Published in final edited form as:

Structure. 2016 July 6; 24(7): 1201–1208. doi:10.1016/j.str.2016.04.019.

Structural Mechanism of Transcriptional Regulator NSD3 Recognition by the ET Domain of BRD4

Qiang Zhang[1,2,5], **Lei Zeng**[1,2,5], **Chen Shen**[3,4], **Ying Ju**[2], **Tsuyoshi Konuma**[1], **Chengcheng Zhao**[2], **Christopher R. Vakoc**[3], and **Ming-Ming Zhou**[1]

[1]Department of Structural and Chemical Biology, Icahn School of Medicine at Mount Sinai, New York, New York, USA

[2]The First Hospital of Jilin University, Changchun, Jilin, China

[3]Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, USA

[4]Molecular and Cellular Biology Program, Stony Brook University, Stony Brook, New York, USA

SUMMARY

The Bromodomains and Extra-Terminal domain (BET) proteins direct gene transcription in chromatin, and represent new drug targets for cancer and inflammation. Here we report that the ET domain of the BET protein BRD4 recognizes an amphipathic protein sequence motif through establishing a two-strand antiparallel β-sheet anchored on a hydrophobic cleft of the three-helix bundle. This structural mechanism likely explains BRD4 interactions with numerous cellular and viral proteins such as Kaposi's sarcoma-associated herpesvirus (KSHV) latency–associated nuclear antigen (LANA), and NSD3 whose interaction with BRD4 via this ET domain mechanism is essential for acute myeloid leukemia (AML) maintenance.

eTOC Blurb

Correspondence: ming-ming.zhou@mssm.edu.
[5]These authors contributed equally to this work.

AUTHOR CONTRIBUTIONS

Q.Z., C.R.V. and M.M.Z. designed experiments. Q.Z., Y.J., C.C.Z., and T.K. performed molecular biology and protein biochemistry experiments and conducted protein/peptide binding study, L.Z. determined the protein/peptide structures by NMR, C.S. performed cellular studies. All authors contributed to the writing of the manuscript.

COMPETING INTERESTS STATEMENT

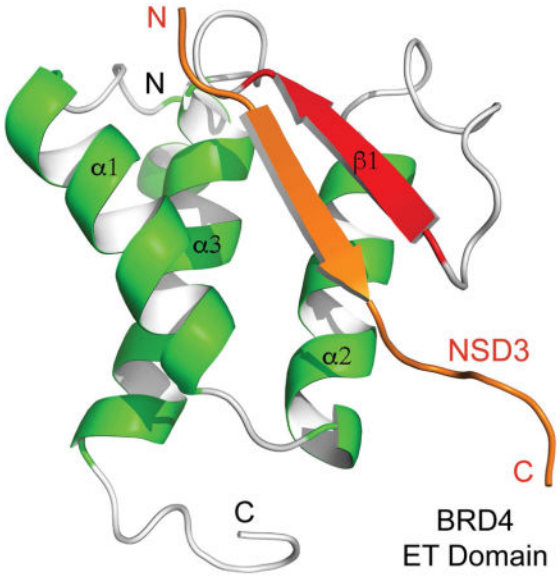
The authors declare no competing financial interests.

ACCESSION NUMBERS

Protein Data Bank: Coordinates for the solution structures of the BRD4 ET domain in complex with NSD3_1, LANA, and NSD3_3 peptides are deposited at Protein Data Bank under PDB ID code 2NCZ, 2ND0, and 2ND1, and the NMR spectral data are deposited at BioMagResBank (BMRB) under BMRB accession numbers 26041, 26402, and 26043, respectively.

SUPPLEMENTAL INFORMATION

Supplementary information includes six figures and two tables can be found with this article online at https://.



Zhang et al. report that the ET domain of BRD4 recognizes an amphipathic motif through a two-strand antiparallel β-sheet. This structural mechanism explains BRD4 association with cellular and viral proteins including Kaposi's sarcoma-associated herpesvirus LANA, and NSD3 whose interaction with BRD4 via the ET domain is essential for AML maintenance.

INTRODUCTION

The BET protein family, consisting of BRD4, BRD3, BRD2, and testis-specific BRDT, are characteristic of two bromodomains (BrDs) followed by the ET domain (Smith and Zhou, 2015; Yang et al., 2005) and function to ensure ordered gene transcription in chromatin (Jang et al., 2005; Schroder et al., 2012; Yang et al., 2005). Specifically, BRD4 functions via association with transcription factors to target gene enhancer and promoter sites, promote the assembly of *cis*-regulatory enhancer elements, and trigger activation of paused RNA polymerase II for productive gene transcription (Chiang, 2009; Roe et al., 2015; Shi et al., 2014; Zhang et al., 2012), all of which relies on the acetyl-lysine binding activity of its BrDs (Dhalluin et al., 1999). The ET domain function as a protein interaction domain was first suggested in human BRD2 and BRD4 interactions with the C-terminal segment of KSHV LANA that facilitates KSHV episome integration into host chromatin for KSHV latency (Ottinger et al., 2006a), and later shown to be important for BRD4 binding to numerous cellular transcription regulators including NSD3, ATAD5, CHD4, GLTSCR1 and JMJD6 (Rahman et al., 2011; Shen et al., 2015). While the structural basis of BRD4 ET domain binding to murine leukemia virus (MLV) integrase, required for viral/host integration (Aiyer et al., 2014; Larue et al., 2014), was recently reported (Crowe et al., 2016), the consensus ET domain binding motif and specificity in cellular proteins are yet to be defined.

Many protein interaction domains in regulation of gene transcription in chromatin can have moderate binding affinity for their biological ligands when measured *in vitro* and in isolation. Therefore, we reasoned that to determine truly consensus binding motif and

specificity of the ET domain, we would need to define conserved features of molecular recognition from both the protein domain and its biological targets. Such features likely manifest in ET domain recognition of multiple different biological ligands that can unravel from structural and biochemical analysis, and shall be validated experimentally in a functional context. In this study, we report the three-dimensional structures of the ET domain of BRD4 in complex with peptides derived from KSHV LANA and NSD3, and describe the consensus ET domain binding motif identified in numerous cellular proteins associated with gene transcription in chromatin.

RESULTS AND DISCUSSION

Defining BRD4 ET Domain Binding Motif in the C-Terminal Region of LANA

To determine the molecular basis of protein recognition by the ET domain, we first set out to define a minimal region in the LANA C-terminal segment (Figure 1A) that was suggested to interact with the BRD4 ET domain (Ottinger et al., 2006b; You et al., 2006), and confirmed by our 2D ^{15}N-HSQC NMR spectroscopy study (Figure 1B). Our observation of very narrow ^{1}H chemical shift dispersion of merely 0.9 ppm (7.7–8.6 ppm) in the ^{15}N-HSQC spectrum of LANA C-terminal fragment (residues 1110–1162) indicates that this region is unstructured (Figure S1A). We thus constructed a BRD4 ET domain fused at its N-terminus to LANA comprising residues 1110–1162 via a flexible linker (GSGSGS), and determined by NMR analysis that the minimal ET binding site in LANA is confined to residues 1131–1149 (NLQSSIVKFKKPLPLTQPG) (data not shown). Notably, the ^{15}N-HSQC spectra of the ET domain displayed nearly identical protein backbone resonance perturbation patterns induced by adding a LANA segment consisting of residues 1110–1162 or 1131–1149 (Figure S1A), or a LANA peptide of residues 1133–1144 (Figure 1B, red *vs.* blue signals), thereby defining the ET domain binding site in LANA to residues 1133–1144. We further determined binding affinity of the ET domain to the LANA peptide by NMR titration to be approximately K_d = 635 μM (Figure S1A).

Structure of the BRD4 ET Domain in Complex with a LANA Peptide

We next solved the three-dimensional structure of the BRD4 ET domain bound to the LANA peptide (residues 1133–1144) using heteronuclear multidimensional NMR spectroscopy (Clore and Gronenborn, 1994) (Figure 1C and Table 1). The structure reveals that the viral peptide establishes a two-strand inter-molecular antiparallel β-sheet with the protein. The aromatic side chain of Phe1139 inserts into a hydrophobic pocket in the ET domain formed within the three-helix bundle whereas two flanking positively charged Lys1138 and Lys1140 form an electrostatic zip with the negatively charged Asp650, Glu651 and Glu653 of the protein (Figure 1D). From site-directed mutagenesis and NMR binding studies (Figure 1E, and Figures S2 and S3), we determined Lys1138 and Phe1139 of LANA and Ile652, Ile654, Glu651 and Glu653 of the ET domain to be critical for the BRD4 ET-LANA recognition, thereby establishing a putative motif of ΦKΦK (where Φ is a hydrophobic or aromatic residue) for the ET domain recognition.

Identification of the Consensus BRD4 ET Domain Binding Motif in NSD3

This ET domain binding site in LANA (QSSIV**K**FKKPLP, residues 1133–1144) has only minimal resemblance to what was defined recently in the MLV integrase (Crowe et al., 2016). The latter involves a much longer sequence of TWRVQRSQNPLKI**R**LTR (residues 389–405), which itself forms a β-hair and establishes a three-strand β-sheet with the ET domain (Crowe et al., 2016). Notably, our detailed structural analysis of these two complexes reveals that the minimal commonality of ligand binding by the ET domain involves the anchoring electrostatic recognition of a Lys or Arg sandwiched between two highly conserved Glu651 and Glu653 in β1 of the protein, which are reinforced by interactions of the flanking hydrophobic or aromatic residues with the residues in the hydrophobic core of the protein. From structure-based protein sequence analysis, we detected this motif or its closely related variants in many of the known ET domain binding proteins (Rahman et al., 2011; Shen et al., 2015) including NSD3, ATAD5, CHD4, GLTSCR1 and JMJD6, and also other transcriptional proteins such as ASH1L, BRG1, INO80, SETD3, HDAC4, CDK8 (Figures 2A and 2B, and Table S1). We confirmed that the peptides containing this motif from some of these proteins indeed bind to the ET domain of BRD4, as demonstrated using 2D ^{1}H-^{15}N HSQC NMR spectra with estimated binding affinities comparable or better than that of LANA (Figure 2B, and Figure S4).

To determine functional relevance of these interactions in biology, we focused our study on BRD4 ET domain binding to the NSD3 short isoform, which has recently been shown to be essential for acute myeloid leukemia (AML) maintenance by Shen and colleagues (Shen et al., 2015). We co-transfected HEK293T cells with GST-BRD4 ET domain and FLAG-NSD3-short isoform constructs carrying Ala mutations at three putative ET binding motifs (KI motifs) at K156/I157, K215/I216 and K598/I599 (Figure 2C). The GST-BRD4 ET domain pull-down experiment showed that only the K156/I157 site mutation (M1) caused a nearly complete loss of BRD4 ET/NSD3 binding, which is consistent with the prior mapping of the BRD4 ET-interacting region of NSD3 to the region of residues 100–263(Shen et al., 2015). This BRD4 ET domain's ability to bind NSD3 was abolished with introduction of double mutations of Glu651 and Glu653 to Ala, as demonstrated by 2D ^{1}H-^{15}N HSQC spectral titration of the ET domain with a NSD3 peptide (residues 152–163) (Figure 2D). These results help us conclude that the K156/I157 site in NSD3 represents the major binding site for the BRD4 ET domain.

Determination of the Structural Basis of NSD3 Recognition by the BRD4 ET Domain

We next solved the NMR structure of the BRD4 ET domain bound to the NSD3 peptide (EIKL**KI**TKTIQN, residues 152–163) (Figure S5A and Table 1) whose affinity to the BRD4 ET domain was determined by isothermal titration calorimetry (ITC) to be K_d = 140.5 μM (Figure S5B). The NMR structure confirms that Lys156 of the core binding motif indeed forms anchoring electrostatic interactions between Glu651 and Glu653 on the surface of the two-strand intermolecular β-sheet (Figure 2E). This core KI (156/157) binding is boosted by another set of electrostatic interactions of Lys154 with Glu653 and Asp655, as well as Leu155 binding to Ile622 and Ile654 in the hydrophobic core of the protein. In addition, Ile153 of NSD3 interactions with Tyr612 located at the N-terminus of α1 of the protein further contributes to an extended two-strand β-sheet in this complex (Figure 2E). Given that

the mutations of the hydrophobic core residues would likely disrupt the protein structure (Shen et al., 2015), we focused on protein surface residues and showed that point mutation of either Lys156 or Lys154 to Ala nearly abolished protein/peptide binding (Figure 2F, and Figure S5C), whereas Ala mutation of Leu155 and Ile157, which inset side chains into the hydrophobic core of the protein, only caused a noticeable reduction in binding to the protein. This data explain the binding results of the ET domain to its putative binding sites in various proteins (Figure 2B), and highlight the importance of the electrostatic nature of the ET domain binding to its effector proteins.

Our NMR structural analysis of the BRD4 ET domain in complex with another NSD3 peptide derived from the K599/I600 (M3) site (VVPKK**K**IKKEQVE, residues 594–606; K_d of ~500 μM as estimated by NMR) further reveals that its unfavorable Lys598 interactions with the hydrophobic core residues in the protein as compared to Leu155 at the K156/I157 (M1) site likely responsible for less favorable formation of an intermolecular β-sheet (Figures S6A and S6B). Finally, in a cDNA/shRNA rescue study, we observed that the K156A/I157A (M1) mutation rendered NSD3-short incapable of supporting the growth of MLL-AF9/Nras^{G12D} acute myeloid leukemia cells (Figure 2G). Collectively, these results confirmed the importance of this amphipathic feature of the ET domain binding motif KΦKΦ in NSD3 and likely in other cellular proteins.

Conclusions

In summary, our study provides the structural mechanism of NSD3 recognition by the ET domain of BRD4, which is functionally important for NSD3/BRD4 association in acute myeloid leukemia maintenance. The consensus binding motif revealed from our structural analysis is likely conserved for the ET domain of BRD4, and represents a fundamental function of the BET proteins in regulation of gene transcription in chromatin through protein-protein interactions with cellular transcriptional regulators as well as viral proteins such as LANA of KSHV. This conserved structural mechanism further provides a foundation in future development of small-molecule therapeutic agents for new treatment of Karposi's sarcoma and acute myeloid leukemia through interfering with KSHV episome maintenance and replication during latency and AML maintenance, respectively.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification

The histidine-tagged-YFP LANA fragment (residues 1110–1162) and the human histidine-tagged Brd4 ET domain (601–683) were subcloned into a modified pET 21b vector (Novagen) Nde I-Bam HI sites and modified pET-32a (Novagen) Bam HI-Xho I sites, respectively. The coding DNA of LANA-BRD4 ET domain was prepared by overlapping PCR method and sub-cloned into modified pET32a Bam HI-Xho I sites. The histidine-tagged LANA, the BRD4 ET domain and LANA-BRD4 ET domain were expressed in *E. Coli* BL21(DE3) cell induced with 0.3 mM isopropyl-β-D-thiogalactopyranoside at 18°C. The His-tagged LANA, BRD4 ET domain and LANA-BRD4 ET domain were purified with Hitrap^{TM} chelating column (GE Healthcare) followed by thrombin cleavage. The LANA fragment was further purified using a Superdex 75 column and a Mono S column (GE

Healthcare), and the BRD4 ET domain as well as LANA-BRD4 ET domain was purified using a Superdex 75 column. Uniformly ^{15}N- and $^{15}N/^{13}C$-labeled proteins were prepared from cells grown in the minimal medium containing $^{15}NH_4Cl$ with or without $^{13}C_6$-glucose in H_2O. The mutants of the BRD4 ET domain were generated using QuikChange site-directed mutagenesis kit (Agilent Technologies), and the presence of appropriate mutations was confirmed by DNA sequencing.

Protein Structure Determination by NMR

NMR samples of the Brd4 ET domain (0.5 mM) in complex with a LANA peptide (residues 1131–1149) or LANA fragment (residues 1110–1162) of 0.5 mM were prepared in 100 mM phosphate buffer (pH 6.5) containing 5 mM perdeuterated DTT and 0.5 mM EDTA in $H_2O/^2H_2O$ (9/1) or ^2H_2O. All NMR spectra were collected at 30°C on NMR spectrometers of 800, 600, or 500 MHz. The ^1H, ^{13}C, and ^{15}N resonances of a protein of the complex were assigned by triple-resonance NMR spectra collected with a $^{13}C/^{15}N$-labeled and 75% deuterated protein bound to an unlabeled peptide(Clore and Gronenborn, 1994). The distance restraints were obtained in three-dimensional ^{13}C- or ^{15}N-NOESY spectra. Slowly exchanging amides, identified in two-dimensional ^{15}N-HSQC spectra recorded after a H_2O buffer was changed to a ^2H_2O buffer, were used with structures calculated with only NOE distance restraints to generate hydrogen-bond restraints for final structure calculations. The intermolecular NOEs were detected in ^{13}C-edited (F_1), $^{13}C/^{15}N$-filtered (F_3), three-dimensional NOESY spectrum.

Structure Calculations

Structures of the Brd4 ET/LANA complex were calculated with a distance geometry-simulated annealing protocol using the X-PLOR program (Brünger AT, 1998) Manually assigned NOE-derived distance restraints were used to calculate initial structures. ARIA (Nilges and O'Donoghue, 1998) assigned distance restraints agree with structures calculated using only the manually determined NOE restraints. Ramachandran plot analysis of the final structures was performed using Procheck-NMR program (Laskowski et al., 1996).

Identification of Putative ET Domain Binding Motif in Cellular Proteins

ScanProsite program (http://prosite.expasy.org) was used to identify putative ET domain binding sites of the ΦKΦKΦ motif (where Φ is a bulky hydrophobic amino acid of Leu, Ile, Val or Met) in 684 cellular proteins. NSD2/3 and CHD4 and other epigenetic proteins that were previously reported as Brd4 ET domain associated proteins by Rahman *et al.* (Rahman et al., 2011) are amongst the list. A selected list of the candidate proteins are shown in Figure 2B and Table S1. The former were further evaluated in NMR binding study using 2D ^{15}N-HSQC spectra.

Measurement of Protein/Ligand Binding Affinity by NMR

NMR titrations of the BRD4 ET domain with peptide ligands were followed with 2D ^1H-^{15}N HSQC spectra. Chemical shift perturbation was calculated using the following equation: $\Delta\delta_{\mathrm{ppm}}=[(\Delta\delta_{\mathrm{HN}})^2+(\Delta\delta_{\mathrm{N}}/5)^2]^{1/2}$, where $\Delta\delta_{HN}$ and $\Delta\delta_N$ are the chemical shift changes in the proton and nitrogen dimensions, respectively (Tochio et al., 1999).

Dissociation constants for the Brd4 ET domain interaction with LANA or NSD3 peptides were determined by a global fit of the chemical shift changes assuming a one-site binding model with Igor (WaveMetrics, Inc.). The one-site binding model assumes $\Delta\delta_{ppm} = \Delta\delta_{FB}\{(K_d + P + L) - [(K_d + P + L)^2 - 4PL]^{1/2}\}/(2P)$, where $\Delta\delta_{FB}$ is the chemical shift difference between the free and bound states, K_d is the dissociation constant, P and L are the total concentrations of the protein and peptide, respectively.

Isothermal Titration Calorimetry

Experiments were carried out on a MicroCal auto-ITC_{200} instrument at 20°C while stirring at 750 rpm in the ITC buffer of pH 6.5, consisting of 50 mM sodium phosphate, 2 mM EDTA and 2 mM β-mercaptoethanol as described previously (Gacias et al., 2014; Zhang et al., 2012). Peptide concentration was determined by weight and confirmed by NMR, and protein concentrations by A_{280} measurements. The protein sample (0.5 mM) was placed in the cell, whereas the micro-syringe was loaded with a peptide (7.5 mM) in the ITC buffer. The titrations were conducted using 17 successive injections of 2.4 μL (the first at 0.4 μL and the remaining 16 at 2.4 μL) with a duration of 4 sec per injection and a spacing of 180 sec between injections. The collected data was processed using the Origin 7.0 software program (OriginLab) supplied with the instrument according to the “one set of sites” fitting model.

Cell Culture and Plasmids

Murine AML cell line RN2 was derived from a MLL-AF9/$Nras^{G12D}$ transplantation-based animal model and cultured in RPMI1640 supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin (Zuber et al., 2011). HEK293T and ecotropic Plat-E viral packaging cells were cultured in DMEM supplemented with 10% FBS and 1% penicillin/streptomycin. All retroviral packaging was performed with Plat-E cells according to established procedures(Morita et al., 2000). For shRNA-based competition assays in RN2 cells, the LMN-mCherry shRNA retroviral vectors were used (MSCV-miR30-shRNA-PGKp-NeoR-IRES-mCherry). For retroviral and transfection-based expression of NSD3-short, a PIG vector was used containing the human NSD3-short cDNA with a C-terminal 3XFLAG (Shen et al., 2015, Molecular Cell). Point mutations were generated by overlap PCR. For bacterial expression of GST-ET domain, cDNA sequences were PCR cloned into a pGEX-4T1 vector (#28-9545-49; GE Healthcare). All of the cloning procedures were performed using the In-Fusion cloning system (#638909; Clontech)

Immunoprecipitation

GST tagged BRD4 ET domain (3 ug), purified from *E. coli,* were incubated with GST-sepharose 4B beads (#17-0756-01; GE Healthcare) for 2 hours at 4°C. Nuclear extracts were prepared from HEK293T cells after transient transfection with an MSCV vector expressing FLAG tagged NSD3-short wild-type or mutant for 48 hours. PBS-washed cell pellets were resuspended in Buffer A2 (10 mM Hepes-KOH pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl) and incubated on ice for 30 minutes to allow hypotonic cell membrane lysis. Nuclei were spun down at 4900 rcf for 5 minutes and resuspended in Buffer C2 (20 mM Hepes-KOH pH 7.9, 25% glycerol, 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA) and incubated on ice for 30 minutes. Samples were then centrifuged at 18000 rcf for 10 minutes and the supernatant

(nuclear extract) was then diluted with Buffer C2_No salt (20 mM Hepes-KOH pH 7.9, 25% glycerol, 1.5 mM $MgCl_2$, 0.2 mM EDTA) to reduce NaCl concentration to 150 mM. 1 mg nuclear extract was incubated with immobilized GST-sepharose 4B beads for 2 hours at 4°C. After three washes with 1 ml TBS (50 mM Tris-Cl, pH 7.5, 150 mM NaCl) plus 0.5% NP-40 and one additional wash with TBS (no NP-40), material was eluted from beads by adding Laemmli Sample Buffer (#161-0737, BIO-RAD) and boiling at 95°C for 5 minutes.

Competition Assay to Measure Cell Proliferation

To evaluate effects of cDNAs on shRNA-induced phenotypes in leukemia cells, RN2 cells were first retrovirally transduced with PIG (empty or with a cDNA), followed by puromycin (1 μg/ml) selection for 3–7 days. Subsequently, LMN-shRNAs-mCherry vectors were retrovirally transduced and the GFP+mCherry+ double positive population of cells were tracked over time using a BD LSR II flow cytometer.

Whole Cell Lysate Preparation and Western Blotting

500,000 live cells were collected and lysed by 2x Laemmli Sample Buffer, supplemented with Beta-mercaptoethanol. Cells were resuspended using 1 ml syringe and 26 ½ gauge needle until smooth and then denatured at 95 degree for 7 min. About 10% of extract was loaded into each well. Protein extracts were resolved by SDS-polyacrylamide gel electrophoresis and then transferred to nitrocellulose membrane for immunoblotting.

RT-qPCR

Total RNA was extracted from PBS-washed cell pellets using TRIzol reagent (Invitrogen) following the manufacturer's instructions. DNase I treatment was performed to eliminate contaminating genomic DNA after RNA isolation. cDNA was synthesized using the Q-Script cDNA SuperMix (Quanta BioScience), followed by qPCR with SYBR green (ABI) on an ABI 7900HT. All results were quantified using the delta Ct method with Gapdh as the control gene for normalization.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Kenneth M. Kaye of Department of Medicine, Harvard Medical School for kindly providing pSG5 F-LANA cDNA, and acknowledge the access to the New York Structural Biology Center's NMR facilities. This work was supported by grants from the National Institutes of Health (C.R.V. and M.-M. Z.)

References

Aiyer S, Swapna GV, Malani N, Aramini JM, Schneider WM, Plumb MR, Ghanem M, Larue RC, Sharma A, Studamire B, et al. Altering murine leukemia virus integration through disruption of the integrase and BET protein family interaction. Nucleic Acids Res. 2014; 42:5917–5928. [PubMed: 24623816]

Brünger ATAP, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A

new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr. 1998; 54:905–921. [PubMed: 9757107]

Chiang C-M. Brd4 engagement from chromatin targeting to transcriptional regulation: selective contact with acetylated histone H3 and H4. F1000 Biol Rep. 2009; 1

Clore GM, Gronenborn AM. Multidimensional heteronuclear nuclear magnetic resonance of proteins. Methods in enzymology. 1994; 239:349–363. [PubMed: 7830590]

Crowe BL, Larue RC, Yuan C, Hess S, Kvaratskhelia M, Foster MP. Structure of the Brd4 ET domain bound to a C-terminal motif from gamma-retroviral integrases reveals a conserved mechanism of interaction. Proc Natl Acad Sci U S A. 2016

Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou MM. Structure and ligand of a histone acetyltransferase bromodomain. Nature. 1999; 399:491–496. [PubMed: 10365964]

Gacias M, Gerona-Navarro G, Plotnikov AN, Zhang G, Zeng L, Kaur J, Moy G, Rusinova E, Rodriguez Y, Matikainen B, et al. Selective Chemical Modulation of Gene Transcription Favors Oligodendrocyte Lineage Progression. Chem Biol. 2014

Jang MK, Mochizuki K, Zhou M, Jeong HS, Brady JN, Ozato K. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. Mol Cell. 2005; 19:523–534. [PubMed: 16109376]

Larue RC, Plumb MR, Crowe BL, Shkriabai N, Sharma A, DiFiore J, Malani N, Aiyer SS, Roth MJ, Bushman FD, et al. Bimodal high-affinity association of Brd4 with murine leukemia virus integrase and mononucleosomes. Nucleic Acids Res. 2014; 42:4868–4881. [PubMed: 24520112]

Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR. 1996; 8:477–486. [PubMed: 9008363]

Morita S, Kojima T, Kitamura T. Plat-E: an efficient and stable system for transient packaging of retroviruses. Gene therapy. 2000; 7:1063–1066. [PubMed: 10871756]

Nilges M, O'Donoghue S. Ambiguous NOEs and automated NOE assignment. Prog NMR Spectroscopy. 1998; 32:107–139.

Ottinger M, Christalla T, Nathan K, Brinkmann MM, Viejo-Borbolla A, Schulz TF. Kaposi's Sarcoma-Associated Herpesvirus LANA-1 Interacts with the Short Variant of BRD4 and Releases Cells from a BRD4- and BRD2/RING3-Induced G1 Cell Cycle Arrest. J Virol. 2006a; 80:10772–10786. [PubMed: 16928766]

Ottinger M, Christalla T, Nathan K, Brinkmann MM, Viejo-Borbolla A, Schulz TF. Kaposi's sarcoma-associated herpesvirus LANA-1 interacts with the short variant of BRD4 and releases cells from a BRD4- and BRD2/RING3-induced G1 cell cycle arrest. J Virol. 2006b; 80:10772–10786. [PubMed: 16928766]

Rahman S, Sowa ME, Ottinger M, Smith JA, Shi Y, Harper JW, Howley PM. The Brd4 extraterminal domain confers transcription activation independent of pTEFb by recruiting multiple proteins, including NSD3. Molecular and cellular biology. 2011; 31:2641–2652. [PubMed: 21555454]

Roe JS, Mercan F, Rivera K, Pappin DJ, Vakoc CR. BET Bromodomain Inhibition Suppresses the Function of Hematopoietic Transcription Factors in Acute Myeloid Leukemia. Mol Cell. 2015; 58:1028–1039. [PubMed: 25982114]

Schroder S, Cho S, Zeng L, Zhang Q, Kaehlcke K, Mak L, Lau J, Bisgrove D, Schnolzer M, Verdin E, et al. Two-pronged binding with bromodomain-containing protein 4 liberates positive transcription elongation factor b from inactive ribonucleoprotein complexes. J Biol Chem. 2012; 287:1090–1099. [PubMed: 22084242]

Shen C, Ipsaro JJ, Shi J, Milazzo JP, Wang E, Roe JS, Suzuki Y, Pappin DJ, Joshua-Tor L, Vakoc CR. NSD3-Short Is an Adaptor Protein that Couples BRD4 to the CHD8 Chromatin Remodeler. Mol Cell. 2015; 60:847–859. [PubMed: 26626481]

Shi J, Wang Y, Zeng L, Wu Y, Deng J, Zhang Q, Lin Y, Li J, Kang T, Tao M, et al. Disrupting the Interaction of BRD4 with Diacetylated Twist Suppresses Tumorigenesis in Basal-like Breast Cancer. Cancer Cell. 2014; 25:210–225. [PubMed: 24525235]

Smith SG, Zhou MM. The Bromodomain: A New Target in Emerging Epigenetic Medicine. ACS Chem Biol. 2015

Tochio H, Zhang Q, Mandal P, Li M, Zhang M. Solution structure of the extended neuronal nitric oxide synthase PDZ domain complexed with an associated peptide. Nat Struct Biol. 1999; 6:417–421. [PubMed: 10331866]

Yang Z, Yik JH, Chen R, He N, Jang MK, Ozato K, Zhou Q. Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. Mol Cell. 2005; 19:535–545. [PubMed: 16109377]

You J, Srinivasan V, Denis GV, Harrington WJ Jr, Ballestas ME, Kaye KM, Howley PM. Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen interacts with bromodomain protein Brd4 on host mitotic chromosomes. J Virol. 2006; 80:8909–8919. [PubMed: 16940503]

Zhang G, Liu R, Zhong Y, Plotnikov AN, Zhang W, Zeng L, Rusinova E, Gerona-Nevarro G, Moshkina N, Joshua J, et al. Down-regulation of NF-kappaB transcriptional activity in HIV-associated kidney disease by BRD4 inhibition. J Biol Chem. 2012; 287:28840–28851. [PubMed: 22645123]

Zuber J, McJunkin K, Fellmann C, Dow LE, Taylor MJ, Hannon GJ, Lowe SW. Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. Nature biotechnology. 2011; 29:79–83.

- 3D structures of the BRD4 ET domain in complex with NSD3 and LANA peptides.
- The ET domain consists of a conserved three-helix bundle fold.
- The ET domain recognizes an amphipathic motif in a two-strand antiparallel β-sheet.
- The ET domain interacts with numerous transcription-associated proteins.

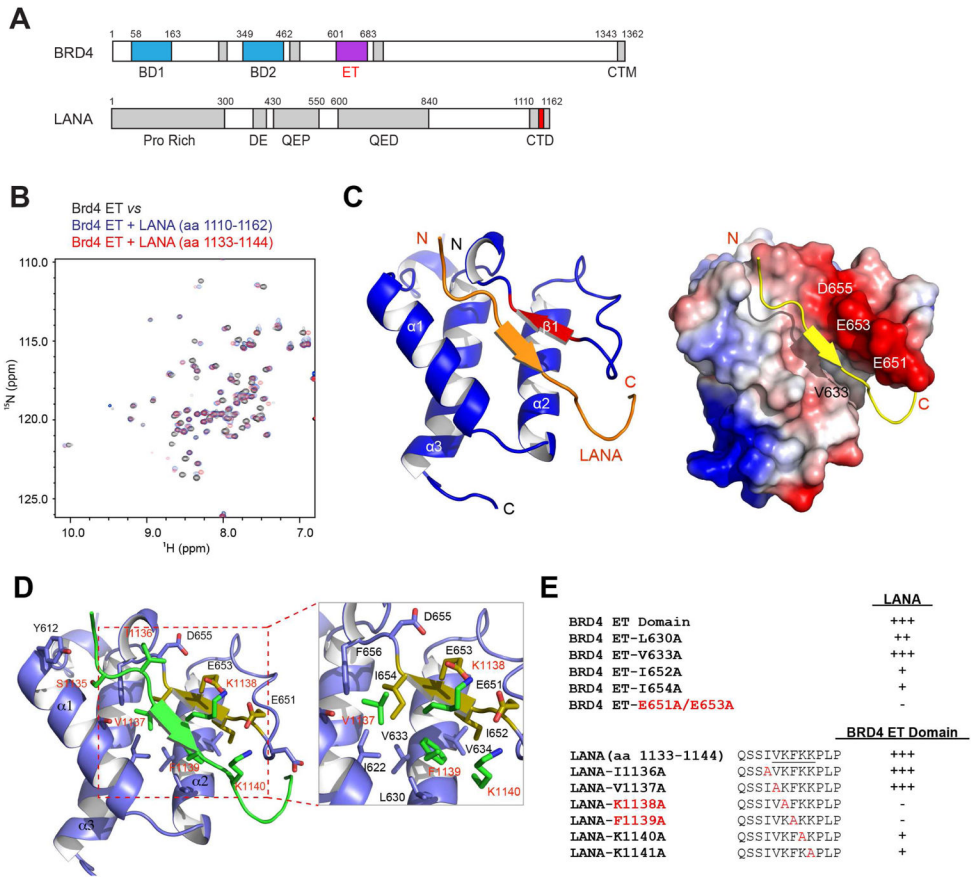


Figure 1. NMR structural analysis of the BRD4 ET domain recognition of LANA

(**A**) Protein domain organization of BRD4 and LANA. Domain annotation is described as follows: Pro rich, Proline-rich region; DE, aspartic acid and glutamic repeat region; QEP, glutamine, glutamic acid and proline-rich repeat; QDE, glutamic, aspartic acid and glutamic acid-rich repeat; BD I, bromodomain I; BD II, bromodomain II; The C-Terminal fragment of LANA and the ET domain of BRD4 are highlighted in red and light green, respectively.

(**B**) Identification of the ET domain binding site in the C-terminal region of LANA by NMR. 2D ^{1}H-^{15}N HSQC spectra of the BRD4 ET domain in the free state (black) or in the presence of peptides derived from the C-terminal region of LANA (blue and red).

(**C**) Left, ribbon representation of average minimized NMR structure of the Brd4 ET domain/LANA peptide complex. Right, electrostatic potential surface representation of the Brd4 ET bound to LANA peptide in cartoon depiction.

(**D**) LANA peptide binding site in the ET domain, showing key residues engaged in intermolecular interactions. The key residues involved in protein/peptide interactions are color-coded by atom type.

(**E**) Effects of mutations of individual residues in the ET domain or the LANA peptide on the protein/peptide binding, as assessed by NMR titration of ^{1}H-^{15}N HSQC spectra on a scale of 1 to 5 (as + sign) where 5 represents K_d <150 μM, 4 for K_d = 150–500 μM, 3 for K_d > 500 μM, and "−" for no binding. The binding affinity of the wild type ET domain and LANA peptide was determined by NMR titration to be K_d of 635 μM.

See also Figures S1, S2 and S3.

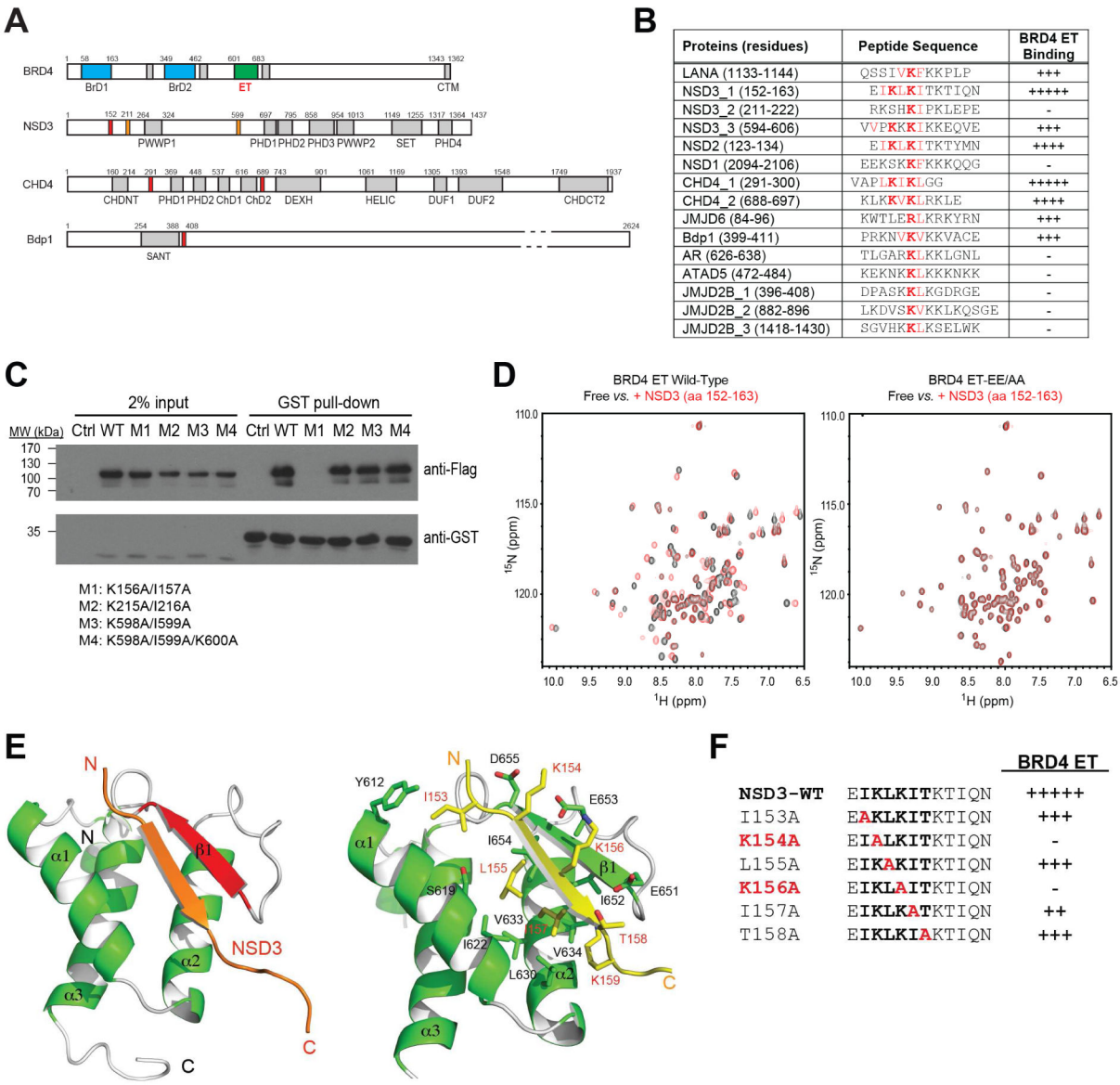


Figure 2. Structural mechanism of NSD3 recognition by the BRD4 ET domain

(A) Protein domain organization of BRD4, NSD3, CHD4 and Bdp1. Domain annotation is described similar to that in Figure 1a. The putative ET domain binding motifs in the epigenetic proteins are highlighted in red. The binding affinity of different peptides to the BRD4 ET domain was shown as assessed by NMR titration of ^{1}H-^{15}N HSQC spectra on a scale of 1 to 5 (as + sign), defined same as in Figure 1E.

(B) A list of putative ET domain binding motifs in selected epigenetic proteins assessed by using 2D ^{1}H-^{15}N HSQC NMR titration, as indicated.

(C) Identification of the BRD4 ET domain binding site in NSD3-short by GST-BRD4 ET domain pull-down assays evaluating interactions with various FLAG-NSD3-short constructs carrying different mutations, as indicated. GST-ET domain immobilized beads were incubated with HEK293T cells expressing the wild-type FLAG-NSD-short or corresponding

mutant fragments of M1, M2, M3, and M4 that contain double or triple Ala mutations of K156A/I157A, K215A/I216A, K598A/I599A, and K598A/I599A/K600A, respectively.
(D) 2D ^{1}H-^{15}N HSQC spectra showing the wild-type BRD4 ET domain (0.1 mM), but not the EE/AA mutant binding to a NSD3 peptide (residues 152–163) (0.5 mM).
(E) Ribbon representation of average minimized NMR structure of the BRD4 ET domain/NSD3 peptide complex, highlighting detailed intermolecular interactions at the protein/peptide interface.
(F) Effects of Ala-scan mutations of the NSD3 peptide on the BRD4 ET domain, as assessed by NMR titration of ^{1}H-^{15}N HSQC spectra on a scale of 1 to 5 (as + sign) where 5 is the highest affinity binding as the wild type.
See also Figures S4, S5 and S6.

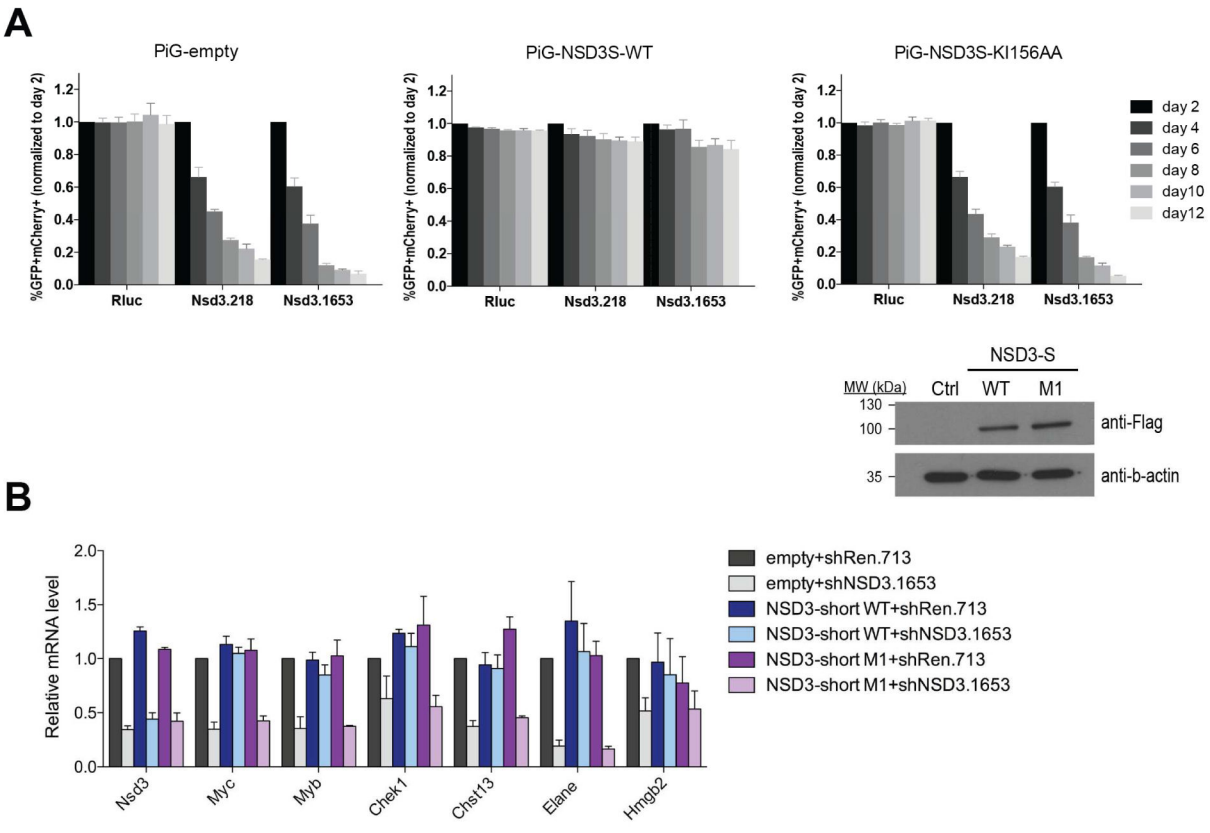


Figure 3. Functional analysis of NSD3 recognition by the BRD4 ET domain

(A) K156A/I157A (M1) impairs NSD3-short function in sustaining leukemia cell proliferation. Upper right insert, western blotting analysis of whole cell lysates prepared from RN2 cells transduced with the indicated PIG retroviral expression constructs. Competition-based assay tracking the abundance of GFP+mCherry+ cells during culturing of transduced RN2 cells. GFP is linked to the indicated cDNA and mCherry is linked to the indicated LMN shRNA. Plotted is the average of three independent biological replicates, normalized to d2. All error bars represent SEM for n=3.

(B) RT-qPCR analysis to evaluate effects of NSD3 knockdown on indicated gene expression in RN2 cells transduced with the indicated the FLAG-NSD3-short construct or an empty vector. Indicated TRMPV-Neo shRNAs were induced by dox for 48 hours. Results are normalized to *Gapdh*.

Table 1

Summary of Statistics of NMR Structures of BRD4 ET Domain/Peptide Complexes

	LANA	NSD3-1	NSD3-3
Protein NMR distance and dihedral constraints			
Distance constraints			
Total NOE	1633	1947	1834
Intra-residue	681	749	724
Inter-residue	952	1198	1110
Sequential (|i − j| = 1)	306	391	345
Medium-range (1 < |i − j| ≤ 5)	305	435	403
Long-range (|i − j| > 5)	341	372	362
Inter-molecular constraints	112	147	82 *
Hydrogen bonds	40	40	40
Total dihedral angle restraints			
Phi angle	71	71	71
Psi angle	71	71	71
Ramachandran Map Analysis (%)			
Most favored regions	98.1	98.1	92.3
Additional allowed regions	1.9	1.9	7.7
Generally allowed regions	0.0	0.0	0.0
Disallowed regions	0.0	0.0	0.0
Structure statistics			
Violations (mean ± s.d.)			
Distance constraints (Å)	0.038 ± 0.0043	0.038 ± 0.0047	0.041 ± 0.0014
Dihedral angle constraints (°)	0.25 ± 0.075	0.22 ± 0.083	0.59 ± 0.105
Max. dihedral angle violation (°)	0.40	0.41	0.74
Max. distance constraint violation (Å)	0.053	0.052	0.044
Deviations from idealized geometry			
Bond lengths (Å)	0.0037 ± 0.0001	0.0042 ± 0.0001	0.0044 ± 0.0001
Bond angles (°)	0.50 ± 0.011	0.58 ± 0.015	0.58 ± 0.010
Impropers (°)	1.2 ± 0.065	1.4 ± 0.062	1.5 ± 0.066

	LANA	NSD3-1	NSD3-3
Average pairwise r.m.s. Deviation [**] (Å)			
Heavy	0.65 ± 0.100	0.53 ± 0.074	0.50 ± 0.060
Backbone	0.26 ± 0.057	0.18 ± 0.041	0.20 ± 0.040

[a] Procheck calculation was done for protein residues 8–40, 53–76.

[b] The residue number ranges used in full molecule pairwise root-mean-square (r.m.s.) deviation calculations consists of 9–76

[c] Pairwise r.m.s. deviation was calculated among top 20/200 lowest energy structures.

[*] Of 82 inter-molecular NOEs, 34 NOEs are from residue V595 in the NSD3-3 peptide (VVPKKKIKKEQVE, residues 594-606).