Discovery of a selective inhibitor for the YEATS domains of ENL/AF9

Authors
Thomas Christott[1], James Bennett[1], Carmen Coxon[1], Octovia Monteiro[1], Charline Giroud[1], Viktor Beke[1], Suet Ling Felce[2], Vicki Gamble[2], Carina Gileadi[2], Gennady Poda[3,4], Rima Al-awar[3,5], Gillian Farnie[2], Oleg Fedorov[1,6]

[1] Structural Genomics Consortium, Nuffield Department of Clinical Medicine, University of Oxford, Target Discovery Institute (TDI), Oxford OX3 7FZ, UK

[2] Structural Genomics Consortium, Nuffield Department of Clinical Medicine, University of Oxford, Botnar Research Centre, Oxford OX3 7LD, UK

[3] Drug Discovery Program, Ontario Institute for Cancer Research, Toronto, ON, Canada.

[4] Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, ON, Canada.

[5] Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON, Canada.

[6] Corresponding author: **Oleg.Fedorov@sgc.ox.ac.uk**

Key Words:

YEATS domain; MLLT1; ENL; MLLT3; AF9; small molecule inhibitor

Publication details: SLAS Discovery: Advancing Life Sciences R&D

slasdisc.sagepub.com
doi.org/10.1177/2472555218809904

Abstract

Eleven-nineteen leukaemia (ENL) contains an epigenetic reader domain (YEATS domain) which recognises lysine acylation on histone 3 and facilitates transcription initiation and elongation through its interactions with the super elongation complex (SEC) and the histone methyl transferase DOT1L. Although it has been known for its role as a fusion protein with MLL in mixed lineage leukaemia, overexpression of native ENL and thus dysregulation of downstream genes in acute myeloid leukaemia (AML) has recently been implicated as a driver of disease that is reliant on the epigenetic reader activity of the YEATS domain. We developed a peptide displacement assay (histone 3 tail with acylated lysine) and screened a small molecule library totalling over 24,000 compounds for their propensity to disrupt the YEATS domain-histone peptide binding. Among these, we identified a first in class dual inhibitor of ENL (K_d = 745 ± 45 nM) and its paralog AF9 (K_d = 523 ± 53 nM) and performed "SAR by catalogue" with the aim of starting the development of a chemical probe for ENL.

Introduction

The YEATS domain (Yaf9, ENL, Af9, Taf41, Sas5) containing paralogs ENL and AF9, also known as MLLT1 and MLLT3, respectively, have long been known for their role in mixed lineage leukaemia (MLL) initiation and progression as fusion partners of MLL. The

interaction of their ANC1 homology domain (AHD) with DOT1L drives over-expression of DNA binding and HOX proteins, and the over-expression of RNA binding and ribosomal synthesis genes via the interaction of ENL/AF9 with p-TEFb[1–4].

Native ENL and AF9 are epigenetic reader and scaffold proteins in the super elongation complex (SEC) and partners of the histone methyl transferase DOT1L and are thus involved in transcription activation and elongation[5]. Even though ENL and AF9 have a high grade of sequence identity (74%), they only share 40% of target genes and show a stark difference in chromatin location. ChIPseq experiments have shown 78.2% of total ENL to be located in promoter regions, compared to only 21.3% of total AF9[5,6].

In recent years several studies have shone a light on the role of functional human YEATS domains in cancer. Native ENL appears to be required for the maintenance of MLL rearranged acute myeloid leukaemia (AML). CRISPR-Cas9 mediated knockout of ENL[6,7] as well as targeted degradation of ENL[7] via the dTAG system[8] showed a dependence of AML cell lines on the ability of ENL to recognise histone acylation for disease maintenance. It has also been shown that preventing native ENL from recognising, or lowering its affinity for acylated histone marks, is also responsible for leukogenesis[9]. Impaired binding to histone peptides is also linked to the development of Wilms tumours, a type of infant renal neoplasm[10].

While AF9 has been shown to play a role in embryonal development[11,12] and, like ENL, does bind to CBX8 (a component of the Polycomb-group repressive complex 1)[13], no compelling biology for the YEATS domain of AF9 in leukaemia has emerged to date.

In order to investigate the druggability of ENL/AF9 YEATS domain and identify starting points for drug discovery, we developed a peptide displacement assay and screened the SGC Oxford and Ontario Institute for Cancer Research (OICR) diversity set of compounds. Several genuine inhibitors were identified and followed up with purchased analogues to investigate SAR. Here, we report the identification of a first in class inhibitor of ENL/AF9 with sub-micromolar potency in peptide displacement assays as well as biophysical methods.

Materials and Methods

Protein Expression and Purification

Sequences for the wild type YEATS domains of ENL, AF9, GAS41 and YEATS2 were cloned into expression vectors (**Table S1**). *E. coli* Rosetta cells were grown in TB medium at 37°C. Overexpression was induced after five hours by addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM and cells were incubated at 17°C overnight.

Cells were harvested by centrifugation and the cell pellet was resuspended in cold binding buffer (50 mM Tris at pH 7.5, 500 mM NaCl, 20 mM imidazole, 2 mM dithiothreitol [DTT]) on a magnetic stirring plate at 4°C. The suspension was lysed using an EmulsiFlex-C5 (Avestin, Inc.) homogeniser (1 pass without pressure, four passes with pressure between 1000-1500 bar). The lysate was clarified via centrifugation at 16,000 rpm at 4°C for 1 hour (JLA-16 rotor in Avanti J-26S XP centrifuge, Beckmann Coulter, Atlanta, GA, USA). The supernatant was loaded onto a standard NiNTA column (HisTrap FF, 5 ml, GE Healthcare Lifesciences) on an ÄKTAxpress system (GE Healthcare Lifesceinces). After washing with binding buffer, the target protein was eluted at 300 mM imidazole. The eluted peak fractions were pooled and concentrated to 5 ml (Amicon concentrators, 10 kDa molecular weight cut off (MWCO)) and then separated via size exclusion chromatography (buffered with 20 mM Tris at pH 7.5, 500 mM NaCl, 2 mM DTT) using a GE Superdex 75 column on an ÄKTAxpress system. Protein concentration was quantified via extinction at 280 nM (NanoDrop ND-1000 SpectrophotometerNanoDrop Technologies, Inc, Wilmington, DE, USA) and protein identity was verified via SDS-PAGE and LC/MS.

Pooled fractions were then again concentrated to at least 2 mg/ml and stored at -80°C.

OICR 24k Diversity Set

OICR 24K Diversity Set was created from the HTS Compound Collection of 142K compounds. We have applied three sets of filters that we developed previously in BIOVIA Pipeline Pilot[14]:

1) OICR HTS Filters (developed in house) to eliminate reactive and undesirable compounds (such as acid halides, aldehydes, catechols, epoxides, haloalkyls, hydrazines, etc.), and also compounds with more than one nitro- and more than two nitrile groups.

2) PAINS Filters: Pan Assay Interference Compounds (PAINS) appear as frequent hitters in biochemical assays. Following Jonathan Baell's work[15,16], we added 480 PAINS SMARTS to our Pipeline Pilot set of filters.

3) Physical Property Filters: we have applied filters based on calculated physical properties to eliminate the undesirable compounds as well as compounds that might have issues with human intestinal absorption. These properties included molecular weight of the parent structure $300 \leq MW \leq 550$, Polar Surface Area ($6 \leq PSA \leq 250$), calculated 1-octanol-water partition coefficient, AlogP [17] ($-3 \leq AlogP \leq 6.5$), number of H-bon donors ($0 \leq HBDon \leq 10$), number of H-bond acceptors ($1 \leq HBAcc \leq 15$), number of freely rotatable bonds ($Nrot \leq 16$), number of rings ≥ 1, number of aromatic rings ≤ 7, number of atoms carrying a negative formal charge at pH 7 ≤ 2, number of atoms carrying a positive formal charge at pH 7 ≤ 2, zero metal atoms, F atom count ≤ 7, Cl atom count ≤ 4, Br atom count ≤ 1, I atom count ≤ 1, zero B atom count.

After application of the aforementioned filters, we derived a dissimilarity set of 25,000 compounds using the ECFP_4 substructural fingerprints implemented in BIOVIA Pipeline Pilot[14]. These compounds were visually inspected by experienced medicinal chemists and 411 additional compounds were removed. From the remaining set, we derived a diversity set of 24,320 compounds. The resulting OICR 24K Diversity Set of high-quality compounds was used in this screening campaign.

Peptide Displacement Assay Development

The assay was set up as a competition assay where the acylated histone tail peptide (purchased from LifeTein; **Table 1**) is displaced from the YEATS domain by an active inhibitor. Detection of YEATS-bound peptide is performed using the AlphaScreen Histidine (Nickel Chelate) Detection Kit (Perkin Elmer, 6760619M) where hexahistidyl tagged YEATS domains are bound by the Ni^{2+} chelating AlphaScreen donor beads and the biotinylated peptide is bound by the streptavidin coated acceptor beads. Upon excitation with light at a wavelength of 680 nm, the donor beads release a singlet oxygen, which triggers the emission of light between 520-620 nm by the acceptor beads. The assay is dependent on the proximity of the beads due to the half-life of the singlet oxygen. Disruption of YEATS-peptide binding causes a loss of signal. Plates were read using a Pherastar FSX plate reader (BMG Labtech) with the appropriate AlphaScreen optics module. All steps involving the AlphaScreen beads were performed under low light conditions. As AlphaScreen based assays have previously

Table 1: Protein:peptide combinations used in the peptide displacement assays

YEATS domain	Peptide shorthand	Boundaries	Peptide sequence
ENL	H3K18ac	12-30	GGKAPR(K-acetyl)QLATKAARKSAPY(K-btn)
AF9	H3K9ac	2-20	ARTKQTAR(K-acetyl)STGGKAPRKQLY(K-btn)
YEATS2	H3K27cro	15-32	btn-GKPRKQLATAAR(K-crotonyl)SAPAT
GAS41	H3K27cro	15-32	btn-GKPRKQLATAAR(K-crotonyl)SAPAT

btn: biotin

generated a very high window between no inhibition and full inhibition of binding with steep dose response curves in bromodomain inhibitor screening campaigns in our lab (data not shown), we opted for this assay technique over other technologies.

Based on recent literature[6,18–20], H3K9x, H3K18x and H3K27x peptides (where x stands for either acetylation or crotonylation) were tested and the strongest binders chosen for the assay. Assay concentrations of protein and peptide were determined via a titration of YEATS domain against peptide in a 16x16 well grid on 384 well ProxiPlates (Perkin Elmer). Protein and peptide were each titrated from 3.2 µM to 0.2 nM final assay concentration, covering all possible combinations. AlphaScreen donor-acceptor bead mix was added to a final assay concentration of 3.3 µg/ml. Plates were incubated for 1 h at room temperature (RT) before being read on the plate reader. For the final ratio of protein and peptide to use in the assay, the point representing the EC_{90} in the two-dimensional titration was chosen. Assay buffer throughout was 25 mM HEPES at pH 7.4, 100 mM NaCl, 0.1% bovine serum albumin (BSA) and 0.05% CHAPS.

Single Shot Screen

Compounds were dispensed into 384 well ProxiPlates (Perkin Elmer) for a final assay concentration of 50 µM in two replicates. On each plate, two columns were reserved for controls (DMSO addition, no addition). Protein and detection reagents were added in the concentrations determined above and plates were incubated for 30 min with protein/peptide and compounds alone before the addition of detection reagent. Signals were determined for each replicate and expressed as per-cent change relative to the mean DMSO control signal.

Determination of IC_{50} Values

Single shot "hits" to be carried forward for IC_{50} determination were re-purchased and solubilised to 50 mM stocks in DMSO. They were dispensed as an 11-point curve in two-fold serial dilution steps (top concentration 100 µM) in duplicate with the difference in dispensing volume between each step backfilled with DMSO to keep the final DMSO concentration consistent at a maximum of 0.5 %. DMSO only (n = 16) and no-addition (n = 16) reference wells were also added on each plate. The assays were performed as described as above and change in signal was also calculated as described above. The inflection point (IC_{50}) of the dose response curve was calculated in GraphPad Prism 7 using a sigmoidal dose response curve with a variable slope:

$$Y(c) = Y_{bottom} + (Y_{top}-Y_{bottom})/(1+10\text{^}(\text{LogIC}_{50}-c)-H))$$

where Y is the per-cent inhibition, Y_{top} and Y_{bottom} are the top (total loss of signal) and bottom (no loss of signal) of the curve, respectively, c is the compound concentration and H is the Hill slope.

At this stage, a counter screen with a biotinylated hexahistidyl peptide instead of the YEATS domain/histone peptide pair was also performed to eliminate compounds that simply interfere with the assay chemistry.

Thermal Shift Assay

Thermal melting experiments were carried out using a Lightcycler480 (Roche), running a temperature gradient from 20 to 95°C with 3 acquisitions per cycle (0.19°C/s). Compounds were dispensed into 384 well PCR plates (20 µl assay volume) at 50 µM in duplicate in three independent experiments. DMSO was backfilled for a final DMSO content of 0.1% (v/v). Measurements from DMSO only and no-addition controls (n = 16 each per assay plate) were also collected. Protein was added at a final concentration of 10 µM buffered with 20 mM HEPES at pH 7.5 and 500 mM NaCl. SYPRO orange (Life Technologies) was used as fluorescent dye at a 1:1000 dilution from purchased stock solution. The midpoints of the unfolding process were estimated using a sigmoidal Boltzmann equation:

$$F(T) = F_{native} + (F_{unfolded} - F_{native})/(1 + e\text{^}((T_{50} - T)/S))$$

where T_{50} is the temperature at the midpoint of unfolding, $F_{unfolded}$ and F_{native} are the

fluorescence intensities of the dye in presence of native or fully unfolded protein, and *S* is the slope of the curve.

The shift in unfolding was then calculated as the difference in temperature between the midpoints of unfolding for each compound sample to the mean DMSO reference.

Peptide affinity

The affinities of peptides for the individual YEATS domains were determined using an Octet RED384 system (FortéBio). Peptides (**Table 1**) were immobilised onto 8 streptavidin coated sensors (SSA) each to saturation of the sensor. An additional set of 8 "free" reference sensors was used. Proteins were titrated in 1:3 dilutions from 100 µM down to 137 nM (7 data points plus one "blank") in 25 mM HEPES, pH 7.5, 150 mM NaCl, 0.05 % (v/v) TWEEN-20. Each peptide was dipped into each protein (baseline in buffer 60 s, association in protein 120 s, dissociation in buffer 120 s). Data were analysed using the FortéBio Data Analysis software (version 9.0) supplied with the instrument. The data from the reference sensograms were subtracted from the peptide bound sensograms to correct for unspecific binding. Report points for each sensor were exported to GraphPad PRISM and fitted to a one site binding model:

$$Y(X) = (B_{max} * X)/(K_d + X)$$

where *Y* is the response, B_{max} the maximum response and *X* the concentration of protein.

Isothermal Titration Calorimetry

Protein was prepared for ITC by dialysis against a ~2000 times excess of ITC buffer (20 mM Tris at pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, 2 mM DTT) using SnakeSkin® Dialysis Tubing (Thermo Scientific) with a 7 kDa MWCO and then concentrated to 300 µM with DMSO added to 0.1 % (v/v) to minimise buffer mismatch between cell and syringe. Compounds were diluted to 50 µM in dialysis buffer. ITC was performed on a NanoITC Standard Volume device (TA Instruments) using reverse titration (compound in cell, protein in titration syringe) at 20°C with an initial injection of 3.7 µl and 30 injections of 7.96 µl at a stir rate of 350 rpm. Data were analysed with the NanoAnalyze software (TA Instruments) using an independent fit model.

NanoLuciferase Bioluminescent Resonance Energy Transfer (NanoBRET) Assay

Cellular activity against AF9 was assessed using a NanoBRET assay[21]. HEK293 cell (8×10^5) were plated in each well of a 6-well plate after 6h cells were co-transfected with C-terminal HaloTag-Histone 3.3 (NM_002107) and an N-terminal NanoLuciferase fusion of AF9 (original AF9 WT sequences from Promega HaloTag® human ORF in pFN21A and AF9 MUT - Y78A Tyrosine is changed to an Alanine) at a 1:10 (NanoLuc® to HaloTag®) ratio respectively with FuGENE HD transfection reagent.

Sixteen hours post-transfection, cells were collected, washed with PBS, and exchanged into media containing phenol red-free DMEM and 4% FBS in the absence (control sample) or the presence (experimental sample) of 100 nM NanoBRET 618 fluorescent ligand (Promega). Cells were then re-plated in a 96-well assay white plate (Corning Costar #3917) at $2x10^4$ cells per well. Compounds were then added directly to media (in the presence of SAHA 2.5 µM) at final concentrations 0-10 µM or an equivalent amount of DMSO as a vehicle control, and the plates were incubated for 24 h at 37°C in the presence of 5% CO_2.

NanoBRET Nano-Glo substrate (Promega) was added to both control and experimental samples at a final concentration of 10 µM. Readings were performed within 10 minutes using a ClarioSTAR (BMG Labtech) equipped with 460 nm and 610 nm filters. A corrected BRET ratio was calculated and is defined as the ratio of the emission at 610 nm/460 nm for experimental samples minus the emission at 610 nm/460 nm for control samples (without NanoBRET fluorescent ligand). BRET ratios are expressed as milliBRET units (mBU), where 1 mBU corresponds to the corrected BRET ratio multiplied by 1000.

Table 2: Final assay concentrations (FAC) of YEATS domains and histone peptides

YEATS domain	FAC [nM]	Histone peptide	FAC [nM]
ENL	50	H3K18ac	100
AF9	100	H3K9ac	100
YEATS2	50	H3K27cro	25
GAS41	50	H3K27cro	12.5

Results and Discussion

Protein production

After initially expressing ENL/AF9 and YEATS2 in TEv-cleavable N-terminally

hexahistidyl tagged form, switching to non-cleavable C-terminally tagged protein greatly increased yield for ENL/F9 (~4 fold from ~10 mg per 12 l culture), allowing for easier conduction of high protein consumption assays such as ITC. No notable change was observed for YEATS2 (yielding ~10 mg per 12 l culture) whereas GAS41 could only be expressed in sufficient quantities for any assay in C-terminally tagged form (~50 mg per 12 l culture). Differences in behaviour in the assays between the N- and C-terminally tagged proteins could not be observed. The affinities of the produced proteins to histone peptides as determined by biolayer interferometry were in good agreement with the literature data obtained by ITC[6,18–20] (**Figure S1**). Importantly, affinities of the H3K9ac and H3K918ac peptides to AF9 were significantly higher than to ENL.

Initial Screening Campaign and Biophysical Characterisation of Hits

We established assay conditions for all four human YEATS domains, resulting in a robust assay with a substantial assay window and a Z' = 0.94 (**Figure S2**) with the outer wells of the plates excluded due to edge effects. The effect of DMSO on the assay appears to be not significant, with the signal from the DMSO-only-addition wells (n = 16 per plate each) being close to 100 % of the signal from no-addition wells for all four proteins (**Table S2**). The final assay concentrations of protein and peptide are shown in **Table 2**.

In the first instance, we screened two libraries from the Ontario Institute of Cancer Research (OICR) as well as internal libraries of acetyl lysine mimetic compounds against the YEATS domain of ENL at a single concentration (50 µM) and obtained a large number of hits which included a sizable portion of known pan-assay inhibiting substances (PAINS). After eliminating compounds that interfere with the assay chemistry, such as obvious metal chelators, we followed up 16 compounds with full dose response curves for all four YEATS domains to experimentally determine potency and selectivity and obtained IC_{50} values for 7 compounds (**Figure 1**). After re-purchasing the compounds, we repeated the IC_{50} experiments (**Table 3**, number of replicates each in **Table S3**) and performed thermal melting experiments with these compounds against all four YEATS domains and only one compound (XS018661) showed selectivity of ENL/AF9 (**Table 3**, original hits denoted with "h"). While the compound caused a larger shift in thermal unfolding for AF9 and a lower K_d value, it produced a higher IC_{50} compared to ENL (**Figure 2B-D**), reflecting the energetic cost of overcoming the higher affinity of AF9 for histone peptides compared to ENL[5,7,22] under the same the same assay conditions (ionic strength, pH, present detergent).

For this compound, we then determined a K_d value of 754 ± 26 nM for ENL and K_d value of 523 ± 53 nM for AF9 via ITC in three independent experiments (**Figure 2E,F**). To further investigate the selectivity of the compound over other acetyllysine readers, we tested this compound in a dose response experiment against a number of bromodomains (BRD4(first domain), BRD9, CECR2, CREBBP, FALZ, TAF1 and BAZ2B) benchmarked against the known pan-

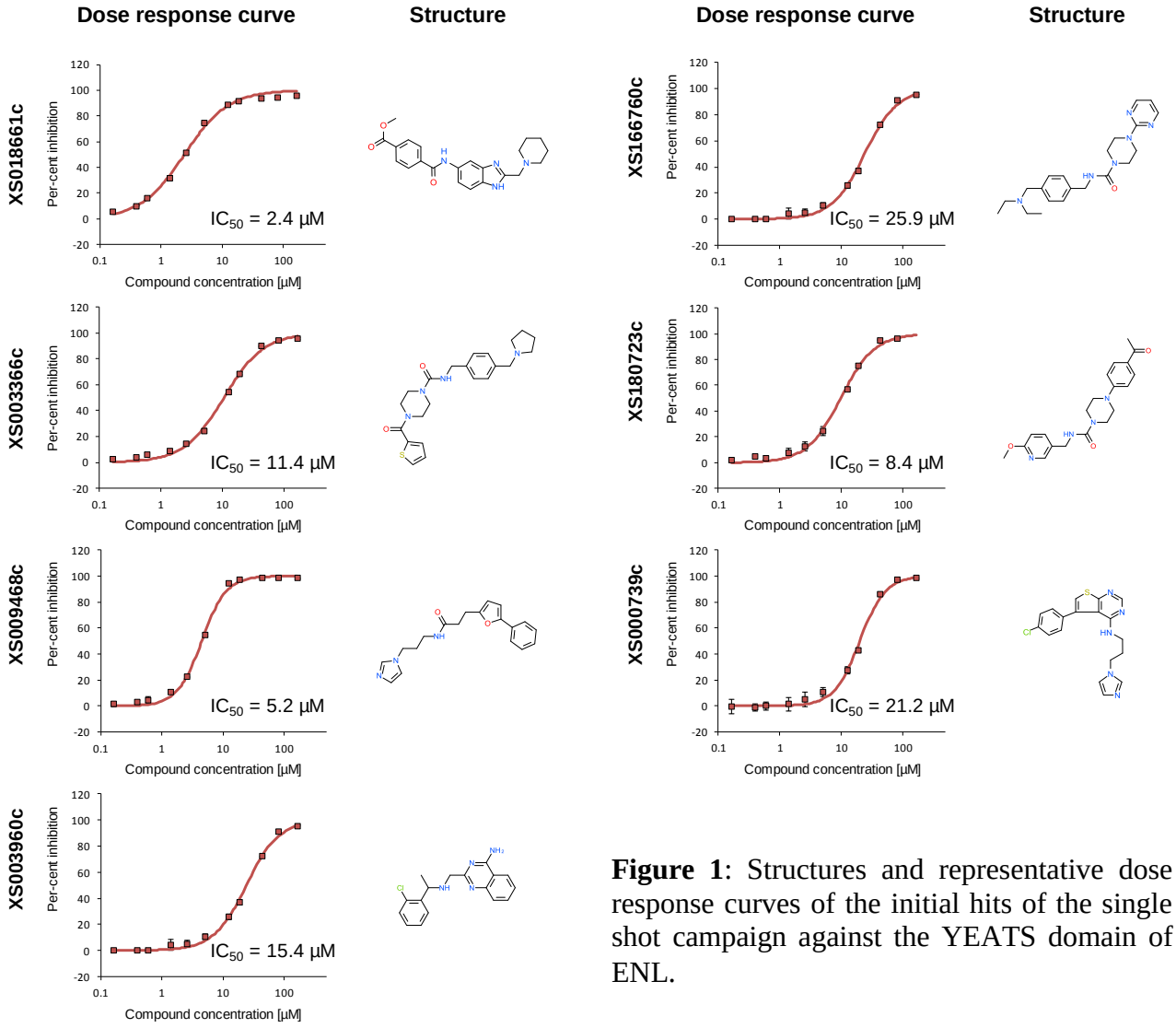


Figure 1: Structures and representative dose response curves of the initial hits of the single shot campaign against the YEATS domain of ENL.

bromodomain inhibitor bromosporine[23] (**Figure S3**). The compound showed no significant inhibition below 20 µM for any of the domains tested, while bromosporine showed IC_{50} values of 2.4 µM for BRD4(1), 29.4 nM for CECR2, 14.0 µM for CREBBP, 22.4 nM for TAF1 and 155.9 nM for BRD9, all in good agreement with previous experiments carried out in our lab for other screening campaigns (data not shown).

SAR by catalogue

We purchased a number of additional compounds to investigate the SAR of the scaffold of XS018661. However, none of the purchased compounds showed inhibition of protein-peptide-binding or a shift in thermal unfolding greater than the originator (denoted with "p" in **Table 3**, **Figure S4**) so that we were unable to glean any initial SAR information from these compounds.

Intracellular activity

We were able to demonstrate moderate cellular target engagement of XS018661 against AF9 in the micromolar range (**Figure 3**). A mutant of

AF9 with a deficiency in the binding pocket (Y78A) was not significantly affected by the inhibitor. The effect of SAHA on the binding of histone peptides to YEATS domains was tested in a dose response curve in the AlphaScreen assay (top concentration 50 µM) and no inhibition was observed.

Table 3: IC_{50} [μM] and ΔT_m [°C] values of original library hits tested.

	ENL		AF9		YEATS2		GAS41		
Compound	IC_{50}	ΔT_m	IC_{50}	ΔT_m	IC_{50}	ΔT_m	IC_{50}	ΔT_m	Note
Bromosporine	>100	0.1 ± 0.2	>100	0.1 ± 0.2	>100	0.1 ± 0.2	>100	0.3 ± 0.2	c
XS018661	1.6 ± 0.9	1.9 ± 0.6	3 ± 3.1	3.7 ± 0.9	75.5 ± 73.2	0.4 ± 0.2	>100	0.1 ± 0.1	h
XS003366	5.9 ± 5.6	0.5 ± 0.2	33 ± 39	0.7 ± 0.4	>100	0.2 ± 0.2	>100	0.2 ± 0	h
XS009468	5 ± 2.2	0.2 ± 0.4	13.4 ± 7	0.2 ± 0.2	7.2 ± 0.3	0.2 ± 0.2	12.1 ± 10	0 ± 0.2	h
XS003960	9.7 ± 5.1	0.1 ± 0.3	21.5 ± 11.9	0.2 ± 0.2	13.9 ± 8.8	0.3 ± 0.2	17.8 ± 3.5	0.1 ± 0.1	h
XS009757	13.6 ± 2.4	0.2 ± 0.2	>100	0.1 ± 0.3	>100	0 ± 0.3	>100	0 ± 0.2	h
XS166760	25.7 ± 0.3	0.6 ± 0.4	>100	0.5 ± 0.2	>100	0.2 ± 0.2	>100	-0.1 ± 0.1	h
XS180723	5.5 ± 3.8	0.3 ± 0.2	>100	0.4 ± 0.5	>100	0.5 ± 0.5	>100	-0.1 ± 0	h
XS000739	18.9 ± 10.5	0.2 ± 0.2	40.7 ± 11.1	0 ± 0.4	22.9 ± 1.4	0.1 ± 0.5	33.9 ± 20.3	0 ± 0.2	h
XS043798	6.7 ± 3.1	0.8 ± 0.3	13.7 ± 12	1.5 ± 0.3	>100	0.3 ± 0.3	>100	-0.1 ± 0.1	p
XS043798	6.7 ± 3.1	0.8 ± 0.3	13.7 ± 12	1.5 ± 0.3	>100	0.3 ± 0.3	>100	-0.1 ± 0.1	p
XS096172	77.4 ± 60.7	0.3 ± 0.3	33.7 ± 8.7	0.5 ± 0.2	>100	0.2 ± 0.2	>100	0 ± 0.1	p
XS098176	68.5 ± 38.5	0.2 ± 0.2	>100	0.2 ± 0.2	>100	0.4 ± 0.2	>100	0.1 ± 0.1	p
XS102315	15.8 ± 6.1	0.3 ± 0.1	26.5 ± 22.3	0.4 ± 0.4	56.3 ± 38.1	0.2 ± 0.2	>100	0.1 ± 0.2	p
XS102728	24.3 ± 8.5	0.2 ± 0.3	53.8 ± 3.5	0 ± 0.1	17.5 ± 1.5	0.3 ± 0.3	59.2	0 ± 0.1	p
XS171208	10.2 ± 4.7	0.1 ± 0.2	34.7 ± 11.2	-0.1 ± 0.1	8.1 ± 1.9	0 ± 0.3	45.6 ± 2.3	-0.1 ± 0.2	p
YT000270	21.1 ± 6.4	0.1 ± 0.2	42.7 ± 14.4	0.6 ± 0.7	43.4 ± 30.5	0.3 ± 0.4	37.1	0.4 ± 0.6	p
YT000272	30.5 ± 18.8	0 ± 0	70.1	0.1 ± 0.2	28.2 ± 8.1	-0.2 ± 0.2	30.4	0.3 ± 0.2	p
YT000289	19.6 ± 7.3	0.4 ± 0.2	>100	0.3 ± 0	>100	0 ± 0.5	>100	0.4 ± 0.1	p
YT000290	26.1 ± 7.7	0.1 ± 0.2	>100	0.5 ± 0.2	>100	0.5 ± 0.2	>100	0.9 ± 0.6	p
YT000291	19.1 ± 6.5	0.3 ± 0	>100	0.7 ± 0.2	>100	0.3 ± 0	>100	0.3 ± 0.2	p
YT000294	>100	0.4 ± 0.5	>100	0.5 ± 0.2	>100	1.1 ± 0.6	>100	0.7 ± 0.3	p
YT000295	70.4 ± 86.7	0.5 ± 0.6	>100	0.4 ± 0.3	>100	0.7 ± 0.3	57.7 ± 0.2	0.6 ± 0.2	p
YT000309	>100	0.8 ± 0.4	>100	0.3 ± 0	>100	0.5 ± 0.5	>100	0.6 ± 0.2	p
YT000330	0.8 ± 0.5	0 ± 0	3.5 ± 0.1	0.1 ± 0.2	1.6	0 ± 0	3.71	0.3 ± 0.2	p
YT000333	46.3 ± 17.7	0.2 ± 0.2	>100	0 ± 0	>100	0 ± 0	>100	0.3 ± 0.2	p

Compounds without standard error were only tested once. c: Negative control compound; h: Original hit from library; p: Purchased follow up compound

Summary and Conclusions

Here we report the discovery of a potent first in class inhibitor, XS018661, of the YEATS domains of ENL (IC_{50} = 1.6 ± 0.9 μM, K_d = 754 ± 26 nM) and AF9 (IC_{50} = 3.0 ± 3.1 μM, K_d = 523 ± 53 nM) through a library screen that already displays moderate cellular activity. The identified scaffold will serve as the template for the development of a chemical probe. Knockout of AF9 has shown no major impact on cell viability and invasiveness in any AML cell line to date[6,7]. This, taken together with the lack of biology of AF9 in cancer and the absence of AF9 on promoter regions relative to ENL should make even a dual inhibitor of ENL and AF9 a useful tool compound. Indeed, CRISPR-Cas9 mediated knockout of AF9 was used as a control experiment by Wan et al[6] showed no adverse effect on cell viability in a number of

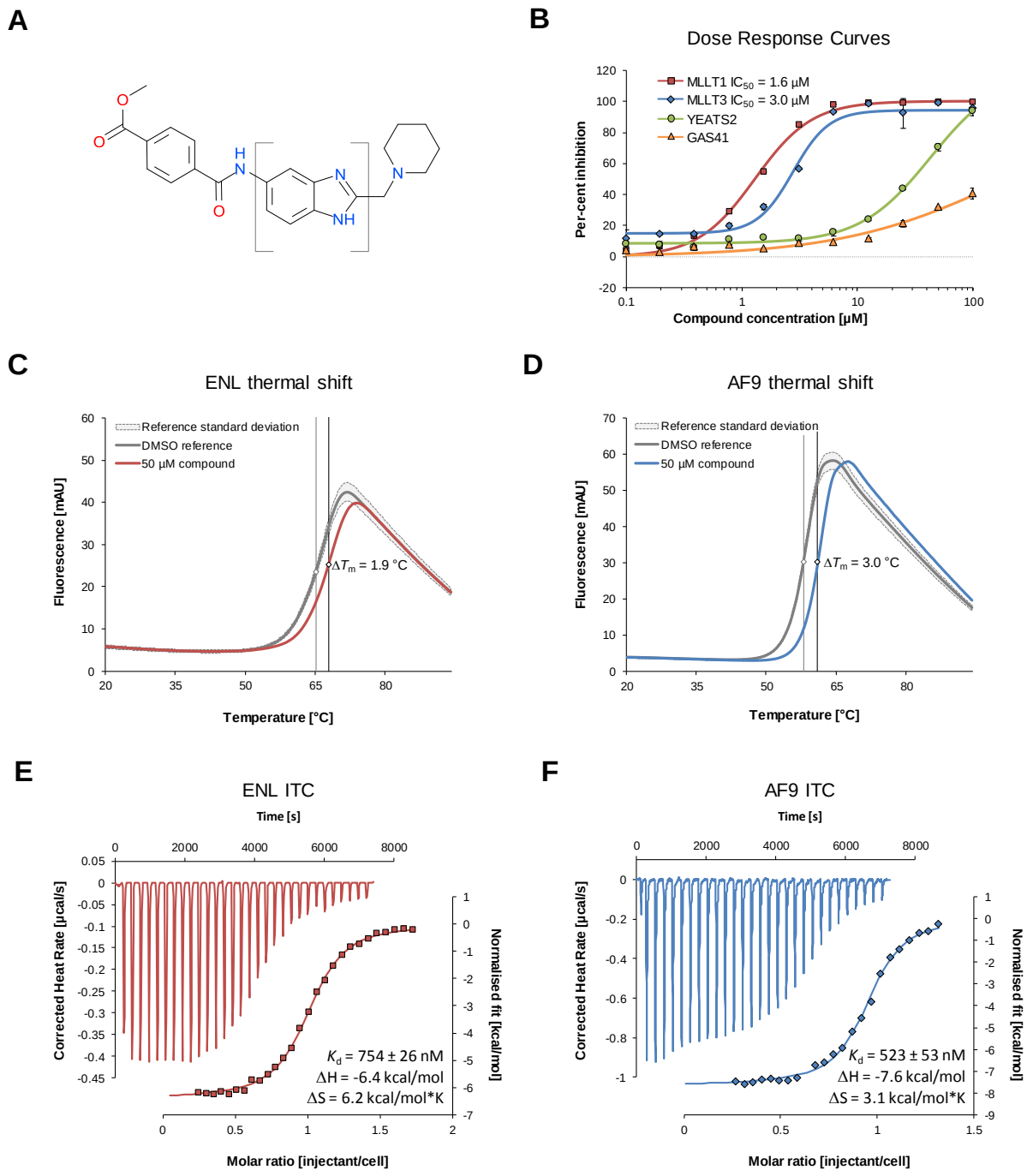


Figure 2: Biophysical characterisation of most potent compound. **A** Structure of XS018661. The core scaffold around which analogues were purchased later is marked in square brackets. **B** Dose response curves of XS018661 against all four human YEATS domains. **C, D** Thermal shift profile of XS018661 against the YEATS domains of ENL and AF9, respectively. Inflection points of the curves are marked. **E, F** Representative ITC measurement of the binding of XS018661 to ENL and AF9, respectively.

MLL rearranged leukaemia cell lines as well as HeLa and U2Os cell lines.

While the biology of ENL/AF9 as fusion partners of MLL has been intensively investigated in mixed lineage/acute leukaemia, there is no consensus about the role of the native protein in leukaemia, yet. It has recently been shown that binding of ENL to PAF1 might be responsible for leukaemogenesis in non-MLL-ENL rearranged leukaemia[9], slightly contradicting the evidence for the role of its

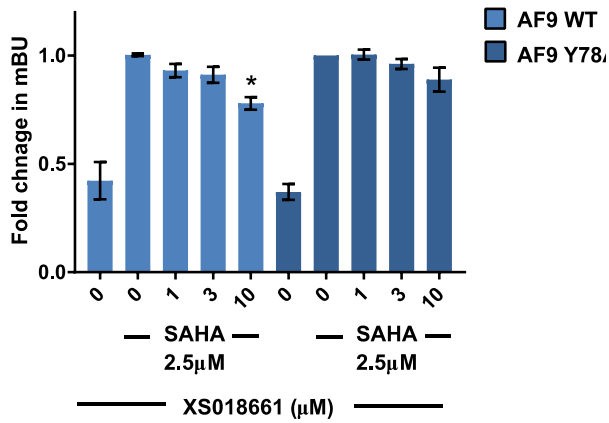


Figure 3: NanoBRET assay with wild type and mutant AF9. Bar chart represents fold change in mBU NanoBRET showing the displacement of N-terminal-nanoLuc-WT AF9 (AF9 WT) or N-terminal-nanoLuc-MUT AF9 (AF9 Y78A) from C-terminal HaloTag-H3.3 after 24h treatment with XS018661 (1-10 µM) in the presence of 2.5 µM SAHA. Graph represent n=3 independent biological replicates, with n=4 technical replicates. Mean±SD, mBU – BRET units. One-way anova AF9 WT $P \leq 0.003$; post-hoc Dunnett's multiple comparisons test (using SAHA 2.5µM control from AF9 WT or Y78A respectively), * P=0.0073, AF9 Y78A Not significantly different.

canonical activity of ENL in the SEC and with DOT1L in AML[6,7]. There is also evidence that ENL/AF9 are involved in trans

criptional repression in double strand repair processes. ENL/AF9 is phosphorylated by ATM and can then interact with the polycomb group complex PRC1 whose E3-ubiquitin ligase RING1B ubiquitinates H2A, thereby stalling transcription until the double strand break has been repaired[24]. Based on this information, one could conceivably envision that inhibition of ENL/AF9 function might sensitise cancer cells with impaired double strand break repair machinery to the conventional chemotherapy. Furthermore, all studies investigating the role of ENL in AML so far have relied on techniques involving the complete abolishment of ENL from the cellular context and/or ectopic expression of native or mutant protein.

In light of this seemingly contradictory background, a potent small molecule inhibitor will enable probing the biology of ENL and AF9 without perturbing the interaction of the protein with its partners *in vivo* other than the recognition of histone marks. The compound we present here will make an excellent starting point for the further development of highly active and selective inhibitors for the YEATS domains of ENL and AF9.

Acknowledgements

The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada, Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, Merck KGaA Darmstadt Germany, MSD, Novartis Pharma AG, Ontario Ministry of Economic Development and Innovation, Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and Wellcome [106169/ZZ14/Z].

References

1. Mueller D, Bach C, Zeisig D, et al. A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. *Blood*. **2007**, *110*, 4445-4454

2. Mueller D, García-Cuéllar MP, Bach C, et al. Misguided transcriptional elongation causes mixed lineage leukemia. *PLoS Biol*. **2009**, 7

3. Yokoyama A, Lin M, Naresh A, et al. A Higher-Order Complex Containing AF4 and ENL Family Proteins with P-TEFb Facilitates Oncogenic and Physiologic MLL-Dependent Transcription. *Cancer Cell*. **2010** *17*, 198-212

4. Biswas D, Milne TA, Basrur V, et al. Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes. *Proc Natl Acad Sci*. **2011**, *108*, 15751-15756.

5. Li Y, Wen H, Xi Y, et al. AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation. *Cell*. **2014**, *159*, 558-571

6. Wan L, Wen H, Li Y, et al. ENL links histone acetylation to oncogenic gene expression in acute myeloid leukaemia. *Nature*. **2017**, 543, 265-269

7. Erb MA, Scott TG, Li BE, et al. Transcription control by the ENL YEATS domain in acute leukaemia. *Nature*. **2017**, 543, 270-274

8. Nabet B, Roberts JM, Buckley DL, et al. The dTAG system for immediate and target-specific protein degradation. *Nat Chem Biol*. **2018**, *14*, 431-441

9. Hetzner K, Garcia-Cuellar M-P, Büttner C, et al. The interaction of ENL with PAF1 mitigates polycomb silencing and facilitates murine leukemogenesis. *Blood*. **2017**, *131*, 662-373

10. Perlman EJ, Gadd S, Arold ST, et al. MLLT1 YEATS domain mutations in clinically distinctive Favourable Histology Wilms tumours. *Nat Commun*. **2015**, *6*, 10013.

11. Collins EC, Appert A, Ariza-McNaughton L, et al. Mouse Af9 Is a Controller of Embryo Patterning, Like Mll, Whose Human Homologue Fuses with AF9 after Chromosomal Translocation in Leukemia. *Mol Cell Biol*. **2002**, *22*, 7313-7324

12. Qiao Y, Wang X, Wang R, et al. AF9 promotes hESC neural differentiation through recruiting TET2 to neurodevelopmental gene loci for methylcytosine hydroxylation. *Cell Discov*. **2015**, *1*, 1-19

13. Malik B, Hemenway CS. CBX8, a component of the Polycomb PRC1 complex, modulates DOT1L-mediated gene expression through AF9/MLLT3. *FEBS Lett*. **2013**, *587*, 3038-3044

14. Dassault Systèmes BIOVIA. BIOVIA Pipeline Pilot. http://accelrys.com/products/collaborative-science/biovia-pipeline-pilot/.

15. Baell JB, Holloway GA. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem*. **2010**, *53*, 2719-2740

16. Saubern S, Guha R, Baell JB. KNIME workflow to assess PAINS filters in SMARTS format. Comparison of RDKit and indigo cheminformatics libraries. *Mol Inform*. **2011**, *30*, 847-850

17. Ghose AK, Viswanadhan VN, Wendoloski JJ. Prediction of hydrophobic (lipophilic) properties of small organic molecules using fragmental methods: An analysis of ALOGP and CLOGP methods. *J Phys Chem A*. **1998**, *102*, 3762-3772

18. Li Y, Wen H, Xi Y, et al. AF9 YEATS Domain Links Histone Acetylation to DOT1L-Mediated H3K79 Methylation. *Cell*. **2014**, *159*, 558-571

19. Zhao D, Guan H, Zhao S, et al. YEATS2 is a selective histone crotonylation reader. *Cell Res*. **2016**, *26*, 629-632

20. Hsu C, Shi J, Yuan C, et al. Recognition of histone acetylation by the GAS41 YEATS domain promotes H2A.Z deposition in non-small cell lung cancer. *Genes Dev*. **2018**, *32*, 58-69

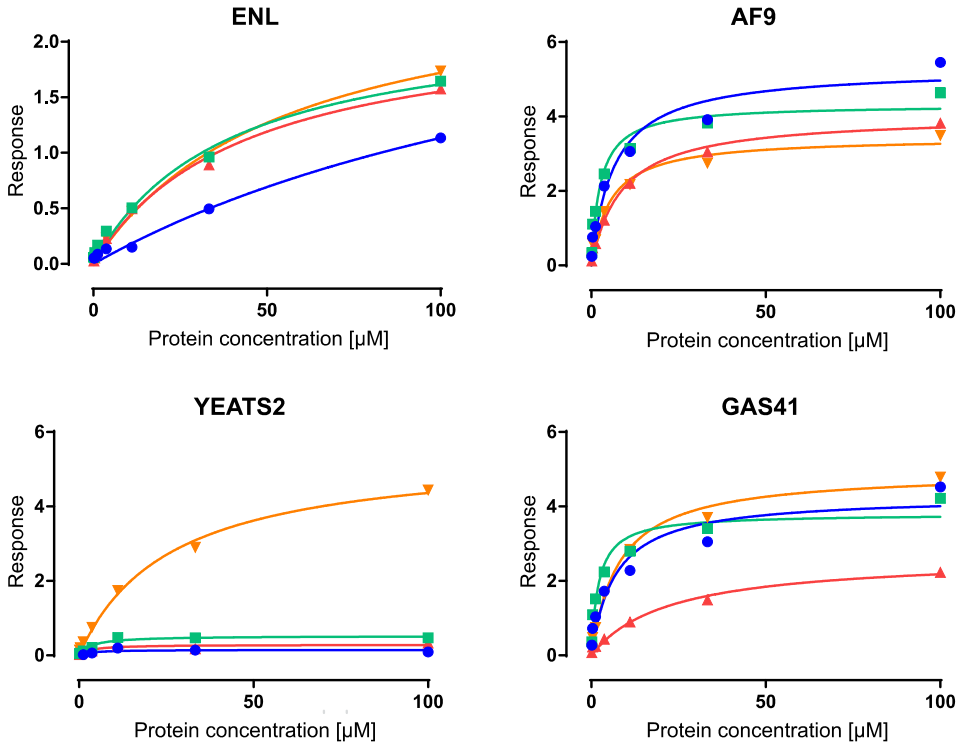
21. Machleidt T, Woodroofe CC, Schwinn MK, et al. NanoBRET-A Novel BRET Platform for the Analysis of Protein-Protein Interactions. *ACS Chem Biol*. **2015**, *10*, 1797-1804

22. Li Y, Sabari BR, Panchenko T, et al. Molecular Coupling of Histone Crotonylation and Active Transcription by AF9 YEATS Domain. *Mol Cell*. **2016**, *62*, 181-193

23. Picaud S, Leonards K, Lambert JP, et al. Promiscuous targeting of bromodomains by bromosporine identifies BET proteins as master regulators of primary transcription response in leukemia. *Sci Adv*. **2016**, *2*, 15-20

24. Ui A, Nagaura Y, Yasui A. Transcriptional elongation factor ENL phosphorylated by ATM recruits polycomb and switches off transcription for DSB repair. *Mol Cell*. **2015**, *58*, 468-482

Supplementary Material



K_d in µM

	H3K9ac	H3K9cro	H3K18ac	H3K27cro
MLLT1	90.9	10.0	8.1	9.1
MLLT3	2.0	0.7	1.4	1.3
YEATS2	N/D	N/D	N/D	3.6
YEATS4	2.9	0.7	4.9	1.3

Figure S1: Affinity of the YEATS domains of ENL, AF9, YEATS2 and GAS41 for histone peptides determined with biolayer interferometry.

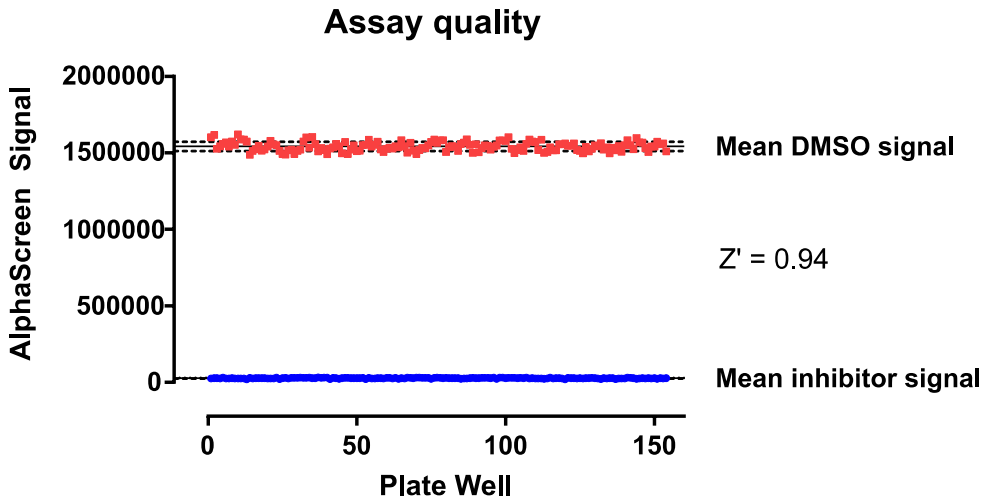


Figure S2: AlphaScreen assay quality.

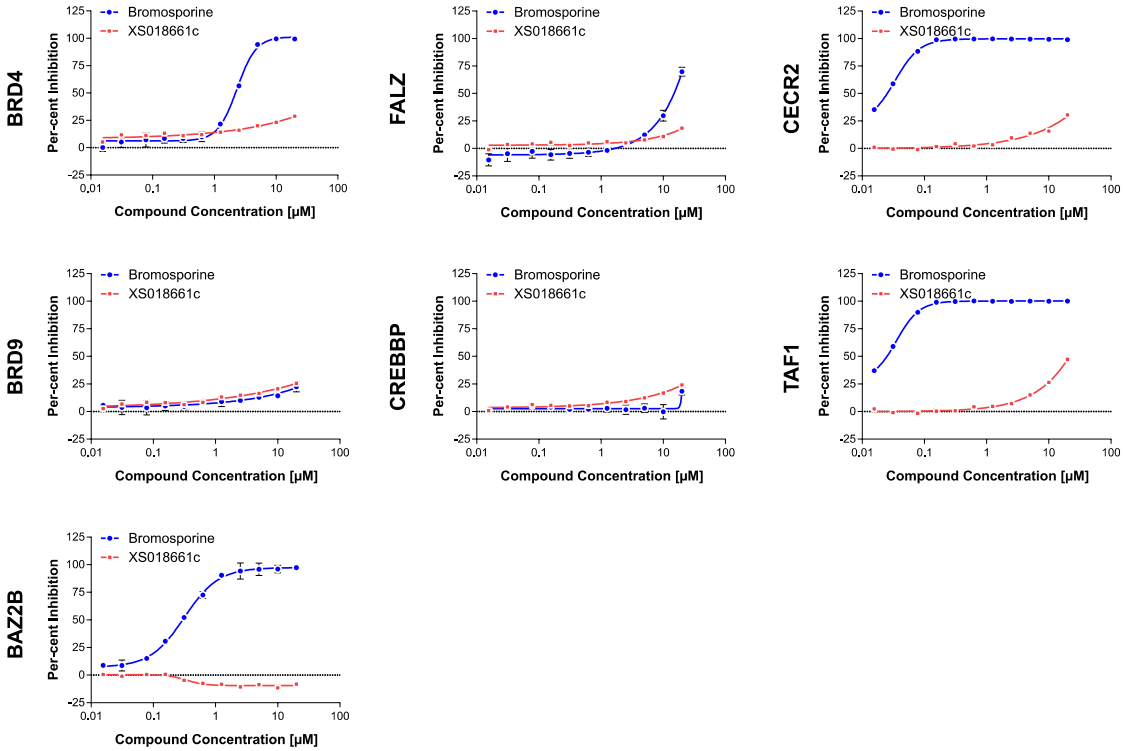


Figure S3: Dose response curves for XS018661 against selected bromodomains

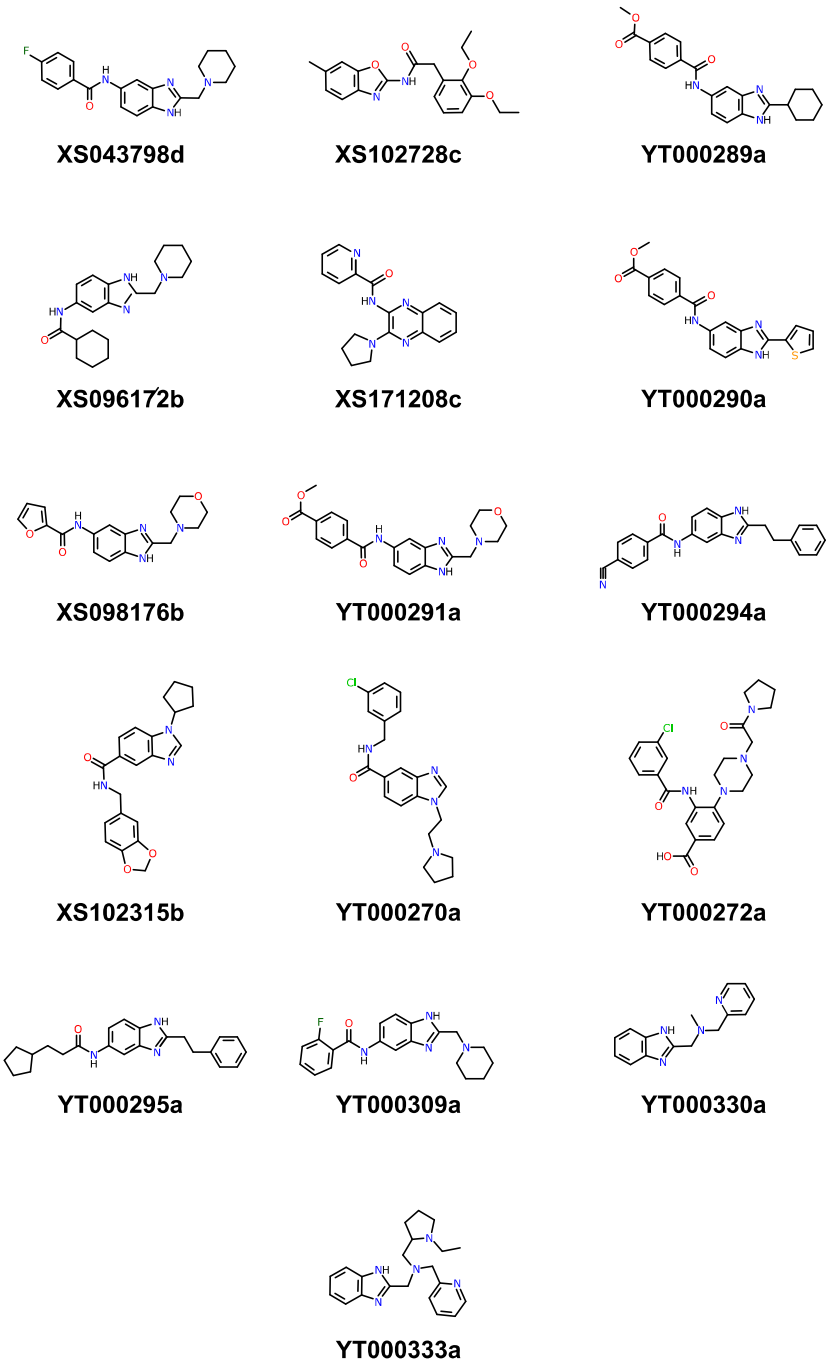


Figure S4: Structures of purchased follow up compounds

Table S1: Constructs and vectors used

Protein	Vector	Start	End	Hexahistidyl tag	Use
ENL	pET28a	D2	M144	N-terminal	AS
	pNIC-CH	M1	A148	C-terminal	AS, ITC, TM
AF9	pNIC-Bsa4	M1	D141	N-terminal	AS
	pNIC-CH	M1	D141	C-terminal	AS, ITC, TM
YEATS2	pNIC-Bsa4	S202	E345	N-terminal	AS, TM
GAS41	pNIC-CH	V16	K225	C-terminal	AS, TM

AS: AlphaScreen

ITC: Isothermal Titration Calorimetry

TM: Thermal shift assays

Table S2: DMSO effect on AlphaScreen signal in peptide displacement assay. Signal of no-compound-addition wells ("buffer") expressed as per-cent signal of DMSO-only-addition wells (0.5 % v/v) is shown alongside the number of well pairs and number of corresponding plates.

Protein	DMSO relative to buffer (%)	Well pairs	Assay plates
ENL	96.1 ± 8.4	206	26
AF9	98.1 ± 5.3	204	26
YEATS2	94.6 ± 6.2	208	26
GAS41	97.0 ± 5.8	207	26

Table S3: Number of replicate IC_{50} experiments

	Replicate experiments			
Compound	ENL	AF9	YEATS2	GAS41
Bromosporine	1	1	1	1
XS018661	9	4	5	8
XS003366	4	3	2	2
XS009468	5	3	2	2
XS003960	4	3	2	2
XS009757	3	3	2	2
XS166760	2	1	2	2
XS180723	4	3	2	2
XS000739	4	3	2	2
XS043798	8	3	5	5
XS096172	4	2	1	1
XS098176	4	3	2	2
XS102315	4	3	2	1
XS102728	4	3	2	1
XS171208	4	3	2	2
YT000270	4	3	2	1
YT000272	2	1	2	1
YT000289	4	3	2	2
YT000290	4	3	2	2
YT000291	4	3	2	2
YT000294	4	3	2	2
YT000295	4	3	1	2
YT000309	6	3	4	4
YT000330	5	2	1	1
YT000333	4	2	1	1