Check for updates

The FEBS Journal

FEBS PRESS science publishing by scientists

Nizp1 is a specific NUP98-NSD1 functional interactor that regulates NUP98-NSD1-dependent oncogenic programs

Andrea Berardi[1], Oronza A. Botrugno[2], Giacomo Quilici[1], José Manuel Garcia Manteiga[3], Angela Bachi[4], Giovanni Tonon[2] and Giovanna Musco[1]

1 Biomolecular NMR, Division of Genetics and Cell Biology, IRCCS Ospedale San Raffaele, Milan, Italy
2 Functional Genomics of Cancer, Division of Experimental Oncology, IRCCS Ospedale San Raffaele, Milan, Italy
3 Center for Omics Sciences, IRCCS Ospedale San Raffaele, Milan, Italy
4 Functional Proteomics Group, IFOM-FIRC Institute of Molecular Oncology, Milan, Italy

Keywords
acute myeloid leukaemia; histone-methyltransferase; Nizp1; NMR; NSD-family; PHD finger

Correspondence
G. Tonon, Functional Genomics of Cancer, Division of Experimental Oncology, IRCCS Ospedale San Raffaele, Milan 20132, Italy
Tel: +00390226435624
E-mail: tonon.giovanni@hsr.it
G. Musco, Biomolecular NMR, Division of Genetics and Cell Biology, IRCCS Ospedale San Raffaele, Via Olgettina, 58, Milan 20132, Italy
Tel: +00390226434824
E-mail: musco.giovanna@hsr.it

Andrea Berardi and Oronza A. Botrugno contributed equally to this article

(Received 5 August 2022, revised 27 September 2022, accepted 21 October 2022)

doi:10.1111/febs.16664

NSD1, NSD2 and NSD3 proteins constitute a family of histone 3 lysine 36 (H3K36) methyltransferases with similar domain architecture, but diversified activities, in part, dependent on their non-enzymatic domains. These domains, despite their high sequence identity, recruit the hosting proteins to different chromatin regions through the recognition of diverse epigenetic marks and/or associations to distinct interactors. In this sense, the PHDvC5HCH finger tandem domain represents a paradigmatic example of functional divergence within the NSD family. In this work, we prove and give a structural rationale for the uniqueness of the PHDvC5HCH domain of NSD1 in recognizing the C2HR Zinc finger domain of Nizp1 (NSD1 interacting Zn finger protein). Importantly, we show that, in a leukaemogenic context, Nizp1 is pivotal in driving the unscheduled expression of *HoxA* genes and of genes involved in the type I IFN pathway, triggered by the expression of the fusion protein NUP98-NSD1. These data provide the first insight into the pathophysiological relevance of the Nizp1-NSD1 functional association. Targeting of this interaction might open new therapeutic windows to inhibit the NUP98-NSD1 oncogenic properties.

Introduction

The nuclear receptor-binding SET (suppressor of variegation 3–9, enhancer of zeste and Trithorax), domain (NSD) family of histone H3 lysine 36 methyltransferases (HMKT) consists of three phylogenetically distinct members: NSD1, NSD2 (MMSET/WHSC1) and NSD3 (WHSC1L1). These histone methyltransferases catalyse the deposition of methyl groups to specific lysine (K) residues on histone tails herewith fine-regulating gene expression [1,2]. They also constitute a biomedically relevant protein family, as their overexpression, engagement in chromosomal translocations or mutations are associated to different developmental disorders (Sotos

Abbreviations
C5HCH, Cys-His-rich; CSP, chemical shift perturbation; HSQC, heteronuclear single quantum coherence; Nizp1, NSD1 interacting Zn finger protein; NMR, nuclear magnetic resonance; NSD, nuclear receptor-binding SET (suppressor of variegation 3–9, enhancer of zeste and Trithorax) domain; NUP98, nucleoporin 98kD; PHD, plant homeo-domain; shRNA, small hairpin RNA.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[3], Weaver [4] and Wolf-Hirschhorn syndromes [5]) and cancers (e.g. multiple myeloma [6–8], breast cancer [9], lung cancer [10], acute myeloid leukaemia [11]).

The NSD proteins are large multidomain proteins containing a catalytic SET domain, responsible for dimethylation of lysine 36 on histone H3 tail [12], and several repeated chromatin binding modules, including two PWWP (proline–tryptophan–tryptophan–proline) [13] and six Zn-binding PHD (plant homeo-domain) finger domains [14,15] (Fig. 1A). The NSD proteins normally play non-redundant roles during development, as shown by the lethality caused by genetic deletion of either NSD1 [16] or NSD2 in mice [17]. Their diversified activities in part also rely on their non-enzymatic domains that recruit the hosting protein to different chromatin regions through the recognition of diverse epigenetic marks and/or interactions with distinct protein partners. As a matter of fact, these non-enzymatic domains are actively involved in the transcriptional regulation mechanisms mediated by the NSD family and are emerging as potential pharmacological targets in malignant conditions, as shown for the PWWP domain [18–20]. Also the PHD fingers, and in particular the tandem domain composed by the fifth PHD finger and the adjacent cysteine rich C5HCH domain (PHDvC5HCH), seem to play a fundamental pathophysiological role within the hosting proteins. In NSD1, PHDvC5HCH ($PHDvC5HCH_{NSD1}$) is a hot spot for point mutations causing Sotos syndrome, a developmental overgrowth disease [21]. In the context of acute myeloid leukaemia (AML), where NSD1 is fused to the Nucleoporin 98 (NUP98; Fig. 1A), deletion of the very same domain reduces the expression of the *HoxA* oncogenes [11]. In another haematological cancer, multiple myeloma (MM), where NSD2 is translocated in a subset of patients, deletion of

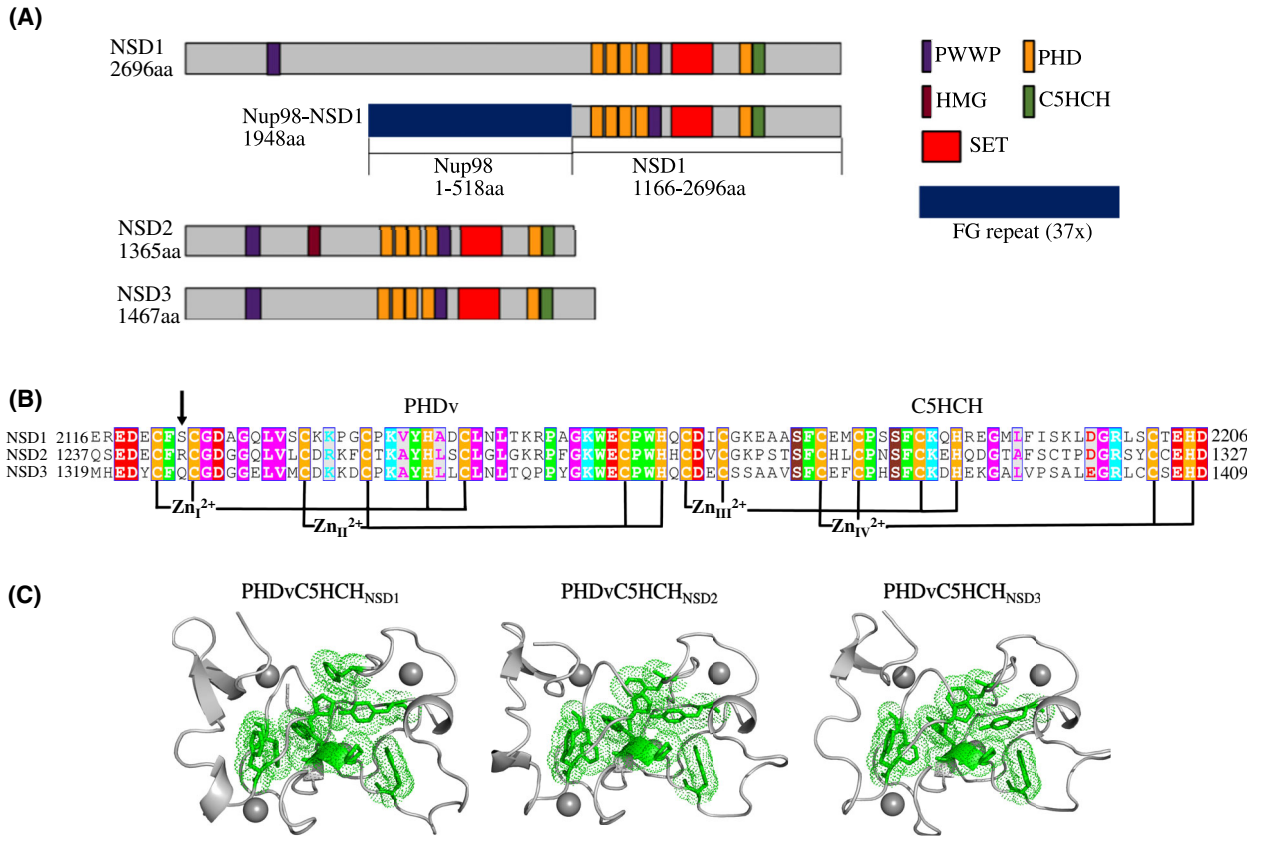


Fig. 1. The NSD family and the NSD-PHDvC5HCH domains (A) domain architecture of NSD1, NSD2 and NSD3. (B) Multiple sequence alignment (performed with Epspript [52]) of the human PHDvC5HCH domains of NSD family (NSD1 NP_071900.2; NSD2 096028.1; NSD3 sequence NP_075447.1, conserved residues are coloured; note that the human and murine sequences of $PHDvC5HCH_{NSD2}$ and $PHDvC5HCH_{NSD3}$ are identical; the human and the murine sequence of $PHDvC5HC_{NSD1}$ are 99% identical, with Val2166 in the murine sequence being substituted by I2167 in the human one). The arrow indicates residues $S2123_{NSD1}$ $R1244_{NSD2}$, $Q1326_{NSD3}$ (C) cartoon representation of $PHDvC5HCH_{NSD1}$ (PDB code: 2NAA), $PHDvC5HCH_{NSD2}$, (ALPHAFOLD2 model) and $PHDvC5HCH_{NSD3}$ (PDB code: 4GND) with conserved aromatic residues represented in green sticks and dots. Molecular images were generated by PYMOL molecular graphics system, version 2.0 Schrödinger, LLC.

The FEBS Journal (2022) © 2022 The Authors. *The FEBS Journal* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the corresponding domain, PHDvC5HCH$_{NSD2}$, results in the reduction of the activation of cell adhesions genes (e.g. *JAM2*) and of MM cells proliferation [8,22]. Finally, in NSD3, several cancer-related missense or frameshift mutations target PHDvC5HCH$_{NSD3}$ [23], further supporting the relevant pathophysiological role of this domain in the NSD family.

The functional versatility of PHDvC5HCH, showing high sequence identity conservation (~ 60%) within the NSD members (Fig. 1B,C), is also testified by the diversified histone and non-histone recognition ability within this protein family. PHDvC5HCH$_{NSD3}$ specifically decodes the H3K4me0K9me3 signature [24], whereas PHDvC5HCH$_{NSD2}$ seems to prefer the H3K4me0K9me0 mark [24]. PHDvC5HCH$_{NSD1}$ diverges from the classical epigenetic reader activity: on the one hand, it displays only weak interactions (mainly of electrostatic nature) for these and other histone epigenetic marks, on the other hand, we have previously shown that it binds with low micromolar affinity (Kd ~ 4 μM) to the degenerate Zn-finger domain (~ 40 aminoacids) of Nizp1 (C2HR$_{Nizp1}$) [25]. Nizp1, also known as Zfp494 [26], derives its name from its interaction with NSD1 (NSD1 interacting Zinc finger protein) via its degenerate Zn-finger domain; it is a poorly characterized multidomain repressor, harbouring other four Zn-fingers domains, a SCAN (named after SRE-ZBP, CTfin51, AW-1 and Number 18 cDNA) and the repressive KRAB (Krüppel-associated box)-domain [27]. Nizp1 repressor activity appears to be reinforced by NSD1 recruitment [27,28], even though the real physiological relevance (if any) of this interaction remains elusive. This lack of knowledge might be also ascribable to the context dependent transcriptional activity of the cognate protein NSD1, that can work both as transcriptional activator and repressor [29–31]. Overall, this finger–finger interaction appears to be a peculiarity of NSD1, as GST-pulldown assays [27] on PHDvC5HCH$_{NSD3}$ and yeast two hybrid assays using fragments corresponding to the C5HCH domain of both NSD2 and NSD3 did not support a direct interaction with C2HR$_{Nizp1}$ [24]. However, these yeast two hybrid experiments and their interpretation should be taken with caution, as they have been performed on a fragment that only included a part of the entire tandem, that should be actually considered as an indivisible structural and functional unit.

Thus, it remains unclear whether the interaction of Nizp1 with the NSD family is biologically relevant and whether also NSD2 and NSD3 are potential interactors of Nizp1 via PHDvC5HCH and C2HR association. Thus in this work, we studied the functional role of Nizp1/NUP98-NSD1 interaction and compared at atomistic level by nuclear magnetic resonance (NMR) the ability and specificity of the three NSD-PHDvC5HCH domains to interact with C2HR$_{Nizp1}$.

Our results provide first insights into the functional role of Nizp1, showing that, in a cellular model context of acute myeloid leukaemia (AML), the fusion protein NUP98-NSD1 and Nizp1 genetically interact to regulate the expression of *HoxA* genes and of genes involved in type I Interferon (IFN) pathway, thus revealing a hitherto unknown pathophysiological relevance for this interaction. Additionally, we prove and give a structural rationale for the diverging ability of the NSD members in recognizing C2HR$_{Nizp1}$.

Results

Nizp1 and NUP98-NSD1 genetically interact to regulate HoxA genes and type I interferon transcriptional program

To determine whether the interaction between NSD1 and Nizp1 has any functional relevance, we focused on the oncogenic fusion protein NUP98-NSD1 that contains all the chromatin related domains present in NSD1, including the tandem PHDvC5HCH (Fig. 1A). Notably, NUP98-NSD1 translocation has been identified in ~ 16% of paediatric cases of acute leukaemias [32], and previous studies have proposed an essential role for PHDvC5HCH in the inappropriate *HoxA5* and *HoxA9* genes activation and oncogenesis [11].

Thus, to verify the contribution of PHDvC5HCH to tumorigenesis, we took advantage of a well-recognized cellular assay to study the leukaemogenic properties of specific targets, based on isolated mouse c-Kit$^+$/Sca-1$^+$/Lin$^-$ bone marrow progenitors [33]. We retrovirally infected these cells with plasmids either coding for NUP98-NSD1 or for a deletion mutant of NUP98-NSD1 lacking the PHDvC5HCH domain (NUP98-NSD1Δ$_{PHDvC5HCH}$) and subsequently analysed *HoxA* genes expression by RT-qPCR. In line with a previous report [11], NUP98-NSD1 overexpression increased the expression of *HoxA5* and *HoxA9* genes as compared to the control. Conversely, transduction with NUP98-NSD1Δ$_{PHDvC5HCH}$ failed to induce the expression of these oncogenes, confirming a direct involvement of the PHDvC5HCH domain in the leukaemogenic properties of NUP98-NSD1 [11] (Fig. 2A).

Next, inspired by a previous model, that suggested the existence of a yet unidentified adaptor protein(s) to recruit NUP98-NSD1 via the PHDvC5HCH domain to specific gene loci [11], we asked whether Nizp1 could mediate NUP98-NSD1-induced transcriptional deregulation of *HoxA5-A9*. To investigate this

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

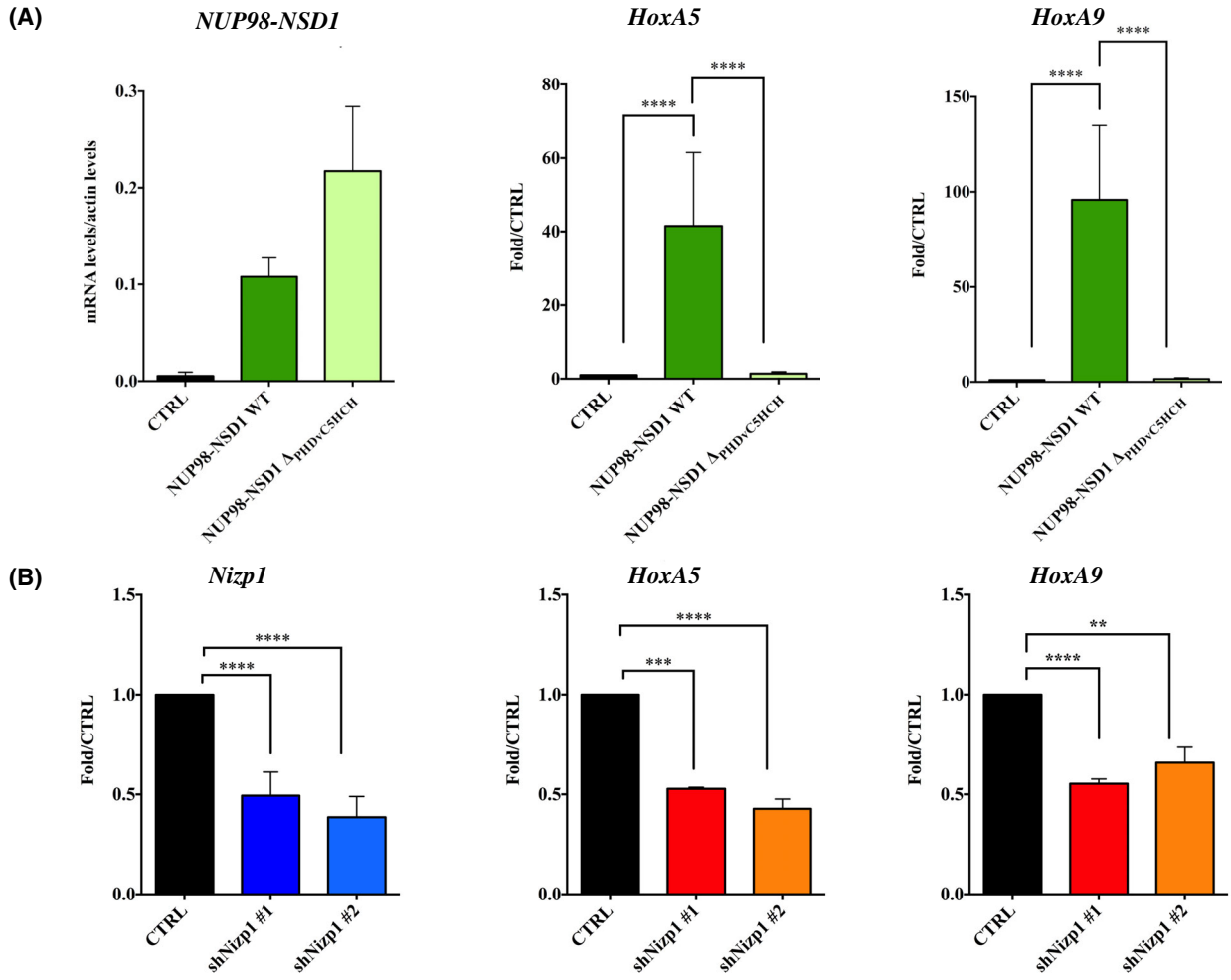


Fig. 2. *Nizp1* And *NUP98-NSD1* genetically interact to regulate *HoxA* genes (A) c-kit$^+$/Sca-1$^+$/Lin$^-$ cells were transduced with an empty vector (CTRL) or vectors expressing the full-length form of NUP98-NSD1 (WT) or a deletion mutant lacking the PHDvC5HCH domain (NUP98-NSD1 $\Delta_{PHDvC5HCH}$). The histograms show the mRNA expression levels of *NUP98-NSD1* (normalized to Actin mRNA levels), *HoxA5* and *HoxA9* genes (normalized to Actin mRNA levels and reported as fold relative to CTRL). (B) NUP98-NSD1 expressing cells were transduced with a control shRNA (CTRL) or two different *Nizp1*-specific shRNA (shNizp1 #1 and #2). The mRNA levels of *Nizp1* are normalized to Actin mRNA levels and reported as fold relative to CTRL. The mRNA levels of *HoxA5* and *HoxA9* genes are normalized to Actin mRNA levels and reported as fold relative to CTRL. Results are averaged from three independent experiments (mean ± standard error of the mean [SEM]). Statistical analysis was performed using two-way ANOVA followed by Dunnett correction for multiple comparisons. Adjusted *P* value: **** < 0.0001; *** = 0.0002; ** = 0.0011.

hypothesis, we silenced *Nizp1* in c-Kit$^+$/Sca-1$^+$/Lin$^-$ cells expressing NUP98-NSD1 using two different *Nizp1*-specific shRNAs (sh*Nizp1* #1 and #2), each effective in silencing *Nizp1*. Importantly, its downregulation impaired the induction of NUP98-NSD1-dependent *HoxA* genes expression with respect to the control cells, as assessed by RT-qPCR, thus indicating that Nizp1 strongly contributes to NUP98-NSD1 oncogenes expression (Fig. 2B).

We then sought to obtain a more comprehensive perspective on the role of Nizp1 in regulating genes and pathways modulated by NUP98-NSD1. We performed RNA-seq analysis on c-Kit$^+$/Sca-1$^+$/Lin$^-$ cells expressing or not NUP98-NSD1 and cells expressing NUP98-NSD1 but depleted for Nizp1 (sh*Nizp1* #1 and #2; Table S1). Gene Set Enrichment Analysis (GSEA) analysis revealed that overexpression of NUP98-NSD1 significantly upregulated the Type I Interferon (*IFN*) and inflammatory response pathways (Fig. 3A,B, Table S2). Importantly, the concomitant downregulation of Nizp1 in the presence of NUP98-NSD1 overexpression ablated the expression of genes involved in these pathways, including *Rsad2*, *Ifit2 and Cxcl10* (Fig. 3C,D).

The FEBS Journal (2022) © 2022 The Authors. *The FEBS Journal* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

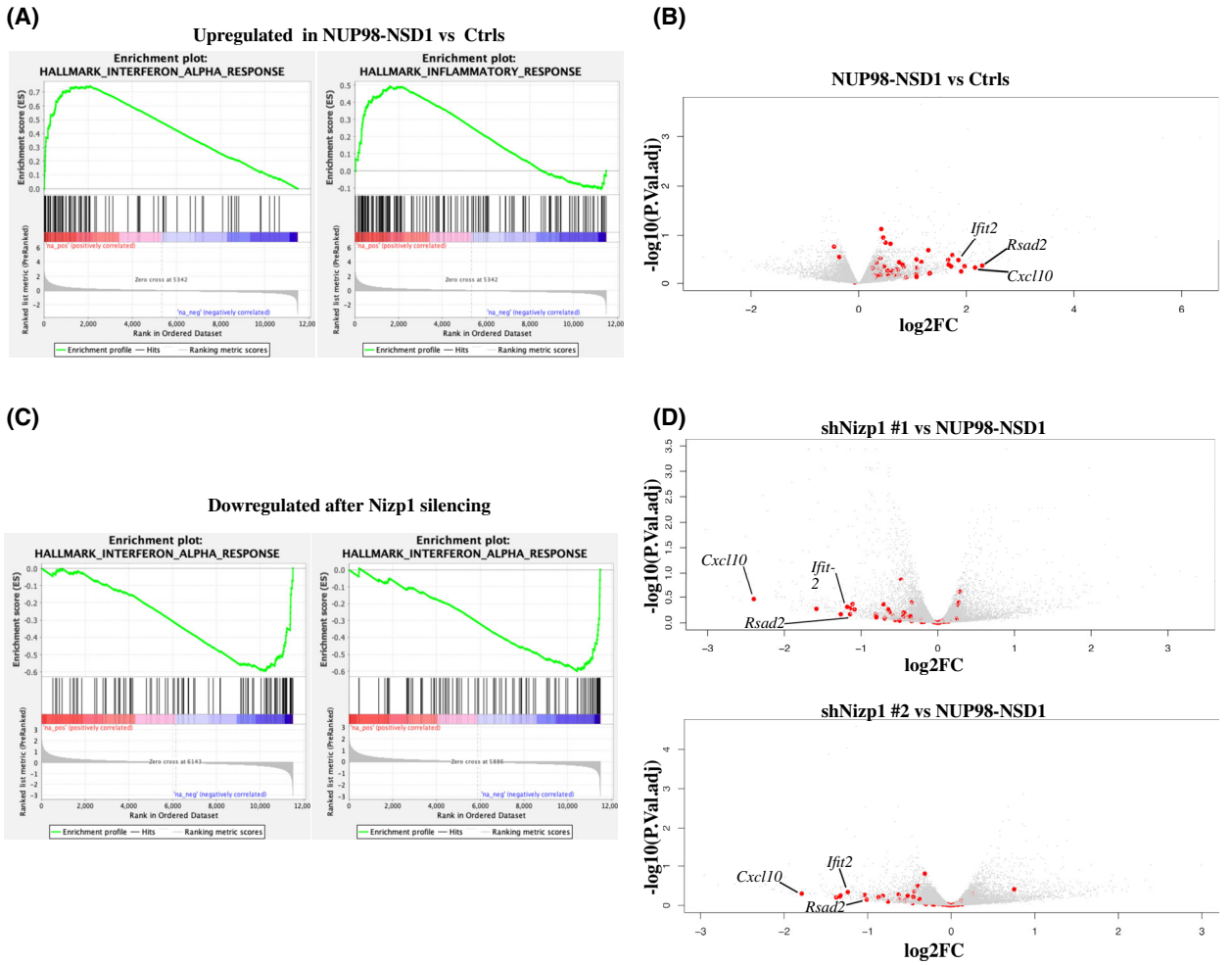


Fig. 3. RNA-seq analysis of NUP98-NSD1 and Nizp1 expression in c-kit$^+$/Sca-1$^+$/Lin$^-$ cells (A) gene set enrichment analysis (GSEA) of MSigDB (molecular signatures database) Hallmark pathways of top significant upregulated pathways in c-kit$^+$/Sca-1$^+$/Lin$^-$ cells expressing NUP98-NSD1 vs. CTRL cells (log2FC was calculated across three independent experiments) (B) volcano plot of RNA-seq differential gene expression (DGE) of NUP98-NSD1 vs. CTRL cells (n = 3). (C) Gene set enrichment analysis (GSEA) of MSigDB (molecular signatures database) for type I interferon Hallmark pathway in downregulated pathways for NUP98-NSD1 expressing cells upon transduction with sh*Nizp1* #1 (left) and sh*Nizp1* #2 constructs (right). (log2FCs were calculated across three independent experiments). Complete tables with GSEA results are reported in Table S2. (D) volcano plot of RNA-seq differential gene expression of sh*Nizp1* #1 (top) and sh*Nizp1* #2 (bottom) vs. NUP98-NSD1 cells (n = 3). Complete tables with DGE results are reported in Table S1. In volcano plots, red circles highlight the leading edge GSEA genes, i.e. the ones that mostly contribute to the enrichment in type I IFN hallmark pathway. (GSEA was performed with 1000 permutations).

Taken together, these results show for the first time that Nizp1 plays a fundamental role in fostering NUP98-NSD1-dependent oncogenes expression, shedding new light on the functional relevance of Nizp1/NUP98-NSD1 interaction.

The finger–finger interaction between C2HR$_{Nizp1}$ and PHDvC5HCH is specific for NSD1

Nizp1 and NSD1 interact through the C2HR domain and PHDvC5HCH tandem domain, respectively [28], and in the past, we have characterized at molecular level this finger–finger interaction [25]. To identify residues driving binding specificity and to obtain an exhaustive overview on the actual ability (or inability) of CH2R$_{Nizp1}$ to specifically recognize the PHDvC5HCH tandem domain of NSD2 and NSD3, we expressed in *Escherichia coli* and purified the corresponding domains (PHDvC5HCH$_{NSD2}$, PHDvC5HCH$_{NSD3}$; Fig. 1B) in unlabeled and isotopically labelled form (^{15} N/^{13}C). We assigned by classical NMR triple resonance experiments their backbone resonances (Fig. 4A,B) and investigated

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

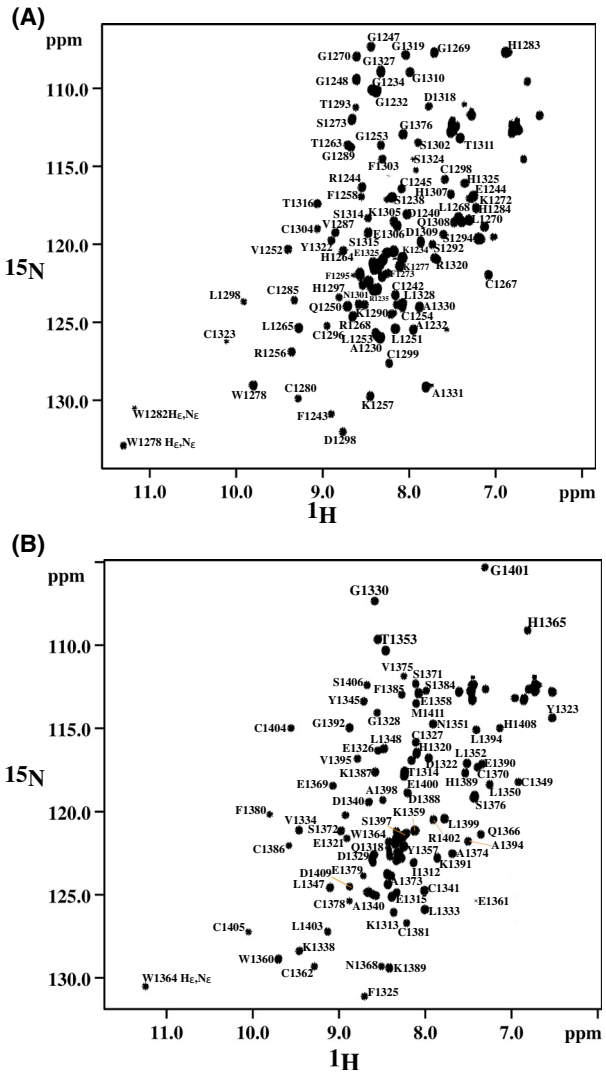


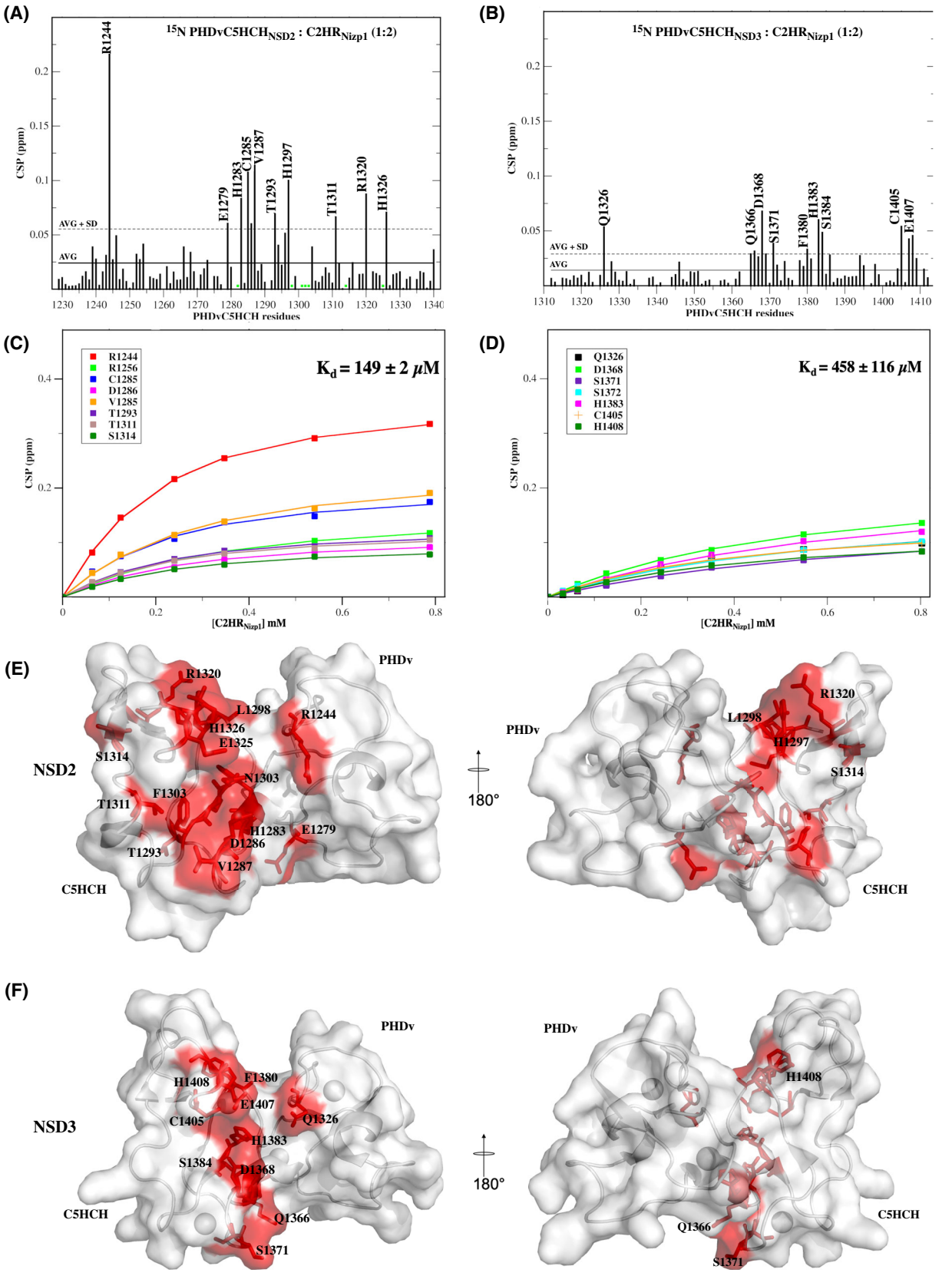
Fig. 4. ^{1}H-^{15}N HSQC spectra of ^{15}N PHDvC5HCH$_{NSD2}$ and ^{15}N PHDvC5HCH$_{NSD3}$ (A) ^{1}H-^{15}N HSQC spectrum of ^{15}N PHDvC5HCH$_{NSD2}$ with assignments, at T = 295 K, phosphate buffer 20 mM, pH 6.3, 150 mM NaCl, 1 mM TCEP, 10 μM $ZnCl_2$. (B) ^{1}H-^{15}N HSQC spectrum of ^{15}N PHDvC5HCH$_{NSD3}$ with assignments, at T = 295 K, phosphate buffer 20 mM, pH 6.3, 150 mM NaCl, 1 mM TCEP, 10 μM $ZnCl_2$.

their interaction with C2HR$_{Nizp1}$ by NMR-based chemical shift perturbation (CSP) assays. To this end, 2D ^{1}H-^{15}N HSQC (heteronuclear single quantum coherence) spectra of ^{15}N labelled PHDvC5HCH$_{NSD2}$ and PHDvC5HCH$_{NSD3}$ were recorded to monitor changes in their ^{1}H-^{15}N chemical shifts upon successive additions of unlabelled C2HR$_{Nizp1}$ (Fig. 5A,B). For both PHDvC5HCH$_{NSD2}$ and PHDvC5HCH$_{NSD3}$, the complexes were in the fast exchange regime on the NMR chemical-shift time scale with dissociation constants of ~ 150 and ~ 400 μM, respectively, as measured by fitting the shifts of selected peaks as a function of C2HR$_{Nizp1}$ concentration (Fig. 5C,D). Notably, the analogous titration performed using ^{15}N PHDvC5HCH$_{NSD1}$ was in the slow-intermediate exchange, consistent with the stronger affinity (~ 4 μM) [25]. In agreement with the weaker interaction, the entity of the CSPs of both domains (with CSP$_{AVG}$ of 0.025 and 0.012 ppm for PHDvC5HCH$_{NSD2}$ and PHDvC5HCH$_{NSD3}$, respectively) was by far lower as compared to PHDvC5HCH$_{NSD1}$ (CSP$_{AVG}$ = 0.100 ppm) [25]. Notably, residues exhibiting significant amide chemical shift changes (Fig. 5A,B) once mapped on PHDvC5HCH$_{NSD2}$ ALPHAFOLD2–model (Fig. 5E; R1244, E1279, H1283, C1285, V1287, T1293, H1297, T1311, R1311, R1320, H1326) and on PHDvC5HCH$_{NSD3}$ crystallographic structure (Fig. 5F; Q1326, Q1366, D1368, S1371, F1380, H1383, S1384, 1405, E1407) identified the interdomain region between PHDv and C5HCH for the finger-finger interaction (Fig. 5E,F). A similar binding surface was previously identified for PHDvC5HCH$_{NSD1}$ [25]. Reversed titrations of unlabelled PHDvC5HCH$_{NSD2}$ and PHDvC5HCH$_{NSD3}$ into ^{15}N labelled C2HR$_{Nizp1}$ were also performed, confirming the weak interaction, as assessed by the smaller CSPs (PHDvC5HCH$_{NSD2}$, CSP$_{AVG}$ = 0.035 ppm; PHDvC5HCH$_{NSD3}$, CSP$_{AVG}$ = 0.028 ppm; Fig. 6A,B) with respect to PHDvC5HCH$_{NSD1}$ (CSP$_{AVG}$ = 0.095 ppm) [25]. In all the ^{15}N C2HR$_{Nizp1}$ titrations, the residues comprising part of the α-helix (V418, F420, R422) and the evolutionary conserved RWR-loop

Fig. 5. Interaction of ^{15}N PHDvC5HCH$_{NSD2}$ and ^{15}N PHDvC5HCH$_{NSD3}$ with C2HR$_{Nizp1}$ (A) Bar graph showing residue-specific amide chemical shift perturbations (CSPs) of ^{15}N-labelled PHDvC5HCH$_{NSD2}$ (0.12 mM) upon addition of 2-fold excess of C2HR$_{Nizp1}$. Missing residues are prolines or are absent because of exchange with the solvent. The continues line represents the average CSP (AVG), the dashed line represents AVG + one standard deviation (SD), green asterisks indicate residues whose resonances disappear upon binding. (B) Bar graph showing residue-specific amide CSPs of ^{15}N PHDvC5HCH$_{NSD3}$ (0.12 mM) upon addition of 2-fold excess of C2HR$_{Nizp1}$. Missing residues are prolines or are absent because of exchange with the solvent. The continuous line represents the average CSP (AVG), the dashed line represents AVG + one standard deviation (SD). (C) Weighted average of selected amides ^{1}H and ^{15}N chemical shift changes of PHDvC5HCH$_{NSD2}$ in the presence of increasing concentration of C2HR$_{Nizp1}$. Nonlinear curve fitting yields Kd = 149 ± 2 μM. (D) Weighted average of selected amides ^{1}H and ^{15}N chemical shift changes of PHDvC5HCH$_{NSD3}$ in the presence of increasing concentrations of C2HR$_{Nizp1}$. Nonlinear curve fitting yields Kd = 458 ± 116 μM. (E) Cartoon and surface representation of PHDvC5HCH$_{NSD2}$. Residues showing significant CSP are represented in red sticks. (F) Cartoon and surface representation of PHDvC5HCH$_{NSD3}$. Residues showing significant CSP are represented in red sticks. Molecular images were generated by PYMOL molecular graphics system, version 2.0 Schrödinger, LLC.

The FEBS Journal (2022) © 2022 The Authors. *The FEBS Journal* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

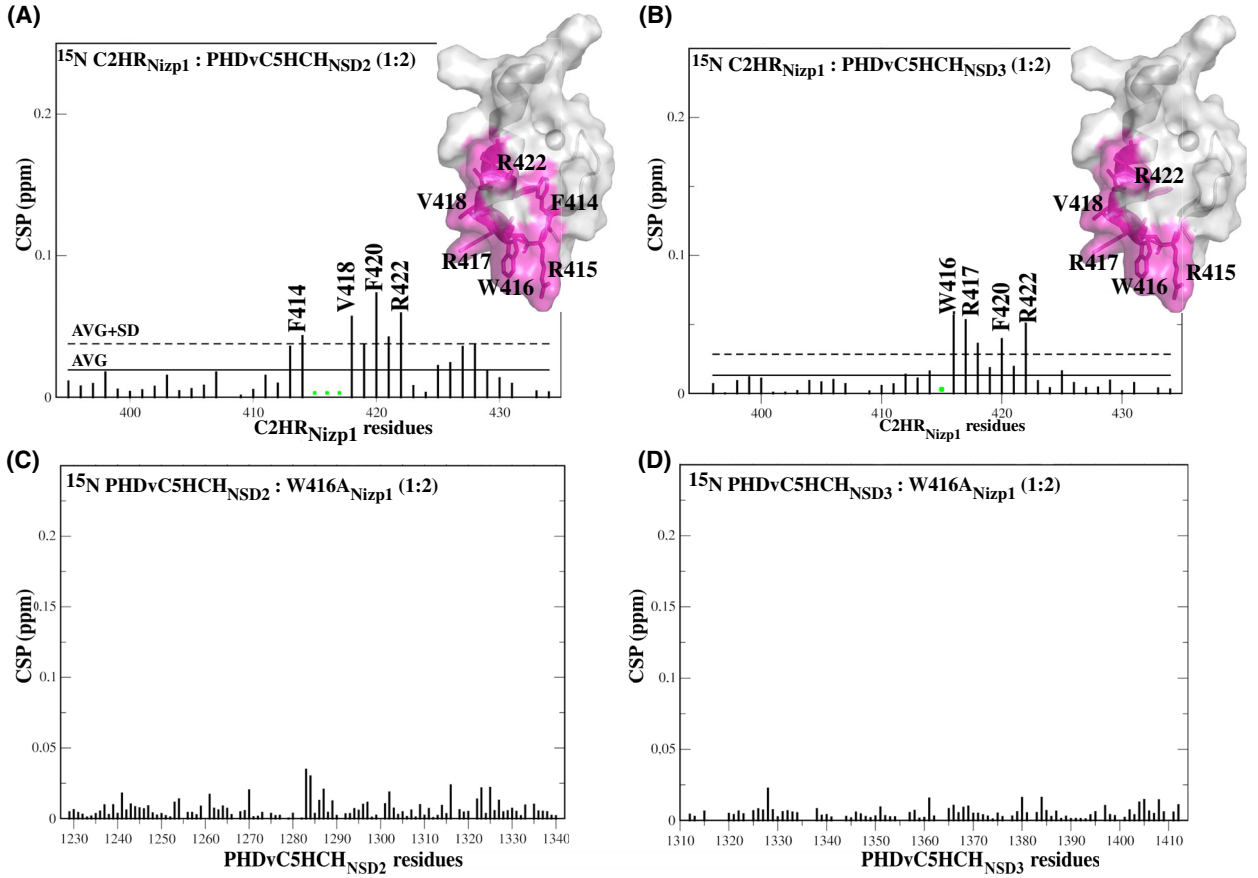


Fig. 6. Interaction of $C2HR_{Nizp1}$ and of C2HR-$W416A_{Nizp1}$ mutant with $PHDvC5HCH_{NSD2}$ and $PHDvC5HCH_{NSD3}$. (A) Bar graphs showing residue-specific amide chemical shift perturbations (CSPs) of ^{15}N $C2HR_{Nizp1}$ (0.2 mM) upon addition of a 2-fold excess of $PHDvC5HCH_{NSD2}$. In the inset cartoon and surface representation of $C2HR_{Nizp1}$(grey). Residues showing significant CSP or peaks disappearance upon binding are represented in sticks and are coloured in magenta. (B) Bar graphs showing residue-specific amide chemical shift perturbations (CSPs) of ^{15}N $C2HR_{Nizp1}$ (0.2 mM) upon addition of a 2-fold excess of $PHDvC5HCH_{NSD3}$. In the inset cartoon and surface representation of $C2HR_{Nizp1}$ (grey). Residues showing significant CSP or peaks disappearance upon binding are represented in sticks and are coloured in magenta. In all the histograms, the continuous line represents average AVG, the dashed line represents AVG + one standard deviation (SD), green asterisks indicate residues whose resonances disappear upon binding. Molecular images were generated by PYMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. (C) Bar graphs showing residue-specific amide chemical shift perturbations (CSPs) of ^{15}N $PHDvC5HCH_{NSD2}$ (0.2 mM) upon addition of a 2-fold excess of C2HR-$W416A_{Nizp1}$ mutant ($W416A_{Nizp1}$). (D) Bar graphs showing residue-specific amide chemical shift perturbations (CSPs) of ^{15}N $PHDvC5HCH_{NSD3}$ (0.2 mM) upon addition of a 2-fold excess of C2HR-$W416A_{Nizp1}$ mutant ($W416A_{Nizp1}$). Missing bars correspond to either Prolines or amides that are not visible because of exchange with the solvent.

(R415, W416, R417) [25] were the ones mostly affected (in terms of CSPs or peaks disappearance) by the interaction with the PHDvC5HCH tandem domain. This suggests an important role for the RWR signature in the interaction (Fig. 6A,B), in agreement with previous observations for $PHDvC5HCH_{NSD1}$ [25]. Accordingly, fold-preserving mutation of W416 into Alanine [25] ($C2HR_{Nzip1}$-W416A mutant) resulted in almost total abrogation of the binding, as indicated by the negligible CSPs induced in the ^{1}H-^{15}N HSQC spectra of both $PHDvC5HCH_{NSD2}$ and $PHDvC5HCH_{NSD3}$ (Fig. 6C,D). Collectively, these results indicate that the interaction of $C2HR_{Nizp1}$ with the tandem domains $PHDvC5HCH_{NSD2}$ and $PHDvC5HCH_{NSD3}$ is reminiscent of the one with $PHDvC5HCH_{NSD1}$. It relies on the interaction between the RWR-loop and the shallow groove at the interface of PHDv and C5HCH (Fig. 5E,F); however, the binding is two orders of magnitude smaller (Fig. 5C,D) as compared to $PHDvC5HCH_{NSD1}$, thus supporting the specificity of $PHDvC5HCH_{NSD1}$ in recognizing $C2HR_{Nizp1}$ [25].

The FEBS Journal (2022) © 2022 The Authors. *The FEBS Journal* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

A unique serine present in PHDvC5HCH$_{NSD1}$ confers specificity to the interaction with C2HR$_{Nizp1}$

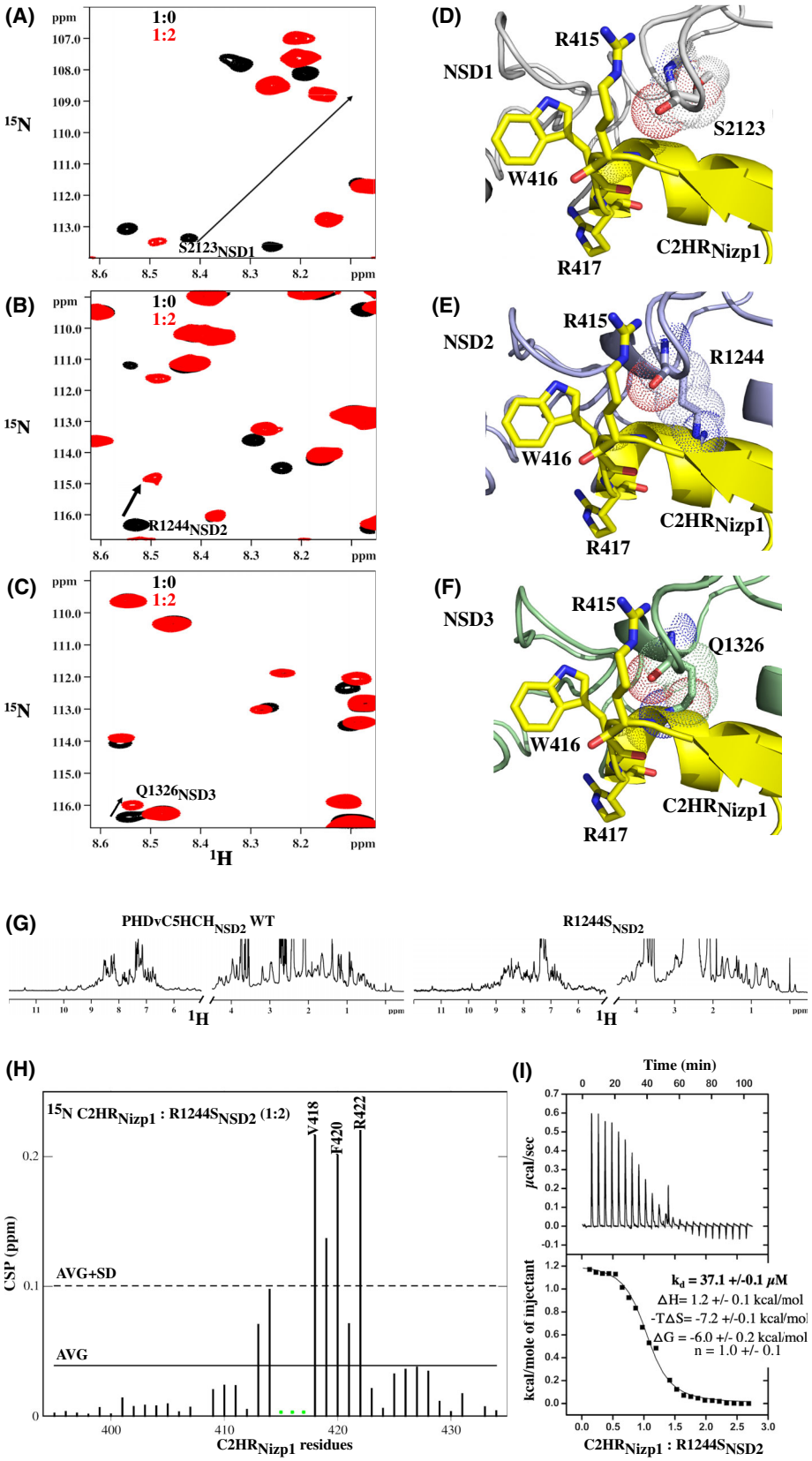
To gain insight on the residues responsible of the binding preference of C2HR$_{Nizp1}$ for PHDvC5HCH$_{NSD1}$, we explored the multiple sequence alignment of the three NSD tandem PHDvC5HCH domains (Fig. 1B) and the interaction surface with C2HR$_{Nizp1}$, located on the conserved hydrophobic inter-domain groove (Fig. 1C). We were intrigued by a Serine that is present only in NSD1 (S2123$_{NSD1}$; Fig. 1B), whose amide resonances displayed the highest CSP in titrations with C2HR$_{Nizp1}$ (Fig. 7A, [25]). The equivalent position is occupied by an Arginine (R1244$_{NSD2}$) and a Glutammine (Q1326$_{NSD3}$) in PHDvC5HCH$_{NSD2}$ and PHDvC5HCH$_{NSD3}$, respectively (Fig. 1B). Even though the corresponding amide resonances displayed the highest CSPs upon titration of C2HR$_{Nizp1}$ the entity of the perturbations was by far smaller (Figs 5A,B and 7B,C), as compared to PHDvC5HCH$_{NSD1}$ (Fig. 7A [25]), in agreement with the weaker interaction. Indeed, superposition of PHDvC5HCH$_{NSD2}$ and PHDvC5HCH$_{NSD3}$ structures on the PHDvC5HCH$_{NSD1}$/C2HR$_{Nizp1}$ complex model [25] (Fig. 7D) suggests that residues R1244$_{NSD2}$ (Fig. 7E) and Q1326$_{NSD3}$ (Fig. 7F) might create steric clashes with the α-helix of C2HR$_{Nizp1}$, thus weakening the interaction. To confirm the role of this serine in the binding, we generated the R1244S-PHDvC5HCH$_{NSD2}$ and Q1326S-PHDvC5HCH$_{NSD3}$ mutants to be used in NMR titrations of ^{15}N C2HR$_{Nizp1}$. Only the NSD2 mutant was soluble, well folded (Fig. 7G) and suitable for NMR binding experiments. Importantly, when titrated into ^{15}N C2HR$_{Nizp1}$ it induced larger CSPs (CSP$_{AVG}$ = 0.039 ppm; Fig. 7H) and higher affinity (37 μM), as compared to wild-type PHDvC5HCH$_{NSD2}$ (CSP$_{AVG}$ = 0.019 ppm). Interestingly, similarly to PHDvC5HCH$_{NSD1}$ [25], the reaction was endothermic with a favourable entropic contribution, suggesting an analogous binding mechanism (Fig. 7I). Taken together, these data indicate that a non-conserved Serine located at the interface of the PHDvC5HCH$_{NSD1}$ domain is indispensable for complex formation and is at the basis of the binding specificity between PHDvC5HCH$_{NSD1}$ and C2HR$_{Nizp1}$.

Discussion

NSD proteins are a family of methyltransferases, whose dysfunctions result in epigenomic aberrations leading to different types of cancer [1]. Although the three family members share the same conserved modular architecture, they exert diversified activities that most likely are also linked to small sequence differences in their non-catalytic domains. These domains, as observed in other chromatin related proteins, work as versatile structural platforms that not only direct the hosting protein to specific chromatin modifications [34], but also contribute to the recruitment of different cofactors contributing to the activity of the hosting protein [35]. In this sense, the PHDvC5HCH tandem domain of the NSD family represents a paradigmatic example of the multifaceted role of these non-catalytic domains, whose activities range from epigenetic readers to protein–protein interaction hubs as demonstrated for PHDvC5HCH$_{NSD3}$ [24] and PHDvC5HCH$_{NSD1}$ [25], respectively. The latter has been shown in the past to be indispensable for the interaction with the transcriptional co-repressor Nizp1 via its C2HR Zn-finger domain [27];

Fig. 7. A non conserved serine confers specificity to PHDvC5HCH$_{NSD1}$/C2HR$_{Nizp1}$ interaction (A) zoom into the superposition of ^{1}H–^{15}N HSQC spectrum of PHDvC5HCH$_{NSD1}$ without (black) and with unlabeled C2HR$_{Nizp1}$ (red) [25]. (B) Zoom into the superposition of ^{1}H–^{15}N HSQC spectrum of PHDvC5HCH$_{NSD2}$ without (black) and with (red) unlabeled C2HR$_{Nizp1}$. (C) Zoom into the superposition of ^{1}H–^{15}N HSQC spectrum of PHD$_{V}$C5HCH$_{NSD3}$ without (black) and with (red) unlabeled C2HR$_{Nizp1}$. (D) Representative pose of the HADDOCK model of PHDvC5HCH$_{NSD1}$ (grey) in complex with C2HR$_{Nizp1}$ (yellow) [25]. S2123$_{NSD1}$ is explicitly shown. (E) PHDvC5HCH$_{NSD2}$ ALPHAFOLD2 model (light purple) superimposed on HADDOCK model of PHD$_{V}$C5HCH$_{NSD1}$ (not shown) in complex with C2HR$_{Nizp1}$ (yellow). The side chain of R1244$_{NSD2}$ creates steric clashes with C2HR$_{Nizp1}$. (F) PHD$_{V}$C5HCH$_{NSD3}$ X-ray structure (grey; PDB code: 4GND) superimposed on the HADDOCK model of PHDvC5HCH$_{NSD1}$ (not shown) in complex with C2HR$_{Nizp1}$ (yellow). The side chain of Q1326$_{NSD3}$ creates steric clashes with C2HR$_{Nizp1}$. The side chains of C2HR$_{Nizp1}$ RWR motif are explicitly shown. Molecular images were generated by PYMOL molecular graphics system, version 2.0 Schrödinger, LLC. (G) 1D-^{1}H NMR spectra of PHDvC5HCH$_{NSD2}$ WT (left) and R1244S-PHDvC5HCH$_{NSD2}$ (R1244S$_{NSD2}$) (right). The good peak dispersion and line width of the R1244S-PHDvC5HCH$_{NSD2}$ spectrum indicate that the mutation preserves the domain fold. (H) Bar graph showing residue-specific amide chemical shift perturbations (CSPs) of ^{15}N C2HR$_{Nizp1}$ (0.2 mM) upon addition of a 2-fold excess of R1244S-PHDvC5HCH$_{NSD2}$ mutant. The continuous line represents the average CSP (AVG), the dashed line represents AVG + one standard deviation (SD), green asterisks indicate residues whose resonances disappear upon binding. (I) ITC-binding curves of C2HR$_{Nizp1}$ titrated into R1244S-PHDvC5HCH$_{NSD2}$. The upper panel shows the sequential heat pulses for domain–domain binding, and 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 (line) (molar ratio, C2HR$_{Nizp1}$: R1244S-PHDvC5HCH$_{NSD2}$). The experiment has been repeated twice, one representative experiment is shown with the fitting error of the curve.

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



The FEBS Journal (2022) © 2022 The Authors. *The FEBS Journal* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

however, the pathophysiological relevance (if any) for the hosting protein of this finger–finger interaction has remained until now elusive. To tackle this issue, we focused on NUP98-NSD1 onco-fusion protein and asked whether Nizp1 could work as adaptor, modulating NUP98-NSD1-dependent unscheduled expression of oncogenes. We first concentrated on *HoxA* genes whose expression had been previously proposed to depend also on PHDvC5HCH [11]. Indeed, our data obtained in mouse c-Kit$^+$/Sca-1$^+$/Lin$^-$ bone marrow progenitors transduced with NUP98-NSD1 and with a deletion mutant lacking PHDvC5HCH confirmed the contribution of the tandem domain to the oncogenic potential of NUP98-NSD1. This result is important, as it shows that, besides the documented pathophysiological relevance of the NUP98 moiety [36–39] and of the catalytic SET domain [11], also other non-catalytic domains within the C-terminal fusion partner may actively participate in oncogenes expression. Notably, using the same experimental set-up we could show for the first time that the *HoxA5-9* transcription was strongly dependent not only on the PHDvC5HCH tandem domain but also on the presence of Nizp1, as shown by the drastic reduction of *HoxA* genes mRNA levels upon silencing of *Nizp1* in mouse c-Kit$^+$/Sca-1$^+$/Lin$^-$ bone marrow progenitors transduced with NUP98-NSD1. Of note also the expression of genes involved in type I INF pathway program (such as *Rsad2*, *Ifit2* and *Cxcl10*), which are usually upregulated in oncogenic NUP98-fusion proteins [38], appeared to be fostered by the presence of Nizp1, as assessed by RNA-seq analysis. Intriguingly, while these results highlight an unprecedented role for Nizp1 in promoting the activation of NUP98-NSD1-dependent oncogenes, previous reporter assays suggested co-repressor activity of Nizp1 on NSD1 transcriptional activity via its C2HR domain [27]. Collectively, our results reveal a context-dependent activity for Nizp1 that can result either in co-repression or co-activation.

Having established for the first time a functional pathophysiological interaction between Nizp1 and the onco-fusion protein NUP98-NSD1, we asked whether at molecular level the finger-finger interaction between C2HR$_{Nizp1}$ and PHDvC5HCH was restricted to NSD1 or whether the other members of the NSD family could potentially associate to Nizp1. By NMR binding assays and mutagenesis we have demonstrated that, despite the high sequence identity and a similar interaction pattern with C2HR$_{Nizp1}$, they bind with reduced affinities to C2HR$_{Nizp1}$. The interactions of both PHDvC5HCH$_{NSD2}$ and PHDvC5HCH$_{NSD3}$ are more than two orders of magnitude weaker (millimolar range) than PHDvC5HCH$_{NSD1}$ (low micromolar range) [25], thus definitively establishing a binding preference for NSD1 and possibly excluding a functional interaction of Nizp1 with the other NSD members. Multiple sequence alignment of the three NSD-PHDvC5HCH domains and inspection of the binding surfaces, as identified by CSPs, suggest that the specificity of the interaction strongly relies on a non-conserved Serine (S2123$_{NSD1}$) that is present only in PHDvC5HCH$_{NSD1}$ sequence. The equivalent position is replaced by an Arginine (R1244$_{NSD2}$) and a Glutammine (Q1326$_{NSD3}$) in PHDvC5HCH$_{NSD2}$ and PHDvC5HCH$_{NSD3}$, respectively. These residues are predicted to create steric clashes with the RWR loop of C2HR$_{Nizp1}$. As a matter of fact mutation of R1244 into serine (R1244S) in PHDvC5HCH$_{NSD2}$ partially reverts the binding ability of PHDvC5HCH$_{NSD2}$ increasing the affinity for C2HR$_{Nizp1}$ from 150 to 37 μM, thus confirming the fundamental role of this residue in binding specificity.

The discovery that Nizp1 is relevant for NUP98-NSD1 oncogenic properties suggests that the inhibition of the specific C2HR$_{Nizp1}$/PHDvC5HCH$_{NSD1}$ interaction could represent a new valuable pharmacological strategy against AML. Modulators of this finger–finger interaction through small molecules [19], or design of PROTAC molecules targeting this tandem domain, as similarly proposed for the PWWP domain of NSD3 [39] might create new opportunities for the development of inhibitors of aberrant NSD1 function(s) and complement the classical blockade of its histone methyltransferase activity [40].

Materials and methods

Purification and transduction of c-Kit$^+$/Sca-1$^+$/Lin$^-$ bone marrow progenitors

Lin$^-$ cells were purified from the bone marrow of 8- to 10-week-old BALB/c mice. Mice (Charles River Laboratories, Milan, Italy) were housed and bred in the institutional pathogen-free animal facility, treated in accordance with the European Union guidelines and with the approval of the San Raffaele Scientific Institute Institutional Ethical Committee. Mice were euthanized by CO_2 inhalation before surgical extraction of femurs and tibias. After bone crashing, erythrocytes were lysed with red blood cell lysis buffer (Merck, Darmstadt, Germany), and mononucleate cells enriched for progenitors by depletion of cells presenting myeloid, erythroid, and lymphoid differentiation markers using commercially available reagents (Stem Cell Technologies, Vancouver, BC, Canada). Purified cells were grown for 72 h in medium containing IL-3 (20 ng·mL^{-1}), IL-6 (20 ng·mL^{-1}), and SCF (100 ng·mL^{-1}; all from PrepoTech, London, UK) and then attached to non-tissue culture-

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

treated plates coated with Retronectin (Takara Bio Europe, Saint-Germain-en-Laye, France). Cells were incubated with the supernatant from Phoenix packaging cells transfected with the retroviral constructs expressing murine NUP98-NSD1 wildtype, NUP98-NSD1$\Delta_{PHDvC5HCH}$ (kind gift of Professor Kamps), and MSCV as control (Takara Bio Europe). Transduced cells were selected with geneticin (1000 $\mu g \cdot mL^{-1}$) for 72 h and allowed to recover for another 72 h in the absence of antibiotic before knockdown of *Nizp1* expression and collection for RNA analysis. Lentiviral production was performed as already published [41]. For knockdown of murine *Nizp1* expression, c-Kit$^+$/Sca-1$^+$/Lin$^-$ cells expressing *NUP98-NSD1* were attached to non-tissue culture-treated plates coated with Retronectin before incubation with shRNA expressing lentiviruses. Constitutive knockdown of *Nizp1* expression was achieved using the gene-specific pLKO.1-puro based shRNA constructs TRCN0000084369, TRCN0000084370 (Horizon, Cambridge, UK). The pLKO.1-puro non-mammalian shRNA plasmid (CAACAAGATGAAGAGCACCAA) from Merck was used as control.

The animal licence number under which we conducted the experiments was approved by the Italian Ministery of Health (Authorization no. 10/2017-PR, answer to prot.6EEAF.21 of 10/10/2016).

RNA extraction, reverse transcription and real-time PCR

Total RNA was isolated using RNeasy Plus Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. cDNA was synthesized with random primers using the Promega Reverse Transcription System (Promega, Madison, WI, USA). Quantitative real-time PCR was performed using SYBR Green Mastermix (Thermofisher, Waltham, MA, USA) in combination with specific primer pairs on the ViiA™7 Real Time PCR System (Thermofisher). mRNA levels were normalized against actin mRNA and referred to control Primer pairs were as follow:

NUP98-NSD1	
	Forward: ACTACGACAGCCACTTTGGG
	Reverse: TCTGTATATTCCAGAAGCCACT
Nizp1	
	Forward: TGTCCAGCAGGAGGTGAC
	Reverse: ATGGGCTGTGCTGGGAGT
actin	
	Forward: GATCATTGCTCCTCCTGAGC
	Reverse: ACATCTGCTGGAAGGTGGAC
HoxA9	
	Forward: TGGTTCTCCTCCAGTTGATAG
	Reverse: AGAAACTCCTTCTCCAGTTCC
HoxA5	
	Forward: GCAAGCTGCACATTAGTCAC
	Reverse: GCATGAGCTATTTCGATCCT

RNA seq

Samples were processed using the 'mRNA Stranded TruSeq' library protocol and sequenced through Illumina NextSeq 500 in order to obtain 10.0 M clusters of 1×75 bp fragments on average. Reads were further trimmed to remove adapters and low quality bases using TRIMMOMATIC [42], aligned to mouse genome GRCm38 using STAR aligner [43]. Gene counts were generated using featureCounts (part of the SUBREAD package [44]), based on GENCODE gene annotation version M22. In order to discard genes highly expressed by one sample only, a filter of cpm (counts per million) ≥ 1 in at least three samples was added. Read counts were normalized with the trimmed mean of M-values (TMM) method [45] using calcNormFactors function and then VOOM [46] was applied. Tables with DGE results and volcano plots were generated using the LIMMA package [47].

Gene set enrichment analysis

Gene set enrichment Analysis (GSEA) was performed using the GSEAPRERANKED JAVA Tool [48] with pre-ranked log2 fold changes calculated using the LIMMA package between the indicated conditions in expressed genes for enrichment in Hallmark Pathways.

Sample preparation for NMR and binding assays

The human sequences (100% identical with the murine sequences) of PHD_vC5HCH_{NSD2} domain (1228–1340) and PHD_vC5HCH_{NSD3} (1310–1413) were cloned into pETM11-SUMO3 vector (EMBL). PHD_vC5HCH_{NSD2} domain contains two free cysteine residues ($C1304_{NSD2}$ and C1324 $_{NSD2}$) and to improve solubility we mutated $C1324_{NSD2}$ into Serine ($C1324S_{NSD2}$). Site-directed mutagenesis (C1324S-PHD_vC5HCH_{NSD2}, R1244S-$PHDvC5HCH_{NSD2}$, Q1326S-$PHDvC5HCH_{NSD3)}$ was performed by standard overlap extension methods. The DNA constructs were sequenced by Eurofins (Milan, Italy). The $His_{(6)}$SUMO3 tagged domains were expressed in *E. coli* BL21 (DE3) after overnight induction with 1 mM of isopropyl thio-ß-D-galactoside (IPTG) at 28 °C, in LB or in minimal medium containing $^{15}NH_4Cl$ with or without ^{13}C-D-glucose supplemented with 0.2 mM $ZnCl_2$. The expressed proteins were purified using a nickel-chelating column (Qiagen), followed by SENP2 protease overnight cleavage, size-exclusion chromatography on a Hiload 16/60 Superdex 75 column (GE Healthcare, Milan, Italy) in NMR buffer containing 20 mM NaH_2PO_4/Na_2HPO_4 pH 6.3, 150 mM NaCl, 1 mM TCEP, 10 µM $ZnCl_2$. Murine $C2HR_{Nizp1}$ (residues E397–K434, corresponding to E398–K435 in the human sequence NM 032752.3) were expressed in *E. coli* and purified as described in [25]. In NMR experiments the NMR buffer

The FEBS Journal (2022) © 2022 The Authors. *The FEBS Journal* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

was supplied with 0.15 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) and D2O (10% v/v).

NMR experiments

NMR experiments to assign backbone resonances were performed at 295 K on a Bruker Avance 600 MHz equipped with inverse triple-resonance cryoprobe and pulsed field gradients (Bruker, Karlsruhe, Germany). Typical sample concentration was 0.3–0.4 mM. Data were processed using or TOPSPIN 3.2 (Bruker) and analysed with CCPNMR ANALYSIS 2.1 [49]. The ^{1}H, ^{13}C, ^{15}N chemical shifts of the backbone atoms of PHD_vC5HCH_{NSD2}, (BMRB ID 51637) and PHD_vC5HCH_{NSD3} (BMRB ID 18664 [24]) were obtained/confirmed through three-dimensional HNCA, CBCA(CO)NH, CBCANH, HNCO experiments.

For NMR titrations, we followed the protocol described in [19], whereby at each titration point a 2D water-flipback ^{1}H-^{15}N-edited HSQC spectrum was acquired with 2048 (160) complex points for ^{1}H (^{15}N), respectively, apodized by 90° shifted squared (sine) window functions and zero filled to 256 points for indirect dimension. Assignment of the labelled proteins in the presence of the ligands ($C2HR_{Nizp1}$ or $PHDvC5HCH_{NSD2–3}$) was obtained following individual cross-peaks through the titration series. For each residue the weighted average of the ^{1}H and ^{15}N chemical shift perturbation (CSP) was calculated as $CSP = [(\Delta^2HN + \Delta^2N/25)/2]^{1/2}$. The apparent dissociation constants of the ligands were estimated from least-squares fitting of CSPs as a function of total ligand concentration according to the following equation:

$$\delta_i = \frac{b - \sqrt{b^2 - 4ac}}{2a}$$

with $a = (K_a/\delta_b)\cdot[P_t]$, $b = 1 + K_a\cdot([L_{ti}] + [P_t])$ and $c = \delta_b\cdot K_a\cdot[L_{ti}]$, $[L_{ti}]$ is the total ligand concentration at each titration point, $[P_t]$ is the total protein concentration, $K_a = 1/K_d$ is the association constant, and δ_b is the chemical shift of the resonance in the complex, δ_i is the absolute change in chemical shift for each titration point [50].

The K_d were obtained from the average of the fitting of 8 and 7 residues for $PHDvC5HCH_{NSD2}$ and $PHDvC5HCH_{NSD3}$, respectively. K_d and δ_b were used as fitting parameters using the XMGRACE program (http://plasma-gate.weizmann.ac.il/Grace/). Molecular images were generated by PYMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC (New York, NY, USA).

$PHDvC5HCH_{NSD2}$ ALPHAFOLD model and data visualization

The ALPHAFOLD [51] structure prediction of the entire NSD2 protein (residue 1–1365, uniProt code O96028) has been download from https://alphafold.ebi.ac.uk/. In this paper we selected the region corresponding to residues K1235-A1330. PHD_vC5HCH_{NSD2} and molecular images were generated by PYMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

Isothermal titration calorimetry thermodynamic analysis

ITC titration was performed using a VP-ITC isothermal titration calorimeter (MicroCal LLC, Northampton, MA, USA), similarly as described in [25] Recombinant R1244S-$PHDvC5HCH_{NSD2}$ and $C2HR_{Nizp1}$ were dialyzed against the same buffer (20 mM NaH_2PO_4/Na_2HPO_4 pH 7.2, 150 mM NaCl, 1 mM TCEP) at 23 °C. Step by step injections of the 1.5 mM titrant ($C2HR_{Nizp1}$) solution into a cell containing a 100 μM R1244S-PHD_vC5HCH_{NSD2} were performed to finally reach a 2.5-fold molar excess of $CH2R_{Nizp1}$ with respect to the protein concentration. The quantity of heat absorbed or released in the process was measured. Control experiments were performed under identical conditions to determine the dilution heat of the buffer into protein samples. The final data were analysed using the software ORIGIN 7.0 (OriginLab Corp., Northampton, MA, USA).

Acknowledgements

This work was supported by Italian Ministry of Health (RF-2013-02354880 to GM, GT, ABa) by the Italian Association for Cancer Research (AIRC) (IG-21440) (to GM) and by Fondazione Umberto Veronesi (to ABe). Open access funding provided by BIBLIOSAN.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

All authors participated at conceptualization of the study and data analyses, ABe expressed and purified recombinant proteins, performed NMR analyses and ITC experiments, GQ performed NMR experiments, OAB performed purification of bone marrow progenitors, RNA extraction and RT-qPCR. JMGM performed RNAseq analysis. GM and GT wrote the original draft, all authors contributed to writing the manuscript. GM, GT, ABa provided resources and ideas. GM supervised the overall study and provided funding.

Peer review

The peer review history for this article is available at https://publons.com/publon/10.1111/febs.16664.

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Data availability statement

RNA-seq data generated during this study have been deposited in the Gene Expression Omnibus (GEO) with the accession code GEO: GSE209850 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE209850; token code for reviewer: mfkpckwclvqnxal).

References

1 Bennett RL, Swaroop A, Troche C, Licht JD. The role of nuclear receptor – binding SET domain family histone lysine methyltransferases in cancer. *Cold Spring Harb Perspect Med*. 2017;**7**:1–18.
2 Morishita M, di Luccio E. Cancers and the NSD family of histone lysine methyltransferases. *Biochim Biophys Acta*. 2011;**1816**:158–63.
3 Kurotaki N, Imaizumi K, Harada N, Masuno M, Kondoh T, Nagai T, et al. Haploinsufficiency of NSD1 causes Sotos syndrome. *Nat Genet*. 2002;**30**:365–6.
4 Douglas J, Hanks S, Temple IK, Davies S, Murray A, Upadhyaya M, et al. NSD1 mutations are the major cause of sotos syndrome and occur in some cases of weaver syndrome but are rare in other overgrowth phenotypes. *Am J Hum Genet*. 2003;**72**:132–43.
5 Stec I, Wright TJ, Van Ommen GJB, De Boer PAJ, Van Haeringen A, Moorman AFM, et al. WHSC1, a 90 kb SET domain-containing gene, expressed in early development and homologous to a drosophila dysmorphy gene maps in the wolf-Hirschhorn syndrome critical region and is fused to IgH in t (4;14) multiple myeloma. *Hum Mol Genet*. 1998;**7**:1071–82.
6 Kuo AJ, Cheung P, Chen K, Zee BM, Kioi M, Lauring J, et al. NSD2 links Dimethylation of histone H3 at lysine 36 to oncogenic programming. *Mol Cell*. 2011;**44**:609–20.
7 Lauring J, Abukhdeir AM, Konishi H, Garay JP, Gustin JP, Wang Q, et al. The multiple myeloma associated MMSET gene contributes to cellular adhesion, clonogenic growth, and tumorigenicity. *Blood*. 2008;**111**:856–64.
8 Popovic R, Martinez-Garcia E, Giannopoulou EG, Zhang Q, Zhang Q, Ezponda T, et al. Histone methyltransferase MMSET/NSD2 alters EZH2 binding and reprograms the myeloma epigenome through global and focal changes in H3K36 and H3K27 methylation. *PLoS Genet*. 2014;**10**:e1004566.
9 Angrand PO, Apiou F, Stewart AF, Dutrillaux B, Losson R, Chambon P. NSD3, a new SET domain-containing gene, maps to 8p12 and is amplified in human breast cancer cell lines. *Genomics*. 2001;**74**:79–88.
10 Tonon G, Wong KK, Maulik G, Brennan C, Feng B, Zhang Y, et al. High-resolution genomic profiles of human lung cancer. *Proc Natl Acad Sci USA*. 2005;**102**:9625–30.
11 Wang GG, Cai L, Pasillas MP, Kamps MP. NUP98–NSD1 links H3K36 methylation to Hox-a gene activation and leukaemogenesis. *Nat Cell Biol*. 2007;**9**:804–12.
12 Li W, Tian W, Yuan G, Deng P, Sengupta D, Cheng Z, et al. Molecular basis of nucleosomal H3K36 methylation by NSD methyltransferases. *Nature*. 2021;**590**:498–503.
13 Qin S, Min J. Structure and function of the nucleosome-binding PWWP domain. *Trends Biochem Sci*. 2014;**39**:536–47.
14 Sanchez R, Zhou M-M. The PHD finger: a versatile epigenome reader. *Trends Biochem Sci*. 2011;**36**:364–72.
15 Musselman CA, Kutateladze TG. Handpicking epigenetic marks with PHD fingers. *Nucleic Acids Res*. 2011;**39**:9061–71.
16 Rayasam GV, Wendling O, Angrand P, Mark M, Niederreither K, Song L, et al. NSD1 is essential for early post-implantation development and has a catalytically active SET domain. *EMBO J*. 2003;**22**:3153–63.
17 Nimura K, Ura K, Shiratori H, Ikawa M, Okabe M, Schwartz RJ, et al. LETTERS a histone H3 lysine 36 trimethyltransferase links Nkx2-5 to wolf – Hirschhorn syndrome. *Nature*. 2009;**460**:287–91.
18 Ferreira De Freitas R, Liu Y, Szewczyk MM, Mehta N, Li F, McLeod D, et al. Discovery of Small-molecule antagonists of the PWWP domain of NSD2. *J Med Chem*. 2021;**64**:1584–92.
19 Berardi A, Ghitti M, Quilici G, Musco G. In silico derived small molecules targeting the finger-finger interaction between the histone lysine methyltransferase NSD1 and Nizp1 repressor. *Comput Struct Biotechnol J*. 2020;**18**:4082–92.
20 Böttcher J, Dilworth D, Reiser U, Neumüller RA, Schleicher M, Petronczki M, et al. Fragment-based discovery of a chemical probe for the PWWP1 domain of NSD3. *Nat Chem Biol*. 2019;**15**:822–9.
21 Tatton-Brown K, Douglas J, Coleman K, Baujat G, Cole TRP, Das S, et al. Genotype–phenotype associations in Sotos syndrome: an analysis of 266 individuals with NSD1 aberrations. *Am J Hum Genet*. 2005;**77**:193–204.
22 Huang Z, Wu H, Chuai S, Xu F, Yan F, Englund N, et al. NSD2 is recruited through its PHD domain to oncogenic gene loci to drive multiple myeloma. *Cancer Res*. 2013;**73**:6277–88.
23 Rathert P. Structure, activity and function of the NSD3 protein lysine methyltransferase. *Life*. 2021;**11**:726.
24 He C, Li F, Zhang J, Wu J, Shi Y. The methyltransferase NSD3 has chromatin-binding motifs,

The FEBS Journal (2022) © 2022 The Authors. *The FEBS Journal* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

PHD5-C5HCH, that are distinct from other NSD (nuclear receptor SET domain) family members in their histone H3 recognition. *J Biol Chem*. 2013;**288**:4692–703.
25 Berardi A, Quilici G, Spiliotopoulos D, Corral-Rodriguez MA, Martin-Garcia F, Degano M, et al. Structural basis for PHD V C5HCH NSD1 –C2HR Nizp1 interaction: implications for Sotos syndrome. *Nucleic Acids Res*. 2016;**44**:3448–63.
26 Mysliwiec MR, Kim T-G, Lee Y. Characterization of zinc finger protein 496 that interacts with Jumonji/Jarid2. *FEBS Lett*. 2007;**581**:2633–40.
27 Nielsen AL, Jørgensen P, Lerouge T, Cerviño M, Chambon P, Losson R. Nizp1, a novel multitype zinc finger protein that interacts with the NSD1 histone lysine methyltransferase through a unique C2HR motif. *Mol Cell Biol*. 2004;**24**:5184–96.
28 Losson R, Nielsen AL. The NIZP1 KRAB and C2HR domains cross-talk for transcriptional regulation. *Biochim Biophys Acta*. 2010;**1799**:463–8.
29 Huang N, Vom Baur E, Garnier JM, Lerouge T, Vonesch JL, Lutz Y, et al. Two distinct nuclear receptor interaction domains in NSD1, a novel SET protein that exhibits characteristics of both corepressors and coactivators. *EMBO J*. 1998;**17**:3398–412.
30 Lucio-Eterovic AK, Singh MM, Gardner JE, Veerappan CS, Rice JC, Carpenter PB. Role for the nuclear receptor-binding SET domain protein 1 (NSD1) methyltransferase in coordinating lysine 36 methylation at histone 3with RNApolymerase II function. *Proc Natl Acad Sci USA*. 2010;**107**:16952–7.
31 Berdasco M, Ropero S, Setien F, Fraga MF, Lapunzina P, Losson R, et al. Epigenetic inactivation of the Sotos overgrowth syndrome gene histone methyltransferase NSD1 in human neuroblastoma and glioma. *Proc Natl Acad Sci USA*. 2009;**106**:21830–5.
32 Shiba N, Ichikawa H, Taki T, Park MJ, Jo A, Mitani S, et al. NUP98-NSD1 gene fusion and its related gene expression signature are strongly associated with a poor prognosis in pediatric acute myeloid leukemia. *Genes Chromosomes Cancer*. 2013;**52**:683–93.
33 Minucci S, Monestiroli S, Giavara S, Ronzoni S, Marchesi F, Insinga A, et al. PML-RAR induces promyelocytic leukemias with high efficiency following retroviral gene transfer into purified murine hematopoietic progenitors. *Blood*. 2002;**100**:2989–95.
34 Franklin KA, Shields CE, Haynes KA. Beyond the marks: reader-effectors as drivers of epigenetics and chromatin engineering. *Trends Biochem Sci*. 2022;**47**:417–32.
35 Gaetani M, Matafora V, Saare M, Spiliotopoulos D, Mollica L, Quilici G, et al. AIRE-PHD fingers are structural hubs to maintain the integrity of chromatin-associated interactome. *Nucleic Acids Res*. 2012;**40**:11756–68.
36 Jevtic Z, Matafora V, Casagrande F, Santoro F, Minucci S, Garre' M, et al. SMARCA5 interacts with NUP98-NSD1 oncofusion protein and sustains hematopoietic cells transformation. *J Exp Clin Cancer Res*. 2022;**41**:34.
37 Terlecki-Zaniewicz S, Humer T, Eder T, Schmoellerl J, Heyes E, Manhart G, et al. Biomolecular condensation of NUP98 fusion proteins drives leukemogenic gene expression. *Nat Struct Mol Biol*. 2021;**28**:190–201.
38 Gough SM, Slape CI, Aplan PD. NUP98 gene fusions and hematopoietic malignancies: common themes and new biologic insights. *Blood*. 2011;**118**:6247–57.
39 Xu C, Meng F, Park KS, Storey AJ, Gong W, Tsai YH, et al. A NSD3-targeted PROTAC suppresses NSD3 and cMyc oncogenic nodes in cancer cells. *Cell Chem Biol*. 2022;**29**:386–97.e9.
40 Shrestha A, Kim N, Lee S-J, Jeon YH, Song J-J, An H, et al. Targeting the nuclear receptor-binding SET domain family of histone lysine methyltransferases for cancer therapy: recent Progress and perspectives. *J Med Chem*. 2021;**64**:14913–29.
41 Botrugno OA, Bianchessi S, Zambroni D, Frenquelli M, Belloni D, Bongiovanni L, et al. ATR addiction in multiple myeloma: synthetic lethal approaches exploiting established therapies. *Haematologica*. 2020;**105**:2440–7.
42 Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;**30**:2114–20.
43 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;**29**:15–21.
44 Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res*. 2019;**47**:e47.
45 Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010;**11**:R25.
46 Law CW, Chen Y, Shi W, Smyth GK. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;**15**:R29.
47 Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;**43**:e47.
48 Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

genome-wide expression profiles. *Proc Natl Acad Sci USA*. 2005;**102**:15545–50.
49 Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins Struct Funct Genet*. 2005;**59**:687–96.
50 Williamson MP. Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Magn Reson Spectrosc*. 2013;**73**:1–16.
51 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;**596**:583–9.
52 Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. 2014;**42**:W320–4.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. RNA-seq analysis of c-Kit$^+$/Sca-1$^+$/Lin$^-$ cells expressing or not NUP98-NSD1 and cells expressing NUP98-NSD1 but silenced for Nizp1 (shNizp1 #1 and #2).

Table S2. Gene set enrichment analysis (GSEA) of Hallmark pathways in DGE of c-Kit$^+$/Sca-1$^+$/Lin$^-$ cells expressing or not NUP98-NSD1 and cells expressing NUP98-NSD1 but silenced for Nizp1 (shNizp1 #1 and #2).

The FEBS Journal (2022) © 2022 The Authors. *The FEBS Journal* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies.

17424658, 0, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1111/febs.16664 by IRCCS Ospedale San Raffaele, Wiley Online Library on [25/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License