

The C-terminal PHD_VC5HCH tandem domain of NSD2 is a combinatorial reader of unmodified H3K4 and tri-methylated H3K27 that regulates transcription of cell adhesion genes in multiple myeloma

Andrea Berardi^{1,†}, Charlotte Leonie Kaestner^{2,†}, Michela Ghitti^{1,†}, Giacomo Quilici¹, Paolo Cocomazzi³, Jianping Li^{2,4}, Federico Ballabio¹, Chiara Zucchelli¹, Stefan Knapp⁵, Jonathan D. Licht^{2,*} and Giovanna Musco^{1,*}

¹Biomolecular NMR Laboratory, Division of Genetics and Cell Biology c/o IRCCS Ospedale San Raffaele Via Olgettina 58, 20132 Milan, Italy

²Division of Hematology/Oncology, The University of Florida Health Cancer, 2033 Mowry Road, Gainesville, FL 32610, USA

³Biophysics Institute, CNR-IBF, Via Corti 12, 20133 Milan, Italy

⁴Department of Pharmacology, Physiology and Cancer Biology, Thomas Jefferson University, Philadelphia, PA 19107, USA

⁵Institute for Pharmaceutical Chemistry, Johann Wolfgang Goethe-University, 60438 Frankfurt am Main, Germany

*To whom correspondence should be addressed. Tel: +39 02 26434824; Fax: +39 02 26436352; Email: musco.giovanna@hsr.it

Correspondence may also be addressed to Jonathan Licht. Email: jdlicht@ufl.edu

†The first three authors should be regarded as Joint First Authors.

Present addresses:

Andrea Berardi, IFOM ETS, Fondazione AIRC per la Ricerca sul Cancro ETS, Milan, Italy.

Andrea Berardi, The AIRC Institute of Molecular Oncology | Via Adamello 16, 20139 Milan, Italy.

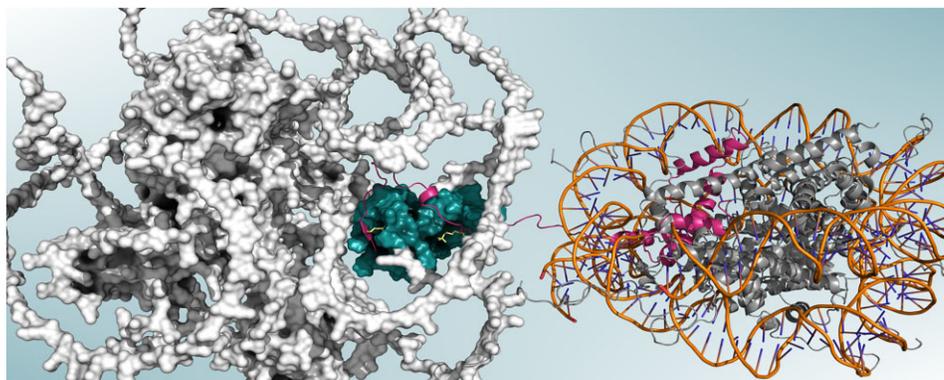
Giacomo Quilici, Centro di Servizi e Misure, University of Parma, Parco Area Delle Scienze, 23/A, Parma, Italy.

Federico Ballabio, Department of Biosciences, University of Milan, Via Celoria 26, 20133 Milan, Italy.

Abstract

Histone methyltransferase NSD2 (MMSET) overexpression in multiple myeloma (MM) patients plays an important role in the development of this disease subtype. Through the expansion of transcriptional activating H3K36me2 and the suppression of repressive H3K27me3 marks, NSD2 activates an aberrant set of genes that contribute to myeloma growth, adhesive and invasive activities. NSD2 transcriptional activity also depends on its non-catalytic domains, which facilitate its recruitment to chromatin through histone binding. In this study, using NMR, ITC and molecular dynamics simulations, we show that the tandem PHD domain of NSD2 (PHD_VC5HCH_{NSD2}) is a combinatorial reader of unmodified histone H3K4 and tri-methylated H3K27 (H3K27me3). This is the first PHD tandem cassette known to decode the methylation status of H3K27. Importantly, in a NSD2-dependent MM cellular model, we show that expression of NSD2 mutants, engineered to disrupt the interaction between H3K27me3 and PHD_VC5HCH, display in comparison to wild-type NSD2: incomplete loss of H3K27 methylation throughout the genome, decreased activation of adhesive properties and cell adhesion genes, and a decrease of the corresponding H3K27ac signal at promoters. Collectively, these data suggest that the PHD_VC5HCH domain of NSD2 plays an important role in modulating gene expression and chromatin modification, providing new opportunities for pharmacological intervention.

Graphical abstract



Received: August 22, 2024. Revised: September 27, 2024. Editorial Decision: October 22, 2024. Accepted: October 28, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the

original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other

permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Introduction

The Nuclear receptor-binding SET (Su(var) 3–9, Enhancer of zeste, Trithorax) domain protein 2 NSD2 (also known as multiple myeloma SET domain, MMSET, or Wolf-Hirschhorn syndrome candidate 1 WHSC1) is a member of the structurally conserved Nuclear Receptor Binding SET Domain (NSD) protein family, composed by the mammalian histone lysine methyl transferases (HKMT) NSD1, NSD2 and NSD3 (1,2). The NSD proteins catalyze via their SET domain the mono- and di-methylation of Lysine 36 of histone H3 (H3K36me1 and H3K36me2, respectively), which serves to exclude the Polycomb Repressive Complex 2 (PRC2) and its repressive H3K27me3 modification, thereby facilitating activation of gene expression (2,3). All three NSD members when aberrantly expressed, amplified or somatically mutated give rise to a variety of genetic diseases and different cancer types (1,2). In particular, the t(4;14)(p16;q32) chromosomal rearrangement, which places the NSD2 gene under the control of the immunoglobulin heavy chain promoter has been associated with a high risk form of multiple myeloma (MM) (4,5). This translocation induces aberrant over-expression of NSD2, that leads to a genome wide 6–8 fold increase of H3K36 dimethylation (H3K36me2) across the genome. In turn, this is associated to a global reduction of the repressive chromatin mark H3K27me3 (6–8), ultimately resulting in the de-repression of normally silenced genes and in the consequent activation of tumorigenesis pathways (7,9). Furthermore, NSD2 overexpression is accompanied by alteration of EZH2 (Enhancer of Zeste Homolog 2)-induced H3K27 methylation across the myeloma genome and misregulation of specific Polycomb target genes (such as *GATA3*, *HOXA2* genes), that contribute to myelomagenesis (7). In accordance with the oncogenic potential of NSD2, its knock-down in t(4;14)+ MM cell lines reduces proliferation, cell-cycle progression, and DNA repair (6,10–12). NSD2 tumorigenic potential is strongly driven by its histone methyltransferase activity, as also highlighted by a gain-of-function mutation (p.E1099K) in the NSD2 SET domain of pediatric Acute Lymphoblastic Leukemia (ALL) patients. This mutation enhances NSD2 catalytic efficiency, yielding increased global levels of H3K36me2 and reduced H3K27me3 (13). The molecular basis of NSD2 histone lysine methyltransferase (HKMT) activity by means of the SET domain has been unraveled by the cryo-electron microscopy structure of NSD2 in complex with nucleosome particles (14,15). However, NSD2 pathophysiological function not only depends on its catalytic chromatin ‘writing’ activity, but also relies on its effector domains. As typical member of the NSD family, NSD2 harbors several chromatin-associated domains, including 5 Plant homeodomain (PHD) finger domains, a PHD finger-like Cys–His rich domain (C5HCH), two proline-tryptophan-tryptophan-proline (PWWP) domains and a DNA-binding high mobility group box (HMG-box) domain (Figure 1A) (2). All these domains likely contribute to NSD2 recruitment to specific chromatin loci via direct binding to chromatin or via mediation of adaptor proteins. In fact, the N-terminal PWWP_I domain functions as preferential reader of NSD2 catalytic product, the H3K36me2 mark, supporting the propagation and spreading of H3K36me2 along chromatin (16). As such, PWWP_I has recently emerged as a promising pharmacological target to inhibit the aberrant function of NSD2 in MM (17,18). In addition, the PHD finger domains, versatile small Zn²⁺ binding domains, seem to sustain NSD2 oncogenic activity (7,9). Within

different sequence contexts PHD fingers mainly work in tandem or alone as chromatin readers, decoding various histone marks, including methylated, unmethylated and acetylated lysine residues (19,20). PHD-finger mediated misinterpretation of these specific epigenetic marks is often associated with a wide range of human pathologies, ranging from cancer to developmental diseases (20–22). NSD2 harbors three distinct NSD2 PHD cassettes, including three contiguous N-terminal PHD fingers (PHD_{I-II-III}), a PHD_{IV}PWWP_{II} tandem domain and a C-terminal PHD_VC5HCH tandem domain (Figure 1A). Both PHD_{I-II-III} and PHD_VC5HCH cassettes were shown to contribute to NSD2 tumorigenic potential in non-redundant ways. On the one hand PHD_{I-II-III} contributes to NSD2 recruitment to chromatin and HKMT activity, as a deletion mutant of this domain fails both to bind chromatin and cause the methylation switch between H3K36me2 and H3K27me3 (7,9). On the other hand, the PHD_VC5HCH tandem domain contributes to H3K27 hypomethylation, as its deletion induces incomplete loss of H3K27 methylation, resulting in only partial transcriptional locus de-repression and modest clonogenic growth stimulation (7,9). However, despite the relevant role of PHD fingers in recruiting NSD2 to oncogenic loci driving MM (9), their actual histone binding ability and specificity towards distinctive marks are still poorly explored at molecular level. In this study we examined the histone tail binding activity of the PHD_VC5HCH tandem domain of NSD2 (PHD_VC5HCH_{NSD2}) and elucidated the molecular and structural determinants underlying its ability to interpret the dual H3K4K27me3 epigenetic mark (i.e. unmethylated H3K4 and tri-methylated H3K27, defined from now on for simplicity as H3K27me3). Furthermore, we demonstrate that within the context of the full-length protein, PHD_VC5HCH_{NSD2} plays a functional role in modulating the cell adhesion properties of MM cells through the regulation of cell adhesion genes.

Material and methods

Protein expression and purification and synthetic peptides

NSD-PHD_VC5HCH tandem domains

Murine PHD_VC5HCH_{NSD1} (E2117–D2207) and human PHD_VC5HCH_{NSD3} (R1310–P1413) were expressed and purified as described (23). Three different constructs of human PHD_VC5HCH_{NSD2} domain (Uniprot code: o96o28-1, residues R1228-K1340 and R1228-A1331 for NMR studies; R1228-R1334 for X-ray studies) were cloned into pETM11-SUMO3 vector (EMBL) expressed and purified as described previously (24). In all these constructs C1324_{NSD2} (a non-Zn binding Cysteine) was mutated into Serine (C1324S_{NSD2}) to improve solubility. Site-directed mutagenesis of all the mutants analyzed by NMR were performed by standard overlap extension methods. The DNA constructs were sequenced by Eurofins (Milan, Italy).

Proteins concentrations were determined by UV spectroscopy using predicted extinction coefficients (ϵ) for PHD_VC5HCH_{NSD1} ($\epsilon=11\ 490\ \text{M}^{-1}\text{cm}^{-1}$), PHD_VC5HCH_{NSD2} ($\epsilon=13\ 490\ \text{M}^{-1}\text{cm}^{-1}$), PHD_VC5HCH_{NSD3} ($\epsilon=15\ 470\ \text{M}^{-1}\text{cm}^{-1}$),

Recombinant histone H3₁₋₄₄

The human histone H3 fragment containing residues 1–44 (H3₁₋₄₄) was cloned into pETM11-His₆-SUMO3 vector

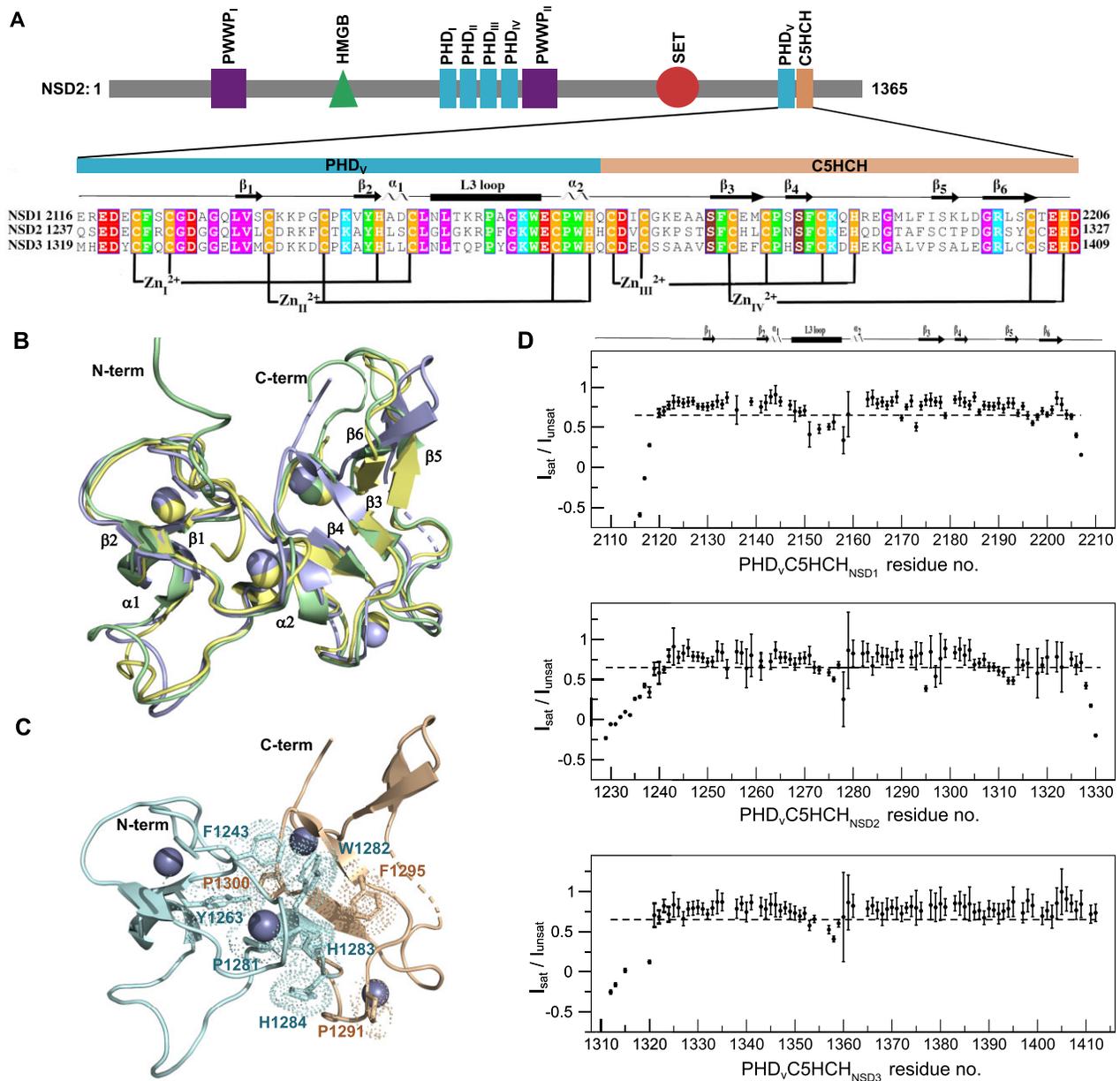


Figure 1. The three PHD_VC5HCH domains of the NSD family share similar structure and dynamics. **(A)** Domain architecture of NSD2. Multiple sequence alignment (performed with ESPrInt.3 (79)) of the human PHD_VC5HCH domain of the NSD family (NCBI codes, NSD1: NP_071900.2; NSD2: NP_579890.1; NSD3: NP_075447.1, conserved residues are colored. Secondary structure elements and Zn²⁺ coordination sites are reported. **(B)** Superimposition of PHD_VC5HCH_{NSD1} (PDB code: 2NAA, yellow), PHD_VC5HCH_{NSD2} (PDB code: 9GBF, chain B, violet) and PHD_VC5HCH_{NSD3} (PDB code: 4gnd, chain A, green). Elements of secondary structure are labeled. **(C)** Cartoon representation of PHD_VC5HCH_{NSD2} tandem domain (PDB code: 9GBF, chain B), with PHD_V and C5HCH colored in pale cyan and wheat, respectively. Zn²⁺ ions are represented in spheres and aromatic residues involved in interdomain interactions are shown as sticks and dots. **(D)** Backbone dynamics of PHD_VC5HCH NSD family. The dotted line indicates the ¹H-¹⁵N heteronuclear NOE value threshold of 0.65. Elements of secondary structure are indicated on the top.

(EMBL), expressed in BL21 (DE3) at 23°C after induction with 1mM of IPTG, in LB medium or in minimal medium containing ¹⁵NH₄Cl with or without ¹³C-D-glucose. The H3 fragment was purified from the soluble fraction using HisTrap HP affinity column (GE Healthcare) with a linear gradient of imidazole. The eluted fusion protein was digested with His₆-SEN2 (25) at 4°C for 4–5 h, followed by affinity chromatography on a Ni-NTA column and size-exclusion chromatography on Superdex 75 prep grade column (GE Healthcare) in a buffer containing 20 mM NaH₂PO₄/Na₂HPO₄, pH 6.3, 150 mM NaCl.

Recombinant histone H3K_C27me₃₁₋₄₄

To generate a recombinant mimic of H3K27me₃ histone tail, a single Lysine-to-Cysteine mutation (K27C) was introduced into the H3₁₋₄₄ sequence by site-directed mutagenesis (H3K_C27₁₋₄₄) and produced in labeled and unlabeled form, as described for recombinant H3₁₋₄₄. H3K_C27₁₋₄₄ was then alkylated as described (26). Briefly, purified recombinant H3K_C27₁₋₄₄ was reduced with DTT before addition of a 50-fold molar excess of (2-bromoethyl) trimethylammonium bromide (Sigma 117196–25G). The reaction mixture was incubated for 4 h at 50°C before quenching with

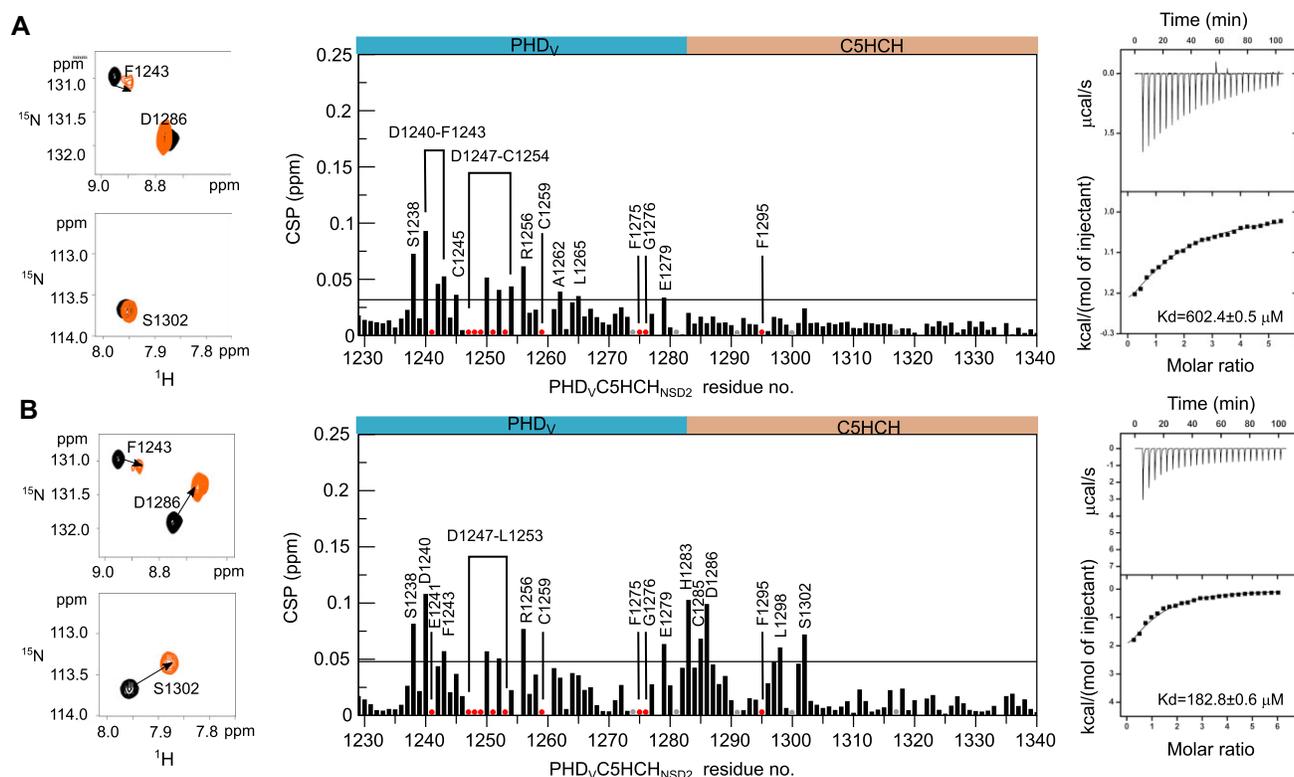


Figure 2. PHDV₅C5HCH_{NSD2} two histone H3 binding sites. NMR and ITC titrations of (A) H3₁₋₁₅ and (B) H3₁₋₃₇ peptides into PHDV₅C5HCH_{NSD2}. On the left superposition of selected regions of ¹H-¹⁵N HSQC spectrum of 0.13 mM PHDV₅C5HCH_{NSD2} without (black) and with 1.5 mM histone peptides (orange). Center panel, bar graph showing residue-specific amide CSPs of ¹⁵N PHDV₅C5HCH_{NSD2} (0.15 mM) upon addition of a 12-fold excess of histone peptides. Residues with CSP > avg + σ (black line) or disappearing upon binding (red dots) are labeled. Grey dots represent prolines. On the right, ITC-binding curves of histone peptides titrated into PHDV₅C5HCH_{NSD2}. In the ITC curves the upper panel shows the sequential heat pulses for peptide–domain binding, the lower panel shows the integrated data, corrected for heat of dilution and fit to a single-site-binding model using a nonlinear least-squares method.

5 mM β-mercaptoethanol. The modified protein was separated and desalted using a PD-10 desalting column (GE Healthcare) pre-equilibrated in water supplemented with 2 mM β-mercaptoethanol and lyophilized. The alkylation was confirmed by western blot analysis.

Synthetic peptides

Synthetic histone H3 peptides (H3₁₋₁₅, H3K4me3₁₋₁₅, H3₁₋₃₇, H3K27me3₁₋₃₇, H3K9me3₁₋₃₇, H3K27me3₁₇₋₂₈) were C-amidated and purchased from Caslo Lyngby, Denmark. Peptide purity (>98%) was confirmed by HPLC and mass spectrometry. Peptide concentrations were estimated from their dry weight.

Crystallization conditions, data collection and structure determination

Well-diffracting crystals were obtained for the PHDV₅C5HCH_{NSD2} construct comprising residues R1228-R1334, using the sitting drop vapor diffusion method at 4°C at a protein concentration of 18 mg/ml. The crystallization buffer contained 28% PEG Smear Medium, 0.1 M HEPES pH 7.5 and 0.05 M magnesium sulfate. Diffraction data were collected at SLS X06SA, processed with XDS (27). Since collected data revealed anisotropic diffraction, they were anisotropically truncated and scaled with STARANISO (28). Molecular replacement was performed with PHASER from CCP4 suite (29) using a homology model of PHDV₅C5HCH_{NSD2} gen-

erated by SWISS-MODEL (30) using PHDV₅C5HCH_{NSD3} apo structure (PDB code: 4gnd) as template. The model was rebuilt in COOT (31) and refined using phenix.refine (32). Water molecules were added using phenix.refine and manually inspected. The model was inspected and validated using MolProbity (33).

NMR spectroscopy

NMR experiments were performed at 295 K on a Bruker Avance III 600 MHz equipped with inverse triple-resonance cryoprobe and pulsed field gradients (Bruker, Karlsruhe, Germany). Typical sample concentration was 0.1–0.4 mM. Data were processed using NMRPipe (34) or Topspin 3.2 (Bruker) and analyzed with CCPNmr Analysis 2.5 (35). The backbone ¹H, ¹³C, ¹⁵N chemical shifts of the PHDV₅C5HCH_{NSD2} were obtained from three-dimensional HNCA, CBCA(CO)NH, CBCANH, HNCO experiments (BMRB code: 51 637). The backbone ¹H, ¹⁵N chemical shift of the PHDV₅C5HCH_{NSD1} and PHDV₅C5HCH_{NSD3} were obtained from the BMRB database (BMRB code: 25 933 and 18 664, respectively) and confirmed by HNCA, CBCA(CO)NH, CBCANH, HNCO experiments. The ¹H and ¹⁵N chemical shifts of H3₁₋₄₄ and H3K27_Cme3₁₋₄₄ were obtained from (36) and confirmed by HNCA, CBCA(CO)NH, CBCANH, HNCO experiments. NMR titrations and relaxation experiments were performed on the constructs R1228-K1340 and R1228-A1331, respectively. The final buffer contained 20 mM

NaH₂PO₄/Na₂HPO₄, pH 6.3, 150 mM NaCl, 1 mM TCEP, 10 μM ZnCl₂ supplied with 0.15 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) and D₂O (10% v/v).

Relaxation experiments

Heteronuclear ¹H-¹⁵N nuclear Overhauser enhancement, longitudinal and transversal ¹⁵N relaxation rates (R₁, R₂) were performed using standard 2D methods (37). Duty-cycle heating compensation was used for both T₁ and T₂ relaxation experiments (38). T₁ and T₂ decay curves were sampled at 12 and 11 different time points respectively (T₁ from 50 to 2000 ms and T₂ from 12 to 244 ms) collected in random order, with 2.5 seconds recovery delay. Two delays were acquired twice for evaluation of ⟨T₁⟩, ⟨T₂⟩ and the corresponding standard deviations. The ¹H-¹⁵N NOEs were measured recording HSQC spectra with and without proton saturation in an interleaved fashion using a 4 second recycle delay/saturation. The standard deviation of the noise of both saturated and unsaturated spectra were used to estimate I_{sat}/I_{unsat} uncertainty via error propagation formula. R₁, R₂ and ¹H-¹⁵N NOEs values have been obtained using the CcpNmr2.5 fitting routine (35) and internal scripts.

NMR binding assays

To minimize dilution and NMR signal loss, titrations were carried out by adding to protein samples (typically 0.1–0.2 mM) small aliquots of concentrated (15 mM) peptide stock solutions in 20 mM phosphate buffer, 150 mM NaCl, at pH 6.3. For each titration point (typically 0.5, 1, 2, 3, 6, 8, 10, 12 equivalents of ligand) a 2D water-flip-back ¹⁵N-edited HSQC spectrum was acquired with 2048 (180) complex points, 100 ms (42 ms) acquisition times, apodized by 60° shifted squared (sine) window functions and zero filled to 2048 (512) points for ¹H (¹⁵N). Assignment of each titration spectrum in the presence of histone peptides was made by following individual cross-peaks through the titration series. For each residue the weighted average of the ¹H and ¹⁵N chemical shift perturbation (CSP) was calculated as described in (1) (39).

$$CSP = \sqrt{\frac{\Delta\delta_H^2 + \left(\frac{\Delta\delta_N}{5}\right)^2}{2}} \quad (1)$$

Isothermal titration calorimetry

ITC titrations were performed using a VP-ITC isothermal titration calorimeter (MicroCal LLC, Northampton, MA, USA). Recombinant proteins (typically 0.2–0.3 mM) and synthetic peptides were dissolved against the same buffer (20 mM NaH₂PO₄/Na₂HPO₄, pH 7.2, 150 mM NaCl, 2 mM β-mercaptoethanol, 10 μM ZnCl₂) at 23°C. Stepwise injections of the 0.5–6 mM histone peptides solution into a cell containing a 0.1–0.2 mM of PHD_VC5HCH_{NSD2} (R1228-K1340) or PHD_VC5HCH_{NSD3} were performed to finally reach 6-fold molar excess of histone peptides with respect to the protein concentration. The heat released in the process was measured. Control experiments were performed under identical conditions to determine the dilution heat of the titrant peptides into buffer and of the buffer into protein samples. The final data were analyzed using the software ORIGIN 7.0 (Origin-Lab Corp., Northampton, MA, USA).

PHD_VC5HCH_{NSD2}-H3K27me3₁₋₃₄ complex model

PHD_VC5HCH_{NSD2}. The PHD_VC5HCH_{NSD2} structure (chain B of the crystallographic structure of PHD_VC5HCH_{NSD2} (PDB code:9GBF) was prepared with the Protein Preparation Wizard module of the Schrödinger Maestro suite, release 2020–4 (40). The crystallographic water molecules and HET groups were removed, missing sidechains were added using Prime and terminal residues were capped. The orientation of the hydroxyl groups of Ser, Thr and Tyr, the side chains of Asn and Gln residues, and the protonation state of His residues were optimized at pH = 7 ± 2. Three independent 0.5 μs Molecular Dynamics (MD) simulations of PHD_VC5HCH_{NSD2} were performed as detailed below.

PHD_VC5HCH_{NSD2}-H3K27me3₁₋₃₄

Based on CSPs data and on sequence conservation, we hypothesized that PHD_VC5HCH_{NSD2} and PHD_VC5HCH_{NSD3} similarly recognized the first nine residues of unmodified histone H3 (H3₁₋₉) through the formation of a third beta-strand on the PHD_V surface. Thus, we modelled PHD_VC5HCH_{NSD2} in complex with H3₁₋₉ using as template the coordinates of H3 in complex with PHD_VC5HCH_{NSD3} (PDB code: 4gnf) (41). Next, the structure of H3₁₀₋₃₄ was extracted from the nucleosome structure (PDB code:1kx5, chain E) (42) and linked to residue H3K9. To generate three different plausible models, we used three different allowed Ψ_{H3S10} dihedral angle values. Next, we introduced the tri-methylation on the amino-group of H3K27. To drive H3K27me3 into the aromatic cage composed by W1282, H1283 and F1295, we performed three independent 1.5 μs long Steered MD simulations, followed by 0.5 μs long plain MD simulations as described in the next paragraph. The production was carried out on the CINECA Marconi 100 supercomputer.

Molecular dynamics

The MD simulations were performed using Gromacs2020.2 (43) and PLUMED2 (44), with Amber14SB force-field (FF) (45) and TIP3P water model (46). The AmberTools20 LEaP module (47) was used to integrate into the FF the Zhang et al. (48) tri-methyl lysine parameters, for the atomic description of Kme3, and with the Zinc AMBER Force Field (ZAFF) (49) to describe the His/Cys zinc ions coordination. The system was solvated into a cubic box of TIP3P water molecules (edge defined as the largest distance between system atoms plus 3.0 nm) and NaCl (150 mM). After energy minimization with steepest descent and conjugate gradient algorithms, the system was equilibrated. A 25 ps long NVT simulation was performed with heavy atoms restrained to their initial positions, followed by a 200 ps long simulation in NPT conditions at 300K. Temperature and pressure were controlled using the velocity rescale algorithm (50) and the Parrinello-Rahman barostat (51), respectively. Short-range Coulomb interactions were calculated using a potential-shift Verlet ‘cut-off’, while long-range Coulomb interactions were computed using the particle mesh Ewald method (52). All the simulations were performed at 300 K with 0.2 fs time steps. Steered MD simulations (3 × 1.5 μs) were performed by applying a harmonic restraint with a force constant of 4 kJ/mol to the distance between the Nζ atom of the histone K27me3 residue and the center of mass of the Cε3 atom of W1282 and the Cε2 atom of F1295 of PHD_VC5HCH_{NSD2}. After the steered simulations, the system was relaxed for 500 ns with a plain MD simulation (one for

each replica). In addition, we perform 3 independent classical MD simulations of PHD_VC5HCH_{NSD2} in apo form. The last 200 ns of each plain MD simulation were analyzed. The root mean square deviation (RMSD) was calculated on the N, C α , C, and O backbone atoms of PHD_VC5HCH_{NSD2} (residues 5–95) in the apo form and in complex with the histone H3K27me3 (residues 1–30). The RMSD was also calculated for H3K27me3 (residues 1–30) bound to PHD_VC5HCH_{NSD2} (residues 5–95). The system coordinates extracted from each trajectory at 300 ns were used as the reference. The root mean square fluctuation (RMSF) was computed on the C α atoms of PHD_VC5HCH_{NSD2} (residues 5–95) in apo form and bound to the H3K27me3 tail, and on the C α atoms of H3K27me3 (residues 1–30) in complex with PHD_VC5HCH_{NSD2}. The RMSF was calculated on a combined trajectory constructed by concatenating the three single replicas belonging to the same condition, using the coordinates at 300 ns of the first replica as a reference. H-bond analysis was performed on each trajectory using the VMD 1.9.3 software (53), with 3.5 Å and 35° as the distance and angle ‘cut-off’, respectively. Only H-bonds that persisted for more than 25% of the simulation time in two of the three models are reported. Molecular images were generated with VMD 1.9.3 software (VMD is developed with NIH support by the Theoretical and Computational Biophysics group at the Beckman Institute, University of Illinois at Urbana-Champaign) (53), the open source version of PyMOL (The PyMOL Molecular Graphics System, Version 2.5.0 Schrödinger, LLC). Figures were prepared using Inkscape 1.3.2.

Protein expression in mammalian cells

Cloning and site directed mutagenesis

NSD2 wild type (WT) construct was cloned into pLVX-IRES-mCherry. Site directed mutagenesis of the C-terminal PHD finger mutants (D1240A, D1255A and F1295A) was performed using Q5 Site-Directed Mutagenesis Kit (NEB). Primers with desired mutations were designed with NEBaseChanger. The primers were used to PCR-amplify the plasmid and the product was transformed into NEB 5-alpha competent cells. Single colonies were selected, and mutations confirmed with sanger sequencing. Colonies with confirmed wild-type or mutant plasmids were expanded, and their plasmid DNA purified using Plasmid Maxi Kit (Qiagen) (Supplementary Table S1).

Lentiviral vector production

HEK293T cells were cultured in T25 flasks and co-transfected with 3.4 µg of NSD2^{WT} or NSD2^{D1240A}, NSD2^{D1255A}, NSD2^{F1295A} plasmid, 2.6 µg of the packaging plasmid (psPAX2) and 1.7 µg of the envelope plasmid (pMD2.G). Lipofectamine 2000 and Plus reagent (Thermo Fisher Scientific) were used as transfection agents. The virus-containing media of each flask was collected 60 h after transfection and cell debris separated by centrifugation at 1500 g for 10 min. The lentiviral-containing supernatant was filtered through a Steriflip-HV 0.45 µm low protein-binding PVDF membrane (Millipore) and concentrated 50-fold using Lenti-X Concentrator (Clontech). All viral solutions were aliquoted and stored at – 80°C until transduction.

Lentiviral transduction

NSD2 TKO cells (10) were seeded in media supplemented with 8µg/mL polybrene at a density of 1 × 10⁶ cells per well

in a 12-well plate. 5 µL of concentrated virus harboring wild-type or mutant NSD2 was added to each well and mixed with the cell suspension. Plates were centrifuged at 1000 g for 2 h at 33°C, and transduced cells were subsequently resuspended in fresh media and recovered for 48h. After recovery, cells were sorted for mCherry to achieve a >99% pool of transduced cells

Protein extraction and western blotting

Nuclear extracts were prepared using the Nuclear Complex Co-IP Kit (Active Motif), following the manufacturer’s protocol. Protein concentrations of all samples were determined with Bicinchoninic acid (BCA) assay (Thermo Fisher Scientific) and equal amounts of protein were loaded onto NuPAGE 4–12% Bis-Tris gels (Invitrogen). Proteins were separated using SDS-PAGE and transferred to nitrocellulose membranes using iBlot 2 (Invitrogen). Histone modifications and protein levels were determined using antibodies including anti-H3K36me2 (Abcam, RRID: AB_1 280 939), anti-H3 (Cell Signaling Technology, RRID: AB_2 756 816), anti-HDAC2 (Cell Signaling Technology, RRID: AB_2 116 822) and anti-NSD2 (Abcam, RRID: AB_1 310 816).

Colony formation and proliferation

Cells were grown in MethoCult H4100 (Stem Cell Technologies) supplemented with 20% FBS, 1% Pen Strep (Life Technologies) and 1% GlutaMax-I (Life Technologies). Cells were seeded at a density of 500 cells/mL in a 12-well plate and allowed to grow for 14 days. Cells were plated at a density of 0.2 × 10⁶ cells/mL. Every 3 days cells were counted with trypan blue dye and replated at the same density.

RT-qPCR

RNA was extracted from cells using the RNeasy Plus Mini Kit (Qiagen). cDNA was synthesized from total RNA using SuperScript IV VILO Master Mix (Invitrogen) and genomic DNA was removed with ezDNase™ Enzyme. Quantitative RT-PCR was performed using a TaqMan Fast Advanced Master Mix (Applied Biosystems) and a predesigned TaqMan assay (Applied Biosystems) for JAM3 (Hs00230289_m1), which was normalized to a GAPDH (Hs02786624_g1) control. qRT-PCR was performed in the CFX96 Touch Real-Time PCR Detection System (Biorad).

Adhesion Assay

12-well plates were coated with 1 mL of 150 µg/mL media-diluted Matrigel solution (Corning) and incubated at 37°C for 1 h to create a thin coat of Matrigel at the bottom of the wells. Each cell line was fluorescently labeled with CytoLight Rapid Red (Sartorius), following manufactures instructions. The Matrigel-coated wells were rinsed gently with serum-free media and all cells were plated at a concentration of 5 × 10⁵ cells/well. Cells were allowed to adhere for 16 h and were then imaged for an initial reading using the Incucyte S3 Live-Cell Analysis System (Sartorius). Then, the plates were flipped to remove non-adherent cells and imaged again. Lastly, additional non-adherent cells were removed with a gentle PBS wash and the plates were imaged a third time. Adherent cells from the second and third reading were normalized to the initial reading to account for differences in cell plating.

RNA-seq

Total RNA was extracted with the Direct-zol RNA MiniPrep Plus Kit. Library preparation and Poly(A) RNA sequencing were performed by the University of Florida Interdisciplinary Center for Biotechnology Research core. Short reads were trimmed using trimmomatic (v 0.36) (54) and quality control on the original and trimmed reads was performed using FastQC (55) (v 0.11.4) and MultiQC (56). The reads were aligned to the GRCh38 transcriptome using STAR version 2.7.9a (58), and transcript abundance was quantified using RSEM (RSEM v1.3.1)(59). PCA and MDS plots produced using custom R scripts were used to evaluate similarity between replicates. Differential expression analysis was performed using DESeq2 (59) with an FDR-corrected *P*-value threshold of 0.05. The output files were further filtered to extract transcripts showing a 1.5-fold change in either direction. Results were reported for protein-coding genes only, and for all transcript types.

ChIP-seq

ChIPmentation was performed as previously described (60), TKO, WT, D1240A and F1295A cell lines were cross-linked with 0.8% formaldehyde and quenched with 0.125 mol/L glycine. Using the Covaris E220 Focused-ultrasonicator (Covaris), the extracted chromatin was sonicated and precleared with Prot.A/G Dynabeads (Invitrogen, #10001D and 10003D). The DNA was immunoprecipitated with anti-H3K27ac (Active Motif, RRID:AB_2561016) overnight at -4°C . The precipitated chromatin was washed and incubated in Tagmentation buffer containing Tagment enzyme (Illumina, #20 034 210) and DNA was purified after reverse cross-linking. The purified DNA samples were amplified for the library with index primers using the KAPA HiFi Hot-Start Kit (Roche, #7 959 028 001). The libraries were cleaned and size-selected using Ampure XP Beads (Beckman Coulter, #A63880), and pooled for Novaseq 6000 S4 2 \times 150 flow cell (ICBR Next-Gen Sequencing Core, University of Florida). Short reads were trimmed using trimmomatic (v 0.36) (54) and quality control on the original and trimmed reads was performed using FastQC (55) (v 0.11.4) and MultiQC (56). Reads were aligned to the GRCh38 genome using Bowtie2 version 2.4.5 (57) with default parameters. Peak detection was performed using MACS (57) version 2.2.7.1. Differential peak analysis was performed with the DASA pipeline (61). DASA identifies the set of consensus peaks in replicates of the same condition, and computes the set of all comparable peak regions, that is, the peaks that appear in at least one condition. Comparable peaks were quantified using bedtools (62) producing a matrix containing the amount of signal under each peak in each condition. The matrix was then analyzed with DESeq2 (59) to determine peaks showing a significant difference between conditions. A similar procedure was employed to identify differential signal at gene transcription start sites (\pm 2KB around the TSS), gene bodies, and known enhancers (retrieved from the enhanceratlas.org site). Results were visualized using deeptools (63) and custom Python and R scripts.

CUT & RUN sequencing and data analysis

Cells (5×10^5) were harvested for Cleavage Under Targets & Release Using Nuclease (CUT&RUN) following the instructions of ChIC/CUT&RUN Kit (EpiCypher, #14–1048).

Briefly, the four cell lines (TKO, TKO repleted with NSD2^{WT}, NSD2^{D1240A}, NSD2^{D1255A}, NSD2^{F1295A}) were incubated in activated beads for 10 min to allow cells to bind to beads. After adding the spike-in SNAP-CUTANA™ K-MetStat Panel, the activated beads with cells were incubated with 0.5 μg antibodies (anti-IgG, anti-H3K4me3, and anti-H3K27me3) on nutator at 4°C overnight. The cells were permeabilized by 5% Digitonin for binding of pAG-MNase to cleave target-DNA complex. The targeted chromatin was digested in 100 mM calcium chloride for 2 h at 4°C . The chromatin was then combined with E. coli Spike-in DNA and released at 37°C for 10 min to the supernatants. The CUT&RUN-enriched DNA was purified for library preparation. Paired-end sequencing libraries were prepared from CUT&RUN enriched DNA (\sim 5 ng) using the CUTANA™ CUT&RUN Library Prep Kit (EpiCypher, #14–1001&14–1002). End repair, adapter ligation and U-Excision were performed to the CUT&RUN enriched DNA. After clean-up with SPRIselect reagent (Beckman Coulter), a unique pair of i5 & i7 Indexing Primers was added to enriched DNA for indexing PCR. The libraries were cleaned up using SPRIselect reagent and enrichment of mononucleosomal fragments (\sim 300 bp, including sequencing adapters) was confirmed by TapeStation. The libraries were pooled for Novaseq 6000 S4 2 \times 150 flow cell (ICBR Next-Gen Sequencing Core, U Florida). The input sequences were trimmed using Trimmomatic. Quality control was performed before and after trimming using FastQC (55). The input sequences were then aligned to the genome using Bowtie (v 2.4.5). For control and experimental reactions spiked with the SNAP-CUTANA K-MetStat Panel, the occurrences of DNA barcodes were counted using both R1 and R2 paired-end reads for assessment of K-MetStat Spike-in Control data. The E. coli Spike-in DNA was used for NGS normalization. Peak detection was performed using SEACR (64).

Results

The PHD_VC5HCH_{NSD2} tandem domain shares similar structure and dynamics with the other members of the NSD-PHD_VC5HCH family.

The NMR and crystallographic structures of PHD_VC5HCH_{NSD1} (23) and PHD_VC5HCH_{NSD3} (41) have been previously determined. To obtain a comprehensive experimental understanding of the NSD family, we determined the crystallographic structure of the PHD_VC5HCH_{NSD2} tandem domain (Table 1). Like the other NSD members, the two Zn-binding domains form an indivisible structural entity, with PHD_V and C5HCH arranged in a *face-to-side* orientation (Figure 1B). PHD_V adopts the classical PHD finger architecture, with an initial flexible loop (R1228-K1235, not visible in the electron density map), a two-stranded anti-parallel β -sheet

(β 1, β 2) followed by two short 3–10 helices α 1 and α 2 linked by the so-called variable L3 Loop (65). α 2 connects PHD_V to the C5HCH domain, an atypical PHD finger, consisting of two orthogonally oriented antiparallel β -hairpins (β 3- β 6). The tandem domain is held together by four Zn²⁺ ions, that are coordinated in an interleaved fashion by conserved Cysteines and Histidines) (Figure 1A, B). Hydrophobic interdomain interactions involving conserved residues (F1243, Y1263, P1281, W1282, H1283, H1284, P1291, F1295, P1300) stabilize the fold, creating a shallow

Table 1. Data collection and refinement statistics for X-ray diffraction data collected on a single crystal of PHD_VC5HCH^a

PHD _V C5HCH _{NSD2}			
Data collection		Refinement	
Space group	P 1 2 ₁ 1	Resolution (Å)	1.60
Cell dimensions	No. reflections	17 295	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	36.13, 45.25, 60.65	<i>R</i> _{work} / <i>R</i> _{free}	0.1870/0.2248
α , β , γ (°)	90.00, 99.65, 90.00	No. atoms	2641
Resolution (Å)	36.08–1.60 (4.97–1.60)	Protein	171
Total reflections	56 250 (1825)	Ligand/ion	12
Unique reflections	17 296 (866)	Water	75
<i>R</i> _{merge}	0.083 (0.611)	<i>B</i> -factors	
<i>R</i> _{meas}	0.099 (0.837)	Protein	32.58
<i>I</i> / σ <i>I</i>	9.7 (1.5)	Ligand/ion	28.42
Completeness	87.3 (42.1)	Water	32.94
(%) _{ellipsoidal}			
Redundancy	3.3 (2.1)	Clash scores	5
CC _{1/2}	0.997 (0.483)	R.m.s. deviations	
		Bond lengths (Å)	0.019
		Bond angles (°)	1.6
		Ramachandran ^b	
		Allowed (%)	1.8
		Favored (%)	98.2

^aValues in brackets correspond to the high-resolution shell. For cross-validation, 10% experimental reflections were randomly selected to calculate the *R*_{free}.

^bMolProbity

hydrophobic groove (Figure 1C). PHD_VC5HCH_{NSD2} aligns with a RMSD of 0.9 Å and 1.2 Å on the C α atoms of the equivalent domains in NSD1 (64% sequence identity) and NSD3 (58.4% sequence identity), respectively. The highest differences are observed in the L3 loop of PHD_V and in the β 4- β 5 loop, showing the greatest sequence divergence among the three sequences. The connecting turn between β 5 and β 6 also varies, likely due to the exclusive presence of a Proline residue (P1317) in NSD2 (Figure 1A,C). In solution, PHD_VC5HCH_{NSD2} shares with the other two family members similar sub-nanosecond internal motions, as indicated by comparable steady-state heteronuclear ¹H-¹⁵N NOE profiles (Figure 1D). Most residues exhibit values around 0.8, except for the N- and C-termini, the L3-loop within PHD_V, and the β 5- β 6 loop, which display lower values indicating higher mobility in the picosecond-nanosecond time range (Figure 1D). This is in line with the low electron density observed in the very same regions. The uniform *R*₂/*R*₁ ratio of residues in both PHD_V and C5HCH suggests that the PHD_VC5HCH_{NSD2} behaves as a monomer, tumbling as a single unit with a calculated rotational correlation time (τ_c) of 8.75 \pm 0.21 ns, in strong analogy to the cognate domains (PHD_VC5HCH_{NSD1}, τ_c = 7.96 \pm 0.41 ns and PHD_VC5HCH_{NSD3}, τ_c = 8.38 \pm 0.24 ns). Some local variations are observed in the *R*₂ and *R*₁ values of PHD_VC5HCH_{NSD2}, particularly in the α 2 helix connecting PHD_V and C5HCH of PHD_VC5HCH_{NSD2}. The higher *R*₂ values suggest the presence of slow conformational exchange phenomena that likely reverberate on neighboring residues (R1256-C1259, C1296-C1299, D1318, Y1322-E1325) (Supplementary Figure S1). Taken together, these findings indicate that the structural and dynamic characteristics of the PHD_VC5HCH tandem domain are highly conserved within the NSD family.

PHD_VC5HCH_{NSD2} is a combinatorial reader of unmodified H3K4 and tri-methylated H3K27

Despite similar structure and dynamics, the three NSD-PHD_VC5HCH tandem domains have evolved to different histone binding abilities. PHD_VC5HCH_{NSD1} shows only weak/non-specific binding towards histone H3 tail peptides (66), whereas PHD_VC5HCH_{NSD3} specifically recognizes unmethylated Lysine 4 and tri-methylated Lysine 9 on histone H3 (H3K9me3) (41). Previously, the binding of PHD_VC5HCH_{NSD2} to the histone H3 tail has been suggested (41), yet the specificity and the molecular details underlying this interaction remain unexplored. We have thus tested by NMR ¹H-¹⁵N heteronuclear single quantum coherence (HSQC) titration experiments and Isothermal Titration Calorimetry (ITC) the binding of PHD_VC5HCH_{NSD2} to different histone H3 peptides with different lengths and modifications. Gradual addition to ¹⁵N-PHD_VC5HCH_{NSD2} of an unlabeled peptide, corresponding to the first 15 unmodified residues of histone H3 tail (H3₁₋₁₅), induced significant chemical shift perturbations (CSP > avg+ σ , i.e. average + standard deviation) and line broadening effects of the amide resonances of residues mainly located at the N-terminus of PHD_V (S1238, D1240-F1243, C1245, D1247-C1254, R1256, C1259, A1262, L1265, F1275, G1276, E1279, F1295). Remarkably, this binary interaction was significantly weaker (K_d \sim 600 μ M) in comparison to higher binding affinities (K_d in the low micromolar range) typically observed in chromatin readers (67,21) (Table 2, Figure 2A). We next tested whether PHD_VC5HCH_{NSD2} sensed the methylation status of histone H3 at position K4. The presence of the N-terminal acidic hallmark (Figure 1A) suggested a preference for binding to unmethylated H3K4. Indeed, tri-methylation of H3K4 (H3K4me₃₋₁₅) had a detrimental effect on the interaction with PHD_VC5HCH_{NSD2} (K_d not measurable by ITC) (Table 2). Next, we wondered whether a C-terminal extension of the histone H3 peptide increased the interaction strength. Intriguingly, NMR and ITC experiments using a longer synthetic histone H3 peptide (residues 1–37, H3₁₋₃₇) revealed perturbations of additional amide resonances on the C5HCH region (H1283, C1285, D1286, F1295, L1298, S1302) (Figure 2B) and a three-fold increase in affinity (K_d \approx 180 μ M, Table 2). Taken together, these results highlighted the existence of a second histone H3 binding site on PHD_VC5HCH_{NSD2} which recognizes residues beyond position 15 of the H3 tail, that contribute to binding affinity. Indeed, reversed titrations performed with recombinant ¹⁵N labelled histone H3₁₋₄₄ (residues 1–44) clearly pinpointed two distinct regions affected by the interaction with PHD_VC5HCH_{NSD2}, characterized by significant CSPs or reduced peaks intensity. The two distinct regions encompassed respectively, the very end of the N-terminal tail (H3T3-H3S10) and residues within Q19-A29 (Supplementary Figure S2). Prompted by the observation that the binding affected the H3K27 amide resonances and inspired by the documented alteration of H3K27 methylation across the myeloma genome (7), we investigated whether PHD_VC5HCH_{NSD2} could decode this specific mark.

The addition of PHD_VC5HCH_{NSD2} to ¹⁵N H3Kc27me₃₋₄₄ (a recombinant histone H3 peptide mimicking H3K27me3) (Figure 3A) confirmed that two different histone regions (T3-T11; A24-Kc27me3) were involved in the binding, as assessed by significant CSPs (Figure 3B, C). The interaction, in agreement with the low micromolar affinity, was in the

Table 2. Thermodynamic parameters, stoichiometry (n) and dissociation constants (K_d) measured by ITC^a

PHD _V C5HCH	H3	ΔG (kcal/mol)	ΔH (kcal/mol)	T ΔS (K [*] kcal/mol)	n	K_d (μ M)
WT	H3 ₁₋₁₅	-4.3 ± 0.7	-0.5 ± 0.1	3.8 ± 0.1	1.2 ± 0.2	602.4 ± 0.5
WT	H3K4me3 ₁₋₁₅	nd ^b	nd ^b	nd ^b	nd ^b	nd ^b
WT	H3 ₁₋₃₇	-5.1 ± 0.8	-4.5 ± 0.8	0.6 ± 0.1	0.9 ± 0.1	182.8 ± 0.6
WT	H3K27me3 ₁₋₃₇	-7.4 ± 0.1	-3.9 ± 0.1	3.5 ± 0.1	1.0 ± 0.1	3.3 ± 0.3
WT	H3K9me3 ₁₋₃₇	nd ^b	nd ^b	nd ^b	nd ^b	nd ^b
F1295L	H3K27me3 ₁₋₃₇	-4.5 ± 0.6	-0.7 ± 0.1	3.8 ± 0.1	1.3 ± 0.1	483.1 ± 0.4
F1295A	H3K27me3 ₁₋₃₇	-4.5 ± 0.9	-1.0 ± 0.1	3.5 ± 0.1	1.2 ± 0.1	465.1 ± 0.2
D1240A		nd ^b	nd ^b	nd ^b	nd ^b	nd ^b
D1247A		nd ^b	nd ^b	nd ^b	nd ^b	nd ^b
D1255A		nd ^b	nd ^b	nd ^b	nd ^b	nd ^b
WT	H3K27me3 ₁₇₋₂₈	nd ^b	nd ^b	nd ^b	nd ^b	nd ^b

^a Experiments have been performed (T = 293 K) between wild type (WT) or mutant forms of PHD_VC5HCH_{NSD2} and histone H3 peptides.

^b and not detectable as binding is too weak to be reliably quantified.

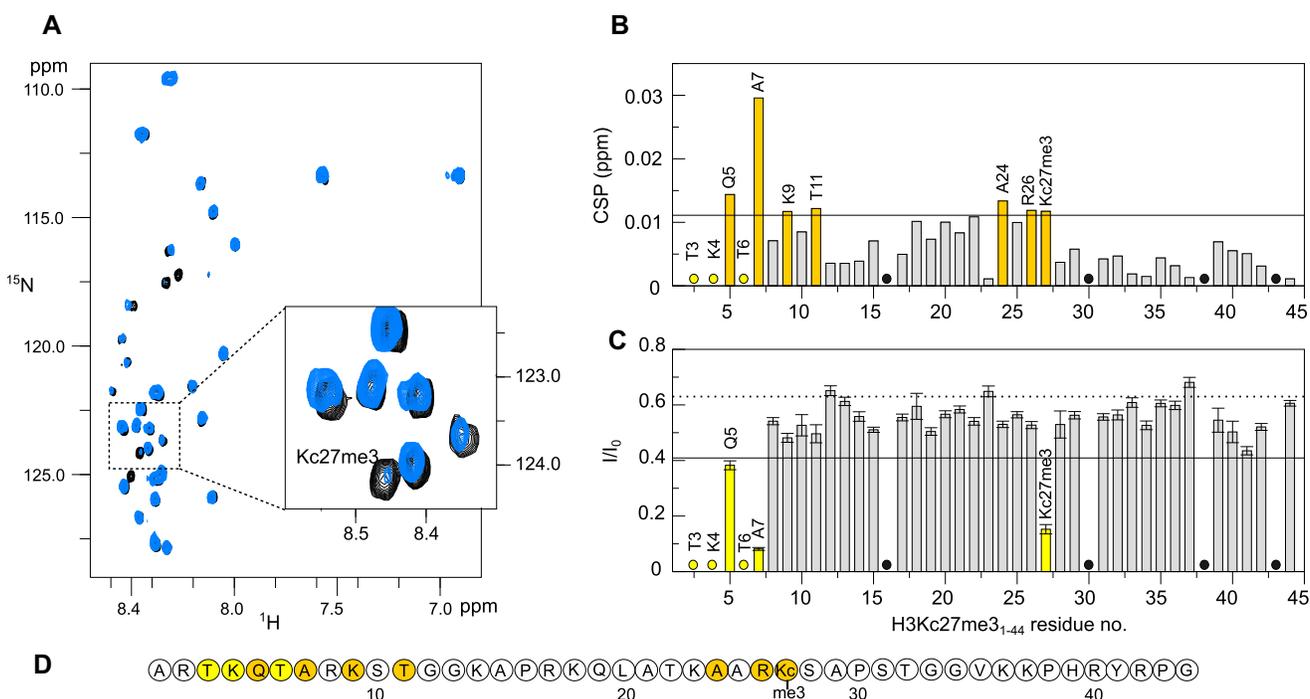


Figure 3. ¹⁵N H3Kc27me3₁₋₄₄ interacts with PHD_VC5HCH_{NSD2}. **(A)** ¹H-¹⁵N HSQC spectrum and selected region of ¹⁵N H3Kc27me3 (0.1 mM) without (black) and with three-fold excess PHD_VC5HCH_{NSD2} (blue). **(B)** Bar graphs showing residue-specific amide CSPs (upper panel) and **(C)** peak intensity ratio (I/I_0) of ¹⁵N H3Kc27me3 (0.1 mM) upon addition of a three-fold excess of PHD_VC5HCH_{NSD2}. Residues with CSP > avg + σ^c (black line and dark yellow bars, σ^c corresponds to corrected standard deviation as defined in (80)), and with I/I_0 < avg - σ (black line and light-yellow bars) or disappearing upon binding (yellow dots) are labeled and plotted on **(D)** the H3Kc27me3₁₋₄₄ (one-letter code) amino-acids sequence. Black dots represent prolines. The dotted line represents the expected decrease of intensity due to dilution. A1 and R2 amide resonances are missing in the ¹H-¹⁵N HSQC reference spectra, because of fast exchange with the solvent.

intermediate exchange regime on the NMR timescale, with several amide resonances decreasing in intensity or disappearing upon binding (Figure 3B-D). The reverse titration of ¹⁵N-PHD_VC5HCH_{NSD2} with unlabeled synthetic H3 peptide methylated at position K27 (H3K27me3₁₋₃₇) resulted in a similar pattern and direction of CSPs in the ¹H,¹⁵N HSQC spectrum, as observed with H3₁₋₃₇, suggesting a comparable mode of interaction. However, the CSPs affecting residues on C5HCH_{NSD2} (H1283, C1285, D1286, V1287, G1289, F1295, H1297, L1298, N1301, S1302) were more pronounced as compared to the unmodified histone peptide, suggesting an increased affinity (Figure 4). Indeed, the methylation of H3K27 substantially enhanced the affinity by two orders of magnitude

compared to the unmodified histone peptide ($K_d \approx 3 \mu$ M) resulting in a value typically observed in PHD fingers and other common epigenetic readers (68). The interaction stoichiometry was 1:1, with both enthalpic and entropic favorable contributions (Table 2).

Decoding of unmodified H3K4 and tri-methylated H3K27 appeared to be a unique feature of PHD_VC5HCH_{NSD2}, as NMR and ITC titrations of H3K27me3₁₋₃₇ into PHD_VC5HCH_{NSD1} and PHD_VC5HCH_{NSD3} revealed a much lower affinity ($K_d \approx 500 \mu$ M and $K_d \approx 120 \mu$ M, respectively, Supplementary Table S2). Interestingly, while PHD_VC5HCH_{NSD3} showed a preference for H3K9me3₁₋₃₇ ($K_d \approx 90 \mu$ M) the same modification abrogated binding to

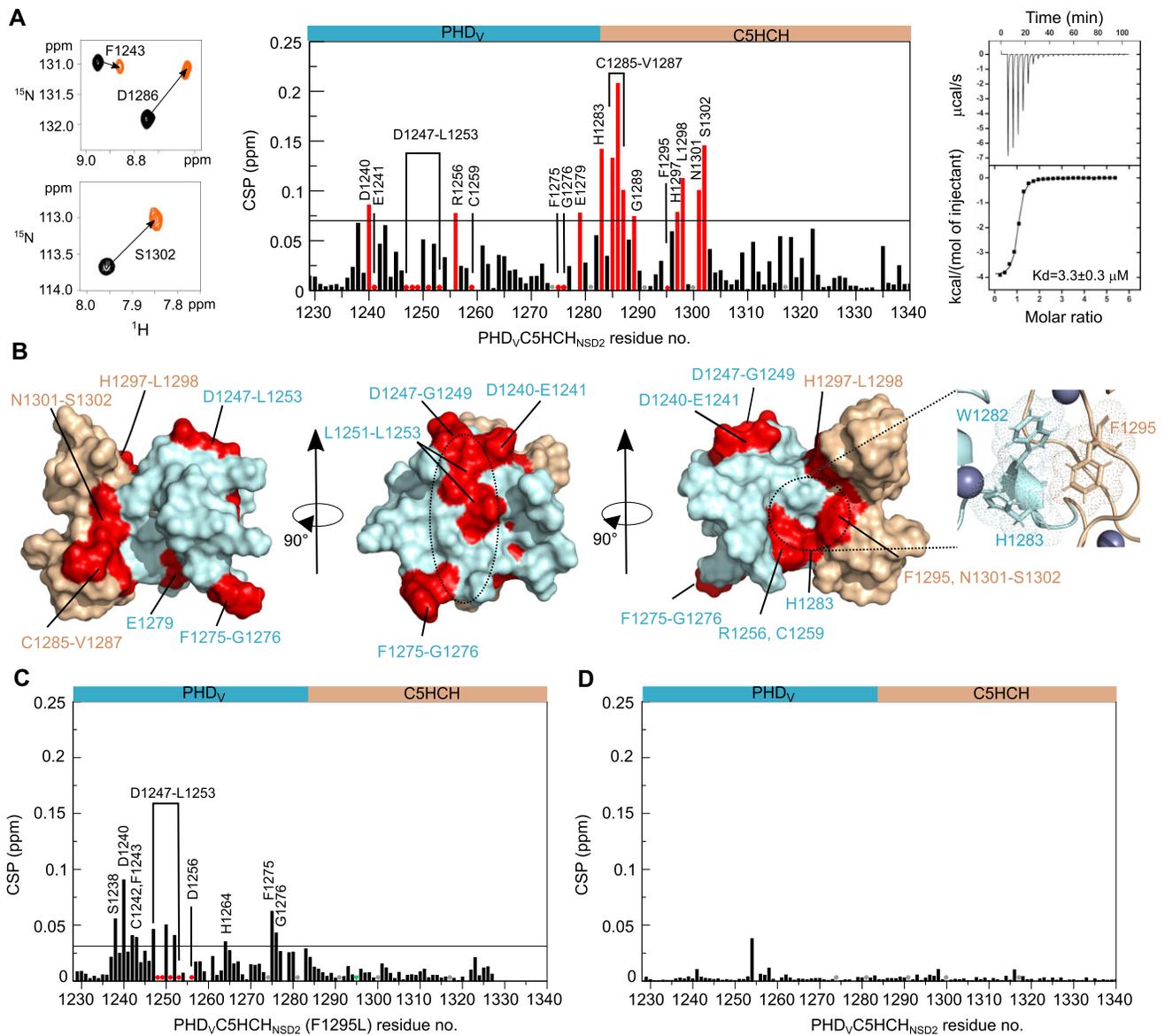


Figure 4. PHDV5HCHNSD2 is a combinatorial reader of unmodified H3K4 and tri-methylated H3K27. **(A)** NMR and ITC titrations of PHDV5HCHNSD2 with H3K27me3₁₋₃₇ peptide. On the left panel superposition of selected regions of ^1H - ^{15}N HSQC spectrum of 0.13 mM ^{15}N PHDV5HCHNSD2 without (black) a 12-fold excess H3K27me3₁₋₃₇ (orange). Central panel, bar graph showing the corresponding residue-specific amide CSPs. Residues with CSP > avg + σ (black line, red bars) or disappearing upon binding (red dots) are labeled. On the right, ITC-binding curves of H3K27me3₁₋₃₇ titrated into PHDV5HCHNSD2. In the ITC curves the upper panel shows the sequential heat pulses for peptide-domain binding, the lower panel shows the integrated data, corrected for heat of dilution and fit to a single-site-binding model using a nonlinear least-squares method. **(B)** Surface representation of PHDV5HCHNSD2 (PHDV, cyan; C5HCHNSD2, wheat; residues with CSP > avg + σ or disappearing upon binding of H3K27me3₁₋₃₇ are colored in red). The canonical histone binding site and the aromatic cage are indicated as dotted oval and circle, respectively. In the inset a zoom of the aromatic cage formed by F1295, W1282 and H1283 is shown in sticks and dots. Bar graphs showing residue-specific amide CSPs of: **(C)** ^{15}N -PHDV5HCHNSD2 (F1295L) (0.13 mM) upon addition of a 12-fold excess of H3K27me3₁₋₃₇ (residues with CSP > avg + σ (black line) or disappearing upon binding are labelled) and **(D)** ^{15}N PHDV5HCHNSD2 upon addition of 12-fold excess of H3K27me3₁₇₋₂₈. Grey (A, C) and the green (C) dots represent prolines and F1295L, respectively.

PHDV5HCHNSD2 (Supplementary Table S2). Collectively, these findings indicated that PHDV5HCHNSD2 is a novel specific combinatorial reader of both the unmethylated H3K4 and H3K27me3 epigenetic marks.

Two distinct binding surfaces of PHDV5HCHNSD2 cooperate to the combinatorial read out of unmodified H3K4 and tri-methylated H3K27

Mapping on the crystallographic structure of PHDV5HCHNSD2 of the residues whose amide reso-

nances mostly shifted and/or disappeared in the presence of H3K27me3₁₋₃₇ highlighted two distinct binding surfaces of the tandem domain (Figure 4B). The first one, corresponding to the classical histone binding site (Figure 4B, center) comprised the acidic residues (D1240, E1241) responsible for the preferential binding to unmodified H3K4, the β 1-strand (L1251-L1253), and the P1274xGxW1278 pocket, that usually accommodates the N-terminal Alanine of histone H3 (H3A1). The second interaction surface was in the interdomain groove between PHDV and C5HCH and featured an aromatic cage formed by W1282, H1283 and F1295, well

suiting for π -cation interactions with H3K27me3 (Figure 4B, right). Indeed, partial disruption of the aromatic cage through the fold preserving mutants F1295L and F1295A (Supplementary Figure S3), resulted in a hundred-fold reduction of the binding affinity (Table 2). Furthermore, analysis of the CSPs induced by the addition of H3K27me3₁₋₃₇ to the ¹⁵N-F1295L mutant revealed that, while the binding to the first site was maintained, the interaction with the second site was significantly impaired (Figure 4C). This observation confirms the essential role of the aromatic cage in recognizing the methylation state of H3K27 and in stabilizing the overall interaction of PHD_VC5HCH_{NSD2} with the histone H3 tail. Importantly, a histone H3 peptide corresponding to residues R17-S28 with tri-methylation on H3K27 (H3K27me3₁₇₋₂₈) was not able to interact with PHD_VC5HCH_{NSD2} (Table 2) and induced only negligible CSPs of ¹⁵N-PHD_VC5HCH_{NSD2} (Figure 4D), indicating that the anchorage of the N-terminal part of histone H3 is indispensable for complex formation and subsequent recognition of the H3K27me3 mark. Collectively, these results reveal that PHD_V and C5HCH constitute a functional and structural unit that orchestrates the combinatorial readout of unmethylated H3K4 and H3K27me3. The recognition of this dual mark follows a hierarchical order in which PHD_V is indispensable to anchor the N-terminal Alanine of histone H3 and H3K4 on the canonical histone binding surface, while the interdomain interface recognizes H3K27me3 via aromatic caging, eventually promoting enhanced binding affinity.

Three-dimensional model of the double-anchor recognition of H3K27me3 by PHD_VC5HCH_{NSD2}

Attempts to obtain crystals of PHD_VC5HCH_{NSD2} in complex with H3K27me3₁₋₃₇ failed, most likely because of the intrinsic dynamics of the system. Moreover, the unfavorable intermediate exchange regime on the NMR time-scale hampered the detection of intermolecular nuclear Overhauser effects (NOEs) and further complex characterization. Thus, to describe the molecular mechanisms underlying recognition of the double epigenetic mark H3K27me3 by PHD_VC5HCH_{NSD2}, we adopted a computational multi-step approach. The protocol included the generation of three starting models of H3K27me3₁₋₃₇ in complex with PHD_VC5HCH_{NSD2}, on which we performed three independent 1.5 μ s Steered MD simulations to drive H3K27me3 into the aromatic cage (Figure 4B). Subsequently, to relax the system, we performed 0.5 μ s plane MD simulations. For comparison we also run three independent 0.5 μ s molecular dynamics simulations on free PHD_VC5HCH_{NSD2}. The RMSD of free and bound PHD_VC5HCH_{NSD2} indicated that the tandem domain was stable and that the binding to H3K27me3 is maintained throughout the simulation time (Supplementary Figure S4). Overall, the root means squared fluctuations (RMSF) profile of bound PHD_VC5HCH_{NSD2} was similar to that of the free form, with smaller values especially in the regions comprising L3 (K₁₂₇₂-G₁₂₇₆), β 4- β 5 (D₁₃₀₉-T₁₃₁₁) and β 5- β 6 (T₁₃₁₆-G₁₃₁₉) loops (Figure 5A), suggesting that histone binding dumps down PHD_VC5HCH_{NSD2} flexibility, with minor effects on its intrinsic dynamic. Notably, the RMSF profile of H3K27me3 was in good agreement with the spectral perturbations observed in ¹⁵N-H3Kc27me3₁₋₄₄ titrations and was in line with the proposed dual recognition mechanism. In fact, the residues with very high RMSF (residues

10–15) corresponded to the resonances that were less affected by the binding, while the two regions with the lowest RMSF values (A1-A7 and R26-S28) showed significant CSPs or peaks disappearance upon binding (Figure 5B), suggesting their direct involvement in the binding. In all the simulations the first nine amino-acids of the histone H3 peptide stably interacted with the canonical ' β 1-surface' of PHD_V forming a third antiparallel β -strand (Figure 5C, Supplementary Figure S5A, C, D) via persistent hydrogen-bonds between the backbone atoms of H3R2, H3K4, H3T6 and L1253, L1251 and G1249 (Supplementary Table S3). H3A1 snugly fit into the so-called PxGxW pocket with the N-terminal amino-group forming hydrogen bonds with both G1276 and R1273 (Supplementary Figure S5C). Charge complementarities and hydrophilic interactions further contributed to complex formation (Supplementary Table S3). D1240, D1247 and D1255 formed stable salt-bridges with H3K4, H3R8 and H3R2, respectively. Accordingly, D1240A, D1247A and D1255A mutants resulted in impaired binding in ITC titrations. Of note, steric exclusion of the methylated ϵ amino group accounts for specificity for non-methylated H3K4 (Supplementary Figure S5C, D, Table 2). The complex models also suggested the presence of persistent hydrogen bonds between, the amino group of H3K9, the carbonyl of G1246 and the carboxylate of E1241 (Supplementary Figure S5C, Supplementary Table S3). Indeed, tri-methylation of H3K9 strongly impaired binding, possibly through steric clashes with E1241 (Table 2). Hydrophobic interactions between H3A1, H3R2, H3K4, H3T6 and P1274, L1251, V1252, L1253 and W1278 also contributed to complex stabilization. Only a few residues beyond H3K9, devoid of secondary structure elements, established stable contacts with the PHD_VC5HCH_{NSD2} (Supplementary Table S3, Supplementary Figure S5B, D). Despite the high flexibility around residue H3K27me3, the tri-methylammonium group appeared stably anchored inside the aromatic pocket, through cation- π , and hydrophobic interactions with W1282, H1283 and F1295 (Figure 5C) Backbone interactions between Arg1320 and Ser1321 also contributed to hold H3K27me3 in place (Supplementary Figure S5B).

Overall, these models were in good agreement with the bipartite interaction surfaces identified by CSPs (Figure 5C), and they provided a structural rationale for the combinatorial readout of unmodified H3K4 and tri-methylated H3K27 by PHD_VC5HCH_{NSD2}.

Mutations within PHD_VC5HCH affect the expression of NSD2 regulated cell adhesion genes

We next asked whether in the context of the full-length NSD2 protein the PHD_VC5HCH domain and its histone H3 tail binding ability played a functional role in NSD2-dependent properties, such as H3K36 di-methylation, cell proliferation, cell colony formation and cell adhesion (10). To answer to this question, we investigated in a NSD2-dependent MM cellular model the effects of site-specific mutants of PHD_VC5HCH (NSD2^{D1240A}, NSD2^{D1255A} and NSD2^{F1295A}), whose *in vitro* binding to H3K27me3 was impaired (Table 2). The t(4;14) KMS11 cell line was previously engineered to create a pair of isogenic cell lines such that either the IgH rearranged overexpressed allele of NSD2 was ablated (targeted knockout cells, TKO; NSD2 low) or the non-rearranged basal level expressed allele was disrupted (NTKO; NSD2 high) (10). KMS11 TKO cells were repleted with NSD2^{WT} or with NSD2^{D1240A},

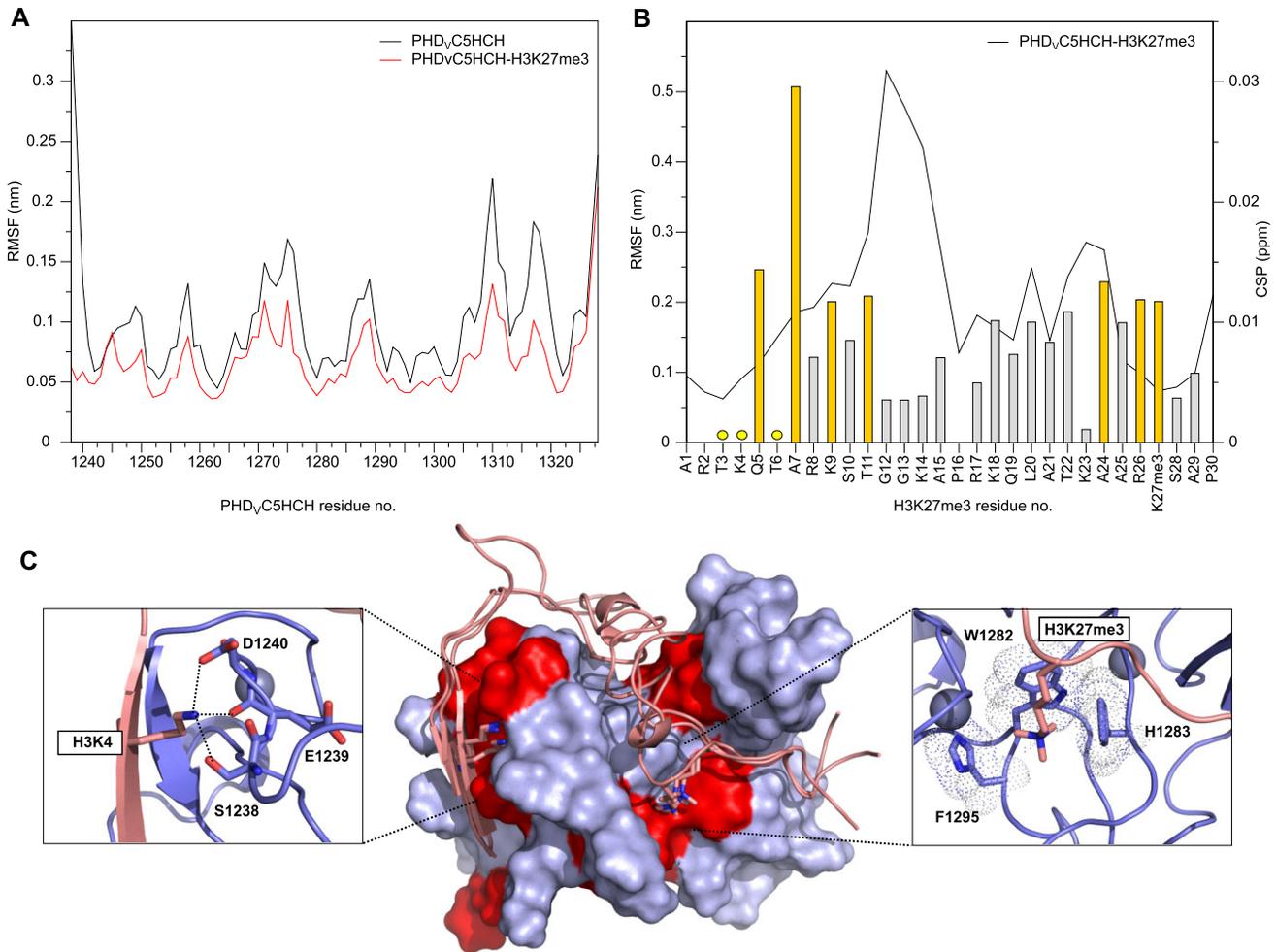


Figure 5. Three-dimensional model of the combinatorial reading of H3K27me3 by PHD_vC5HCH_{NSD2}. **(A)** Root mean square fluctuation (RMSF) of C α atoms of PHD_vC5HCH_{NSD2} (residues 5–95) over the simulation time was calculated on a combined trajectory constructed concatenating three independent replica. The RMSF profiles for the free system and in complex with H3K27me3 are shown in red and black, respectively. **(B)** RMSF of C α atoms of H3K27me3 (residues 1–30) over the simulation time was computed on a combined trajectory constructed concatenating three individual replica. On the same graph the RMSF profile is compared to ¹H-¹⁵N CSPs of ¹⁵N-H3K27me3 residues in the presence of unlabeled PHD_vC5HCH_{NSD2}. Dots indicate peaks disappeared upon binding. Residues showing significant CSP are colored in yellow. **(C)** Superposition of representative snapshots extracted from the three MD simulations of the bound system. PHD_vC5HCH_{NSD2} and H3K27me3 are shown in surface and cartoon representation, respectively, with H3K4 and H3K27me3 highlighted in sticks. Residues of PHD_vC5HCH_{NSD2} mostly affected by the interaction with H3K27me3 (CSP > avg + σ or disappearing upon binding of H3K27me3₁₋₃₇) are colored in red. In the insets, zoom of the intermolecular interactions engaged by H3K4 (left) and H3K27me3 (right).

NSD2^{D1255A} and NSD2^{F1295A} by lentiviral vector transduction (for simplicity defined as NSD2^{WT}, NSD2^{D1240A}, NSD2^{D1255A} and NSD2^{F1295A} cells) (Supplementary Figure S6). After transduction of TKO cells, expression levels of NSD2 (WT and mutants) were higher compared to TKO control cells, and the H3K36me2 levels were comparable to KMS11 t(4;14) NTKO cells (Figure 6A). In addition, the NSD2-dependent oncogenic phenotypes, such as cell colony formation and cell proliferation (10), were preserved in these repleted cell-lines (Supplementary Figure S7A, B). Intriguingly, while repletion of TKO cells with NSD2 mutants did not affect NSD2-dependent catalytic HMT activity, cell colony formation and proliferation, it strongly affected the cell adhesion properties. Indeed, Matrigel adhesion assays performed on TKO cells repleted with NSD2^{WT} led to increased cell adhesion in comparison to TKO control cells, whereas the three cell lines expressing the NSD2 mutants failed to adhere to Matrigel (Figure 6B,

C). We thus investigated whether the mutations affected the expression of adhesion genes previously shown to be upregulated by NSD2 (7). Indeed, we found that the cell lines repleted with mutated NSD2 failed to upregulate junctional adhesion molecule 3 (JAM3) (Supplementary Figure S7C). Prompted by this observation, we further compared by RNA-Seq analysis the global changes in gene expression observed in NSD2^{WT}, NSD2^{D1240A} or NSD2^{F1295A} cells as compared to non-repleted TKO cells (Figure 7A). The NSD2^{D1255A} cells were not further examined due to its low level of expression when expressed in MM cells. Re-expression of NSD2, whether in its WT or mutated forms, lead to dysregulation of gene expression, with a predominance of upregulated genes. NSD2^{WT} and the two mutants cell lines shared 1387 upregulated genes (Figure 7B), indicating that NSD2 mutants retain much of the transcriptional function of the WT protein. Interestingly, the 377 genes exclusively upregulated by NSD2^{WT} and not by the NSD2

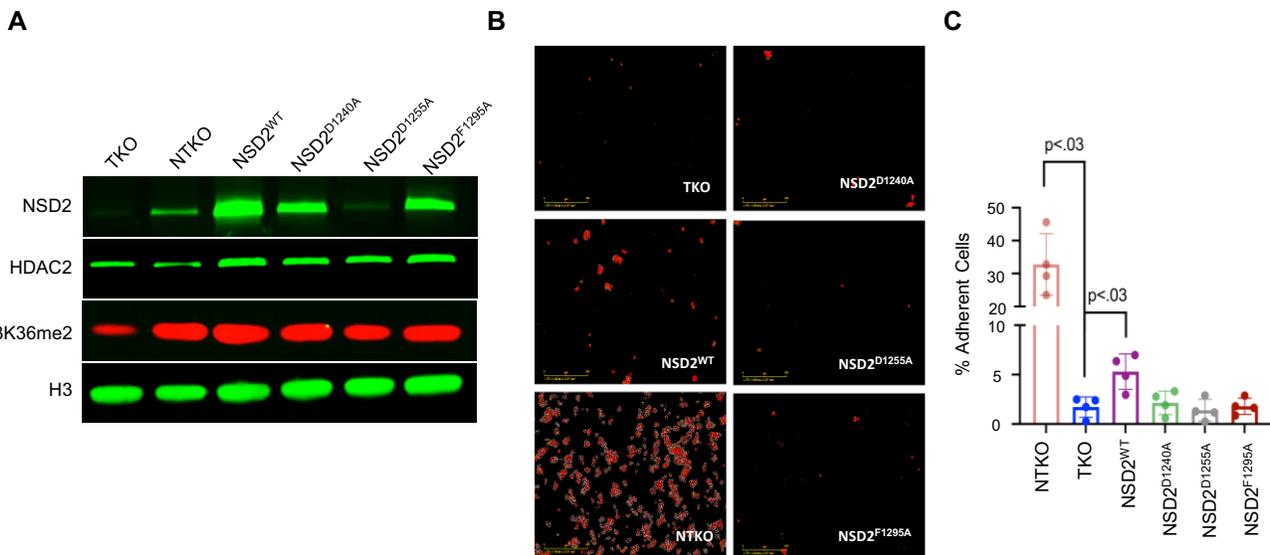


Figure 6. Mutations within the PHD_vC5HCH domain of NSD2 restrict the ability of NSD2 to modulate adhesion of multiple myeloma cells. **(A)** Immunoblot of NSD2 and H3K36me2 levels in cells transduced with wild-type or PHD domain mutant NSD2. HDAC2 and Histone H3 serve as loading controls. **(B)** Adhesion assay of MM cells to Matrigel; and **(C)** corresponding percentage of adherent cells, $n = 4$ biological replicates. Significant P values indicated as calculated by non-parametric t test.

mutants (Supplementary Figure S8A, Supplementary Table S4, Supplementary Table S5) were involved in hematopoietic stem cell migration to the bone marrow (*GAS6*, *JAM3*, *GPLD1* genes), junction formation and cell adhesion (*JAM3*, *PRKCA*, *PXN*) (Figure 7C,D, Supplementary Figure S8B, C). Moreover, KEGG Pathway analysis using gene set enrichment analysis (GSEA) revealed that re-expression of NSD2^{WT} in TKO cells induced significant upregulation of genes belonging to tight junction and focal adhesion subgroups, corroborating the Matrigel adhesion assay results (Figure 7C). Importantly, most of the upregulated genes belonging to these pathways (e.g. *JAM2*, *JAM3*, *PRKCA*, *PXN*, *VEGFC*, *VEGFA*, *DOCK1*) were not as highly expressed in NSD2^{D1240A} or NSD2^{F1295A} cells, in agreement with their reduced adhesion phenotype (Figure 7C). Finally, Hypergeometric Optimization of Motif EnRichment (HOMER) (69) analysis of transcription factor (TF) binding motifs in genes upregulated by either NSD2^{WT} or its mutants revealed that the promoters of genes exclusively upregulated by NSD2^{WT} were enriched with specific TF binding motifs, such as those for *SIX1* and *STAT6* TFs (Supplementary Figure S8C). This observation supports the notion that the PHD_vC5HCH domain plays a crucial role in NSD2-mediated activation of a specific subset of genes.

Taken together, our data suggest that in a MM context the PHD_vC5HCH domain of NSD2 specifically regulates the expression of a subset of genes associated with cell adhesion and that mutations hampering *in vitro* the histone binding activity of this domain, fail to activate the expression of these genes.

Mutations within PHD_vC5HCH of NSD2 result in incomplete loss of H3K27 methylation throughout the genome

Previous studies have shown that in t(4;14) associated MM the increased level of NSD2 induces an increase of the H3K36me2 mark and a depletion of the antagonistic H3K27me3 signature (7). These studies also showed that re-

cruitment of NSD2 to chromatin and the methylation levels of H3K27 were in part dependent on the structural integrity of the PHD domains (7). The latter were also required to recruit NSD2 to specific oncogenic gene loci (9). It was therefore hypothesized that the PHD_vC5HCH domain might modulate NSD2 binding to specific chromatin sites. Indeed, our biophysical analysis indicates that PHD_vC5HCH_{NSD2} domain is a reader of the H3K27me3 mark. Intriguingly, the 377 genes exclusively upregulated by NSD2^{WT} were enriched in target genes (e.g. *JAM3*, *GAS6*, *PRKCA*) of SUZ12, the catalytic subunit of the PRC2 complex responsible of H3K27me3 deposition (Supplementary Figure S8D). Thus, prompted by these observations we asked whether genes upregulated in NSD2^{WT} cells showed differential H3K27me3 levels as compared to TKO control cells. Using CUT&RUN sequencing (Cleavage Under Targets and Release Using Nuclease), a recent method for ultra-sensitive genomic mapping of chromatin targets, we scrutinized the four cell lines (TKO, NSD2^{WT}, NSD2^{D1240A}, NSD2^{F1295A}) for differences in the level and distribution of H3K27me3. In agreement with previous findings (7), we found that chromatin at the promoters and gene bodies of genes activated in the presence of NSD2^{WT}, as determined by RNA-Seq, had a significantly decreased H3K27me3 read density compared to TKO control cells. Conversely, in NSD2^{D1240A} or NSD2^{F1295A} cells H3K27me3 read densities on these genes did not decrease and were comparable to those of TKO control cells (Figure 8A). In contrast, for those genes whose expression is downregulated by NSD2^{WT} there was no difference in H3K27me3 read count when comparing TKO cells with NSD2^{WT} and mutant cells (Figure 8B). On a global scale, differential peak analysis of NSD2^{WT} cells showed more than 2000 differential H3K27me3 peaks compared to TKO control cells, whereas NSD2^{D1240A} and NSD2^{F1295A} cells displayed less than 200 differential peaks relative to TKO cells (Figure 8C). Taken together, these data suggested that the two mutants are functionally impaired in regulating H3K27me3 throughout the genome. Intriguingly, the corresponding mutants in PHD_vC5HCH_{NSD2} had a re-

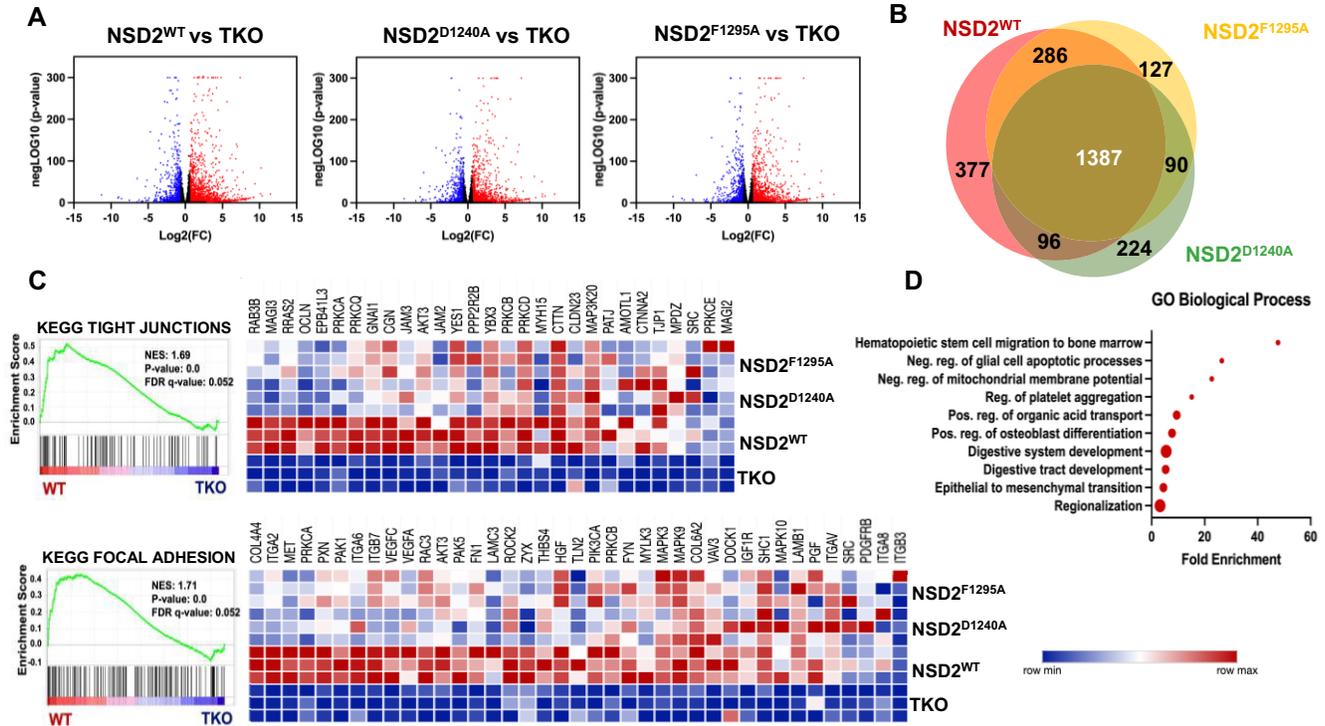


Figure 7. Mutations within the PHD_vC5HCH domain of NSD2 fail to upregulate focal adhesion and tight junction gene expression signatures. **(A)** Volcano plot of dysregulated genes between cells replated with NSD2^{WT} or mutants and NSD2 control cells; red: upregulated genes (>1.5-fold), blue: downregulated genes (>1.5-fold); **(B)** Overlap of upregulated genes from **(A)**. **(C)** On the left, gene set enrichment analysis of upregulated KEGG pathways in NSD2^{WT} replated vs TKO cells. On the right, heat map of normalized expression of genes upregulated in KEGG pathways. Upper panel: tight junction, lower panel: focal adhesion. **(D)** GO Biological Processes of 377 genes only upregulated in NSD2^{WT} cells, ShinyGO 0.77 (81). Dots size is proportional to the number of genes.

duced ability to recognize the H3K27me3 mark *in vitro*. This suggests that recognition of the methylation status of H3K27 via the PHD_vC5HCH of NSD2 plays a role in the subsequent removal of this mark to activate gene expression. The intersection between genes associated with decreased H3K27me3 peaks in NSD2^{WT} cells and genes upregulated in NSD2^{WT} cells, as indicated by RNA-Seq analysis, shows that more than 20% of the upregulated genes also have decreased H3K27me3 peaks (Figure 8D, Supplementary Table S6). This supports the notion that upregulation of gene expression by NSD2 is linked to the removal of H3K27me3 marks. Of the 377 genes exclusively activated by NSD2^{WT} 100 (27%) failed to show decreased H3K27me3 upon mutation of the PHD_vPHD5 domain (Figure 8E, Supplementary Table S6). These include genes involved in the neural and adhesive pathways activated by NSD2, such as *NEO1* and *JAM3* that code for surface adhesive proteins.

Collectively, these results indicate that the non-catalytic PHD_vC5HCH domain of NSD2 is critical for the removal of the H3K27me3 signature required to activate the expression of a specific set of genes that are mainly involved in cell adhesion.

Mutations within PHD_vC5HCH of NSD2 fail to produce increased H3K27ac signal within gene bodies and promoters of adhesion genes

Recent work indicated that NSD2 overexpression leads to increased gene expression due to removal of the H3K27me3 repressive mark and acquisition of the activation-associated

H3K27ac modification at promoters and enhancers (70). Our analysis of patterns of H3K27ac across the genome showed that depletion of TKO cells with NSD2^{WT} led to a significant increase of 1313 H3K27ac peaks and a loss of ~300 peaks (Figure 9A). In marked contrast, re-expression of NSD2^{D1240A} and NSD2^{F1295A} resulted in an increase of less than 100 H3K27ac peaks. Analysis of the 1313 upregulated H3K27ac peaks in NSD2^{WT} cells indicated that the average signal across all peaks had a ~3-fold increase compared to TKO cells. By contrast, the average H3K27ac signal across these peaks increased by only 1.4-fold and 1.6-fold in NSD2^{D1240A} and NSD2^{F1295A} cells, respectively (Figure 9B, C). Upon replating of TKO cells either with NSD2^{WT} or with mutants the majority of increased H3K27ac signals mapped to gene bodies (Figure 9D, Supplementary Tables S7–S15). Nearly 33% of the genes showing increased H3K27ac levels (ChIP-seq) were also upregulated (RNA-Seq) in NSD2^{WT} cells (Figure 9E, Supplementary Table S16). In particular, 10/39 genes encoding tight junction proteins and 6/39 encoding focal adhesion proteins displayed increased H3K27 acetylation. Strikingly, there was no increase of H3K27ac peaks within these gene bodies in NSD2^{D1240A} mutant cells and only one significant peak within the gene body of integrin gene *ITGA2* in NSD2^{F1295A} cells. This suggests a correlation between NSD2 overexpression, the change of expression and H3K27 acetylation level of a specific genes subset. For example, mRNA production and H3K27ac level at the promoter of *JAM2* were significantly upregulated in NSD2^{WT} cells, but only moderately increased in NSD2^{D1240A} and NSD2^{F1295A} cells (Figure 9F).

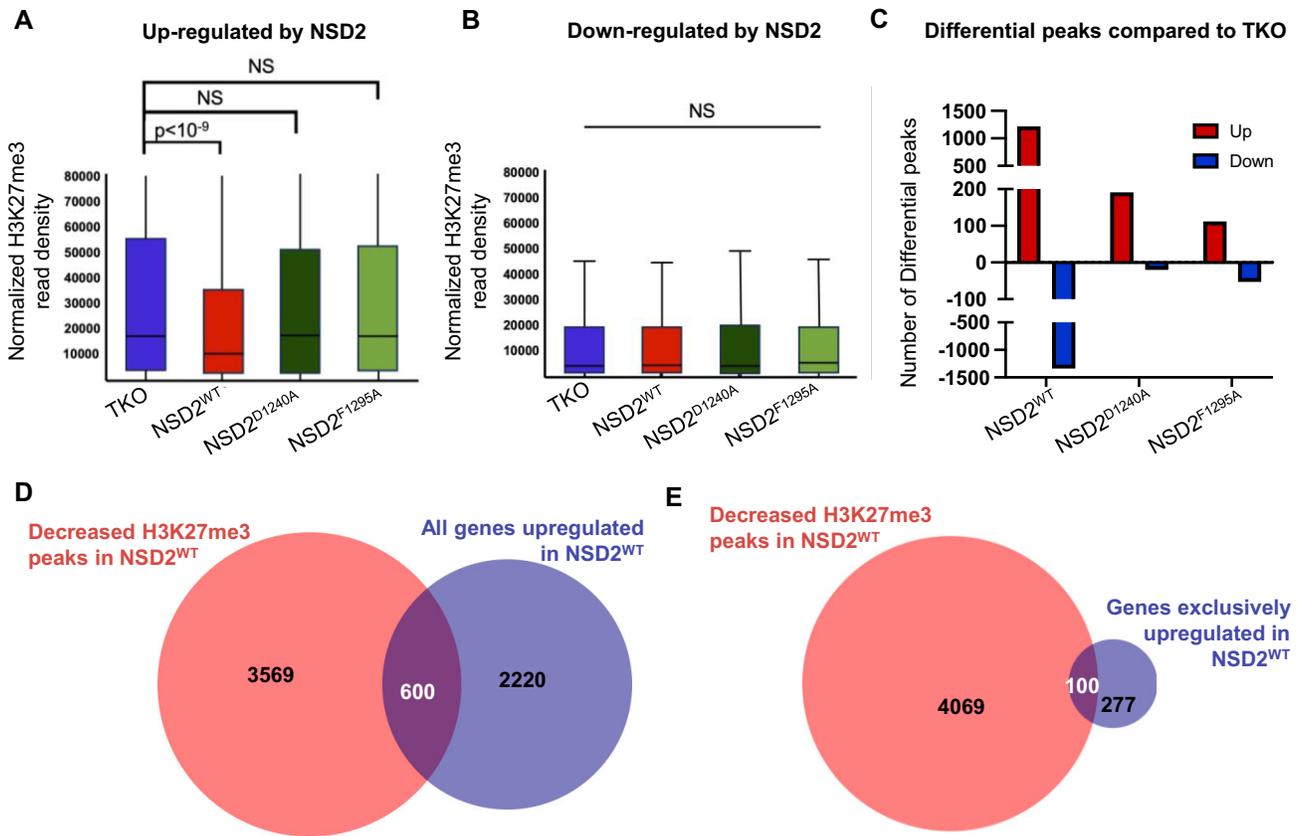


Figure 8. Mutations in PHD_VC5HCH impair the ability of NSD2 to modify chromatin. H3K27me3 read density at genes (A) upregulated and (B) downregulated in NSD2^{WT} or NSD2 mutants cells, as defined by RNA-Seq analysis. (C) Differential H3K27me3 peaks between NSD2^{WT} or NSD2 mutants with respect to TKO control cells, P -value < 0.05, FC > 1.5. (D) Overlap of genes with significantly decreased H3K27me3 peaks and all genes upregulated in NSD2^{WT} cells. (E) Overlap between genes showing a significant decrease in H3K27me3 peaks and genes that are exclusively upregulated by NSD2^{WT}.

Collectively, these data indicate that the PHD_VC5HCH domain contributes to the NSD2-dependent increase of the H3K27ac signal at gene bodies and promoters of adhesion genes.

Discussion

The PHD finger domain is a highly-conserved structural scaffold that alone or in tandem works as versatile protein–protein interaction domain. It plays multifaceted roles in interpreting modifications of histone tails, facilitating the recruitment and retention of host proteins on chromatin. This, in turn, promotes the stabilization of chromatin-bound complexes and/or directs catalytic subunits towards their specific substrates (19–21,71). In this study, we present new evidence revealing the previously unknown ability of the PHD_VC5HCH cassette of HMKT NSD2 to specifically recognize the dual histone mark composed by unmethylated H3K4 and H3K27me3. This interaction contributes to the recruitment of NSD2 to chromatin regions that are transcriptionally repressed and characterized by an abundance of H3K27me3 marks. This is the first evidence of such a specificity for the PHD_VC5HCH cassette in NSD2, expanding our understanding of its functional mechanisms in chromatin regulation. The combined use of X-ray crystallography, NMR, ITC and molecular dynamics simulations provided molecular insights into this double mark recog-

nition. PHD_VC5HCH_{NSD2} forms an indivisible structure that works as combinatorial reader of unmethylated H3K4 and trimethylated H3K27 on the same histone tail. This *in cis* multivalent interaction allows for efficient complex formation with low micromolar affinity. The binding occurs through two distinct binding surfaces within PHD_VC5HCH_{NSD2}, each recognizing different regions of the histone H3 tail. Specifically, the initial 10 amino acids of the histone H3 tail fit into the canonical binding pocket of PHD_V, extending the existing antiparallel β -sheet through the formation of an additional β -strand. This reproduces the well-established interaction pattern between the histone H3 tail and PHD finger domains. Electrostatic interactions between unmodified H3K4, H3R2, H3K9, and D1240, D1255, and D1247, respectively, contribute to the binding specificity. Mutating these Aspartate residues to Alanine results in a loss of binding, supporting their crucial role in complex formation. Importantly, PHD_VC5HCH_{NSD2} is highly sensitive to the methylation status of H3K4 and H3K9, as methylation at these positions significantly impairs binding. NMR spectroscopy and extensive MD simulations indicate that the uttermost N-terminal tail of H3 engages in very stable interactions with the classical β 1 surface. Conversely, residues beyond S10 maintain a high degree of conformational heterogeneity upon binding, with H3K27me3 making crucial and stable π -cation interactions with a conserved aromatic cage at the tandem domain interface, formed by residues

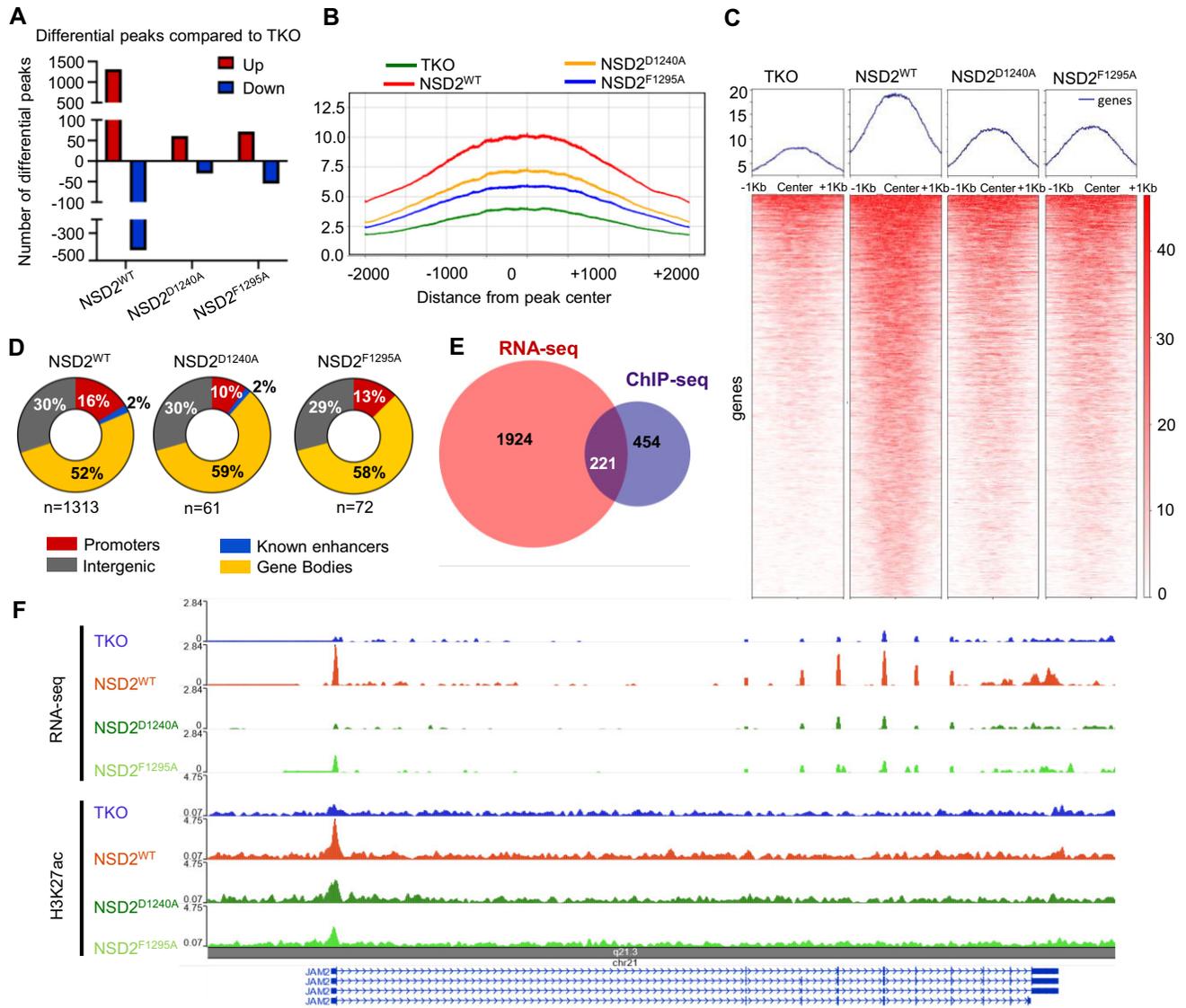


Figure 9. Mutations within PHDV₅HCH of NSD2 fail to produce increased H3K27ac signal at gene bodies and promoters of adhesion genes. **(A)** Differential H3K27ac peaks in NSD2^{WT} or mutants cells with respect to TKO control cells, P -value < 0.05, FC > 1.5. **(B)** Meta-analysis of H3K27ac signal from **(C)**; Y-axis -density of reads. **(C)** plots and heatmaps of H3K27ac ChIP-seq signal at TSS, in NSD2^{WT} or mutants cells; x-axis distance 1kb 5' or 3' from the center of signal. Y-axis -density of reads. **(D)** Genomic features distribution of significantly increased, differential H3K27ac peaks in cells NSD2^{WT} or mutants cells. **(E)** Overlap between significantly increased differential H3K27ac peaks and significantly upregulated genes expressed in NSD2^{WT} cells. **(F)** RNA-Seq and H3K27ac ChIP-seq tracks of JAM2 gene. The Washington University Epigenomics Browser (82).

W1282, H1283 and F1295. Methylation of H3K27 is fundamental for binding specificity, as demethylation of this position results in a 60-fold reduction of the affinity. The recognition mechanism via aromatic caging is typical of methyllysine readers and has been observed in other H3K27me3 binding domains, such as the chromatin modifier organization (chromo) (72) or the (bromo-adjacent homology) BAH domain within the repressive complex BAH-PHD-CPL2 in *Arabidopsis* (73). Interestingly, in this complex the BAH domain of protein APP3 interacts with the PHD finger of the proteins AIPP2/PAIPP2 allowing the bivalent reading of H3K27me3 and H3K4, respectively (73). Intriguingly, while in this latter case the cross-talk among different epigenetic marks occurs through two distinct proteins, in PHDV₅HCH_{NSD2} the combinatorial reading of the distinct marks occurs on the same hi-

stone H3 tail taking advantage of different surfaces within the same molecule. The BAH-PHD cassette of the plant SHORT LIFE (SHL) represents another example of dual reading of histone marks (74). In this case, the adjacent BAH domain and PHD fingers recognize the H3K27me3 and H3K4me3, respectively, and their recognition is independent and mutually exclusive (74). Conversely, in PHDV₅HCH_{NSD2} decoding of the double epigenetic signature follows a hierarchical mechanism that requires the recognition of both marks on the same Histone H3 tail for optimal binding. The uttermost N-terminal histone H3 residues are indispensable to accommodate the unmodified histone H3 tail on the canonical histone binding site of PHDV and to place then H3K27me3 in a suitable position to interact with the aromatic cage at the domains interface. Decoding of H3K27me3 strongly depends on

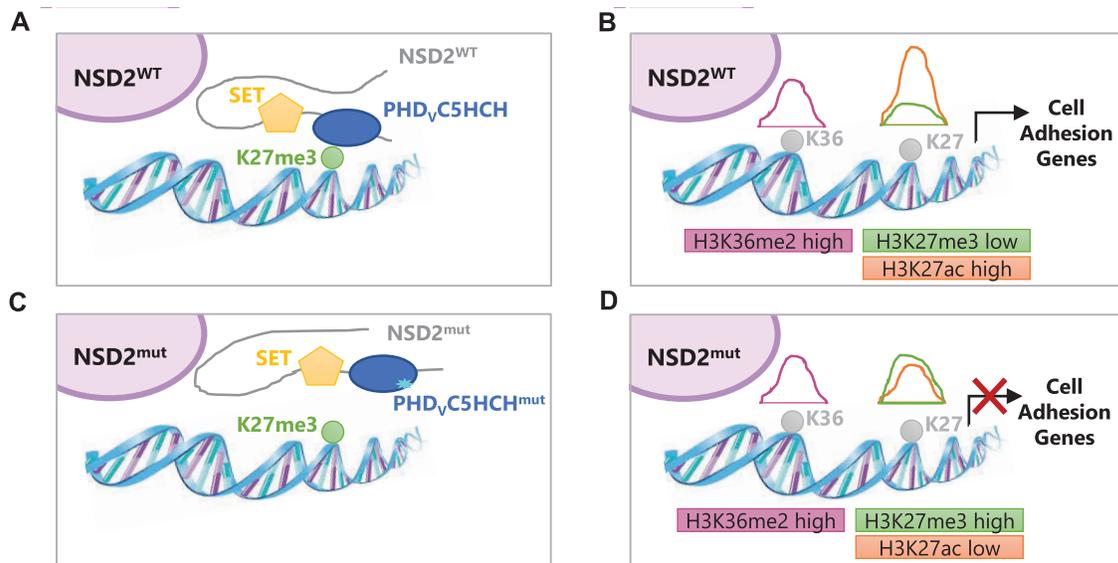


Figure 10. Role of PHD_vC5HCH in the transcriptional activity of NSD2. **(A)** PHD_vC5HCH (blue oval) of NSD2 recognizes H3K27me3 (green circle) on chromatin. **(B)** Cells replated with NSD2^{WT} show high levels of both H3K36me2 (magenta line) and H3K27ac (orange line) and low levels of H3K27me3 (green line), herewith the expression of a specific set of genes (e.g. cell-adhesion genes) is activated. **(C)** Mutations (cyan star) within PHD_vC5HCH impair binding to H3K27me3. **(D)** Cells replated with NSD2^{mut} have similar levels of H3K36me2 as NSD2^{WT}, but for a subset of genes high levels of H3K27me3 and low levels of H3K27ac are observed. As a result, a specific subset of genes (e.g. cell-adhesion genes) is not activated.

unmethylated H3K4 recognition, as a short histone segment harboring H3K27me3 but lacking the N-terminal Histone H3 residues is unable to bind PHD_vC5HCH_{NSD2}. Notably, despite the substantial sequence and structural identity among the NSD family, the three NSD-PHD_vC5HCH proteins have evolved to distinct and non-redundant functions, possibly due to the small sequence variations present on the loops and on β5-β6 of C5HCH (Figure 1A). PHD_vC5HCH_{NSD1} exhibits only weak interactions with the histone H3 tail, regardless of the modifications explored until now (23). We hypothesize that a yet unidentified histone mark (conceivably in the N-terminal residues) is required to increase the affinity of PHD_vC5HCH_{NSD1} for the histone H3 tail. However, it specifically interacts with the Zinc finger domain of the repressor protein Nizp1 (C2HR_{Nizp1}) (23). In contrast, both PHD_vC5HCH_{NSD2} and PHD_vC5HCH_{NSD3} showed low binding affinity to C2HR_{Nizp1} (24), but they clearly function as epigenetic readers. PHD_vC5HCH_{NSD3}, similarly to PHD_vC5HCH_{NSD2}, binds to the unmethylated H3K4 mark through conserved acidic residues, but it specifically recognizes H3K9me3 via a unique Tyrosine (Tyr₁₃₂₃) (41). Our *in vitro* binding assays also suggested that PHD_vC5HCH_{NSD3} has the potential to bind to H3K27me3 through the conserved aromatic cage, with an affinity (high micromolar) which is comparable to the one measured for H3K9me3. Interestingly, the double mark H3K9me3K27me3 did not increase the affinity (Kd 104.9 ± 2.5 μM, Supplementary Table S2), suggesting that additional histone modifications and/or interactors are required to increase the histone H3 binding capacity of PHD_vC5HCH_{NSD3}. Lastly, PHD_vC5HCH_{NSD2}, among all NSD family members, exhibits the strongest affinity for the repressive H3K27me3 mark. Moreover, the sequence-specific histone H3 recognition by PHD_vC5HCH_{NSD2} is negatively modulated by the active mark H3K4me3. The PHD finger (as single domain or in tandem) is generally not sufficient for association of the host complex with chromatin. However,

in many cases it has been shown to be necessary for proper levels of binding, for targeting chromatin enriched in a specific histone mark herewith modulating the transcriptional activity of the host protein (74). This is also the case for the PHD_vC5HCH domain within NSD2. Our study shows that, this tandem domain is not involved in H3K36 di-methylation but is crucial in localizing NSD2 to H3K27me3-enriched regions, herewith controlling the transcription of specific set of genes, in particular cell adhesion genes. Genome wide epigenomic profiling showed that cells replated with NSD2^{WT} had decreased H3K27me3 and increased H3K27ac levels within gene bodies. Conversely, repletion with NSD2 mutants known to impair histone H3 binding of the PHD_vC5HCH domain *in vitro* (D1240A, F1295A) led to less effective gene activation. Herein a subset of genes failed to show decreased H3K27me3 and increased H3K27ac levels (Figure 10). Interestingly, the promoters of NSD2-activated genes were enriched for DNA binding motifs of specific transcription factors, such as SIX1 and STAT6. Of note, SIX1 as member of the SIX family, regulates tumorigenesis, promotes cell proliferation, epithelial-to-mesenchymal transition, and metastasis (75) and STAT6 is associated with cancer cell proliferation, increased malignancy and poor prognosis (75). Conceivably, cooperation with those transcription factors, in addition to proper chromatin recognition, is critical for NSD2 to exert its full oncogenic action in stimulating aberrant gene expression in myeloma. Genes that the NSD2 mutants failed to upregulate are involved in hematopoietic stem cell migration to the bone marrow and in junction formation and cell adhesion, such as *JAM3*, *JAM2*, *GAS6*, *GPLD1*, *PRKCA*. Indeed phenotypically, MM cells expressing the NSD2 mutants have impaired cell adhesion properties. These results are in line with previous studies in which a deletion mutant of NSD2 lacking PHD_vC5HCH (previously defined as PHD4) did not affect H3K36me2 levels but induced incomplete demethylation of H3K27me3, thus allowing for only partial *JAM2* activation (7). Incomplete removal of the

repressive methylation of H3K27 in this case and in the case of the PHD_VC5HCH mutants would preclude its acetylation and the related transcriptional activation. Loss of acetylation can inhibit gene transcription by preventing BRD4 binding which links promoters and enhancers and recruits pTEFB, a factor that stimulates transcriptional elongation (reviewed in (76)). H3K27me3 demethylation of loci in response to NSD2 could be due to rapid deposition of the H3K36me2 mark upon cell replication since this modification prevents the action of PRC2 (via EZH2) to deposit the H3K27me3 modification (77). Alternatively, NSD2 might facilitate the recruitment of H3K27me3 specific demethylases, such as JMJD3 or KDM6A.

In conclusion, in this work we have shown that the PHD_VC5HCH domain of NSD2 is a novel dual reader of the unmethylated H3K4 and H3K27me3 marks that contributes to the transcription of cell adhesion genes induced by the over-expression of NSD2 in pathological MM conditions. As such, its direct pharmacological targeting, might constitute a novel intervention strategy against MM. In analogy to what is currently developed for the PWWP₁ domain of NSD2 (78), the design of small molecule antagonists of PHD_VC5HCH, might represent an additional opportunity for NSD2 inhibition, thus complementing epigenetic approaches based on protein degraders and on inhibitors of writer or eraser enzymes.

Data availability

The data underlying this article are available in the article and in its online supplementary material and will be shared on reasonable request by the corresponding authors. The crystallographic structure of PHD_VC5HCH_{NSD2} has been deposited in the PDB with the following code: 9GBF. The amide chemical shifts of free PHD_VC5HCH_{NSD1}, PHD_VC5HCH_{NSD2} and PHD_VC5HCH_{NSD3} are available at the BMRB databank under the codes 25 933, 51 637 and 18 664, respectively. The ¹H-¹⁵N relaxation data of PHD_VC5HCH_{NSD1}, PHD_VC5HCH_{NSD2} and PHD_VC5HCH_{NSD3} are available at the BMRB databank under the codes, 52 577, 52 576 and 52 578, respectively. The RNA Seq, Chip-Seq and Cut & Run data have been deposited in the GEO repository with the following GEO numbers GSE274233, GSE274234, GSE274250.

Supplementary data

Supplementary Data are available at NAR Online.

Funding

The authors acknowledge the CINECA (Consorzio Interuniversitario per il Calcolo Automatico dell'Italia del Nord Orientale) award under the PRACE (Partnership for Advanced Computing in Europe) initiative for the availability of high performance computing resources and support, proposal number 2020225324 (GM); Associazione Italiana Ricerca sul Cancro [IG-21440 to G.M.]; Florida Department of Health grant [21L05, NCI R01 CA195732 to J.D.L.]; Leukemia and Lymphoma society Specialized Center of Research [grant 7021-20 to J.D.L.]; Myeloma Solution Fund [to J.D.L.]; University of Florida Health Cancer Center, supported in part by state appropriations provided in Fla. Stat. § 381.915; NIH/NCI [P30CA247796 to J.D.L.]; FEBS (Federation of European Biochemical Societies) Short-term Fellow-

ship and EMBO (European Molecular Biology Organization) Short Term Fellowship-7094 [to A.B.].

Conflict of interest statement

None declared

References

- Husmann,D. and Gozani,O. (2019) Histone lysine methyltransferases in biology and disease. *Nat. Struct. Mol. Biol.*, **26**, 880–889.
- Bennett,R.L., Swaroop,A., Troche,C. and Licht,J.D. (2017) The role of nuclear receptor – Binding SET domain Family histone lysine methyltransferases in cancer. *Cold Spring Harb. Perspect. Med.*, **7**, a026708.
- Morishita,M. and di Luccio,E. (2011) Cancers and the NSD family of histone lysine methyltransferases. *Biochim. Biophys. Acta. (BBA)*, **1816**, 158–163.
- Keats,J.J. (2005) Overexpression of transcripts originating from the MMSET locus characterizes all t(4;14)(p16;q32)-positive multiple myeloma patients. *Blood*, **105**, 4060–4069.
- Keats,J.J. (2003) In multiple myeloma, t(4;14)(p16;q32) is an adverse prognostic factor irrespective of FGFR3 expression. *Blood*, **101**, 1520–1529.
- Martinez-Garcia,E., Popovic,R., Min,D.-J., Sweet,S.M.M., Thomas,P.M., Zamdborg,L., Heffner,A., Will,C., Lamy,L., Staudt,L.M., *et al.* (2011) The MMSET histone methyl transferase switches global histone methylation and alters gene expression in t(4;14) multiple myeloma cells. *Blood*, **117**, 211–220.
- Popovic,R., Martinez-Garcia,E., Giannopoulou,E.G., Zhang,Q., Zhang,Q., Ezponda,T., Shah,M.Y., Zheng,Y., Will,C.M., Small,E.C., *et al.* (2014) Histone methyltransferase MMSET/NSD2 alters EZH2 binding and reprograms the myeloma epigenome through global and focal changes in H3K36 and H3K27 methylation. *PLoS Genet.*, **10**, e1004566.
- Kuo,A.J., Cheung,P., Chen,K., Zee,B.M., Kioi,M., Lauring,J., Xi,Y., Park,B.H., Shi,X., Garcia,B.A., *et al.* (2011) NSD2 Links dimethylation of histone H3 at lysine 36 to oncogenic programming. *Mol. Cell*, **44**, 609–620.
- Huang,Z., Wu,H., Chuai,S., Xu,F., Yan,F., Englund,N., Wang,Z., Zhang,H., Fang,M., Wang,Y., *et al.* (2013) NSD2 Is recruited through its PHD domain to oncogenic gene loci to drive multiple myeloma. *Cancer Res.*, **73**, 6277–6288.
- Lauring,J., Abukhdeir,A.M., Konishi,H., Garay,J.P., Gustin,J.P., Wang,Q., Arceci,R.J., Matsui,W. and Park,B.H. (2008) The multiple myeloma associated MMSET gene contributes to cellular adhesion, clonogenic growth, and tumorigenicity. *Blood*, **111**, 856–864.
- Brito,J.L.R., Walker,B., Jenner,M., Dickens,N.J., Brown,N.J.M., Ross,F.M., Avramidou,A., Irving,J.A.E., Gonzalez,D., Davies,F.E., *et al.* (2009) MMSET deregulation affects cell cycle progression and adhesion regulons in t(4;14) myeloma plasma cells. *Haematologica*, **94**, 78–86.
- Marango,J., Shimoyama,M., Nishio,H., Meyer,J.A., Min,D.-J., Sirulnik,A., Martinez-Martinez,Y., Chesi,M., Bergsagel,P.L., Zhou,M.-M., *et al.* (2008) The MMSET protein is a histone methyltransferase with characteristics of a transcriptional corepressor. *Blood*, **111**, 3145–3154.
- Li,J., Troche,C., Swaroop,A., Kulis,M., Oyer,J., Will,C., Dupere-Richer,D., Riva,A., Pipe,C., Norton,S., *et al.* (2018) A gain of function mutation in the NSD2 histone methyltransferase drives glucocorticoid resistance of acute lymphoblastic leukemia. *Blood*, **132**, 653–653.
- Li,W., Tian,W., Yuan,G., Deng,P., Sengupta,D., Cheng,Z., Cao,Y., Ren,J., Qin,Y., Zhou,Y., *et al.* (2021) Molecular basis of nucleosomal H3K36 methylation by NSD methyltransferases. *Nature*, **590**, 498–503.

15. Sato,K., Kumar,A., Hamada,K., Okada,C., Oguni,A., Machiyama,A., Sakuraba,S., Nishizawa,T., Nureki,O., Kono,H., *et al.* (2021) Structural basis of the regulation of the normal and oncogenic methylation of nucleosomal histone H3 Lys36 by NSD2. *Nat. Commun.*, **12**, 6605.
16. Sankaran,S.M., Wilkinson,A.W., Elias,J.E. and Gozani,O. (2016) A PWWP domain of histone-lysine N-methyltransferase NSD2 binds to dimethylated lys-36 of histone H3 and regulates NSD2 function at chromatin. *J. Biol. Chem.*, **291**, 8465–8474.
17. Ferreira De Freitas,R., Liu,Y., Szewczyk,M.M., Mehta,N., Li,F., McLeod,D., Zepeda-Velázquez,C., Dilworth,D., Hanley,R.P., Gibson,E., *et al.* (2021) Discovery of small-molecule antagonists of the PWWP domain of NSD2. *J. Med. Chem.*, **64**, 1584–1592.
18. Dilworth,D., Hanley,R.P., Ferreira de Freitas,R., Allali-Hassani,A., Zhou,M., Mehta,N., Marunde,M.R., Ackloo,S., Carvalho Machado,R.A., Khalili Yazdi,A., *et al.* (2022) A chemical probe targeting the PWWP domain alters NSD2 nucleolar localization. *Nat. Chem. Biol.*, **18**, 56–63.
19. Bienz,M. (2006) The PHD finger, a nuclear protein-interaction domain. *Trends Biochem. Sci.*, **31**, 35–40.
20. Sanchez,R. and Zhou,M.-M. (2011) The PHD finger: a versatile epigenome reader. *Trends Biochem. Sci.*, **36**, 364–372.
21. Musselman,C.A. and Kutateladze,T.G. (2011) Handpicking epigenetic marks with PHD fingers. *Nucleic Acids Res.*, **39**, 9061–9071.
22. Baker,L.A., Allis,C.D. and Wang,G.G. (2008) PHD fingers in human diseases: disorders arising from misinterpreting epigenetic marks. *Mutation Res./Fundamental Mol. Mech. Mutagenesis*, **647**, 3–12.
23. Berardi,A., Quilici,G., Spiliotopoulos,D., Corral-Rodriguez,M.A., Martin-Garcia,F., Degano,M., Tonon,G., Ghitti,M. and Musco,G. (2016) Structural basis for PHD V C5HCH NSD1–C2HR Nizp1 interaction: implications for Sotos syndrome. *Nucleic Acids Res.*, **44**, 3448–3463.
24. Berardi,A., Botrugno,O.A., Quilici,G., Manteiga,J.M.G., Bachi,A., Tonon,G. and Musco,G. (2023) Nizp1 is a specific NUP98–NSD1 functional interactor that regulates NUP98–NSD1-dependent oncogenic programs. *FEBS J.*, **290**, 1782–1797.
25. Besir,H. (2017) A generic protocol for purifying disulfide-bonded domains and random protein fragments using fusion proteins with SUMO3 and cleavage by SenP2 protease. *Methods Mol. Biol.*, **1586**, 141–154.
26. Simon,M.D. (2010) Installation of site-specific methylation into histones using methyl lysine analogs. *Curr. Protoc. Mol. Biol.*, **Chapter 21**, 21.18.1–21.18.10.
27. Kabsch,W., Brünger,A., Diederichs,K., Karplus,P., Diederichs,K., McSweeney,S., Ravelli,R., Evans,P., French,S., Wilson,K., *et al.* (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.
28. Vonrhein,C., Tickle,I.J., Flensburg,C., Keller,P., Paciorek,W., Sharff,A. and Bricogne,G. (2018) Advances in automated data analysis and processing within autoPROC, combined with improved characterisation, mitigation and visualisation of the anisotropy of diffraction limits using STARANISO. *Acta Crystallogr. A Found. Adv.*, **74**, a360.
29. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
30. Waterhouse,A., Bertoni,M., Bienert,S., Studer,G., Tauriello,G., Gumienny,R., Heer,F.T., de Beer,T.A.P., Rempfer,C., Bordoli,L., *et al.* (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.*, **46**, W296–W303.
31. Emsley,P. (2017) Tools for ligand validation in Coot. *Acta Crystallogr. D Struct. Biol.*, **73**, 203–210.
32. Adams,P.D., Afonine,P.V., Bunkóczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.-W., Kapral,G.J., Grosse-Kunstleve,R.W., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
33. Williams,C.J., Headd,J.J., Moriarty,N.W., Prisant,M.G., Videau,L.L., Deis,L.N., Verma,V., Keedy,D.A., Hintze,B.J., Chen,V.B., *et al.* (2018) MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.*, **27**, 293–315.
34. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
35. Vranken,W.F., Boucher,W., Stevens,T.J., Fogh,R.H., Pajon,A., Llinas,M., Ulrich,E.L., Markley,J.L., Ionides,J. and Laue,E.D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins: Struct Function Genetics*, **59**, 687–696.
36. Kumar,G.S., Chang,W., Xie,T., Patel,A., Zhang,Y., Wang,G.G., David,G. and Radhakrishnan,I. (2012) Sequence requirements for combinatorial recognition of histone H3 by the MRG15 and Pf1 subunits of the Rpd3S/Sin3S corepressor complex. *J. Mol. Biol.*, **422**, 519–531.
37. Farrow,N.A., Muhandiram,R., Singer,A.U., Pascal,S.M., Kay,C.M., Gish,G., Shoelson,S.E., Pawson,T., Forman-Kay,J.D. and Kay,L.E. (1994) Backbone dynamics of a free and phosphopeptide-complexed src-homology 2 domain studied by ¹⁵N relaxation. *Biochemistry*, **33**, 5984–6003.
38. Yip,G.N.B. and Zuiderweg,E.R.P. (2005) Improvement of duty-cycle heating compensation in NMR spin relaxation experiments. *J. Magn. Reson.*, **176**, 171–178.
39. Grzesiek,S., Stahl,S.J., Wingfield,P.T. and Bax,A. (1996) 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.
40. Madhavi Sastry,G., Adzhigirey,M., Day,T., Annabhimoju,R. and Sherman,W. (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.*, **27**, 221–234.
41. He,C., Li,F., Zhang,J., Wu,J. and Shi,Y. (2013) The methyltransferase NSD3 has chromatin-binding motifs, PHD5-C5HCH, that are distinct from other NSD (Nuclear Receptor SET Domain) Family members in their histone H3 recognition. *J. Biol. Chem.*, **288**, 4692–4703.
42. Davey,C.A., Sargent,D.F., Luger,K., Maeder,A.W. and Richmond,T.J. (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9Å resolution. *J. Mol. Biol.*, **319**, 1097–1113.
43. Abraham,M.J., Murtola,T., Schulz,R., Páll,S., Smith,J.C., Hess,B. and Lindahl,E. (2015) Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, **1–2**, 19–25.
44. Tribello,G.A., Bonomi,M., Branduardi,D., Camilloni,C. and Bussi,G. (2014) PLUMED 2: new feathers for an old bird. *Comput. Phys. Commun.*, **185**, 604–613.
45. Maier,J.A., Martinez,C., Kasavajhala,K., Wickstrom,L., Hauser,K.E. and Simmerling,C. (2015) ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.*, **11**, 3696–3713.
46. Jorgensen,W.L., Chandrasekhar,J., Madura,J.D., Impey,R.W. and Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926–935.
47. Case,D.A., Belfon,K., Ben-Shalom,I.Y., Brozell,S.R., Cerutti,D.S., Cheatham,T.E. III, Cruzeiro,V.W.D., Darden,T.A., Duke,R.E., Giambasu,G., *et al.* (2020) Amber 2020. University of California, San Francisco.
48. Lu,Z., Lai,J. and Zhang,Y. (2009) Importance of charge independent effects in readout of the Trimethyllysine mark by HP1 Chromodomain. *J. Am. Chem. Soc.*, **131**, 14928–14931.
49. Peters,M.B., Yang,Y., Wang,B., Füsti-Molnár,L., Weaver,M.N. and Merz,K.M. (2010) Structural survey of zinc-containing proteins and development of the zinc AMBER force field (ZAFF). *J. Chem. Theory Comput.*, **6**, 2935–2947.

50. Bussi,G., Donadio,D. and Parrinello,M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.*, **126**, 014101.
51. Parrinello,M. and Rahman,A. (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.*, **52**, 7182–7190.
52. Darden,T., York,D. and Pedersen,L. (1993) Particle mesh Ewald: an $N \cdot \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089–10092.
53. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.
54. Bolger,A.M., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.
55. Andrews,S. (2010) FastQC. *Babraham Bioinformatics*. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
56. Ewels,P., Magnusson,M., Lundin,S. and Källér,M. (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, **32**, 3047–3048.
57. Feng,J., Liu,T., Qin,B., Zhang,Y. and Liu,X.S. (2012) Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.*, **7**, 1728–1740.
58. Li,B. and Dewey,C.N. (2011) RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinf.*, **12**, 323.
59. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
60. Li,J., Hlavka-Zhang,J., Shrimp,J.H., Piper,C., Dupéré-Richér,D., Roth,J.S., Jing,D., Casellas Román,H.L., Troche,C., Swaroop,A., et al. (2022) PRC2 Inhibitors overcome glucocorticoid resistance driven by NSD2 mutation in pediatric acute lymphoblastic leukemia. *Cancer Discov.*, **12**, 186–203.
61. Riva,A., Bennett,R.L., K.Venugopal,O.G., Brant,J. and Licht,J.D. (2019) DASA: a computational pipeline for differential ATAC-Seq analysis. *F1000Research*, **1**, 1117187.
62. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
63. Ramírez,F., Ryan,D.P., Grüning,B., Bhardwaj,V., Kilpert,F., Richter,A.S., Heyne,S., Dündar,F. and Manke,T. (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.*, **44**, W160–W165.
64. Meers,M.P., Tenenbaum,D. and Henikoff,S. (2019) Peak calling by Sparse Enrichment analysis for CUT&RUN chromatin profiling. *Epigenetics Chromatin*, **12**, 42.
65. Kwan,A.H.Y., Gell,D.A., Verger,A., Crossley,M., Matthews,J.M. and Mackay,J.P. (2006) Engineering a protein scaffold from a PHD finger. **11**, 803
66. Berardi,A., Quilici,G., Spiliotopoulos,D., Corral-Rodriguez,M.A., Martin-Garcia,F., Degano,M., Tonon,G., Ghitti,M. and Musco,G. (2016) Structural basis for PHDvC5HCHNSD1–C2HRNizp1 interaction: implications for Sotos syndrome. *Nucleic Acids Res.*, **44**, 3448–3463.
67. Morrison,E.A. and Musselman,C.A. (2016) The role of PHD fingers in chromatin signaling. In: *Chromatin Signaling and Diseases*. Elsevier, Amsterdam, The Netherlands, pp. 127–147.
68. Taverna,S.D., Li,H., Ruthenburg,A.J., Allis,C.D. and Patel,D.J. (2007) How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.*, **14**, 1025–1040.
69. Heinz,S., Benner,C., Spann,N., Bertolino,E., Lin,Y.C., Laslo,P., Cheng,J.X., Murre,C., Singh,H. and Glass,C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell*, **38**, 576–589.
70. Lhoumaud,P., Badri,S., Rodriguez-Hernaez,J., Sakellaropoulos,T., Sethia,G., Kloetgen,A., Cornwell,M., Bhattacharyya,S., Ay,F., Bonneau,R., et al. (2019) NSD2 overexpression drives clustered chromatin and transcriptional changes in a subset of insulated domains. *Nat. Commun.*, **10**, 4843.
71. Li,Y. and Li,H. (2012) Many keys to push: diversifying the ‘readership’ of plant homeodomain fingers recognition of histone tail by single. *Acta Biochimica et Biophysica Sinica*, **44**, 28–39.
72. Dong,C., Liu,Y., Lyu,T.-J., Beldar,S., Lamb,K.N., Tempel,W., Li,Y., Li,Z., James,L.I., Qin,S., et al. (2020) Structural basis for the binding selectivity of Human CDY chromodomains. *Cell Chem. Biol.*, **27**, 827–838.
73. Zhang,Y.Z., Yuan,J., Zhang,L., Chen,C., Wang,Y., Zhang,G., Peng,L., Xie,S.S., Jiang,J., Zhu,J.K., et al. (2020) Coupling of H3K27me3 recognition with transcriptional repression through the BAH-PHD-CPL2 complex in Arabidopsis. *Nat. Commun.*, **11**, 6212.
74. Qian,S., Lv,X., Scheid,R.N., Lu,L., Yang,Z., Chen,W., Liu,R., Boersma,M.D., Denu,J.M., Zhong,X., et al. (2018) Dual recognition of H3K4me3 and H3K27me3 by a plant histone reader SHL. *Nat. Commun.*, **9**, 2425.
75. Hu,X., li,J., Fu,M., Zhao,X. and Wang,W. (2021) The JAK/STAT signaling pathway: from bench to clinic. *Signal Transduct. Target Ther.*, **6**, 402.
76. Altendorfer,E., Mochalova,Y. and Mayer,A. (2022) BRD4: a general regulator of transcription elongation. *Transcription*, **13**, 70–81.
77. Li,Y., Trojer,P., Xu,C., Cheung,P., Kuo,A., Iii,W.J.D., Qiao,Q., Neubert,T.A., Xu,R., Gozani,O., et al. (2009) The target of the NSD Family of histone lysine methyltransferases depends on the nature of the substrate. *J. Biol. Chem.*, **284**, 34283–34295.
78. Meng,F., Xu,C., Park,K.-S., Kaniskan,H.Ü., Wang,G.G. and Jin,J. (2022) Discovery of a first-in-class degrader for nuclear receptor binding SET domain protein 2 (NSD2) and ikaros/aiolos. *J. Med. Chem.*, **65**, 10611–10625.
79. Robert,X. and Gouet,P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.*, **42**, W320–W324.
80. Schumann,F.H., Riepl,H., Maurer,T., Gronwald,W., Neidig,K.-P. and Kalbitzer,H.R. (2007) Combined chemical shift changes and amino acid specific chemical shift mapping of protein–protein interactions. *J. Biomol. NMR*, **39**, 275–289.
81. Ge,S.X., Jung,D., Jung,D. and Yao,R. (2020) ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*, **36**, 2628–2629.
82. Li,D., Purushotham,D., Harrison,J.K., Hsu,S., Zhuo,X., Fan,C., Liu,S., Xu,V., Chen,S., Xu,J., et al. (2022) WashU Epigenome Browser update 2022. *Nucleic Acids Res.*, **50**, W774–W781.