A PWWP domain of histone-lysine N-methyltransferase NSD2 binds to dimethylated Lys36 of histone H3 and regulates NSD2 function at chromatin

Saumya M. Sankaran, Alex W. Wilkinson, and Or Gozani#

From the Department of Biology, Stanford University, Stanford, California 94305

Running title: NSD2 PWWP domain binds H3K36me2 and regulates NSD2 function

#To whom correspondence should be addressed: Prof. Or Gozani, Department of Biology, Stanford University, 371 Serra Mall, Stanford, CA. Tel.: 650-736-7639; Fax: 650-725-8309; e-mail: ogozani@stanford.edu.

Keywords: NSD2; MMSET; WHSC1; Histone Methylation; Chromatin; PWWP

ABSTRACT

The readout of histone modifications plays a critical role in chromatin-regulated processes. Dimethylation at Lys 36 on histone H3 (H3K36me2) is associated with actively transcribed genes, and global upregulation of this modification is associated with several cancers. However, the molecular mechanism by which H3K36me2 is sensed and transduced to downstream biological outcomes remains unclear. Here we identify a PWWP domain within the histone lysine methyltransferase and oncoprotein NSD2 that preferentially binds to nucleosomes containing H3K36me2. In cells, the NSD2 PWWP domain interaction with H3K36me2 plays a role in stabilizing NSD2 at chromatin. Furthermore, NSD2’s ability to induce global increases in H3K36me2 via its enzymatic activity, and consequently promote cellular proliferation, is compromised by mutations within the PWWP domain that specifically abrogate H3K36me2-recognition. Together, our results identify a pivotal role for NSD2 binding to its catalytic product in regulating its cellular functions, and suggest a model for how this interaction may facilitate epigenetic spreading and propagation of H3K36me2.

Chromatin dynamics play a critical role in the regulation of diverse cellular functions, the dysregulation of which is linked to the development and progression of human diseases. A major mechanism for regulating chromatin functional states involves the reversible covalent post-translational modification of histone proteins by chemical moieties such as methyl-, acetyl-, and phospho- groups. Histones provide a highly modifiable signaling surface on which these chemical marks combine to define particular chromatin states and regulate the extent of accessibility of DNA to trans-acting factors. In this context, the proteins and domains that recognize histone modification fundamentally influence DNA-templated processes, transducing molecular events at chromatin to biological outcomes. Protein lysine methylation is a principal chromatin-regulatory mechanism. The chemical addition of methyl moieties to lysine residues is catalyzed by lysine methyltransferases (KMTs). Lysine residues can accept up to three methyl groups forming mono-, di-, and tri-methylated derivatives (referred to as me1, me2, and me3, respectively). Histone methylation has been linked via methyllysine-binding proteins to diverse processes, including transcription, DNA recombination, DNA repair and DNA replication.

Methylation of histone H3 at lysine 36 (H3K36) is found at gene bodies of actively transcribed genes, but the state of methylation at this residue defines distinct biological outcomes. Trimethylation of this site (H3K36me3), which is mediated by the KMT SETD2 in humans, is involved in splicing regulation, RNA processing, and DNA damage signaling (1-7). Loss of SETD2 and H3K36me3 is a recurring phenomenon in clear cell renal cell carcinoma (ccRCC) and other cancers, suggesting a tumor suppressor role for SETD2 (8-10). In contrast, the specific molecular functions associated with H3K36me2 are unclear. However, elevated levels of this modification lead to aberrant activation of normally silenced genes (11), and upregulation of this mark is linked with...
numerous cancer types including acute myeloid leukemia, multiple myeloma, lung cancers, breast cancers, and glioblastomas (11-14). Thus, different states of methylation at H3K36 – dimethyl vs. trimethyl – are linked to dramatically different biological and disease-associated readouts.

The bulk of H3K36me2 in a number of cell types is generated by NSD2 (11) a KMT containing a conserved catalytic SET domain and several chromatin-associated domains comprising four PHD fingers, two PWWP domains, and an HMG box. This enzyme is implicated in diverse human diseases. NSD2 haploinsufficiency is associated with the developmental disorder Wolf Hirschhorn syndrome (WHS), which is characterized by growth and mental retardation, congenital heart defects, and antibody deficiencies (15). Indeed, NSD2-deficient mice exhibit a spectrum of defects resembling WHS (16). This enzyme is also implicated in the pathogenesis of the hematologic malignancy multiple myeloma (MM). 15-20% of MM patients carry a translocation between chromosomes 4 and 14 [t(4;14)(p16.3;q32)], which places the transcription of NSD2 under the control of strong IgH intronic Eµ enhancer and leads to aberrant upregulation of this gene (17, 18). Furthermore, a recent study in pediatric ALL cell lines and patient samples revealed a recurrent gain-of-function mutation (p.E1099K) in the NSD2 SET domain that confers increased catalytic efficiency of this enzyme, resulting in elevated levels of H3K36me2 (19). These and additional observations support a role for the NSD2 E1099K mutation as an epigenetic-mediated driver in the development of pediatric ALL. This mutation is also found in other cancers (20). Thus, NSD2 plays an important role during mammalian development and its overexpression or hyperactivity may cause cancer.

While dimethylation at H3K36 is associated with oncogenicity, the molecular mechanism underlying this outcome remains unknown. The readout of histone modifications is mediated by effector proteins containing modules that recognize modifications, often with great state- and sequence-specificity. Such “readers” have been identified for H3K36me3, providing direct links to RNA-processing (1, 2) and DNA damage (21). However, a reader domain that preferentially binds to the dimethyl state at H3K36 has yet to be identified.

In this study, we performed quantitative proteomic experiments with designer nucleosomes carrying installed methylation at K36. These experiments identified PWWP domains as possible binders of H3K36me2. We directly tested over a dozen PWWP domains for binding to H3K36 methylated nucleosomes. We identified the N-terminal PWWP domains of NSD2 as a preferential binder to H3K36me2. This interaction was found to be important for stable association of NSD2 with chromatin and NSD2 molecular and cellular functions. These results suggest a model in which the ability of NSD2 to bind to the mark it generates plays a role in spreading of H3K36me2 along chromatin as well as epigenetic propagation through cell division.

**EXPERIMENTAL PROCEDURES**

**Protein Sequences and Plasmids** – The following full-length and PWWP domain sequences were cloned into pGEX-6p-1 (GE Healthcare) for recombinant protein expression and purification with N-terminal GST fusions: HNRNP A1 (NCBI Accession # NP_002127.1), HNRNP AB (NCBI Accession # NP_112556.2), TEAD1 (NCBI Accession # NP_068780.2), PSIP1 PWWP (M1-N64; NCBI Accession # NP_001121689.1), HDGF2 PWWP (M1-G93; NCBI Accession # NP_057157.1), NSD1 PWWP (S304-D454 and R1696-K1874, respectively; NCBI Accession # NP_071900.2), NSD2 PWWP1 and NSD2 PWWP2 (P208-E368 and T818-K998, respectively; NCBI Accession # NP_001121689.1), MSH6 PWWP (D89-E192; NCBI Accession # NP_001164491.1), MUM1L1 PWWP (W406-K539; NCBI Accession # NP_001164491.1), N-PAC PWWP (M1-S114; NCBI Accession # NP_001164491.1), DNMT3A PWWP (G278-E427; NCBI Accession # NP_072046.2), DNMT3B PWWP (E206-P355; NCBI Accession # NP_0008823.1), MBD5 PWWP (S1371-R1494; NCBI Accession # NP_060798.2), MLL1 PWWP (W406-K539; NCBI Accession # NP_001164491.1), N-PAC PWWP (M1-S114; NCBI Accession # NP_001164491.1). Full-length NSD2 was cloned into Gateway pENTR 3C vector (Life Technologies) then recombined into mammalian expression vector pLenti6.2/V5-DEST (Life Technologies) or into pMSCVpuro (Clontech) modified with an N-terminal FLAG-myc affinity tag and a destination cassette (Life Technologies). Single point mutations were introduced into the
NSD2 PWWP domain binds H3K36me2 and regulates NSD2 function

NSD2 sequence by site-directed mutagenesis.

Cell Culture and Transfections – HeLa, HEK 293T, and HT1080 cells were cultured in DMEM (Life Technologies) supplemented with 10% fetal bovine serum (Gibco/Life Technologies), penicillin-streptomycin (Life Technologies), L-glutamine (Life Technologies), sodium pyruvate (Life Technologies), and MEM non-essential amino acids (Life Technologies). Viral transductions were performed as previously described (11) to generate HT1080 cells stably expressing NSD2 WT or derivative mutants. Cells transduced with pMSCVpuro-FLAG-myc-NSD2 were selected under puromycin (Sigma) at 2 µg/mL, and those with pLenti6.2/v5-DEST-NSD2 were selected under blasticidin (Life Technologies) at 10 µg/mL.

Antibodies – The antibodies used were: GST (generated at Covance), histone H3 (EpiCypher), NSD2 (EpiCypher), β-tubulin (Millipore, catalog no. 05-661), H3K36me1 (Cell Signaling Technology, catalog no. 14111S), H3K36me2 (Cell Signaling Technology, catalog no. 2901S), H3K36me3 (Cell Signaling Technology, catalog no. 4909S), and FLAG M2 (Sigma, catalog no. F1804). Antibodies against methylated H3K36 were validated by probing a custom peptide array as previously described (22).

Peptides, MLA Histones and Recombinant Nucleosomes – Histone H3(21-44) N-terminally biotinylated peptides were synthesized with unmodified, mono-, di- or tri-methylated lysine at K36 (ATKAARKSAPATGGVK meKPHRYRPG) at the Stanford Protein and Nucleic Acid facility. Recombinant X. laevis core histones were expressed, purified, and assembled into mononucleosomes with 5′ biotinylated 186 bp “Widom” 601 sequence DNA as previously described (23). To generate methylated full-length histones, single lysine-to-cysteine point mutations (K4C, K9C, K27C, K36C, K79C) were incorporated into the H3 sequence by site-directed mutagenesis. These H3 mutant constructs were recombinantly expressed, and mono-, di- or tri-methyl modifications were installed on the introduced cysteine residues using MLA chemistry, as previously described (24). Modified histones were verified by LC-ESI/MS with a MicrO-TOF-QII (Bruker) mass spectrometer.

SILAC protein pulldowns and quantitative mass spectrometry – HeLa cells grown in SILAC media with light amino acids (L-lysine/L-arginine, AppliChem) or heavy amino acids (13C15N-L-lysine/13C15N-L-arginine, Thermo Scientific) were used to prepare nuclear extracts as described (25, 26). Nucleosome pulldowns indicated in Figure 1A were performed as described previously (27) using 5 µg of the recombinant nucleosomes described above, 100 µl Dynabeads MyOne Streptavidin T1 (Thermo Fisher Scientific) slurry, and 500 µg nuclear extract per pulldown. Precipitated proteins were separated by SDS-PAGE, subjected to in-gel digest with trypsin (Promega) as previously described (26), and analyzed by LC-MS/MS on an Orbitrap Elite mass spectrometer. Raw mass spectra were analyzed using MaxQuant version 1.3.0.5 (28).

Recombinant protein preparation and pulldown assays – GST-fusion proteins were expressed in BL21 E.coli with 0.1 mM IPTG (Sigma), purified using glutathione-Sepharose 4B (GE Healthcare) beads, and eluted with reduced L-glutathione (Sigma). Protein concentrations were measured using the Coomassie Plus assay (Pierce). For direct pulldown assays, 1 µg biotinylated peptide or 5 µg biotinylated recombinant nucleosome was immobilized on 20 µl Streptavidin Sepharose High Performance beads (GE Healthcare), incubated with 5 µg recombinant protein in binding buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% (v/v) NP-40, 1mM PMSF) overnight at 4°C, and washed three times in buffer before analysis by Western blot.

Electromobility shift assays – Nucleosomes were incubated with recombinant GST-PWWP domains in EMSA buffer (20 mM Tris pH 8.0, 1 mM EDTA, 1 mM DTT, 10% glycerol) for 30 minutes at room temperature and analyzed by native 0.2x TBE PAGE. Each reaction contained 1 pmol of nucleosome with the indicated pmol of domain. Gels were stained with ethidium bromide, and bands were analyzed using an in-house iPython script. Intensity of NCP band was measured, and loss of free NCP was used to calculate fraction of NCP bound.

Cell Fractionation and Lysates – Biochemical fractionation was adapted from a protocol previously described (29). Briefly, 1x10^7 cells per cell line were collected, washed in PBS, and lysed in Buffer A (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM MgCl2, 0.34 M sucrose, 10%glycerol, 1 mM DTT, cOmplete protease

Downloaded from http://www.jbc.org/ by guest on December 5, 2017
inhibitor tablet (Roche)), containing Triton X-100 at a final concentration of 0.1%, for 10 min on ice. Cytoplasmic proteins were separated from nuclei by centrifugation at 1300g for 5 min at 4°C. Nuclei pellets were washed in Buffer A and lysed in Buffer B (3 mM EDTA, 0.2 mM EGTA, 1 mM DTT, complete protease inhibitor) for 30 min on ice, and soluble proteins were separated from chromatin by centrifugation at 1700g for 5 min at 4°C. Nuclear proteins were extracted by incubating nuclei pellets in Buffer B with 150 mM NaCl for 20 min on ice, followed by centrifugation at 1700g for 5 min at 4°C, and the supernatant was collected as the 150mM salt soluble nuclear fraction. The remaining nuclei pellets were incubated in Buffer B with 300 mM NaCl for 20 min on ice, followed by centrifugation at 1700g for 5 min at 4°C, and the resulting supernatant was collected as the 300mM salt soluble nuclear fraction. Chromatin pellets were resuspended in the SDS sample buffer and solubilized by sonication for 10 min in a Bioruptor (Diagenode).

For whole cell extracts, 1x10^7 cells per cell line were collected, washed in PBS, and lysed in a RIPA buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 2 mM EDTA, 1% NP-40, 0.1% SDS, 1 mM DTT, complete protease inhibitor tablet (Roche)) for 10 min on ice followed by sonication for 10 min in a Bioruptor. Samples were clarified by centrifugation at 21,000g for 10 min at 4°C. For normalization, protein concentrations were measured using the DC Protein Assay kit (BioRad).

**RESULTS**

Proteomic screen for candidate H3K36 methyl-binding proteins using modified nucleosome substrates – Reader domains of histone tail PTMs are often identified using peptides chemically synthesized with the modified residue of interest. However, based on the proximity of the histone H3 Lys36 residue to the nucleosome core, recognition of H3K36 modifications by reader domains may be influenced by structural components of the core. Indeed, the majority of H3K36me3 reader domains show stronger binding to substrates comprising both the histone modification and DNA – such as nucleosomes – than to the modification alone (30). We therefore reasoned that any potential H3K36me2-specific binding domain would likely also require reading the modification in a nucleosomal context. To test this hypothesis, we assembled recombinant nucleosomes (23) with histones modified by methyl lysine analog (MLA) chemistry (24) to install the different states of methylation at H3K36 (H3K36me1, H3K36me2, and H3K36me3). The nucleosome DNA was 5'-biotinylated to allow for high-affinity purification of these designer nucleosomes, which coupled with SILAC (Stable Isotope Labelling by Amino acids in Cell culture)-based quantitative proteomic screening can be used to search in an unbiased manner for proteins that preferentially bind to H3K36me2 (27).

We used this system to isolate proteins from nuclear extracts that bound differentially to H3K36me2 or H3K36me3 versus H3K36 unmodified nucleosomes (see experimental outline in Figure 1A). This analysis identified enrichment of several proteins binding to H3K36me2 nucleosomes (Figure 1B). We note that the majority of these hits were also found in the H3K36me3 pulldowns (Figure 1B & Supplementary Tables 1 & 2), and included proteins containing a PWWP domain, a known H3K36me3-binding module (30-32), several proteins in the HNRNP family, RAN-associated...
NSD2 PWWP domain binds H3K36me2 and regulates NSD2 function

proteins, and a few other proteins that did not segregate into a particular group. We tested the top hits for direct binding to H3K36me2/3 nucleosomes. The HNRNP family proteins HNRNPA1 and HNRNPAB, bound nonspecifically to substrates regardless of methyl state, while the transcriptional activator TEAD1 showed no binding to any of the substrates (Figure 1C). In contrast, we found that two PWWP-containing proteins, PSIP1/LEDGF and HDGF2, bound to H3K36 methylated nucleosomes (see Figure 2B). PSIP1 has previously been characterized to bind H3K36me3-containing nucleosomes, though its ability to bind to other states of methylation at H3K36 has not previously been tested(32).

Analysis of human PWWP domains binding to H3K36 MLA-modified nucleosomes – As previous studies of PSIP1 and other PWWP domains binding to H3K36me3-containing nucleosomes did not test H3K36me2-containing nucleosomes as a potential substrate, we undertook a candidate approach and screened the majority of human PWWP domains for preferential binding to H3K36me2 (Figure 2A). We performed binding assays using recombinant PWWP domains and mono-, di-, or tri-methylated H3K36 substrates in the form of biotinylated H3 tail peptides (amino acids 21-44) or the MLA nucleosomes (Figures 2B-D). As shown in Figure 2B, PWWP domains from the proteins PSIP1, DNMT3A, and DNMT3B bound preferentially to H3K36me3-containing nucleosomes as a potential substrate, while the transcriptional activator TEAD1 showed no binding to any of the substrates (Figure 1C). In contrast, we found that two PWWP-containing proteins, PSIP1/LEDGF and HDGF2, bound to H3K36 methylated nucleosomes (see Figure 2B). PSIP1 has previously been characterized to bind H3K36me3-containing nucleosomes, though its ability to bind to other states of methylation at H3K36 has not previously been tested(32).

Notably, our screen revealed two PWWP domains, an N-terminal PWWP of NSD2 (NSD2PWWP1) and a similar PWWP domain within NSD3 (NSD3PWWP1), which bound preferentially to H3K36me3 versus the other K36 methyl states and did so in a nucleosomal context but not on peptides (Figure 2B). In contrast, the first PWWP of NSD1 did not show the same specific binding as the other two NSD family members. An alignment of these three PWWP domains shows that the sequence flanking the canonical PWWP motif comprises aromatic and aliphatic residues that are conserved between NSD2 and NSD3 but contains basic residues in these positions on NSD1 (Figure 2A), which could significantly alter binding affinity for histone substrates. Like the first PWWP of NSD1, no binding or no methyl-specific binding was observed for the second PWWP domains found on NSD1, NSD2 and NSD3 (NSD1PWWP1, NSD2PWWP2 and NSD3PWWP2, respectively), or the PWWP domains from MBD5, MUM1L1 and N-PAC (Figure 2D). Together, our results identify PWWP domains found within NSD2 and NSD3, two H3K36 lysine dimethyltransferases (11, 35), as potential H3K36me2-binding modules. Here we focus our studies on the NSD2PWWP1 as NSD2 plays a clear role in the etiology of several cancers and is important for regulating the bulk of H3K36me2 in multiple cell types(11).

N-terminal PWWP domain of NSD2 preferentially binds H3K36me2-modified nucleosomes – Besides H3K36, histone H3 in humans bears four additional canonical lysine methylation sites: H3K4, H3K9, H3K27 and H3K79. We next tested the sequence specificity of NSD2PWWP1 in binding assays using MLA nucleosomes dimethylated at each of the five canonical lysines on H3. As shown in Figure 3A, NSD2PWWP1 bound to H3Kc36me2 but did not bind to the other dimethyl lysine sites. Based on these data we conclude that NSD2PWWP1 shows high specificity for H3K36me2, making this domain to our knowledge the first known reader to have specificity for H3K36me2.

To assess the binding specificity of NSD2PWWP1 for H3Kc36me2 versus H3Kc36me3 nucleosomes by an independent and more quantitative method, we characterized the PWWP-nucleosome interaction in electrophoretic mobility shift assays (EMSA) (Figure 3B). The apparent $K_d$ of the NSD2PWWP1-H3Kc36me2 interaction was $\sim 0.25 \mu M$ versus $\sim 0.57 \mu M$ for H3K36me3 nucleosomes (Figure 3C-D). As a control, we also analyzed binding of the known H3K36me3 reader PSIP1PWWP, which bound to H3K36me3 nucleosomes with an apparent $K_d$ of $\sim 0.67 \mu M$, similar to previous reports (32) and at $\sim 0.93 \mu M$ for H3Kc36me2 (Figure 3D, data not shown).
Taken together, we conclude that NSD2 PW1 is a reader domain of methylated H3K36, with preference for the dimethyl state.

The PW1-H3K36me2 interaction stabilizes NSD2 at chromatin – To explore the functional implications of the NSD2 PW1-H3K36me2 interaction, we first sought to identify point mutations within the PW1 domain that would specifically abrogate binding. Essential for all known methyllysine-binding domains is the presence of a hydrophobic, aromatic cage to encapsulate the modified residue (36). To identify residues within NSD2 PW1 that potentially form the aromatic cage and mediate H3K36me2 recognition, we aligned the sequence of this domain with that of PSIP1 PW1, for which a structure has been solved (32, 37) (Figure 4A). This analysis identified W236 and F266 on NSD2 as conserved aromatic residues that could form the H3K36 methyllysine-binding pocket. We found that substitution of either of these residues to an alanine abrogated NSD2 PW1 binding to H3K36me2 in direct nucleosome pulldown assays (Figure 4B).

We next investigated the role of H3K36me2-binding by NSD2 PW1 on the function of this enzyme in cells. As the major H3K36 dimethyltransferase in many cell types, NSD2 is generally tightly bound at chromatin (11). We therefore postulated that the NSD2 PW1-H3K36me2 interaction might regulate NSD2 stability at chromatin. To test this idea, we established HT1080 cell lines stably expressing either Flag-myc-NSD2 wild-type or the F266A mutant derivative. Lysates from these cells were used for biochemical fractionation to assess chromatin association using a modified Stillman fractionation protocol (see schematic Figure 4C; (29)). As shown in Figure 4D, the wild-type protein remained largely bound to chromatin through 300mM salt washes, whereas the mutant protein showed weaker chromatin association, with a more equal distribution between the high salt soluble nuclear fraction and the chromatin fraction. These data are consistent with a role for the NSD2 PW1-H3K36me2 interaction in stabilizing NSD2 at chromatin.

We next tested whether the ability of NSD2 to recognize H3K36me2 via its PW1 domain impacts target gene occupancy. Chromatin immunoprecipitation (ChIP) assays with anti-Flag coupled with quantitative PCR was used to assess binding of Flag-myc-NSD2 wild-type or F266A at the CDC42 gene, an NSD2 target identified by ENCODE (38-40). Enrichment of NSD2 signal was far greater at the promoter, TSS-proximal region, and 3' end terminus of CDC42 gene (Figure 4E, top panel) in cells expressing wild-type NSD2 compared to the F266A mutant derivative or the vector alone (Figure 4E, middle panel). Occupancy of the NSD2 H3K36me2-binding mutant was only slightly higher than the vector control (Figure 4E, middle panel) and largely similar to the signal seen with the IgG control ChIP from the three cell lines (Figure 4E, bottom panel). Similar trends were observed at TSS-proximal regions of two additional NSD2 target genes, MYC and TGFA (11) (Figure 4F and 4G, respectively). Together, these data support a role for the interaction between the first PW1 domain of NSD2 and H3K36me2 in stabilizing NSD2 at chromatin and regulating NSD2 occupancy at its gene targets.

Regulation of NSD2 cellular functions by the PW1-H3K36me2 interaction – The major molecular function of NSD2 is to generate the bulk of H3K36me2 in cells, and overexpression of NSD2 alone is sufficient to elicit changes in the global levels of H3K36me2 (11). To test if PW1-recognition of H3K36me2 impacted this NSD2 activity, we compared global levels of H3K36me2 in HT1080 cells moderately overexpressing NSD2 WT, two different PW1 mutants that abrogate H3K36me2-binding, a catalytically dead mutant (NSD2 Y1092A, a previously characterized catalytically dead mutant referred to here as NSD2CDM (11), or the vector alone as a control (Figure 5A). As expected, NSD2 expression specifically increased H3K36me2 levels relative to control and NSD2CDM (Figure 5A). Expression of the two PW1 mutants, NSD2 W236A and NSD2 F266A, also resulted in a global increase in H3K36me2 levels relative to wild-type NSD2 but were impaired in this activity relative to wild-type NSD2 (Figure 5A). Expression of the two PW1 mutants, NSD2 W236A and NSD2 F266A, also resulted in a global increase in H3K36me2 levels relative to control and NSD2CDM but were impaired in this activity relative to wild-type NSD2 (Figure 5A). No changes in H3K36me1 or H3K36me3 levels were observed, in keeping with previous studies of NSD2 regulating H3K36 dimethylation (11). Together, these results suggest that the PW1 domain does not directly impact NSD2 catalytic activity, but the ability to bind to H3K36me2 nonetheless regulates generation of this
NSD2 PWWP domain binds H3K36me2 and regulates NSD2 function

Modification in cells by a non-catalytic mechanism.

Overexpression of NSD2, in addition to causing a global increase in H3K36me2 levels, is also known to increase cellular proliferation rates (11, 39). To test the role of H3K36me2-recognition in this function, we established HT1080 cell lines stably expressing NSD2 and various mutants. As shown in Figure 5B, cells expressing wild-type NSD2 grew faster than the vector control treated cells. In contrast, cells expressing NSD2 derivatives harboring mutations in the PWWP domain that abrogate H3K36me2-binding did not increase the proliferation rate of the cells, displaying growth that was comparable to cells expressing the vector control and the catalytic dead mutant (Figure 5B and 5C). Taken together, these data demonstrate that the ability of the NSD2 N-terminal PWWP domain to bind to H3K36me2 is important for NSD2’s ability to, upon overexpression, generate H3K36me2 and promote cellular proliferation, two activities that are linked to the oncogenic potential of this epigenetic factor.

DISCUSSION

Here we report the N-terminal PWWP domain of the lysine methyltransferase NSD2 as a reader of H3K36me2. While other PWWP domains have been reported as H3K36me3 readers, this is the first instance of a methyllysine binding domain showing preference for the dimethyl state of H3K36 over the trimethyl state (Figures 3C & 3D). Our data further show that this interaction is not detectable on histone tail peptides alone, but occurs robustly and specifically in the physiologically relevant context of a nucleosome substrate (Figure 2C, Figure 3A).

Interestingly, the modification recognized by the NSD2 PWWP domain is the same one deposited by NSD2’s catalytic SET domain. We propose that the NSD2_{PWWP1} and NSD2_{SET} domains cooperate to recognize nucleosomes carrying H3K36me2 and deposit the same modification on neighboring nucleosomes (Figure 6A). Such a mechanism could lead to NSD2-mediated spreading of H3K36me2 across nucleosomes (Figure 6B, top) to generate regions of H3K36me2 occupancy. Propagation of a modification may also function in maintenance after DNA replication (Figure 6B, bottom), whereby the original H3K36me2-bearing nucleosomes that get distributed into the two daughter DNA strands serve as a template to fill in the modification on newly incorporated histones, thus reestablishing the regions of H3K36me2 occupancy in the daughter cells. Indeed, our results show that cells overexpressing NSD2 PWWP mutants do not generate as much global H3K36me2 as cells overexpressing wild-type NSD2 (Figure 5A). These data are consistent with a role for the PWWP-H3K36me2 interaction in propagating this modification.

The presence of modules that both recognize and “write” the same histone modification within a single protein or complex has previously been linked to propagation of two silencing histone modifications. Such a mechanism was first described for H3K9me3 (41-43), and later for H3K27me3 (44). In contrast, our study presents the first instance of such a connection between methyl reader and writer modules for a histone modification associated with actively transcribed genes, H3K36me2.

NSD2 has been characterized as a powerful driver of multiple myeloma and general oncogenic programming (11, 45). Specifically, this oncogenic potential is dependent on NSD2’s catalytic function and consequent upregulation of global H3K36me2 levels (11). This in turn mediates global activation of gene transcription and leads to aberrant cell growth through mechanisms yet to be elucidated. In our study, we note that overexpression of NSD2 PWWP mutants – whose catalytic domains are unaltered – does not lead to increased cell proliferation as does overexpressed wild-type NSD2. This suggests that the NSD2 PWWP-mediated propagation of H3K36me2 may be important in directing this modification to target loci, or creating regions of increased local concentration of H3K36me2, in order to transduce a downstream biological outcome. Elucidating the specific role of the PWWP-H3K36me2 interaction in stabilizing NSD2 at chromatin will be important in further understanding the function of NSD2 and the H3K36me2 modification in normal physiologic conditions as well as aberrant disease states. Furthermore, given the difficulty in generating inhibitors of the catalytic activity of NSD2, our results suggest an alternative strategy of blocking the PWWP-H3K36me2 interaction with small
molecule inhibitors that may lead to therapeutic options for NSD2-driven cancers.

Taken together, our data identify the N-terminal PWWP domain of NSD2 as the reader domain that preferentially binds to H3K36me2. We provide evidence that this molecular interaction is important for NSD2-driven upregulation of H3K36me2 levels and increased cell proliferation, both of which are implicated in this enzyme’s oncogenic potential.

ACKNOWLEDGMENTS
This work was supported in part by grants from the NIH to O.G. (R01 GM079641). A.W.W. was supported by NIH training grant (T32 GM007276). O.G. is a co-founder of EpiCypher, Inc.

AUTHOR CONTRIBUTIONS
S.M.S. performed the experiments and analyzed data. A.W.W. constructed the bacterial expression constructs, generated MLA histones, and provided technical assistance. O.G. supervised the work and S.M.S. and O.G. designed experiments and wrote the paper. All authors reviewed the results and approved the final version of the manuscript.
REFERENCES


NSD2 PWWP domain binds H3K36me2 and regulates NSD2 function

Mol. Genet. 7, 1071–1082


NSD2 PWWP domain binds H3K36me2 and regulates NSD2 function


FOOTNOTES
This work was supported in part by the National Institute of Health Grant to O.G. (R01 GM079641). A.W.W. was supported by NIH training grant (T32 GM007276).

The abbreviations used are: KMT, lysine methyltransferase; H3K36, histone H3 lysine 36; H3K36me2, dimethylated H3K36; H3K36me3, trimethylated H3K36; MLA, methyl lysine analog; PWWP, Pro-Trp-Trp-Pro conserved motif.

FIGURE LEGENDS

FIGURE 1. Proteome-wide screen for H3K36 methyl binding proteins using modified nucleosome substrates.
A, schematic of proteomic approach to screen for candidate H3K36 methyl binders from SILAC nuclear extract. B, enrichment of proteins bound to MLA H3K36me2 (top) or H3K36me3 (bottom) nucleosomes over unmodified nucleosome substrates. Identified proteins are plotted by SILAC ratio in the forward (x-axis) and reverse (y-axis) experiments, with candidate H3K36 methyl binders in the top right quadrant. Proteins enriched with ratio >1.8 (log2 ratio > 0.85) in the forward SILAC experiment and <0.55 (log2 ratio < -0.85) in the reverse SILAC experiment are labeled and categorized as PWWP domain proteins (orange), HNRNP family (cyan), RAN-associated (magenta), or Other (grey). C, Western blot analysis of direct pulldown assays testing binding of the indicated GST fusion proteins to H3 tail (21-44 aa) peptide substrates bearing unmodified (unmod), monomethyl (me1), dimethyl (me2) or trimethyl (me3) H3K36, or the indicated MLA H3K36 nucleosomes.

FIGURE 2. Candidate screen of human PWWP domains reveals NSD2PWWP1 and NSD3PWWP1 as H3K36 methyl readers.
A, sequence alignment of selected human PWWP domains. The PWWP motif is highlighted. Arrows indicate residues that are similar between NSD2 and NSD3 but not NSD1. “*”, conserved residue. “:”, similar residues. B,C,D, Western blot analysis of pulldown assays testing direct binding of the indicated PWWP domains as in Fig. 1C. H3K36me3-binding domains are grouped in B, domains with preferential binding to H3K36me2 are shown in C, and the non-binding or non-specific binding domains are grouped in D.

FIGURE 3. NSD2PWWP1 preferentially binds H3K36me2-modified nucleosomes.
A, Western blot analysis as in Fig. 2B of GST-NSD2PWWP1 binding to nucleosomes carrying dimethyl analogs at the indicated H3 lysine residues. B, titration of GST-NSD2PWWP1 against MLA H3K36me2 (top) or H3K36me3 (bottom) nucleosomes in electrophoretic mobility shift assays (EMSA). Bands indicated correspond to free nucleosome core particle (NCP) and nucleosome bound by one (+1) or two (+2) PWWP molecules. C, binding curves quantified from EMSA experiments of GST-NSD2PWWP1 on MLA H3K36me2 or H3K36me3 nucleosomes. Error bars represent s.e.m from three independent experiments. D, apparent binding affinity (KdAPP) for H3K36me2 and H3K36me3 nucleosomes calculated from binding curves for GST-NSD2PWWP1 in C and for GST-PSIP1PWWP (data not shown). Error indicates s.e.m from three independent experiments.

FIGURE 4. PWWP-H3K36me2 binding stabilizes NSD2 chromatin association.
A, sequence alignment of PSIP1PWWP and NSD2PWWP1. Residues involved in the PSIP1PWWP H3K36me3-binding aromatic cage are highlighted. Boxes indicate residues chosen for mutational analysis. B, mutation of the indicated aromatic cage residues abrogates NSD2PWWP1 binding to H3K36me2 in direct pulldown assays as in Fig. 1C and Fig. 2B-D. C, schematic of modified Stillman fractionation method to biochemically separate chromatin-associated proteins from soluble nuclear proteins in cellular extracts. D, Western blot analysis of biochemical fractions defined in C from HT1080 cells expressing full-length
NSD2 PWWP domain binds H3K36me2 and regulates NSD2 function

FLAG-myc-NSD2 wild-type or F266A PWWP1 mutant shows altered chromatin association for the PWWP1 mutant derivative. Tubulin and H3 blots serve as controls for intact fractions. WCE, whole cell extract. nuc 150, 150mM salt soluble nuclear fraction. nuc 300, 300mM salt soluble nuclear fraction. chromatin, chromatin fraction. E, ChIP analysis of FLAG-myc-NSD2 wild-type or F266A occupancy across CDC42 gene. A schematic of the gene (top panel) indicates the location of primers used for qPCR analysis of FLAG (middle panel) or IgG control (bottom panel) ChIP from each stable cell line. vector, HT1080 cells transduced with empty vector. WCE, whole cell extract. nuc 150, 150mM salt soluble nuclear fraction. nuc 300, 300mM salt soluble nuclear fraction. chromatin, chromatin fraction. E, ChIP analysis of FLAG-myc-NSD2 wild-type or F266A occupancy across CDC42 gene. A schematic of the gene (top panel) indicates the location of primers used for qPCR analysis of FLAG (middle panel) or IgG control (bottom panel) ChIP from each stable cell line. vector, HT1080 cells transduced with empty vector. Error bars indicate s.e.m. from three experiments. F and G, ChIP analysis as in E at TSS-proximal sites (analogous to CDC42 primer 2) of NSD2 gene targets MYC and TGFA, respectively.

FIGURE 5. PWWP-H3K36me2 interaction regulates NSD2 cellular functions.
A, Western blot analysis with the indicated antibodies on lysates from HT1080 cells expressing NSD2 wild-type or mutant derivatives. vector, HT1080 cells transduced with empty vector. CDM, catalytic dead mutant. B, proliferation assays using the cells described in A counted over 8 days. Error bars indicate s.e.m. from three experiments. C, cell counts from Day 8 of the proliferation assay in B. p values were calculated using a two-tailed Student’s t test. ***, p<0.01. n.s., not significant.

FIGURE 6. A model for propagation of H3K36me2 by NSD2.
A, the N-terminal PWWP domain and SET domain of NSD2 cooperate to recognize nucleosomes carrying H3K36me2 and deposit the same mark on neighboring nucleosomes, thus propagating this modification. B, by this mechanism of propagation, NSD2 mediates the spreading of H3K36me2 across a genomic region (top) or maintenance of H3K36me2 patterns in daughter cells after DNA replication, when parental histones have been distributed between daughter cells and new histones have been incorporated (bottom).
Figure 1

A) "light" SILAC cells → "heavy" SILAC cells

- nuclear extract
- MLA-nuc pulldowns
- pool

forward experiment → reverse experiment

mass spectrometry

- intensity
- m/z

nucleosome

- biotin tag
- histone tail
- methyllysine (Kme)

B) 

- Scatterplot w/ Boxplots
- PWWP domain proteins
- Other
- RAN-associated proteins

C) 

<table>
<thead>
<tr>
<th>Input 5% beads only</th>
<th>peptide</th>
<th>nucleosome</th>
</tr>
</thead>
<tbody>
<tr>
<td>HNRNPA1</td>
<td>GST</td>
<td>H3</td>
</tr>
<tr>
<td>HNRNPAB</td>
<td>GST</td>
<td>H3</td>
</tr>
<tr>
<td>TEAD1</td>
<td>GST</td>
<td>H3</td>
</tr>
</tbody>
</table>

- GST
- H3

Reverse H/L (log2) vs. Forward H/L (log2)
Figure 3

A

<table>
<thead>
<tr>
<th>MLA-H3 Nucleosomes</th>
<th>Beads only</th>
<th>WT</th>
<th>Kc4me2</th>
<th>Kc9me2</th>
<th>Kc27me2</th>
<th>Kc36me2</th>
<th>Kc79me2</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSD2 PWWP1</td>
<td>input 5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B

GST-NSD2 PWWP1

C

D

<table>
<thead>
<tr>
<th>MLA nuc</th>
<th>KdAPP (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSD2 PWWP</td>
<td>Kc36me2 0.25 ± 0.06, Kc36me3 0.57 ± 0.06</td>
</tr>
<tr>
<td>PSIP1 PWWP</td>
<td>Kc36me2 0.93 ± 0.04, Kc36me3 0.67 ± 0.05</td>
</tr>
</tbody>
</table>
Figure 4

A

PSIP1 (7) PGDLIFAKKGYHPARVDEPDGAV-------KPTNKLPIFFGTH-RTAFLGPKDIFPYSEN (64)
NSD2_PWWP1 (222) VGDLVWSKVSCYPWPMVSADPLLSYTKLGQIKSARQYHQQFEDAPRPAWIFEKSLVAPEG- (286)

B

<table>
<thead>
<tr>
<th></th>
<th>input 5%</th>
<th>nucleosome</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSD2 WT</td>
<td>GST</td>
<td>H3</td>
</tr>
<tr>
<td>NSD2 W236A</td>
<td>GST</td>
<td>H3</td>
</tr>
<tr>
<td>NSD2 F266A</td>
<td>GST</td>
<td>H3</td>
</tr>
</tbody>
</table>

C

lyse cells

cytoplasmic fraction

nuclei

nuclear extraction

150mM salt

insoluble nuclear fraction

300mM salt

soluble nuclear fraction

chromatin fraction

D

<table>
<thead>
<tr>
<th></th>
<th>NSD2 WT</th>
<th>NSD2 F266A</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCE 25%</td>
<td>nuc 150mM</td>
<td>nuc 300mM</td>
</tr>
<tr>
<td>nuc 150mM</td>
<td>nuc 300mM</td>
<td></td>
</tr>
<tr>
<td>chromatin</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NSD2

tubulin

H3

E

CDC42

1 2 3

40.3 kb

F

MYC

G

TGFA

FLAG (% input)

IgG (% input)

vector NSD2 WT NSD2 F266A

primer 1 primer 2 primer 3

primer 1 primer 2 primer 3

primer 1 primer 2 primer 3

primer 1 primer 2 primer 3
Figure 5

(A) Graph showing expression levels of NSD2, H3, and H3K36me1-3 in HT1080 cells.

(B) Graph showing the number of cells over time for different NSD2 mutants (WT, W236A, F266A, CDM).

(C) Bar graph comparing cell counts for the indicated NSD2 mutants (WT, W236A, F266A, CDM) with vector control. * indicates a significant difference (n.s., not significant).
Figure 6

A

Key:

- H3 tail
- H3K36me2 recognition
- H3K36me2 methyltransferase
- nucleosome
- new histones

B

1. Spreading

2. Epigenetic maintenance

DNA replication

new histones deposited
A PWWP domain of histone-lysine N-methyltransferase NSD2 binds to dimethylated Lys36 of histone H3 and regulates NSD2 function at chromatin

Saumya M. Sankaran, Alex W. Wilkinson and Or Gozani

J. Biol. Chem. published online February 24, 2016

Access the most updated version of this article at doi: 10.1074/jbc.M116.720748

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2016/02/24/M116.720748.DC1