

Fragment-based discovery of a chemical probe for the PWWP1 domain of NSD3

Jark Böttcher^{1*}, David Dilworth^{2#}, Ulrich Reiser^{1#}, Ralph A. Neumüller^{1#}, Michael Schleicher¹, Mark Petronczki¹, Markus Zeeb³, Nikolai Mischerikow¹, Abdellah Allali-Hassani², Magdalena M. Szewczyk², Fengling Li², Steven Kennedy², Masoud Vedadi^{2,4}, Dalia Barsyte-Lovejoy², Peter J. Brown², Kilian V. M. Huber^{5,6}, Catherine M. Rogers^{5,6}, Carrow I. Wells⁷, Oleg Fedorov^{5,6}, Klaus Rumpel¹, Andreas Zoephel¹, Moriz Mayer¹, Tobias Wunberg¹, Dietrich Böse¹, Stephan Zahn¹, Heribert Arnhof¹, Helmut Berger¹, Christoph Reiser¹, Alexandra Hörmann¹, Teresa Krammer¹, Maja Corcokovic¹, Bernadette Sharps¹, Sandra Winkler¹, Daniela Häring¹, Xiao-Ling Cockcroft¹, Julian E. Fuchs¹, Barbara Müllauer¹, Alexander Weiss-Puxbaum¹, Thomas Gerstberger¹, Guido Boehmelt¹, Christopher R. Vakoc⁸, Cheryl H. Arrowsmith^{2,9}, Mark Pearson¹, Darryl B. McConnell¹

¹ Boehringer Ingelheim RCV GmbH & Co KG, Vienna, Austria; ² Structural Genomics Consortium, University of Toronto, Toronto, Ontario, M5G 1L7, Canada; ³ Boehringer Ingelheim Pharma GmbH & Co KG, Biberach, Germany; ⁴ Department of Pharmacology and Toxicology, University of Toronto, Toronto, Ontario, M5S 1A8, Canada; ⁵ Structural Genomics Consortium, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom; ⁶ Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom; ⁷ Structural Genomics Consortium, Division of Chemical Biology and Medicinal Chemistry, UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA; ⁸ Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA; ⁹ Princess Margaret Cancer Centre and Department of Medical Biophysics, University of Toronto, Toronto, Ontario, M5G 2M9, Canada.

#equally contributing authors

*corresponding author, email: jark.boettcher@boehringer-ingelheim.com

Abstract:

Here we report the fragment-based discovery of BI-9321 (**17**), a potent, selective, and cellular active antagonist of the NSD3-PWWP1 domain. The human NSD3 protein is encoded by the *WHSC1L1* gene located in the 8p11-p12 amplicon, frequently amplified in breast and squamous lung cancer. Recently it was demonstrated that the PWWP1 domain of NSD3 is required for the viability of Acute Myeloid Leukemia cells. To further elucidate the relevance of NSD3 in cancer biology, we developed a chemical probe BI-9321 (**17**) targeting the methyl-lysine binding site of the PWWP1 domain with sub-micromolar *in vitro* activity and cellular target engagement at 1 μ M. As a single agent BI-9321 (**17**) downregulates Myc mRNA expression and reduces proliferation in MOLM-13 cells. This first-in-class chemical probe BI-9321 (**17**), together with the negative control BI-9466 (**12**), will greatly facilitate the elucidation of the underexplored biological function of PWWP domains.

Introduction

Post-translational modification (PTM) of histones by histone-modifying enzymes, often referred to as writers and erasers, represents a key element of epigenetic regulation¹. Acetylation, phosphorylation and methylation constitute the most frequent of the many reported histone modifications². These marks are recognized by reader domains that can be part of multi-domain histone-modifying enzymes. Epigenetic reader domains are structurally diverse and classified into Bromodomains, PHD domains, Methyl-lysine- and/or Methyl-arginine-binding domains, frequently co-locating within the same proteins³. By interfering with the function of specific domains, selective chemical probes represent powerful tools to dissect the biological functions of these multi-valent epigenetic regulators^{4,5}.

Although several PWWP (Pro-Trp-Trp-Pro) domain-containing proteins have been reported in context of human diseases, such as DNMT3A and the NSD (Nuclear Receptor Binding SET Domain) family of protein lysine methyltransferases^{6,7}, no PWWP domain-targeting chemical probes have been reported so far. NSD3 (encoded by *WHSC1L1*) is one such protein, suggested to exhibit a tumour-promoting function in a variety of cancers⁸. NSD3 is a multi-domain protein that exists in three isoforms (long, short, and the testis-specific Whistle). Both NSD3 long and Whistle isoforms contain the SET domain, with lysine methyltransferase activity, as well as several chromatin reader domains, including PHD and PWWP domains. The NSD3-short isoform contains only the first PWWP domain and a recent study indicated that it serves as an adaptor protein, which couples BRD4 to the chromatin remodeller CHD8⁹. Importantly, this function is dependent on the PWWP1 domain and required for the viability of Acute Myeloid Leukaemia (AML) cells¹⁰. Besides its function in AML, NSD3 resides in the 8p11-p12 amplicon¹¹, which is

frequently amplified in different tumour types, including 12% of breast cancer and 13% of squamous lung cancer (The Cancer Genome Atlas; TCGA).

Here we describe the discovery of molecules binding to the proposed methyl-lysine binding site of the PWWP1 domain of NSD3 using fragment-based methods. The hits we optimized into the cellular active chemical probe BI-9321 (**17**) using virtual screening and subsequent structure-based optimization. We demonstrate the dislocation of NSD3 from chromatin upon addition of the probe. Moreover, we provide evidence that blocking the PWWP1 domain with BI-9321 (**17**) downregulates Myc mRNA expression in MOLM-13 AML cells and reduces proliferation. By providing access to the chemical probe BI-9321 (**17**) and the negative control BI-9466 (**12**), we provide well characterized tools to elucidate the biological function of the NSD3-PWWP1 domain.

RESULTS

Fragment-Based Screening (FBS) of the PWWP1 domain of NSD3

In order to identify binders to the PWWP1 domain of NSD3, we screened a proprietary fragment library of 1899 compounds against a recombinant construct comprising amino acids 247-398 of NSD3 (**Supplementary Table 1**). Saturation Transfer Difference - Nuclear Magnetic Resonance (STD-NMR) spectroscopy and Differential Scanning Fluorimetry (DSF) served as primary screening techniques. 285 hits showed significant STD effects and 20 hits generated a stabilisation of $>1^{\circ}\text{C}$, in DSF, whereas only three fragments were identified by both techniques. Subsequent hit confirmation was performed through observation of cross peak shifts in 2D $^1\text{H}/^{15}\text{N}$ -TROSY NMR (Transverse Relaxation Optimized Spectroscopy) spectra (**Fig. 1a**). Hits initially found with both techniques could not be confirmed. To discriminate between various potential binding sites we mapped the methyl-lysine binding site by observing a specific chemical

shift perturbation pattern after the addition of a dimethyl histone H3K36 peptide (H-STGGV-Lys(me2)-KPHRY-OH), as weak binding was described previously (**Supplementary Fig. 1a**)¹². Finally, 15 out of 302 hits showed chemical shift perturbation of cross peaks that were also perturbed by the peptide, suggesting that the fragments bind to a similar region of the protein and provide an opportunity for competition with histone peptide binding. This resulted in a low overall hit rate of 0.8%. All 15 confirmed FBS hits were analysed in 2D ¹⁵N TROSY NMR detected titration experiments. Fragments exhibiting an NMR K_d of < 2mM entered co-crystallisation and soaking experiments and only fragments with a binding mode confirmed by protein crystallography were pursued further (**Supplementary Table 2, Supplementary Fig. 1b,c, 2 and 3**).

The crystal structure of the PWWP1 domain of NSD3 (residues 247-398) was solved by experimental phasing using a seleno-methionine labelled variant. The structure revealed a classical PWWP domain fold, as described previously, with an *N*-terminal β -barrel formed of 5 antiparallel β -strands (β 1- β 5), with a short helix insertion between β 4 and β 5, followed by 3 α -helices¹³ (**Fig. 1b, Supplementary Fig. 4**). Based on the observation that *N*-terminal residues were disordered, we designed a new protein construct for x-ray crystallography containing amino acids 263-398. This construct yielded crystals exclusively in the presence of the weak fragment **1** (NMR; K_d = 1700 μ M) (**Supplementary Table 3**) bound to a hydrophobic surface patch comprising the amino acid side chains of Phe313; Leu370, Tyr391 and Phe392. Whereas the carboxylic acid of **1** forms only a water-mediated hydrogen bond to His374, it is involved in charged interactions to Lys277; Lys358 and Arg360 of a symmetry-related molecule forming an attractive crystal packing interface (**Supplementary Fig. 5a**). Soaking of these crystals further enabled the determination of crystal structures of two additional fragments, **2** (NMR; K_d = 330

μM) and **3** (NMR; $K_d = 650 \mu\text{M}$) (**Fig. 1c, d, Supplementary Table 3, Supplementary Fig. 5b,c**).

The two latter fragments **2** and **3** bind to the methyl-lysine binding site formed by the aromatic cage, which is conserved in the majority of PWWP domains¹³. In the case of the PWWP1 domain of NSD3, it is formed by the aromatic amino acids Trp284, Tyr281, and Phe312. Superposition of the fragment structures with the structure of the BRPF1-PWWP domain complexed with a H3K36me3 peptide (PDB ID: 2X4Y) suggests the potential for interference with methyl-lysine binding (**Fig. 1e**)¹⁴. The pyrazole ring of **2** is positioned between the aromatic rings of Tyr281 and Phe312 with a centroid distance of 4.0 and 4.1 Å and a minimal contact of 3.3 and 3.4 Å at the bottom of the pocket respectively, as a result of a 30° deviation from ring to ring planarity. The 3,5-dimethyl-1,2-oxazole moiety is rotated 50° out of plane and the methyl groups fill the remaining space close to Tyr281 and Phe312 (**Fig. 1c**). The pyridine ring addresses a lipophilic plane formed by Trp284 and forms a hydrogen bond to Ser314 with a distance of 2.8 Å. In the case of **3** the core is addressed by the methyl-dihydropyridazine ring exhibiting a more favourable orientation for π - π interactions with Tyr281 and Phe312, when compared to the structure of **2**, with a deviation from planarity of 14° and 19° and a centroid distance of 3.7 and 3.9 Å respectively (**Fig. 1d**). In addition, the methyl substituent is forming attractive CH- π interactions to Trp284 with a distance of 3.4 Å. The phenyl substituent addresses the Trp284 plane and the nitrogen atom of the piperidine substituent forms a direct hydrogen bond to Glu318 with a distance of 3.0 Å.

Based on the observed interactions, **4** was designed in order to introduce the pyridine substituent from **2** into the methyl-dihydropyridazine scaffold to address Ser314. The design concept could be confirmed by protein crystallography and the newly introduced hydrogen bond to Ser314 (2.8

Å) resulted in a significant increase in potency (SPR; $K_d = 47 \mu\text{M}$) (**Supplementary Table 3 and Supplementary Fig. 6a**). **4** was subsequently used as a positive control for SPR measurements and the biotinylated analog **5** (NMR; $K_d = 21 \mu\text{M}$; SPR; $K_d = 177 \mu\text{M}$) (**Supplementary Table 3**) used for the development of the TR-FRET (Time-Resolved Fluorescence Energy Transfer) assay.

The two validated FBS hits **2** and **3** were taken as query templates for SAR by catalogue^{15,16}, using a framework for ligand-based virtual screening developed in-house, Morhits¹⁷. 601 virtual screening hits from the Boehringer Ingelheim compound collection were subjected to SPR single point screening and 108 compounds were identified for K_d determination follow-up. The x-ray structure of the most potent hit **6** (SPR; $K_d = 160 \mu\text{M}$), possessing a methylimidazole core, revealed binding to the same binding site as the query templates and was used as a new starting point for hit optimization (**Fig. 2a, Supplementary Fig. 6b, Supplementary Table 4**).

Structure-based Optimization of Methylimidazoles as NSD3-PWWP1 Antagonists

The structure of the PWWP1 domain in complex with **6** revealed that the methyl group of the imidazole core forms a CH- π interaction with Trp284, with a distance of 3.4 Å. In addition, the *N*-methyl group leads to a shift of the five-membered ring 0.7 Å away from Trp284, when compared to the pyrazole ring of **2**, enabling a more favourable π - π interaction with Tyr281 and Phe312 (deviation from planarity of 16° and 19° and a centroid distance of 3.8 and 3.8 Å respectively). In order to evaluate if this interaction results in improved potency we merged the 3,5-dimethyl-1,2-oxazole moiety of **2** into **6**. The resulting compound **7** (SPR; $K_d = 22 \mu\text{M}$; TR-FRET; $\text{IC}_{50} = 13 \mu\text{M}$), confirmed our hypothesis and encouraged us to continue the structure-based optimization of the methylimidazole class (**Fig. 2b, Supplementary Fig. 6c**). We focused on the two substituents in 4- and 5-position of the methylimidazole core (**Fig. 2c, d, Supplementary Table 4**) using the developed TR-FRET assay as the profiling assay. We tested

2-methyl- and 2,6-dimethyl-phenyl substituents in the 4 position of the methylimidazole core which in contrast to the dimethyl-1,2-oxazole substituent offer the possibility to introduce additional substituents. Both modifications led to moderately improved potency (**8** = 7.6 μ M; **9** = 11 μ M, (TR-FRET; IC₅₀)), when compared to the 4,5-dimethyl-1,2-oxazole analogue **7** (TR-FRET; IC₅₀ = 13 μ M). We then introduced a methylamine substituent in the *para* position to form a charged interaction with Glu318. Both **10** (TR-FRET; IC₅₀ = 3.1 μ M) and **11** (TR-FRET; IC₅₀ = 5.2 μ M) exhibited improved potency, but the anticipated interaction between the amino group and Glu318 was not formed and the amino functionality points into the solvent (**Fig. 2c**, **Supplementary Fig. 7a**). Based on the improved affinity with highly conserved binding modes within the series and the observation that the pyridine substituent in 5-position remained undistorted with a ring-ring angle deviation of 2°, comparing the structures of **7** and **11**, we decided to improve the substituents in the 5-position separately and to subsequently combine the best substituents. The initial fragment SAR guided us toward the hydrogen bond formed by the pyridine nitrogen in the para-position to Ser314 as a favourable interaction; therefore we decided to maintain this pharmacophore during optimization. The introduction of a more basic nitrogen at this position in case of BI-9466 (**12**) (TR-FRET; IC₅₀ = 120 μ M) induced a significant loss in potency. Therefore, we selected the quinoline as a promising substituent to decrease the basicity of the ring nitrogen¹⁸, while, in parallel, filling the lipophilic space close to Trp284. The quinoline-substituted derivative **13** (TR-FRET; IC₅₀ = 3.9 μ M) exhibited improved potency when compared to its pyridine analogue **7** (TR-FRET; IC₅₀ = 13 μ M). This effect is also confirmed by the observed affinity difference of the 7-methoxy quinoline derivative **14** (TR-FRET; IC₅₀ = 21 μ M) with the 7-fluoro quinoline derivative **15** (TR-FRET; IC₅₀ = 1.9 μ M). The co-crystal structure of **15** (**Fig. 2d**, **Supplementary Fig. 7b**) revealed some significant differences to the previous structures. In order to accommodate the large quinoline moiety Tyr281 is rotated 21°

when compared to the structure of the pyridine analogue **6**. This leads to a favourable CH- π interaction of the quinoline CH with the π -system of Tyr281 with a carbon-to-ring centroid distance of 3.7 Å. In addition the side chain of Gln367 is displaced, leading to collapse of a complex hydrogen bond network frequently formed by the amino acids Asn315, Gln316 and Gln367. Besides these changes, the relative orientation of the aromatic planes of methylimidazole substituents in the 4- and 5-position is conserved between both structures (deviation from planarity of 4° and 16° respectively), implying an opportunity to combine the optimal substituents in one structure. The 3-methyl- and 3,5-dimethyl-benzylamine substituents in the 4-position were merged with the 7-fluoro-quinoline derivative in the 5-position yielding **16** (TR-FRET; IC₅₀ = 403 ± 35 nM) and BI-9321 (**17**) (TR-FRET; IC₅₀ = 203 ± 11 nM). The structure of NSD3-PWWP1 in complex with BI-9321 (**17**) revealed that the orientation of the 7-fluoro-quinoline is conserved when compared to the structure of **15** (deviation of 2°) (**Fig. 2e**, **Supplementary Fig. 7c**). The 3,5-dimethyl-benzylamine substituent instead is rotated an additional 11° out of plane addressing space released by the previously described conformation of Tyr281. The distribution of angles between ring planes in o,o-dimethyl-substituted biaryl systems in small molecule crystal structures reveals angles between 60 and 90° to be the most abundant. Additionally, quantum mechanical calculations on ω B97XD/cc-pVDZ level (unrestrained energy minimization and systematic scan of biaryl torsion) identify this region of conformational space as most favoured for BI-9321 (**17**) and **15**. Of note, the placement of the benzene ring in BI-9321 (**17**) allows for charge-assisted hydrogen bonding of the methylene-ammonium substituent with Glu318. Based on the significantly increased potency we decided to profile BI-9321 (**17**) and validate it as a suitable probe for studying the biological function of the PWWP1 domain of NSD3 in a cellular context. The close and significantly less potent analogue

BI-9466 (**12**) (TR-FRET; $IC_{50} = 120 \mu M$) was selected as a negative control to exclude non-target related effects (**Fig 2e**, **Supplementary Fig. 8a**)

Biophysical characterization of the NSD3-PWWP1 probe BI-9321 (17**)**

The binding affinity of BI-9321 (**17**) was confirmed by SPR ($K_d = 166 \pm 3 \text{ nM}$; **Fig. 3a**) and isothermal titration calorimetry (ITC) ($K_d = 445 \pm 8 \text{ nM}$; **Fig. 3b**). In order to rule out the contribution of inhibiting other PWWP domains on the cellular readout, the probe was profiled against 14 other purified PWWP domains using DSF (**Fig. 3c**). Whereas BI-9321 (**17**) leads to a significant stabilisation of $5.5 \text{ }^{\circ}\text{C}$ at $100 \mu M$ in the case of NSD3-PWWP1, no stabilisation $\geq 2 \text{ }^{\circ}\text{C}$ was observed for the other domains, suggesting selectivity. The closest family members NSD2-PWWP1 and NSD3-PWWP2 were additionally tested by ITC K_d determination and binding was not detected (**Supplementary Fig. 8b**). Both BI-9321 (**17**) and BI-9466 (**12**) were tested at a concentration of $500 \mu M$ in a diverse set of methyl-lysine binding domains with DSLS and no significant T_{agg} shift was observed (**Supplementary Table 5**). Inhibition of 35 protein, DNA and RNA methyltransferases was tested and no liabilities at $10 \mu M$ were observed for BI-9321 (**17**) and BI-9466 (**12**) (**Supplementary Fig. 9a, b**). No kinase inhibition was detected in a selection of 31 diverse kinases for BI-9321 (**17**) at $10 \mu M$ (**Supplementary Table 6**). The *in vitro* properties of the chemical probe BI-9321 (**17**) are summarized in **Supplementary Table 7**. Moreover, selectivity profiling by quantitative chemical proteomics revealed NSD3 as main interactor of immobilized BI-9321 (**17**) in high-salt chromatin extracts of Cal-120 cells, which express high levels of NSD3 due to an amplification of the locus (**Fig. 3d**; **Supplementary Datasheet 1**). In addition previously reported NSD3 interactors, KDM1B and BRD4^{10,19}, were identified in the experiment and confirmed by western blotting (**Supplementary Fig. 10a**). No specific interactors of BI-9321 (**17**) were identified at $10 \mu M$ in low-salt whole-cell extract

(**Supplementary Datasheet 2**). In order to exclude a direct interaction of BI-9321 (**17**) with KDM1B and BRD4, co-precipitation experiments were performed with an NSD3 antibody. We observed a significant overlap of enriched proteins in both chemical- and immuno-precipitated materials, in particular, BRD4 and KDM1B, suggesting the interactions observed are mediated by NSD3 and not BI-9321 (**17**). Additionally, the presence of 10 μ M BI-9321 (**17**) did not significantly affect enrichment of co-immunoprecipitated proteins (**Fig. 3e; Supplementary Datasheet 3**), indicating these interactions are not dependent on the PWWP domains' putative methyl-lysine binding pocket. The inhibition of KDM1B and antagonism of BRD4(1) bromodomain was also tested in enzymatic and AlphaLisa assays respectively, revealing no significant inhibition (KDM1B $IC_{50} > 100 \mu$ M; BRD4(1) $IC_{50} = > 1$ mM (**Supplementary Fig. 10b, c**). BI-9321 (**17**) was tested in a broad panel of 48 bromodomains using a thermal shift assay revealing no binding, indicated by no stabilisation $> 2^{\circ}C$ (**Supplementary Table 8**). BI-9321 (**17**) has attractive *in vitro* ADME properties, namely high solubility at pH 6.8, low plasma protein binding, high absorptive permeability, high *in vitro* metabolic stability and profiling of cytochrome P450 inhibition revealed strong inhibition only of the 2C19 isoform ($IC_{50} < 0.2 \mu$ M) (**Supplementary Table 7**). The overall properties confirmed the suitability of BI-9321 (**17**) as a chemical tool to probe the function of NSD3-PWWP1 in cellular experiments.

Cellular target engagement and effects of the chemical probe BI-9321 (17)

To test if BI-9321 (**17**) enters cells and engages its target, we measured bioluminescence resonance energy transfer (BRET) between a transfected Nluc-NSD3-PWWP1 fusion protein (donor) and tracer derivative **18** (acceptor) in U2OS cells (**Synthetic procedures**). We observed a specific and dose-dependent BRET signal ($EC_{50} = 5.5 \pm 0.1 \mu$ M; **Fig. 4a**) that is displaced by BI-9321 (**17**) ($IC_{50} = 1.4 \pm 0.5 \mu$ M; **Fig. 4b**)²⁰. No significant effect was observed with the

negative control BI-9466 (**12**), and no compound toxicity was observed in U2OS cells after 24 hours treatment. To further confirm binding of BI-9321 (**17**) to the PWWP1 domain of NSD3 in cells and demonstrate selectivity over NSD3's closest homolog, NSD2, we next attempted cellular thermal shift assays (CETSA). Despite several attempts, we were unsuccessful in measuring a thermal shift of the endogenous NSD3; an observation in-line with known limitations of the assay for measuring target engagement of large multi-domain chromatin-associated proteins²¹. As an alternative, we performed a NanoLuciferase thermal shift assay (NaLTSA) with the PWWP1 domains of NSD3 and NSD2²². We observed a shift in the thermal stability of NSD3-PWWP1, but not NSD2-PWWP1, in response to 10 μ M BI-9321 (**17**), confirming cellular target engagement and demonstrating the selectivity of BI-9321 (**17**) in living cells (**Supplementary Fig. 11a, b**). The negative control BI-9466 (**12**), did not affect either fusion protein.

The recruitment of epigenetic proteins to chromatin is often mediated by context-dependent multi-valent interactions. Having demonstrated robust target engagement in cells, we next sought to establish the utility of the probe for interrogating PWWP1-dependent interactions of the oncogenic short isoform of NSD3 with chromatin using FRAP (Fluorescence Recovery After Photobleaching) (**Fig. 4c**)²³. U2OS cells transfected with GFP-tagged NSD3-short variants were photobleached using a high-intensity laser pulse and the migration of labelled proteins back into the bleached area were monitored over time. To benchmark the effect of μ M BI-9321 (**17**), we included NSD3-short Trp284Ala and Glu318Lys mutants harbouring inactivating mutations in the PWWP1 domain¹⁰. Both μ M BI-9321 (**17**) treatment (1 μ M) and mutant constructs significantly enhanced fluorescence recovery, further demonstrating target engagement and exemplifying the value of μ M BI-9321 (**17**) as a tool to study PWWP1-dependent recruitment of NSD3-short to chromatin. The reduction in FRAP recovery time was also consistent with

observations in cell fractionation experiments, revealing elevated nucleoplasmic levels of NSD3-short in μM BI-9321 (**17**) treated cells, and those expressing the Glu318Lys mutation (**Fig. 4d**, **Supplementary Fig. 12a**). To quantify PWWP1 antagonism by μM BI-9321 (**17**), but not BI-9466 (**12**), the NanoBRET assay was used to measure disruption of the protein-protein interaction between C-terminally NanoLuc® tagged NSD3-PWWP1 with C-terminally HaloTag®Fusion tagged histone H3 in a dose-dependent manner ($\text{IC}_{50} = 1.2 \pm 0.06 \mu\text{M}$; **Fig. 4e**). The BRET signal of NSD3-PWWP1 double mutant W284A/E318K, as well as NSD2-PWWP1, was not affected by BI-9321 (**17**), indicating that BI-9321 (**17**) specifically disrupts histone interactions of the PWWP1 domain of NSD3 (**Supplementary Fig. 12b**). Together, these results demonstrate the value of BI-9321 (**17**) as a tool to study PWWP1-dependent interactions of NSD3 in cells.

The PWWP1 domain of NSD3 was recently shown to be required for the proliferation of RN-2 cells, a cell line derived from a MLL-AF9/NrasG12D AML mouse model¹⁰. Therefore, we were next interested in evaluating the therapeutic potential of NSD3-PWWP1 antagonism. We first used a domain based CRISPR approach to assess the potential consequence of NSD3-PWWP1 antagonism in the MOLM-13 human MLL-AF9 translocated AML cell line. gRNAs targeting the PWWP1 domain confirmed the requirement for the NSD3-short isoform to sustain proliferation, whereas gRNAs targeting the SET domain, or a control region adjacent to NSD3, did not result in a continued decrease of CRISPR positive cells over time (**Fig. 5a**, **Supplementary Table 9**). Consistent with genetic experiments, the addition of increasing concentrations of BI-9321 (**17**) inhibited proliferation of both MOLM-13 ($\text{IC}_{50} = 26.8 \pm 4.4 \mu\text{M}$) and RN-2 cells ($\text{IC}_{50} = 13 \pm 2 \mu\text{M}$), with no significant effect observed in cells treated with the negative control BI-9466 (**12**) (**Fig. 5b**) (Sequence identity of mouse and human NSD3 short; methyl-lysine binding site = 100%; overall = 97.1%). No effects on proliferation were observed for the AML cell lines HL-60

and MV4;11, suggesting the role of NSD3's PWWP1 domain in the maintenance of AML may be limited in the context of some cell backgrounds.

NSD3 promotes proliferation of RN-2 AML cells by coupling BRD4 to the chromatin remodeler CHD8 to enforce oncogenic transcription, including transcription of the major oncogene MYC¹⁰. Therefore, we next quantified the mRNA level of MYC upon treatment with BI-9321 (**17**) to evaluate changes in its gene expression. Consistent with observed antiproliferative effects, BI-9321 (**17**) (10 μ M) induced a ~30% reduction of MYC mRNA levels in MOLM-13 cells (**Fig. 5c**). This effect was transient, reaching a maximum down-regulation at 6 hours after compound addition. The stability of BI-9321 (**17**) under the cell culture conditions was confirmed with quantitative tandem mass spectroscopy ensuring constant compound concentrations during the experiment (**Supplementary Fig. 13**). Given NSD3's role in regulating bromodomain biology, we next asked, if BI-9321 (**17**) could potentiate the effects of BRD4 antagonism. When tested in combination treatment experiments, BI-9321 (**17**) potentiated the JQ1-dependent proliferation phenotype in the MOLM-13 cell line (**Fig. 5d**). These data suggest that inhibition of the NSD3-PWWP1 as a single agent or in combination with BRD4 antagonists may be a strategy for the treatment of NSD3-dependent tumours.

DISCUSSION

In this study, we describe the development of the chemical probe BI-9321 (**17**) for the PWWP1 domain of NSD3. This first-in-class chemical probe reveals the tractability of PWWP domain modulation by small molecules. In absence of a protein structure and any known ligands at the outset, we applied fragment-based screening and structure-based optimization to discover BI-9321 (**17**). Assays for various potency stages of the compounds and guidance by x-ray crystal structures were the basis for efficient hit optimization, yielding a 1000-fold potency improvement.

With BI-9321 (**17**) we provide a well-characterized small molecule fulfilling state of the art requirements of a chemical probe (**Supplementary Table 7**)²⁴. The *in vitro* potency of 203 ± 11 nM (TR-FRET; IC₅₀) was confirmed with two orthogonal methods SPR ($K_d = 166 \pm 3$ nM) and ITC ($K_d 445 \pm 8$ nM). No significant off-targets could be identified *in vitro* within the PWWP protein family, methyl-lysine binding domains, methyltransferases or kinases. This observation was confirmed with quantitative chemical proteomics. Cellular target engagement was shown at around 1 μ M with separate methods, FRAP, cellular fractionation and nanoBRETTM methods. The provision of the negative control BI-9466 (**12**) enables the stringent analysis of cellular effects, such as PWWP1-dependent protein-protein interactions and the regulation of MYC transcription in MOLM-13 cells. Consistent with the downregulation of MYC, BI-9321 (**17**) reduced proliferation in MOLM-13 and RN-2 cells that are sensitive to genetic loss-of-function of NSD3. This effect was not observed in other cell lines such as HL-60 and MV4;11, excluding a general cytotoxic effect. Thus, BI-9321 (**17**) will enable the scientific community to determine effects on cells upon NSD3-PWWP1 inhibition. Our data in MOLM-13 cells argue that pharmacological inhibition of the NSD3-PWWP1 domain affects cell proliferation as a single agent and could be a therapeutic strategy in AML. In addition, BI-9321 (**17**) potentiates the effects of **JQ1** in the MOLM-13 cell line, suggesting that NSD3-PWWP1 domain antagonism in combination with other targeted therapies might be relevant to cancer therapy.

The chemical probe BI-9321 (**17**) and the related negative control BI-9466 (**12**) will be made available to the scientific community *via* the SGC homepage (<http://www.thesgc.org/chemical-probes>) and *via* the Boehringer Ingelheim open innovation portal opnMe (opnMe.com). The probes can be used to increase understanding of the biological function of the PWWP domain and serve as an additional tool to decode the complex code of multi-valent epigenetic regulators such as NSD3.

Acknowledgements

The Structural Genomics Consortium is a registered charity (no. 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 & Co.; Novartis Pharma AG; Ontario Ministry of Economic Development and Innovation; Pfizer; São Paulo Research Foundation-FAPESP; Takeda; and the Wellcome Trust. We thank Christopher Vakoc for scientific and technical advice and the Expose team for data collection at the Swiss Light Source beamlines X06SA and X06DA. We thank Danette Daniels, Matt Robers and Cesear Corona from Promega for advising on the NanoBRET and target engagement assays, and acknowledge the Natural Sciences and Engineering Research Council of Canada (NSERC) for a postdoctoral fellowship awarded to DD.

Author contributions

J.B. and U.R. supervised the chemistry team. U.R., T.W, S.Z. and D.B designed synthetic strategies. R.N., M.S., M.Pet., D.B-L., M.V., K.V.M.H., C.R.V. and M.Pea. designed biological experiments and supervised the biology teams. M.Z. supervised the DSF, NMR measurements of the FBS screening. J.B., B.M. and A.W. performed structural analysis. K.R., S.K. and S.W. performed SPR assays. F. L and A. A-H performed selectivity assays. A. A-H contributed to performing SPR and ITC experiments. C.M.R. performed FRAP assays. M.Sz. performed NanoBRET assays. D.D. and M.C. performed cellular experiments. CW synthesized tracer ligands. M.M and H.B. performed analytics and wrote the synthesis supporting information. A.Z. supervised protein production and ITC experiments. N.M. performed quantitative proteomics experiments. A.H. performed quantitative tandem mass spectroscopy C.R. A.H. T.K. performed biological experiments. B.S., D.H, and T.G performed biochemical assays. X-L.C. J.E.F provided

Compchem support. G.B. supervised the collaboration with SGC. P.B. managed the project for SGC. C.H.A. provided supervision and funding. D.McC was responsible for the medicinal chemistry strategy. J.B., D.D., R.N, U.R, prepared the manuscript with input from all authors.

Competing interests

The authors declare no competing interests.

References

Uncategorized References

1. Allis, C.D. & Jenuwein, T. The molecular hallmarks of epigenetic control. *Nat Rev Genet* **17**, 487-500 (2016).
2. Bannister, A.J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Research* **21**, 381-395 (2011).
3. Arrowsmith, C.H., Bountra, C., Fish, P.V., Lee, K. & Schapira, M. Epigenetic protein families: a new frontier for drug discovery. *Nat Rev Drug Discov* **11**, 384-400 (2012).
4. Brown, P.J. & Müller, S. Open access chemical probes for epigenetic targets. *Future medicinal chemistry* **7**, 1901-1917 (2015).
5. Huston, A., Arrowsmith, C.H., Knapp, S. & Schapira, M. Probing the epigenome. *Nature Chemical Biology* **11**, 542 (2015).
6. Yang, L., Rau, R. & Goodell, M.A. DNMT3A in haematological malignancies. *Nature Reviews Cancer* **15**, 152 (2015).
7. Vougiouklakis, T., Hamamoto, R., Nakamura, Y. & Saloura, V. The NSD family of protein methyltransferases in human cancer. *Epigenomics* **7**, 863-874 (2015).
8. Kang, D. et al. The histone methyltransferase Wolf–Hirschhorn syndrome candidate 1-like 1 (WHSC1L1) is involved in human carcinogenesis. *Genes, Chromosomes and Cancer* **52**, 126-139 (2013).
9. Angrand, P.O. et al. NSD3, a new SET domain-containing gene, maps to 8p12 and is amplified in human breast cancer cell lines. *Genomics* **74**, 79-88 (2001).
10. Shen, C. et al. NSD3-short is an adaptor protein that couples BRD4 to the CHD8 chromatin remodeler. *Molecular cell* **60**, 847-859 (2015).
11. Gelsi-Boyer, V. et al. Comprehensive Profiling of 8p11-12 Amplification in Breast Cancer. *Molecular Cancer Research* **3**, 655 (2005).
12. Wu, H. et al. Structural and Histone Binding Ability Characterizations of Human PWWP Domains. *PLoS ONE* **6**, e18919 (2011).
13. Qin, S. & Min, J. Structure and function of the nucleosome-binding PWWP domain. *Trends in Biochemical Sciences* **39**, 536-547 (2014).
14. Vezzoli, A. et al. Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1. *Nat Struct Mol Biol* **17**, 617-619 (2010).
15. Hubbard, R.E. Fragment approaches in structure-based drug discovery. *Journal of Synchrotron Radiation* **15**, 227-230 (2008).

16. Baurin, N. et al. Design and Characterization of Libraries of Molecular Fragments for Use in NMR Screening against Protein Targets. *Journal of Chemical Information and Computer Sciences* **44**, 2157-2166 (2004).
17. Bergner, A. & Parel, S.P. Hit Expansion Approaches Using Multiple Similarity Methods and Virtualized Query Structures. *Journal of Chemical Information and Modeling* **53**, 1057-1066 (2013).
18. Hosmane, R.S. & Liebman, J.F. Paradoxes and paradigms: why is quinoline less basic than pyridine or isoquinoline? A classical organic chemical perspective. *Structural Chemistry* **20**, 693-697 (2009).
19. Fang, R. et al. Human LSD2/KDM1b/AOF1 Regulates Gene Transcription by Modulating Intragenic H3K4me2 Methylation. *Molecular Cell* **39**, 222-233 (2010).
20. Machleidt, T. et al. NanoBRET—A Novel BRET Platform for the Analysis of Protein–Protein Interactions. *ACS Chemical Biology* **10**, 1797-1804 (2015).
21. Jafari, R. et al. The cellular thermal shift assay for evaluating drug target interactions in cells. *Nature Protocols* **9**, 2100-2122 (2014).
22. Dart, M.L. et al. Homogeneous Assay for Target Engagement Utilizing Bioluminescent Thermal Shift. *ACS medicinal chemistry letters* **9**, 546-551 (2018).
23. Philpott, M. et al. Assessing cellular efficacy of bromodomain inhibitors using fluorescence recovery after photobleaching. *Epigenetics & Chromatin* **7**, 14 (2014).
24. Frye, S.V. The art of the chemical probe. *Nature Chemical Biology* **6**, 159-161 (2010).

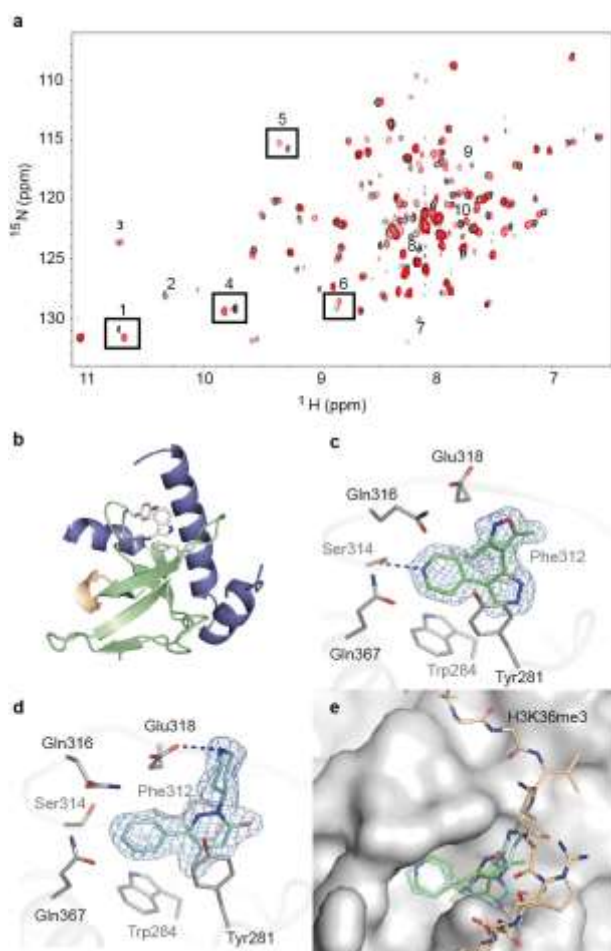


Figure 1: ^{15}N TROSY NMR Confirmation and structural analysis of fragment hits of the NSD3-PWWP1 domain

(a) Superposition of 2D ^{15}N TROSY NMR spectra of 100 μM uniformly ^{15}N labelled NSD3-PWWP1 (residues 247-398) in the absence (black) or presence (red) of 500 μM of the fragment hit **2**. Cross peaks of amino acids potentially forming the histone peptide binding site are highlighted (see **Supplementary Fig. 1**). The experiment was repeated independently with a similar result (n=2) (b) Overall structure of the PWWP1 domain of NSD3. The β -barrel is colored in pale green, the helix insertion between β_4 and β_5 is colored in wheat and the helix bundle shown in light blue. The residues forming the aromatic cage (Tyr 281, Trp 284 and Phe 312) are shown as gray sticks color coded by atom type. (c-d) Binding modes of the FBS hits **2**

and **3** as observed in the crystal structures in complex with the PWWP1 domain of NSD3. The ligands are shown as sticks in green (**2**) and blue (**3**), colour coded by atom type. The refined 2Fo-Fc electron densities are contoured at 1 σ . (f) Superposition of the crystal structures of **2** (C α RMSD = 5.5) and **3** (C α RMSD= 4.0) in complex with the PWWP1 domain of NSD3 with the structure of the BRPF1 PWWP domain in complex with a H3K36me3 peptide (PDB ID: 2X4Y)

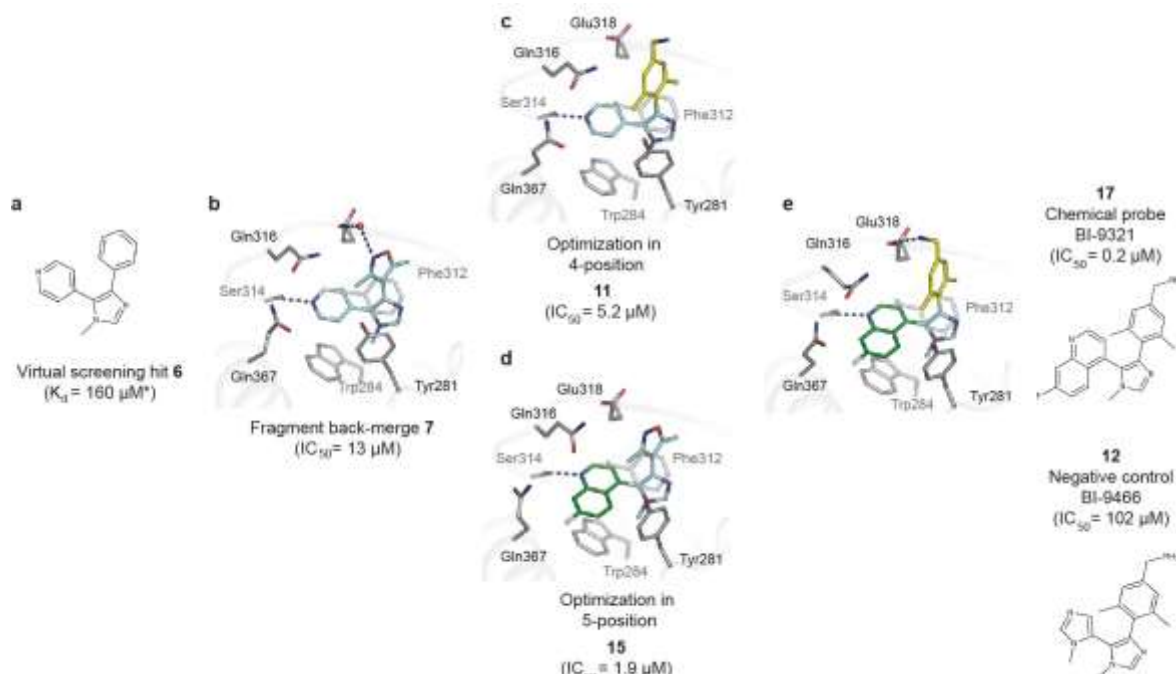


Figure 2 Structure-based optimization of methylimidazoles as NSD3-PWWP1 antagonists:

(a-e) Ligands are colour coded by atom type and hydrogen bonds are shown as blue dashes. Shown IC₅₀ values are originating from the TR-FRET assay except for **6** which has been determined with SPR, indicated with asterisk. (a) Virtual screening hit **6** used as basis for the structure-based optimization of methylimidazoles. (b) Binding mode of **7** obtained by merging the 3,5-dimethyl-1,2-oxazole moiety of **2** into **6** (c) Binding mode of **11** containing the (2,6-dimethyl-phenyl)methanamine substituent in 4-position of the methylimidazole core (d) Binding mode of **15** containing the 7-fluoro quinoline substituent in 5-position of the methylimidazole core. (e) Binding mode of BI-9321 (**17**) containing optimized substituents in 4- and 5-position of

the methylimidazole core as indicated by color coding. Structure and binding affinities of the chemical probe BI-9321 (**17**) and the respective negative control BI-9466 (**12**).

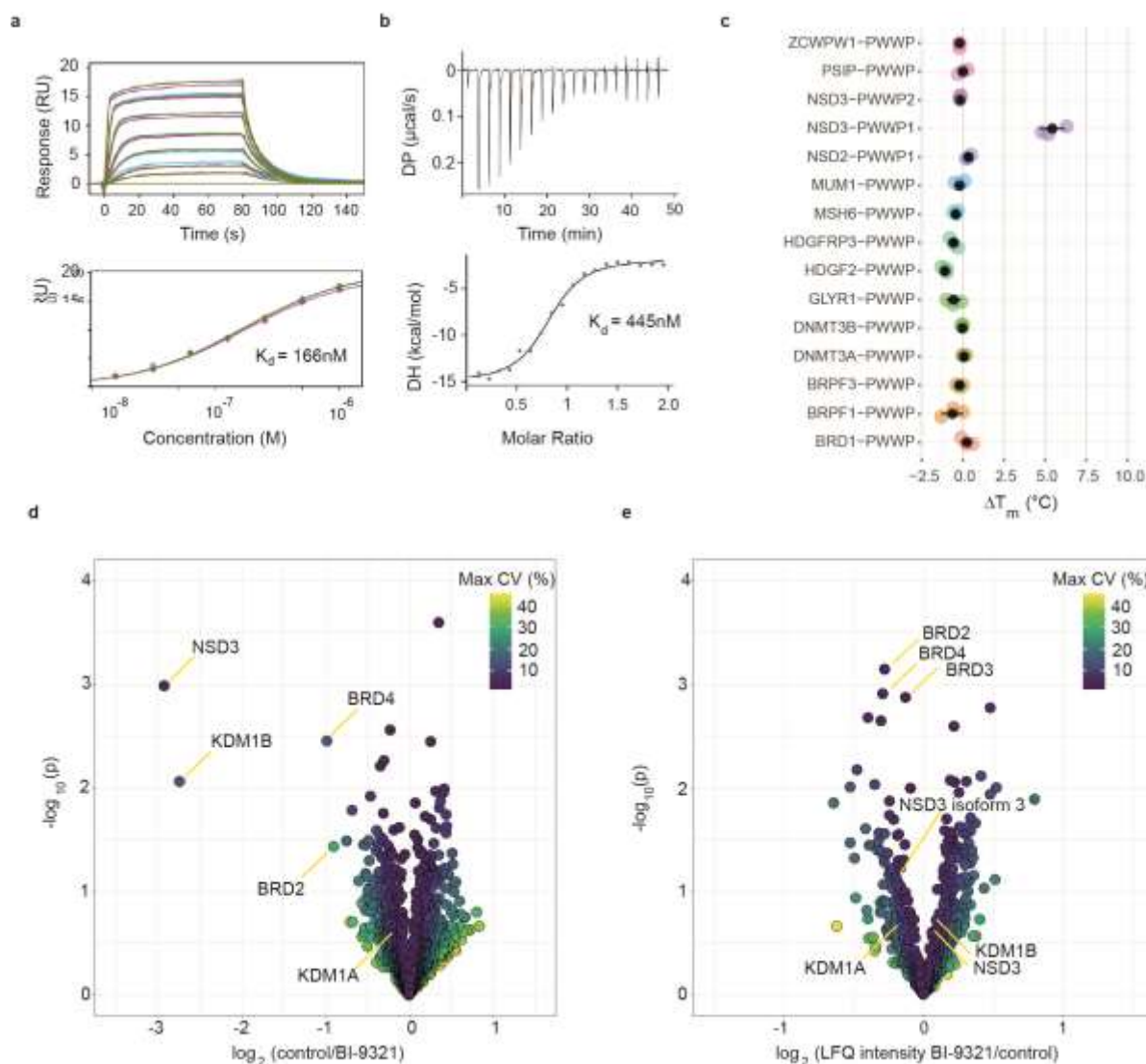


Figure 3: Biophysical validation and selectivity profiling of the chemical probe BI-9321 (17**)**

(a) SPR analysis of BI-9321 (**17**) (top: sensorgrams, bottom: steady-state fit using a 1:1 interaction model, $K_d = 166 \pm 3 \text{ nM}$ as mean value of three independent experiments \pm SD) (b) ITC analysis of BI-9321 (**17**) (ITC; $K_d = 445 \pm 8 \text{ nM}$; $\Delta H = -14.5 \pm 0.63 \text{ kcal/mol}$; $\Delta G = -8.67$; -

T_{ΔS} = 5.77) as mean value of four independent experiments ± standard deviations; **(c)** Histogram representation of the PWWP domain selectivity profile for BI-9321 (**17**) using differential scanning fluorimetry (DSF). Thermal shift difference for 14 PWWP domains upon treatment with 100 μM BI-9321 (**17**) (ΔT_m °C). Data are from three independent experiments and are presented as mean ± SD. **(d)** Selectivity profile of BI-9321 (**17**) in high-salt chromatin extract by quantitative proteomics. Proteins (dots) that are displaced from immobilized BI-9321 (**17**) by 10 μM competing **BI-9321** (**17**) specifically (*versus* control) and significantly (N = 3, Welch's t-test) position in the upper left of the Volcano plot. **(e)** NSD3 interactor changes upon BI-9321 (**17**) treatment by quantitative IP-MS. No significant (N = 4, Welch's t-test) changes among the identified interactors were induced in Cal-120 cells by 10 μM BI-9321 (**17**) (*versus* control).

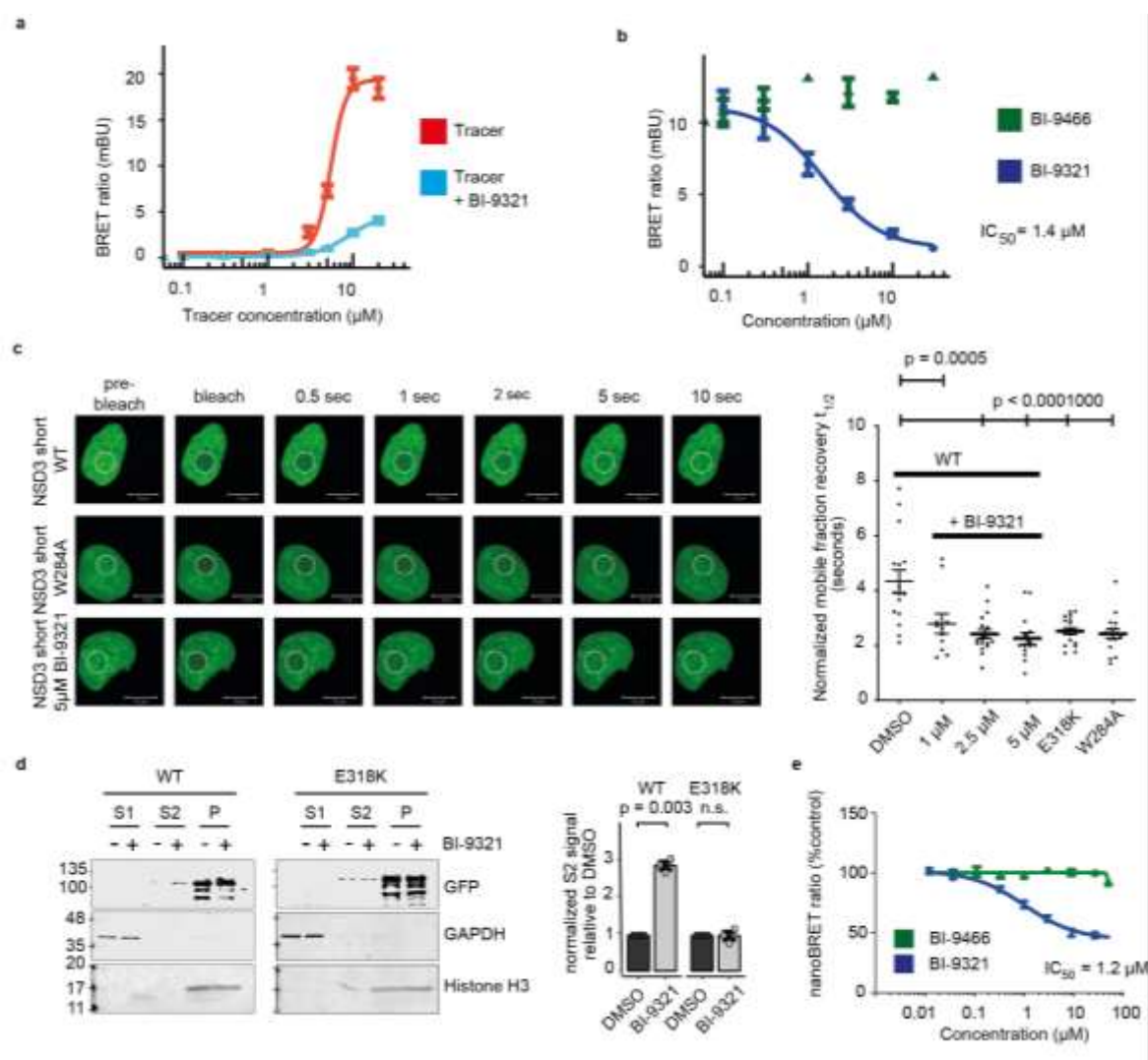


Figure 4: Cellular target engagement of the NSD3-PWWP1 Probe BI-9321 (17) (a) BRET tracer experiments, BRET measurements in response to increasing concentrations of tracer in the presence or absence of excess BI-9321 (17) (20 μM) ($\text{EC}_{50} = 5.5 \pm 0.1 \mu\text{M}$). Shown is the mean \pm SEM of three biological replicates, each measured in quadruplicate. (b) Dose response curve with increasing concentrations of either BI-9321 (17) or control compound BI-9466 (12) with tracer concentration held constant ($\text{IC}_{50} = 1.4 \pm 0.5 \mu\text{M}$). Shown is the mean \pm SEM of three biological replicates, each measured in quadruplicate. (c) Fluorescence Recovery After Photobleaching (FRAP) assay in U2OS cells transfected with GFP-tagged NSD3-short and the NSD3-short Trp284Ala mutant harbouring an inactivating mutation in the PWWP domain. Twenty cells were

imaged for each sample and one-way ANOVA with Dunnett's correction for multiple comparisons was used to detect significant differences ($P < 0.05$) between treatment groups. Data shown as mean \pm SEM. and images are representative of six independent experiments ($n = 6$). (e) Dose response curve with increasing concentrations of either **BI-9321** or control compound **BI-9466** with tracer concentration held constant ($IC_{50} = 1.4 \pm 0.5 \mu M$). $N = 3$ biological replicates each performed in quadruplicate (scale bar $10 \mu m$) (d) Biochemical fractionation experiments in HEK293 cells transfected with GFP-tagged NSD3short WT or E318K expression vectors and treated with BI-9321 (**17**) ($10 \mu M$) or DMSO control for 6 hours. Bar graph depicts the relative mean \pm SEM from three biological replicates of the nucleoplasmic signal normalized to total nuclear signal. P-values calculated using an unpaired two-tailed t-test with Welch's correction. Full western blots images used in the quantification s **Supplementary Figure 12 (d)** NanoBRET based on the interaction of C-terminally NanoLuc® tagged NSD3-PWWP1 with C-terminally HaloTag®Fusion tagged histone H3 (IC_{50} of $1.2 \mu M \pm 0.06$). Shown is a representative dose response curve. IC_{50} calculated from the mean \pm SD of three independent replicates.

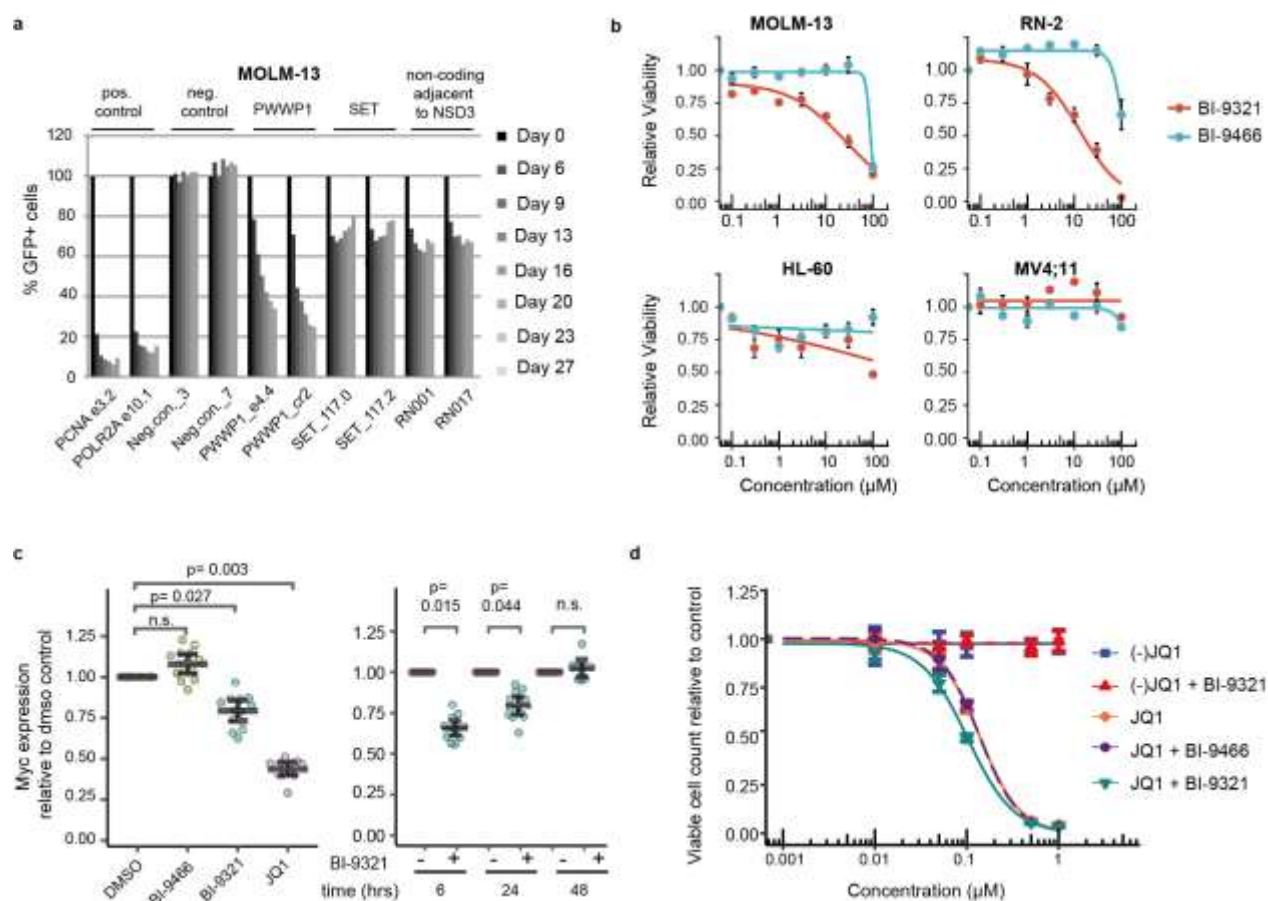


Figure 5: BI-9321 (17) downregulates MYC mRNA and reduces proliferation in MOLM-13 and RN-2 cells (a) A timecourse CRISPR depletion experiment, following the depletion kinetics of GFP+, Cas9 and gRNA expressing cells relative to the GFP- control MOLM-13 AML cells. PCNA and POLR2A serve as core essential control genes. MP049 and MP053 depict non-targeting control gRNAs. Only gRNAs targeting the PWWP1 domain result in depletion of GFP+ cells, whereas gRNAs targeting the SET domain and a non-coding region adjacent to NSD3 have no effects on the proliferation of MOLM-13 cells. (b) Measuring effects of BI-9321 (17) on the viability of AML cell lines by the RealTime-Glow MT Cell Viability Assay (MOLM-13 IC_{50} = 26.8 \pm 4.4 μ M; RN-2 cells IC_{50} = 13 \pm 2 μ M). Shown is the mean \pm SEM of three biological replicates, each measured in quadruplicate (c) NSD3-PWWP1 Antagonism Downregulates Myc mRNA Expression in MOLM-13 Cells. Transcript levels of Myc relative to PBGD and TBP in

MOLM-13 cells treated for 20 h with either **BI-9466** control (10 μ M), **BI-9321** (10 μ M), **JQ1** (0.5 μ M), or DMSO vehicle control. Myc expression relative to PBGD and TBP in MOLM-13 cells treated for 6, 24, or 48 hrs with BI-9321 (**17**) (10 μ M) relative to DMSO vehicle control. Shown is the mean \pm SEM of three biological replicates with four technical replicates each. P-values calculated using an unpaired two-tailed t-test with Welch's correction. **(d)** Representative dose response curve measuring proliferation of MOLM-13 cells in response to **JQ1** treatment alone ($IC_{50} = 139 \pm 6.7$ nM) or in combination with BI-9321 (**17**) (10 μ M; $IC_{50} = 120 \pm 4.5$ nM) or BI-9466 (**12**) (10 μ M; $IC_{50} = 137 \pm 6.5$ nM) control compound over 72 hours. Data points represent the mean \pm 95% CI of four replicates.

Online Methods

Chemical development

The dimethyl histone H3K36 peptide (H-STGGV-Lys(me₂)-KPHRY-OH) was purchased from JPT Peptide Technologies (MW=1257.47, purity >95% HPLC - 220nm - C18 - linear gradient). Compound **1** was purchased from Maybridge (#NRB 01585), **2** and **3** from ASINEX Ltd. (#ART 15393860 and #ART 13280592). Synthesis of all other compounds is shown in **Synthetic Procedures**. Unless otherwise indicated, all reactions were carried out in standard commercially available glassware using standard synthetic chemistry methods. Reagents were obtained from commercial sources and used without additional purification.

Protein purification

Constructs for expression of the NSD3-PWWP1 domain (residues 247-398 and 263-398) were obtained by gene synthesis (GeneArt, Thermo-Fisher) in donor vector (pDONR-221) and transferred by recombinant cloning into pDEST15 vectors. Both constructs included an N-terminal GST tag followed by a TEV cleavage site. Plasmids were used to transform *Escherichia coli*, strain BL21(DE3), in order to express and purify the proteins. For protein expression, 2xYT (Merck) or LB broth (ThermoFisher) containing ampicillin (100 µg.mL⁻¹) was inoculated with transformed BL21(DE3) cells, induced with IPTG (0.5 mM) at an optical density (O.D.) of 0.6 and grown overnight at 18 °C. Cell pellets obtained after centrifugation were stored at -20 °C. For protein purification, cells were solubilized in buffer A (50 mM HEPES, pH 7.2, 500 mM NaCl, 0.1% CHAPS, 5% glycerol, 5 mM DTT) containing a protease inhibitor cocktail (cOmplete, Roche). Cells were lysed by sonication (Sonopuls from Bandelin) for a total of 4 min (20 s pulse followed by 30 s rest on ice). The sonicated lysate was clarified by centrifugation at 10,000 x g for 45 min at 4 °C. The protein was captured by affinity chromatography on GSH-beads

(Glutathione HiCap, Qiagen) (batch mode) equilibrated in buffer B (same as buffer A without CHAPS). The beads were poured into an empty column, washed with buffer B and eluted with 20 mM GSH in buffer B. For crystallization the GST-tag was cleaved by incubation with His-TEV protease (produced in-house) overnight at 4 °C. For NSD3-PWWP1 (247-398) cleavage with His-TEV protease results in protein with an additional glycine at the N-terminus. The NSD3-PWWP1 (263-398) construct starts with the native serine after cleavage. His-TEV protease was removed by passage through a Ni-NTA (HISTrap HP, GE-Healthcare) column. Proteins were further purified by size-exclusion chromatography on HiLoad Superdex 75 pg (16/60, GE-Healthcare) equilibrated in 50 mM HEPES, 100 mM NaCl, 5 mM DTT, 5% glycerol, pH 7.2. Pure NSD3-PWWP1 was concentrated by ultrafiltration (30 kD membrane, Amicon, Millipore) to approximately 14 mg.mL⁻¹. Protein purity was assessed using 12% SDS-polyacrylamide gel (Criterion XT, Bio-Rad). For assays, NSD3 protein (GST-fusion) was purified as above omitting the tag cleavage step. For seleno-methionine labelling, GST-tagged NSD3-PWWP1 was expressed in the methionine auxotrophic E. coli strain B834 (DE3). A starter culture grown in minimal medium in the presence of 50 µg.mL⁻¹ methionine was centrifuged and the pellet resuspended in minimal medium without methionine and grown overnight at 25 °C. On the next day seleno-methionine was added to 50 mg.L⁻¹ and incubated for 30 minutes at 30 °C. After induction with 0.2 mM IPTG the temperature was decreased to 25 °C and cells grown overnight. Cells were harvested the next day and NSD3-PWWP1 purified as described above. For protein observed NMR, cells expressing NSD3-PWWP1 were grown in M9 minimal medium supplemented with ¹⁵NH₄Cl (0.5 g.L⁻¹) and protein purified as specified above.

NMR spectroscopy

NMR experiments for primary fragment screening were performed on a Bruker AVII 600MHz spectrometer equipped with a 5 mm z-gradient QCI cryogenic probe ($^{15}\text{N}/^{13}\text{C}/^{19}\text{F}/^1\text{H}$) and a Bruker Sample Rail. Samples were freshly prepared just-in-time using a Tecan Freedom Evo pipetting robot²⁵) which was in-house customized for 2.5 mm NMR sample tube filling before subsequent fully automated data acquisition. STD NMR²⁶ spectra were recorded in an interleaved fashion whereby protein resonances were saturated for 3 s with a Gaussian pulse train centered at 400 Hz. For the off-resonance spectra irradiation was applied at 40 kHz. Suppression of residual protein resonances was achieved by application of a 30 ms spinlock pulse prior to data acquisition. For binding experiments 10 μM NSD3-PWWP1 (residues 247-398) were incubated with mixes of four fragments (250 μM each) in 25 mM Na-phosphate, 100 mM NaCl pH 7.5 in D_2O at 298 K and a total $\text{d}_6\text{-DMSO}$ concentration of 2%. Data was analyzed using Topspin 3.0 (Bruker Biospin) and binders were detected by positive signals in the difference spectra. Comparison with prior recorded 1D ^1H reference spectra allowed the identification of the individual “active” fragment in the mixture. Hit confirmation was carried out on a fully automated (see above) Bruker AVIII 600MHz spectrometer equipped with a 5 mm z-gradient TCI cryogenic probe ($^{15}\text{N}/^{13}\text{C}/^1\text{H}$). 100 μM uniformly ^{15}N labelled NSD3-PWWP1 (residues 247-398) was mixed with 500 μM fragment in 25 mM Na-phosphate, 100 mM NaCl pH 7.5 with 8% (v/v) D_2O at 298 K and 2D ^1H ^{15}N TROSY^{27,28} experiments were carried out with spectral widths of 16 ppm and 28 ppm respectively. For K_d determinations by NMR, distinct samples of the respective compound concentration at a constant DMSO concentration of 1% were prepared. Chemical shift perturbations were calculated and titration curves were analyzed as described²⁹. and chemical shift perturbation was analysed using FELIX2004 (Felix Inc., San Diego, CA, USA)³⁰. Superpositions of 2D ^{15}N TROSY spectra were done with NMRViewJ, version 8.0.rc51 (One Moon Scientific Inc., Newark, NJ, USA).

Differential Scanning Fluorimetry (DSF)

DSF experiments were performed using a Bio-Rad CFX384 Real-Time System (C1000Touch Thermal Cycler) in sealed Hard-Shell PCR 384 well plates (#HSP3805; PCR Sealers; #MSB1001; Bio-Rad). Compound dilutions of 8 μ L in assay buffer from 50 mM DMSO stock solutions were produced using a Hamilton Microlab Starlet pipetting robot equipped with a CO-RE 384 probe head. Manual addition of 2 μ L of the NSD3-PWWP1 SYPRO Orange stock mixture using a 12 channel matrix multipette (Thermo Fisher) yielded the final assay volume of 10 μ L. Assay optimization resulted in experimental conditions such as 25 mM HEPES pH 7.5, 1 mM TCEP, 500 μ M compound, 1% DMSO with 10 μ M NSD3-PWWP1 and a 5x dye excess (5000x SYPRO Orange stock in DMSO, Invitrogen). Thermal denaturation was achieved by applying a temperature ramp from 25 $^{\circ}$ C to 95 $^{\circ}$ C (1 $^{\circ}$ C/h) with fluorescence readings every 0.5 $^{\circ}$ C.

NSD3-PWWP1 SPR K_d assay

NSD3-PWWP1 (263-398) was immobilized to 3000 - 4000 RUs on flow cells 2 and 4 of a Biacore CM5 sensor chip on a Biacore T200 instrument after predilution to a concentration of 0.1 mg/ml with 20 mM bicine, pH 8.0. Carbonic anhydrase II was used as a reference protein on flow cell 3 and flow cell 1 was kept blank as a reference surface. After changing to assay buffer (50 mM TRIS, pH 8.0; 150 mM NaCl; 1 mM TCEP; 0.005 % Tween 20; 2% DMSO) the chip was pre-equilibrated for several hours before use for K_d measurements. **4** was used as a positive control and experimental procedures followed were as described previously³¹. Biacore T200 evaluation software was used for determining K_d values. Data were fitted by either performing a global fit of the double-referenced association and dissociation kinetics or by fitting steady-state responses at each concentration to a 1:1 interaction model. Average K_d s were calculated from the

data for the two flow cells. Data reported in this study are mean of at least three independent experiments \pm standard deviations.

SPR-based single point screening of Morhits compounds

The single-point measurements were performed on a Biacore 4000 instrument (GE Healthcare). NSD3-PWWP1 was immobilized onto a CM5-chip (GE Healthcare) as described above and the protein was injected over spots 1, 4 and 5 on all 4 flow cells of the chip after activation with EDC/NHS according to the manufacturer's instructions ($T = 25\text{ }^{\circ}\text{C}$). Carbonic anhydrase II (Sigma Aldrich) served as a negative control and was immobilized onto spot 2 (EDC/NHS activation). Buffer was injected over spot 3 of all 4 flow cells to generate a blank reference surface. Approximately 8000 RU of each of the two proteins were immobilized. After switching to screening buffer (50 mM TRIS, pH 8.0; 150 mM NaCl; 1 mM TCEP; 0.005 % Tween 20; 2% DMSO) compounds were screened at a concentration of 0.05 mg/ml with **4** as a positive control. Single point measurements and data analysis were performed following a published procedure³¹. Compounds displaying abnormal binding behaviour were disqualified. In addition, compounds showing binding to carbonic anhydrase II were also not followed up any further. MEDIAN and standard deviation (STDEV) over all compounds analysed were calculated and compounds showing a $\text{FSO} \geq \text{MEDIAN} + 3 \times \text{STDEV}$ were selected for subsequent SPR K_d determination as described above.

NSD3-PWWP1 SureLight Time-Resolved Fluorescence Resonance Electron Transfer (TR-FRET) assay

This assay was used to identify compounds which competitively interact with the binding of a biotinylated probe to NSD3-PWWP1. The biotinylated probe **5** and GST-tagged NSD3-PWWP1

(residues 247-398) were used as binding partners. Test compounds dissolved in DMSO were dispensed onto assay plates (Proxiplate 384 PLUS, white, PerkinElmer) using an Access Labcyte Workstation with the Labcyte Echo 55x. For the chosen highest assay concentration of 500 μ M, 150 nL of solution was transferred from a 50 mM DMSO stock solution. A series of eleven fivefold dilutions per compound was transferred to the assay plate, compound dilutions were tested in duplicates. DMSO was added as backfill to a total volume of 150 nL. The assay runs on a fully automated robotic system. For the assay, 15 μ L of the NSD3-PWWP1 (50 nM final assay concentration), biotinylated probe (**5**) solution (25 nM final assay concentration), Lance Eu-W1024 labelled Streptavidin (PerkinElmer, 2.5 nM final assay concentration) and Anti-GST surelight APC (PerkinElmer, 50 nM final assay concentration) mixed in assay buffer (1x PBS; 0.05% Tween20; 0.1 % BSA; filtered) were added. Plates were kept at room temperature for 60 minutes. After incubation the TR-FRET signal was measured in a PerkinElmer Envision HTS Multilabel Reader using the TR-FRET LANCE Ultra specs of PerkinElmer (used wavelengths: excitation 320 nm, emission1 665 nm, emission2 615 nm). Each plate contains negative controls (DMSO instead of test compound; with biotinylated probe solution) and positive controls (DMSO instead of test compound; without biotinylated probe solution). Negative and positive control values were used for normalization. **4** was used as internal control and IC₅₀ values were calculated and analysed using a 4 parametric logistic model.

Ligand-Based Virtual Screening (Morhits)

Morhits is a complex and integrated workflow for the automated execution of different complementary ligand-based virtual screening methods in parallel in Pipeline Pilot³². Searching is then followed by a sophisticated and dynamic data fusion process³³. This ensures that the outcome not only adequately represents each template molecule and search method but also

covers a diverse representation of the expanded chemical space. The ligand-based virtual screening methods implemented in Morhits include standard substructure searches, 2D fingerprints similarity search employing different fingerprint sets, TurboSim³⁴ search, 2D pharmacophore graph searches Discngine³⁵, shape searches using ROCS³⁶ and BLAZE³⁷ searches. In addition to using the pre-defined query structures, Morhits also constructs virtual template molecules on-the-fly, using, for example, bio-isostere enumeration ring assembly modification rules. The virtual template molecules are subjected to the same panel of similarity searches as the initial template structures.

Isothermal Titration Calorimetry (ITC) assay

Calorimetric experiments of BI-9321 (**17**) were performed on a Microcal PEAQ-ITC calorimeter (MicroCalTM, LLC Northampton, MA). Protein solutions were buffer exchanged by dialysis into buffer 20 mM HEPES, 100 mM NaCl, 3% DMSO, pH 8.0. All measurements were carried out at 25 °C. The cell was loaded with protein solutions in the range of 0 to 10 µM. All injections were performed using an initial injection of 0.5 µL followed by 19 injections of 2 µL of compound in the range of 100-300 µM. The data were analysed with the MicroCal PEAQ-ITC analysis software package. The first data point was excluded from the analysis. Thermodynamic parameters were calculated ($\Delta G = \Delta H - T\Delta S = -RT\ln K_d$ where ΔG , ΔH and ΔS are the changes in free energy, enthalpy and entropy of binding, respectively).

Quantitative Proteomics

Cal-120 cells were cultured under standard medium conditions. Cells were lysed in 20 mM HEPES pH 7.5, 150 mM KCl, 1% Triton X-100 and spun at 5000 x g. The resulting pellet was extracted with lysis buffer containing 350 mM KCl and cleared at 30,000 x g. For selectivity

profiling, the BI-9321 (**17**) affinity matrix was prepared by amide formation between BI-9321 (**17**) and Magnetic NHS Sepharose (GE Healthcare) in DMSO/TEA to a ligand density of 10%. Extract and 200 nL affinity matrix were incubated in the presence of 10 μ M BI-9321 (**17**) or DMSO for 2h. For immunoprecipitations, Cal-120 cells were cultured in the presence of 10 μ M BI-9321 (**17**) or DMSO for 24 h and extracted as described above. Extracts were incubated with anti-NSD3 antibody (polyclonal antibody raised against the peptide: PTDYYHSEIPNTRPHEC) for 1 h and 500 nL Protein A Mag Sepharose (GE Healthcare) for 1 h¹⁰. After incubation the matrices were washed and bound proteins were proteolytically eluted using Trypsin (Pierce) in the presence of 5 mM TCEP and 10 mM 2-chloroacetamide. Peptides were purified by solid-phase extraction (Sep-Pak tC18, Waters) and analyzed on a nanoflow-LC system (Thermo Fisher) hyphenated to a Q-Exactive Plus mass spectrometer (Thermo Fisher). Assignment of tandem mass spectra and label-free quantification was performed using MaxQuant³⁸. Statistical testing was performed using the statistical programming language R.

KDM1B biochemical assay

We optimized a SETD7-coupled scintillation proximity lysine demethylase assay for the full-length recombinant KDM1B (activemotif, cat # 31479) based on the protocol we previously published for other KDMs (KDM1A, KDM3A, and KDM4A)³⁹. We used 60 nM KDM1B for these assays, which was incubated in the presence and absence of compounds with 4 μ M biotin-labeled H3K4me1(1–25) peptide in 50 mM Tris-HCl, pH 8.0, 50 nM FAD, 0.01% Triton X-100 for 1 hour at 23° C. The reaction plate was then heated at 80 °C for 15 min in a water bath to denature and inactivate KDM1B. The samples were then centrifuged at 2000 rpm at 4 °C for 5 min. SETD7 at a final concentration of 0.1 μ M, 5 mM DTT, and 2 μ M ³H SAM were added into sample wells, and incubated for another 1 hour at 23°C. The reactions were stopped by adding 10

μL of 7.5 M guanidine hydrochloride and the CPM counts were measured using Topcount plate reader (Perkin Elmer, www.perkinelmer.com). All enzymatic reactions were performed in triplicate and IC_{50} values were determined by fitting the data to Four Parameter Logistic equation using GraphPad Prism 7 software.

Structural Analysis

NSD3-PWWP1 domain crystals (residues 247-398) were obtained using the hanging drop method by mixing 1 μL protein solution (5 mg.mL^{-1} in 50 mM TRIS, 100 mM NaCl pH 8.0) with 1 μL of reservoir solution (100 mM HEPES pH 7.8, 28% PEG3350, 2% PEG200, 1.5% 1,2-Butandiol) at 20 °C. Rod shaped crystals appeared within days and were flash frozen in reservoir containing 20% EG. Crystals diffracted to 2.5 Å ($P2_1$ $a=45.18$, $b=88.09$, $c=79.99$ $\alpha=90.00$ $\beta=89.88$ $\gamma=90.00$), but the structure could not be solved using publicly available molecular replacement models. Therefore crystallization was repeated with selenomethionine labelled material by mixing 0.2 μL protein solution (15 mg.mL^{-1} in 50 mM HEPES, 100 mM NaCl pH 8.0) with 0.2 μL of reservoir solution (50 mM TRIS pH 9.0 and 23% PEG3350) at 20 °C. Crystals belonged to the same space group $P2_1$ containing 4 monomers in the asymmetric unit ($a=45.04$ Å, $b=88.55$ Å, $c=77.97$ Å, $\alpha=90.00^\circ$, $\beta=90.08^\circ$, $\gamma=90.00^\circ$). Data was collected at a wavelength of 0.9794 Å and the structure was solved by SAD using the phenix software package⁴⁰. The selenomethionine labelled structure was refined at a resolution of 2.8 Å and the native structure at a resolution of 2.5 Å. Both models were validated with MolProbity revealing 99.6% of the residues in Ramachandran favoured and 0.4% in allowed regions.

The structural information was used for design of the construct NSD3-PWWP1 (residues 263-398). Crystals of this protein variant were obtained in presence of 2 mM fragment **1** mixing 0.2 μL protein solution (15 mg.mL^{-1} in HEPES 20 mM, 100 mM NaCl pH 8.0) with 0.2 μL of

reservoir solution (100 mM Morpheus Buffer 3, 30% P550MME_P20K, 10% Morpheus Ethylene glycols) at 4 °C. Cubic crystals appeared within several weeks and grew to a final size of 100-100 μm . Structures of **2**, **3**, **6**, **7**, **11**, **15** and BI-9321 (**17**) were obtained by soaking co-crystals of **1** for 3 to 5 days in the presence of 50 mM ligand. All crystals were flash frozen from reservoir in liquid nitrogen. Crystals belonged to space group $P2_12_12_1$ and contained one monomer per asymmetric unit. Diffraction data was collected at X06SA and X06DA beamlines of the Swiss Light Source (Paul Scherrer Institute, Switzerland). Images were processed with autoPROC⁴¹. Structures were solved by molecular replacement using the previously solved structure as model. Model building and refinement was performed with standard protocols using CCP4, COOT and autoBUSTER v.2.11.2. (<http://www.globalphasing.com>)^{42,43}. Final models were analyzed with MolProbity revealing 100/0/0 (**1**), 100/0/0 (**2**), 98.3/1.7/0 (**3**), 98.3/1.7/0 (**4**), 99.2/0.8/0 (**6**), 100/0/0 (**7**), 99.2/0.8/0 (**11**); 99.2/0.8/0 (**14**) and 100/0/0 BI-9321 (**17**) % of the residues in Ramachandran favoured, allowed and disallowed regions, respectively. Statistics for the respective data collection and refinement can be found in **Supplementary Tables 10, 11, 12, and 13**. Stereo images of the respective binding modes (wall-eye stereo) with the refined $2F_o - F_c$ electron can be found in **Supplementary Fig. 4, 5, 6, and 7**.

Fluorescence Recovery After Photobleaching (FRAP)

U2OS cells were seeded at 40% confluency into 35 mm cell culture dishes (Fluorodish, World Precision Instruments), followed by reverse transfection with mammalian overexpression constructs encoding NSD3-short proteins fused to GFP (wild type or mutants) using Lipofectamine 2000 (Life Technologies). After 6 hours incubation at 37 °C/5% CO₂, the medium containing transfection reagents was removed and replenished with fresh medium and cells were left to grow at 37 °C/5% CO₂ overnight. One hour before imaging, cells were incubated with

compounds at either 1, 2.5 or 5 μ M at 37 °C/5% CO₂. Bleaching and GFP fluorescence imaging were carried out with an argon ion laser (488 nm) and a Zeiss LSM 710 scan head (Zeiss GmbH, Jena, Germany). A circular region (6 μ m²) of the GFP-positive nuclei was selected and bleached with 20-25 cells selected for bleaching for each experimental condition. A time-lapse series was taken to record GFP recovery. The average intensity at each imaging time point was measured for three regions: the bleached region, the total cell nucleus and a random region outside the cell for background subtraction. The image datasets were exported to OriginPro 9.1 software (OriginLab, Northampton, MA, USA) using a custom LabTalk script for analysis. The relative fluorescence signal in the bleached region of each cell was calculated for each time point. . Graphpad Prism was utilised to perform one-way ANOVA with Dunnett's correction for multiple comparisons to detect significant differences ($P < 0.05$) between treatment groups²³.

Cellular Fractionation

HEK293 cells were seeded in 6-well plates at 100 000 cells per well and incubated overnight. The following day, cells were transfected with 2 μ g GFP tagged NSD3-short wild-type or E318K expression vectors with jetPRIME® (Polyplus transfection) following the manufacturer's instruction. Cells were incubated for 48 hours followed by the addition of either **BI-9321** (10 μ M) or DMSO control for 6 hours. Cells were then collected by trypsinization and centrifugation at 400 x g for 5 minutes at 4 °C. Cell pellets were washed in 1x PBS and resuspended in 200 μ L of Hypotonic Buffer A (10 mM HEPES pH 7.5, 10 mM KCl, 1.5 mM MgCl₂, 0.3 M Sucrose, 1 mM TCEP, and protease inhibitors) per 1 x 10⁶ cells. The cell suspension was incubated on ice for 15 minutes followed by centrifugation at 1300 x g for 5 minutes. Supernatant was collected and cleared by centrifugation to produce the cytoplasmic (S1) fraction. The pellet was washed in Buffer A and resuspended in an equal volume of Buffer B (3 mM EDTA, 3 mM EGTA, 1 mM

TCEP, and protease inhibitors). Following incubation on ice for 30 minutes, the nuclear fraction was centrifuged at 1300 x g for 5 minutes. The supernatant was again collected and cleared to yield the nucleoplasmic (S2) fraction. The chromatin-bound or pellet fraction (P) was then resuspended in an equal volume of Histone buffer (20 mM Tris HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 10 mM MgCl₂, 1% Triton X-100, 1 mM TCEP, 30 µg.µ⁻¹ benzonase, and protease inhibitors). Samples were boiled in SDS loading buffer before western blotting with antibodies to GFP (Takara Bio Cat# 632381, RRID:AB_2313808; WB 1:1000), GAPDH (Millipore Cat# MAB374, RRID:AB_2107445; WB 1:5000), and Histone H3 (Abcam Cat# ab1791, RRID:AB_302613; WB 1:10 000) using the NuPAGE electrophoresis and transfer system (Invitrogen). Immunoblots were imaged on a LiCor Odyssey CLx and quantified in Image Studio Lite v5.2.5 (Li-Cor Biosciences).

NanoLuc® Cellular Target Engagement Assays

Tracer target engagement experiments were performed as previously described⁴⁴. Briefly, U2OS cells were plated in 6-well plates (1 x 10⁵/well) in DMEM supplemented with 10% FBS, penicillin (100 U.mL⁻¹) and streptomycin (100 µg.mL⁻¹). 4 hours after plating cells were transfected with 1.8 µg of carrier plasmid and 0.2 µg of C-terminally tagged NSD3-PWWP1-NanoLuc® construct using X-tremeGENE HP DNA Transfection Reagent (Sigma), following manufacturer instructions. Next day cells were trypsinized and resuspended in DMEM/F12 (no phenol red) supplemented with 4% FBS, penicillin (100 U.mL⁻¹) and streptomycin (100 µg.mL⁻¹) at density 1.1x10⁵ mL⁻¹ and plated in 384-well plates. Cells were equilibrated for four hours with energy transfer probe and test compounds. To measure BRET, NanoBRET™ NanoGlo® Substrate and Extracellular NanoLuc® Inhibitor (Promega) were added according to the manufacturer's recommended protocol. Donor emission at 450 nm (filter: 450 nm/BP 80nm) and

acceptor emission at 618 nm (filter: 610nm/LP) was measured within 10 minutes of substrate addition using CLARIOstar microplate reader (Mandel). NanoLuc® Thermal Shift Assays were performed as previously described²². U2OS cells were transfected with either C-terminally tagged NSD2-PWWP1-NanoLuc® or NSD3-PWW1-NanoLuc® (Promega) and incubated overnight. Cells were trypsinized and resuspended at $1.1 \times 10^5 \text{ mL}^{-1}$ in DMEM/F12 (no phenol red) supplemented with 4% FBS and protease inhibitors. 90 μL of cells was dispensed per well of a 96-well PCR plate followed by addition of 10 μL of a 10x compound stock or a DMSO control. Cells were equilibrated for 2 hours prior to heating across a temperature gradient for 3 minutes on a Mastercycler® thermal cycler (Eppendorf) and addition of NanoGlo® Substrate (Promega). Heated cells were transferred to 384-well white plastic plates (Corning) and luciferase signal read at 450 nm (filter: 450 nm/BP 80nm) on a CLARIOstar microplate reader (Mandel). For NanoBRET™ protein-protein interaction assays, U2OS cells were transfected as above with 2 μg of C-terminally tagged histone H3-HaloTag® Fusion Vector DNA and 0.2 μg of C-terminally tagged NSD2-PWWP1- NanoLuc® , NSD3-PWW1-NanoLuc® (Promega) or NSD3-PWWP1 double mutant W284A, E318K-NanoLuc® using X-tremeGENE HP DNA Transfection Reagent (Sigma), following manufacturer instructions. Next day cells were trypsinized and resuspended in DMEM/F12 (no phenol red) supplemented with 4% FBS, penicillin ($100 \text{ U} \cdot \text{mL}^{-1}$) and streptomycin ($100 \mu\text{g} \cdot \text{mL}^{-1}$) at density $1.1 \times 10^5 \text{ mL}^{-1}$. Cells were divided into two pools. To the first pool HaloTag® NanoBRET™ 618 Ligand (Promega) was added and to second pool DMSO, following manufacturer instructions. Cells were plated in 96-well white plastic plates (Greiner) in the presence or absence of compounds for 20 h. Next, NanoBRET™ Nano-Glo® Substrate (Promega) solution was added to each well, following manufacturer instructions. Donor emission at 450 nm (filter: 450 nm/BP 80nm) and acceptor emission at 618 nm (filter: 610nm/LP) was measured within 10 minutes of substrate addition using CLARIOstar microplate reader (Mandel).

For tracer and protein-protein interaction assays, mean corrected NanoBRET™ ratios (mBU) were determined by subtracting mean of 618/460 signal from cells without NanoBRET™ 618 Ligand x 1000 from mean of 618/460 signal from cells with NanoBRET™ 618 Ligand x 1000.

Myc mRNA Expression in MOLM-13 Cells

Total RNA was prepared from MOLM-13 cells by Qiazol extraction (Qiagen, 79306) following the manufacturer's protocol. Isolated RNA was reverse transcribed using the iScript cDNA Synthesis Kit (Bio-Rad, 1708841). RT-qPCR was performed using a CFX384 Touch Real-Time PCR Detection System (Bio-Rad) and PowerUp SYBR Green Master Mix (Thermo Fisher, 100029284). Relative transcript levels were determined by the Ct method and normalized to the geometric mean of PBGD and TBP.

Measuring Viability of AML Cell Lines

1000 cells were plated in 96-well suspension culture plates and treated with several concentrations of NSD3-PWWP1 antagonist BI-9321 (**17**) or negative control BI-9466 (**12**). Following 6 days of treatment, cell viability was measured in 384-well plates using the RealTime-Glow MT Cell Viability Assay (Promega) and luminescence read on a CLARIOstar microplate reader (Mandel).

CRISPR-Cas9 Experiments

All cells used were kept in standard medium conditions as indicated by the vendors. An expression vector encoding Cas9 and carrying a puromycin selection marker was stably integrated into the target cells. Thereafter, cells were virally transduced with constructs encoding for GFP and gRNAs targeting positive, negative controls and NSD3. The fraction of GFP positive cells was determined by FACS measurement. The fraction of GFP positive cells at day 2

or day 3 post-infection was set to 100%. All gRNA sequences are listed in **Supplementary Table 9**.

Data availability

The authors declare that the data supporting the findings of this study have been deposited in the RCSB Protein Data Bank (PDB; <http://www.rcsb.org>) with the accession numbers: Selenomethionine labelled (6G3P); unlabelled (6G3T); **1** (6G24); **2** (6G25); **3** (6G27); **4** (6G29); **6** (6G2B); **7** (6G2C); **11** (6G2E); **15** (6G2F); compound BI-9321 (**17**) (6G2O) or are available within the publication and its supplementary information.

Uncategorized References

10. Shen, C. et al. NSD3-short is an adaptor protein that couples BRD4 to the CHD8 chromatin remodeler. *Molecular cell* **60**, 847-859 (2015).
22. Dart, M.L. et al. Homogeneous Assay for Target Engagement Utilizing Bioluminescent Thermal Shift. *ACS medicinal chemistry letters* **9**, 546-551 (2018).
23. Philpott, M. et al. Assessing cellular efficacy of bromodomain inhibitors using fluorescence recovery after photobleaching. *Epigenetics & Chromatin* **7**, 14 (2014).
25. Ross, A. & Senn, H. Automation of measurements and data evaluation in biomolecular NMR screening. *Drug discovery today* **6**, 583-593 (2001).
26. Mayer, M. & Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angewandte Chemie International Edition* **38**, 1784-1788 (1999).
27. Nietlispach, D. Suppression of anti-TROSY lines in a sensitivity enhanced gradient selection TROSY scheme. *Journal of Biomolecular NMR* **31**, 161-166 (2005).
28. Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. Attenuated T(2) relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 12366-12371 (1997).
29. Grzesiek, S., Stahl, S.J., Wingfield, P.T. & Bax, A. The CD4 Determinant for Downregulation by HIV-1 Nef Directly Binds to Nef. Mapping of the Nef Binding Surface by NMR. *Biochemistry* **35**, 10256-10261 (1996).
30. Peng, C., Unger, S.W., Filipp, F.V., Sattler, M. & Szalma, S. Automated evaluation of chemical shift perturbation spectra: New approaches to quantitative analysis of receptor-ligand interaction NMR spectra. *Journal of biomolecular NMR* **29**, 491-504 (2004).
31. Martin, L.J. et al. Structure-Based Design of an in Vivo Active Selective BRD9 Inhibitor. *Journal of medicinal chemistry* **59**, 4462-4475 (2016).
32. BIOVIA Pipeline Pilot version 9.3. *San Diego: Dassault Systèmes* (Release 2013).
33. Willett, P. Combination of Similarity Rankings Using Data Fusion. *Journal of Chemical Information and Modeling* **53**, 1-10 (2013).
34. Hert, J. et al. New Methods for Ligand-Based Virtual Screening: Use of Data Fusion and Machine Learning to Enhance the Effectiveness of Similarity Searching. *Journal of Chemical Information and Modeling* **46**, 462-470 (2006).
35. Discngine, S.A.S., Paris, France.
36. Hawkins, P.C.D., Skillman, A.G. & Nicholls, A. Comparison of Shape-Matching and Docking as Virtual Screening Tools. *Journal of Medicinal Chemistry* **50**, 74-82 (2007).
37. Cheeseright, T.J., Mackey, M.D., Melville, J.L. & Vinter, J.G. FieldScreen: Virtual Screening Using Molecular Fields. Application to the DUD Data Set. *Journal of Chemical Information and Modeling* **48**, 2108-2117 (2008).
38. Cox, J. et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & cellular proteomics : MCP* **13**, 2513-2526 (2014).
39. Yu, W. et al. A scintillation proximity assay for histone demethylases. *Analytical Biochemistry* **463**, 54-60 (2014).
40. Chen, V.B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography* **66**, 12-21 (2010).
41. Vonrhein, C. et al. Data processing and analysis with the autoPROC toolbox. *Acta Crystallographica Section D* **67**, 293-302 (2011).
42. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography* **66**, 486-501 (2010).

43. Collaborative. The CCP4 suite: programs for protein crystallography. *Acta Crystallographica Section D* **50**, 760-763 (1994).
44. Robers, M.B. et al. Target engagement and drug residence time can be observed in living cells with BRET. *Nature communications* **6**, 10091-10091 (2015).

