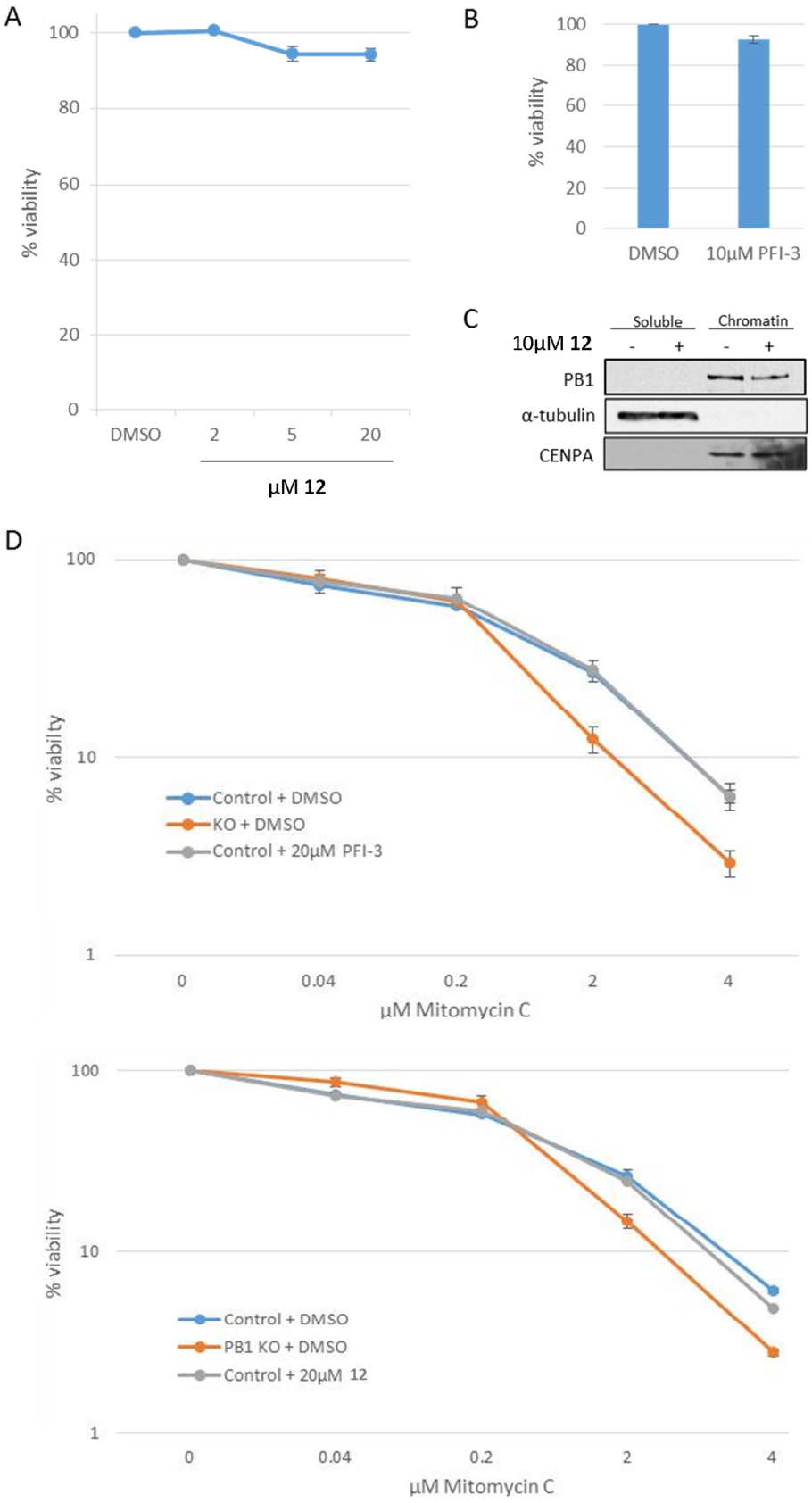

A
100
80
60
40
20
0
% viability
DMSO
2
5
20
μM 12
B
100
80
60
40
20
0
% viability
DMSO
10μM PFI-3
C
Soluble
Chromatin
10μM 12
- + - +
PB1
α-tubulin
CENPA
D
100
10
1
% viability
Control + DMSO
KO + DMSO
Control + 20μM PFI-3
0
0.04
0.2
2
4
μM Mitomycin C
100
10
1
% viability
Control + DMSO
PB1 KO + DMSO
Control + 20μM 12
0
0.04
0.2
2
4
μM Mitomycin C

Figure S2. Compound **12** slightly reduces viability of 1BR-hTERT human fibroblast cells. (A) Viability curve for 1BR-hTERT cells treated with DMSO or **12**. (B) Viability graph for cells as in (A) treated with DMSO or **2**. (C) Cells treated with DMSO or **12** were processed for chromatin fractionation as outlined in the experimental procedures. Samples were blotted for PB1, α-tubulin or CENPA as indicated. (D) Viability curve for U2OS PB1 knock-out cells treated with **2** (upper graph) or **12** (lower graph) in presence of mitomycin C (MMC).