

ARTICLE

DOI: 10.1038/s41467-017-02259-9

OPEN

# H3K14ac is linked to methylation of H3K9 by the triple Tudor domain of SETDB1

Renata Z. Jurkowska<sup>1</sup>, Su Qin<sup>2,3</sup>, Goran Kungulovski<sup>1</sup>, Wolfram Tempel<sup>2</sup>, Yanli Liu<sup>2</sup>, Pavel Bashtrykov<sup>1</sup>, Judith Stiefelmaier<sup>1</sup>, Tomasz P. Jurkowski<sup>1</sup>, Srikanth Kudithipudi<sup>1</sup>, Sara Weirich<sup>1</sup>, Raluca Tamas<sup>1</sup>, Hong Wu<sup>2</sup>, Ludmila Dombrovski<sup>2</sup>, Peter Loppnau<sup>2</sup>, Richard Reinhardt<sup>4</sup>, Jinrong Min<sup>2,5</sup> & Albert Jeltsch<sup>1</sup>

SETDB1 is an essential H3K9 methyltransferase involved in silencing of retroviruses and gene regulation. We show here that its triple Tudor domain (3TD) specifically binds to doubly modified histone H3 containing K14 acetylation and K9 methylation. Crystal structures of 3TD in complex with H3K14ac/K9me peptides reveal that peptide binding and K14ac recognition occurs at the interface between Tudor domains (TD) TD2 and TD3. Structural and biochemical data demonstrate a pocket switch mechanism in histone code reading, because K9me1 or K9me2 is preferentially recognized by the aromatic cage of TD3, while K9me3 selectively binds to TD2. Mutations in the K14ac/K9me binding sites change the sub-nuclear localization of 3TD. CHIP-seq analyses show that SETDB1 is enriched at H3K9me3 regions and K9me3/K14ac is enriched at SETDB1 binding sites overlapping with LINE elements, suggesting that recruitment of the SETDB1 complex to K14ac/K9me regions has a role in silencing of active genomic regions.

<sup>1</sup>Department of Biochemistry, Institute of Biochemistry and Technical Biochemistry, Stuttgart University, Allmandring 31, 70569 Stuttgart, Germany.

<sup>2</sup>Structural Genomics Consortium, University of Toronto, 101 College Street, Toronto, ON M5G 1L7, Canada. <sup>3</sup>Life Science Research Center, Southern University of Science and Technology, 518055 Shenzhen, China. <sup>4</sup>Max-Planck-Genomzentrum Köln, Carl-von-Linné-Weg 10, 50829 Köln, Germany.

<sup>5</sup>Department of Physiology, University of Toronto, Toronto, ON M5S 1A8, Canada. Renata Z. Jurkowska and Su Qin contributed equally to this work.

Correspondence and requests for materials should be addressed to J.M. (email: [jr.min@utoronto.ca](mailto:jr.min@utoronto.ca)) or to A.J. (email: [albert.jeltsch@ibc.uni-stuttgart.de](mailto:albert.jeltsch@ibc.uni-stuttgart.de))

**H**istone posttranslational modifications (PTMs) are essential for the regulation of chromatin states<sup>1,2</sup>. Trimethylation of histone H3 at lysine 9 (H3K9me3) is a hallmark of facultative and constitutive heterochromatin in almost all eukaryotes<sup>3–6</sup>, and it is also enriched in silenced genes<sup>7</sup>. In contrast, monomethylated H3K9 is found in the promoter regions of active genes<sup>7</sup> and histone acetylation is a signal of active transcription and open chromatin<sup>8</sup>. SETDB1 (also known as ESET or KMT1E) is a histone H3K9 methyltransferase that generates H3K9me3 in euchromatic regions<sup>9</sup>. Gene knockout studies have shown that it is essential for early development in mice<sup>10</sup>. SETDB1 plays a critical role in gene regulation<sup>11</sup>, and it forms a complex with KAP-1 (also called TRIM28) that is required in embryonic stem cells<sup>12,13</sup>, primordial germ cells<sup>14</sup>, and B lymphocytes<sup>15</sup>. SETDB1 is also recruited to target sites by the HUSH complex<sup>16</sup> and it associates with additional silencing factors including DNMT3A<sup>17</sup>, other H3K9 PKMTs<sup>18</sup>, and the NuRD HDAC chromatin remodeling complex<sup>19–21</sup>. The KAP-1/SETDB1 complex is required for silencing of LTR retroviruses<sup>12–15</sup>, and it has also been connected to H3K9 methylation of LINE elements<sup>14,15,22–24</sup>.

Histone-modifying enzymes often contain small functional domains for readout of specific histone PTMs that are involved in their recruitment and regulation of their enzymatic activity<sup>25</sup>. SETDB1 contains three Tudor domains (3TD) and a methyl-CpG binding domain (MBD) in its N-terminal part, followed by a split SET domain, which harbors the catalytic center for H3K9 methylation (Fig. 1a). Tudor domains are known to recognize methylated lysine and arginine<sup>26,27</sup>, but the potential binding ligand of the SETDB1 Tudor domains and their role in SETDB1 function remains unclear. We show here that SETDB1 3TD binds to histone H3 tails containing K14 acetylation combined with K9 methylation. Crystal structures revealed that K14ac recognition occurs at the interface between Tudor domain 2 (TD2) and Tudor domain 3 (TD3). Depending on the methylation state of K9, the methyllysine is preferentially bound by the aromatic cage of TD3 (K9me1 and K9me2) or TD2 (K9me3). Ectopic expression of 3TD wildtype and mutants fused to fluorophores demonstrates that the intact K14ac/K9me binding sites have an important role in the sub-nuclear localization of 3TD. We show that K9me3/K14ac regions bound by SETDB1 are enriched in LINE elements and propose that 3TD-mediated recruitment of SETDB1 to chromatin-containing active marks (K14ac) contributes to the efficient silencing of these regions. In summary, we characterize a reader for the combined K14ac/K9me mark, demonstrate the role of Tudor domains in binding acetyllysine, and identify a pocket switch mechanism in histone reading domains. Future work will show if this modification-dependent conformational change of the substrate triggers any downstream effects in SETDB1 or any of its complex partners.

## Results

**3TD binds H3 peptides containing K9me1/2/3 and K14ac.** To study the function of the triple Tudor domain (3TD) of SETDB1, we first investigated the histone peptide binding ability of the GST-tagged 3TD. Using modified histone peptide arrays, we identified strong and selective binding of 3TD to three combinatorially modified H3 peptides: H3K9me1/K14ac, H3K9me2/K14ac, and H3K9me3/K14ac (Fig. 1b). Peptide array binding experiments at 100-fold increased concentration of 3TD detected weaker binding to monomethylated H3K14ac, but no binding to H3K9me (Supplementary Fig. 1a). To further evaluate the relative contribution of each modification to the overall binding, we performed Isothermal Titration Calorimetry (ITC) and NMR chemical shift perturbation binding experiments with different

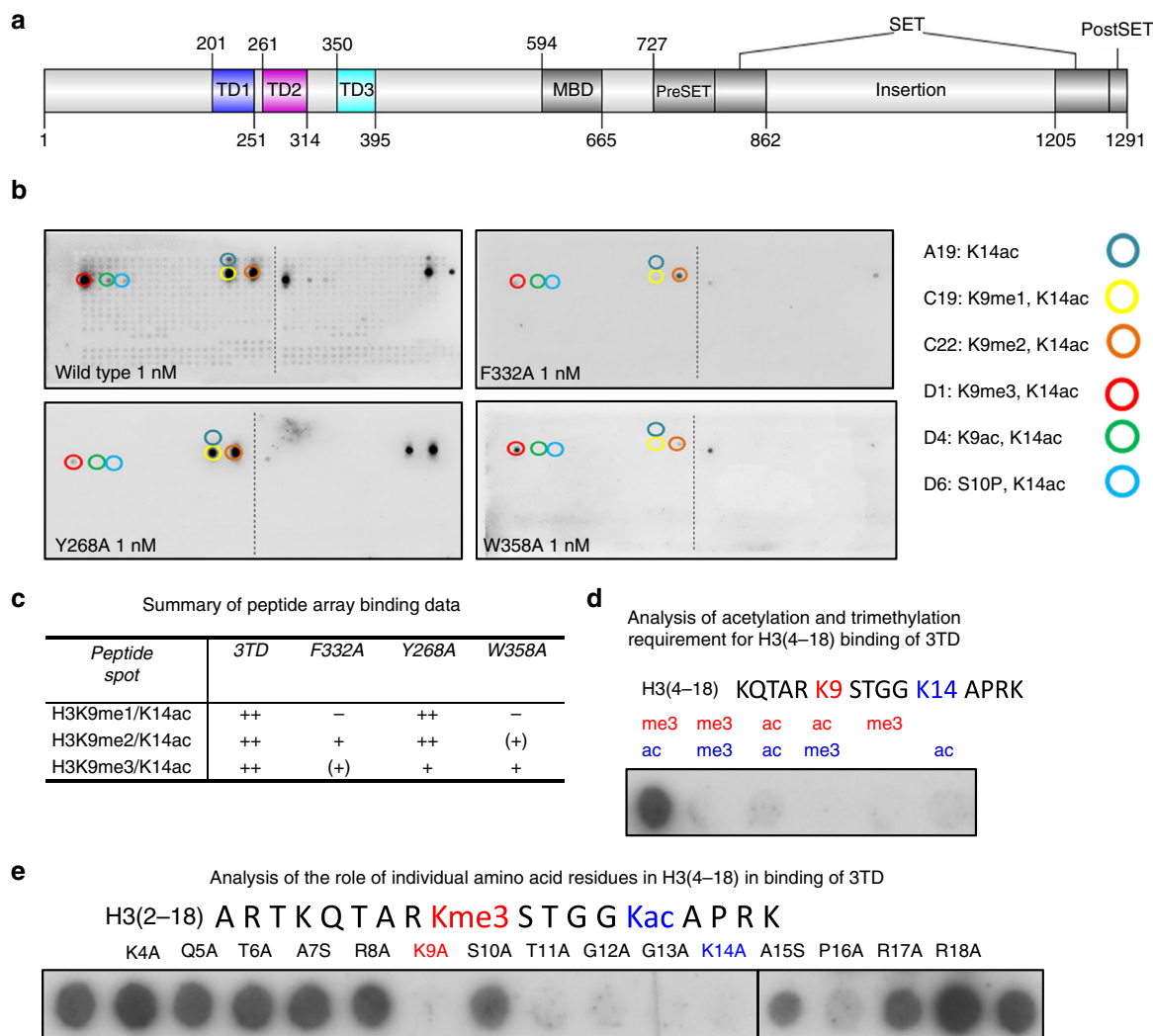
peptides, which confirmed the binding and revealed dissociation constants ( $K_d$ ) in the lower micromolar range for H3K9me/K14ac peptides (Fig. 2a, b). In agreement with the peptide array binding results, these experiments revealed that although the acetylation at H3K14 is indispensable for 3TD binding, binding to monomethylated H3K14ac was very weak. Methylation of H3K9 alone was insufficient for binding, but it strongly enhanced binding when combined with K14ac (Fig. 2b, c, Supplementary Fig. 1b, c). Peptide SPOT array binding experiments confirmed the requirement of K9me/K14ac for 3TD binding (Fig. 1d). In addition, binding of 3TD to an alanine scanning peptide SPOT array showed that the H3 sequence between T11 and P16 is critical for the interaction (Fig. 1e). The binding of the Tudor domains of SETDB1 to methylated histone tails has been investigated previously, but no binding was observed<sup>28</sup>. This could be explained by the fact that the specific K9me/K14ac combination of modifications was not tested in that study. Taken together, we conclude that the 3TD domain of SETDB1 specifically recognizes the acetylation mark of H3K14 and methylation of H3K9 on the same histone tail strongly enhances the binding affinity.

## Mode of histone H3 binding by the 3TD domain of SETDB1.

To understand the molecular basis for the recognition of the combinatorial H3K9me/K14ac histone mark, we solved the crystal structure of 3TD in complex with an H3(4–19) K9me2/K14ac doubly modified peptide (Fig. 3, Supplementary Figs. 2 and 3, Supplementary Table 1). The three Tudor domains compose a stack, in which the individual domains interact with each other face-to-back. Each domain forms a  $\beta$ -barrel-like fold characteristic of Tudor domains<sup>26</sup>, with the barrel bottoms pointing in the same direction. Notably, there are extensive interactions among the Tudor domains and the inter-domain linkers, indicating that 3TD folds and functions as a single unit (Supplementary Fig. 2a). The histone H3 peptide lies in a groove formed by TD2, TD3, and the linker between them (Fig. 3a, b). The specificity of the H3 peptide interaction is determined mainly through (1) recognition of K14ac by an acetyl-lysine binding groove between TD2 and TD3, (2) recognition of K9me2 by an aromatic cage in TD3, and (3) binding of G12 and G13 to a narrow groove between TD2 and TD3. In agreement with this observation, peptide SPOT binding experiments showed that alanine residues at the positions of G12 and G13 of H3 disrupted the interaction (Fig. 1e), indicating that any bigger residue at this position would preclude peptide binding. Moreover, E386 and R394 form a salt bridge inside TD3 and the backbone of G12–G13 is hydrogen-bonded to them (Fig. 3c). Binding to modified histone peptide arrays showed that mutation of E386 or R394 abolished the peptide-binding ability of the Tudor domains (Supplementary Fig. 4a).

The K14ac is recognized by residues from TD2, TD3, and the linker between them (Fig. 3d). The acetyl group forms two hydrogen bonds with the backbone carbonyl-oxygen of G300 and the backbone amine of Y389. The side chains of F332, F296, A302, and Y389 form hydrophobic interactions with the acetyl-lysine. Replacement of F332 by alanine caused a strong reduction in affinity of 3TD to the H3K9me2/K14ac peptide in peptide array and ITC peptide-binding experiments (Figs. 1 and 2, Supplementary Fig. 4a), and mutation of I388, which is a part of the hydrophobic core of TD3, disrupted the binding completely (Supplementary Fig. 4a), confirming the importance of these hydrophobic interactions. In addition, the backbone of the histone peptide forms a series of hydrogen bonds with the residues from TD3 further stabilizing the complex (Fig. 3c).

**TD3 preferentially recognizes dimethylated K9.** The K9me2 side chain is inserted into an aromatic cage of TD3



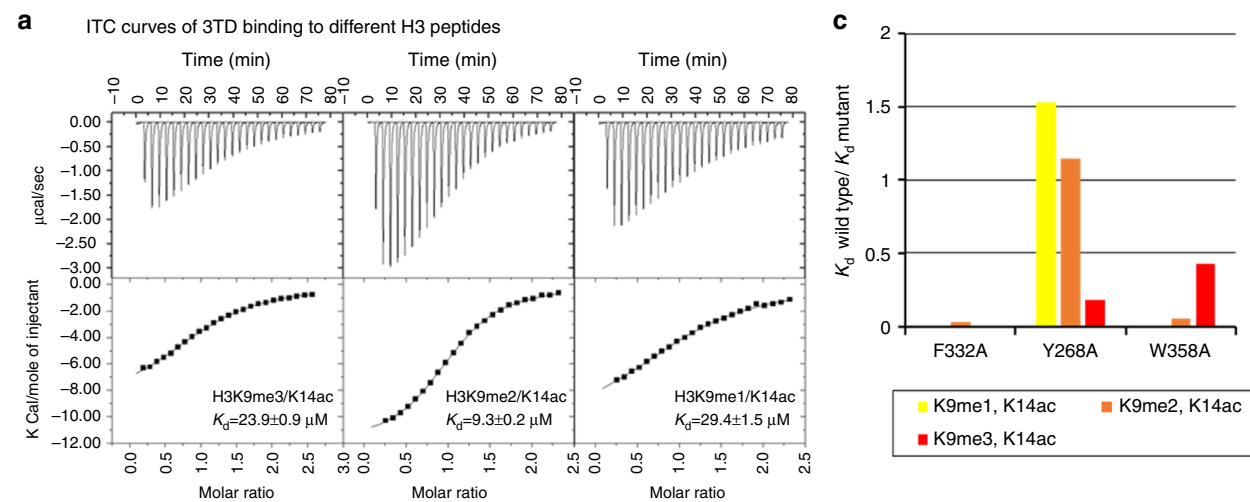
**Fig. 1** Peptide binding of SETDB1 Tudor domains. **a** Domain structure of SETDB1. **b** Modified histone peptide array binding of the 3TD and 3TD mutants. Peptide spots are annotated with the color-coded circles denoting the presence of the designated modifications. The full annotation of all peptide spots is provided in Supplementary Table 2. **c** Summary of the binding of 3TD and 3TD mutants to K9me1/2/3–K14ac peptides on peptide arrays shown in (**b**). **d** Peptide SPOT array binding of 3TD showing the double specificity of binding to K9me3 and K14ac. H3 (4–18) peptides were synthesized containing lysine trimethylation (me3) or acetylation (ac) at K9 and K14 and binding of 3TD analyzed. **e** Peptide SPOT array binding to map the recognition sequence of 3TD. The first and last spots contain the H3 (2–18) K9me3/K14ac peptide. The other spots contain peptides with the indicated amino acid exchanges. The data reveal that 3TD interacts with residues from K9 to P16

formed by W358, W363, and F379 (Fig. 3e). A direct hydrogen bond between the carboxylate of D382 and the  $\epsilon$ -ammonium of K9me2 appears to further stabilize the complex explaining the preference of this aromatic cage for binding of K9me2 over K9me3. Disruption of the aromatic cage by mutation of W358 or W363 abrogated H3K9me1 binding and reduced H3K9me2 peptide binding in peptide array binding (Fig. 1 and Supplementary Fig. 4d) and ITC peptide-binding experiments (Fig. 2, Supplementary Fig. 6c), while H3K9me3 binding was only 2-fold reduced (see below). Similarly, mutation of D382 to alanine or arginine mainly reduced binding of K9me1 to the pocket, but did not affect the binding of K9me2 or K9me3 (Supplementary Fig. 4d). This can be explained by the fact that K9me1 has the weakest interactions with the aromatic cage, so that the D382-mediated hydrogen bond is more critical for K9me1 than for K9me2 binding (for K9me3 binding, also see below). Comparison with the apo-structure of 3TD revealed a conformational change associated with the methyl-lysine binding. In the absence of the bound peptide, R384 is packed into the aromatic cage, blocking

the peptide-binding site (Fig. 3f). Consequently, alanine mutation of R384 enhanced the peptide-binding affinity of 3TD by 3- to 6-fold (Fig. 2b, Supplementary Fig. 6b).

We also solved the structures of 3TD in complex with H3K9me3/K14ac or H3K9me0/K14ac (Supplementary Fig. 3, and Supplementary Table 1). In the H3K9me3/K14ac bound structure, the H3 peptide interacts with 3TD in a similar mode as in the H3K9me2/K14ac complex. The K9me3 is inserted into the aromatic cage of TD3, and only slight rearrangements were observed in the K9–aromatic cage interaction (Fig. 3g). In the H3K9me0/K14ac complex structure, the unmethylated K9 and residues preceding it cannot be tracked clearly suggesting that they are disordered (Supplementary Fig. 3).

**TD1 contains an inactive and incomplete aromatic cage.** Having shown that TD3 preferentially binds K9me2, we next asked whether TD1 and TD2 also have potential roles in histone binding. Sequence and structure alignments revealed that TD1



**b**  $K_d$  values and stoichiometries (n) of 3TD and 3TD mutants binding to H3 (4–19) peptides

| Peptide substrate   | WT                      |                 | F332A<br>(in TD2–TD3 linker) |   | Y268A<br>(in TD2)       |                 | W358A<br>(in TD3)       |                 | R384A<br>(in TD3)       |                 |
|---------------------|-------------------------|-----------------|------------------------------|---|-------------------------|-----------------|-------------------------|-----------------|-------------------------|-----------------|
|                     | $K_d$ [ $\mu\text{M}$ ] | n               | $K_d$ [ $\mu\text{M}$ ]      | n | $K_d$ [ $\mu\text{M}$ ] | n               | $K_d$ [ $\mu\text{M}$ ] | n               | $K_d$ [ $\mu\text{M}$ ] | n               |
| H3K9me1/K14ac       | $29.4 \pm 1.5$          | $1.11 \pm 0.02$ | NB                           | – | $19.2 \pm 0.5$          | $1.05 \pm 0.01$ | NB                      | –               | $5.1 \pm 0.6$           | $1.13 \pm 0.03$ |
| H3K9me2/K14ac       | $9.3 \pm 0.2$           | $1.05 \pm 0.01$ | $350 \pm 70^{**}$            | – | $8.1 \pm 0.2$           | $1.16 \pm 0.01$ | $174 \pm 12$            | $0.99 \pm 0.23$ | $2.6 \pm 0.2$           | $1.05 \pm 0.01$ |
| H3K9me3/K14ac       | $23.9 \pm 0.9$          | $1.07 \pm 0.01$ | NB                           | – | $135 \pm 11$            | $1.34 \pm 0.11$ | $55.6 \pm 5.5$          | $1.04 \pm 0.01$ | $9.2 \pm 1.3$           | $1.18 \pm 0.06$ |
| H3K14ac             | $930 \pm 89^*$          | –               | –                            | – | –                       | –               | –                       | –               | –                       | –               |
| H3K9me3             | NB                      | –               | –                            | – | –                       | –               | –                       | –               | –                       | –               |
| H3(4–19) unmodified | NB                      | –               | –                            | – | –                       | –               | –                       | –               | –                       | –               |

\*Estimation determined by NMR. All other data were determined by ITC. No n-value could be determined. \*\*Estimation. No n-value could be determined. NB, no detectable binding. –, not measured.

**Fig. 2** 3TD binding to different H3 peptides analyzed by ITC. **a** Representative ITC curves and dissociation constants of 3TD binding to different H3 peptides. **b** Compilation of dissociation constants of 3TD and 3TD mutants to different peptides (Supplementary Figs. 1b,c, and 6). **c** Relative binding affinity of F332A, Y268A, and W358A to K9me1/2/3-K14ac peptides. While F332A has generally reduced binding, Y268A shows a specific loss in binding of K9me3-K14ac and W358A shows a preferential reduction of K9me1/2-K14ac binding

has an incomplete aromatic cage consisting of only H214 and F234. The position of the first aromatic residue is occupied by K208, suggesting that TD1 is unable to bind methylated lysine residues (Supplementary Fig. 2b and c). This was confirmed by a mutational study, as substitution of residues in TD1 did not or only weakly affected peptide binding (Supplementary Fig. 4b).

**TD2 preferentially recognizes trimethylated K9.** TD2 has all conserved aromatic residues (Y268, W275, Y277, F297, and Y301), which are positioned proximal to the H3K9me/K14ac peptides mentioned above. To test whether TD2 contributes to peptide binding, we generated a Y268A mutant, which disrupts the aromatic cage in TD2, and used it for both peptide array binding and ITC peptide-binding assays. Interestingly, the Y268A mutant showed a clear reduction in affinity to the H3K9me3/K14ac peptide, while the affinities to H3K9me2/K14ac and H3K9me1/K14ac peptides were not affected (Fig. 1b, c, and Supplementary Fig. 4c). These data suggest that TD2 is involved in H3K9me3/K14ac peptide binding.

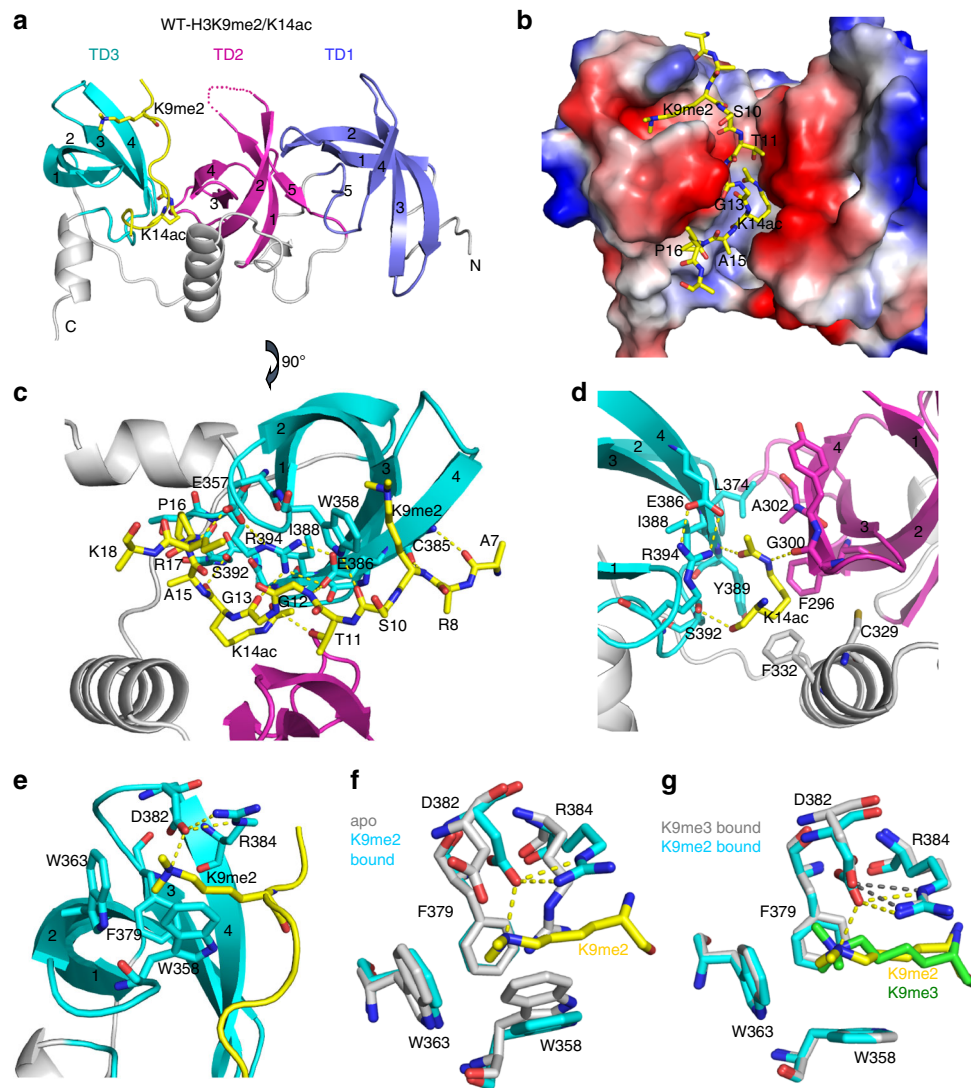
In contrast, the mutation of W358 in TD3 (W358A), an aromatic cage residue of TD3, led to a loss of H3K9me1/K14ac binding and a strong reduction in binding of H3K9me2/K14ac, while binding of H3K9me3/K14ac was only moderately reduced in the peptide array binding (Fig. 1b, c). In the ITC peptide-binding experiments, similar results were obtained; the affinity to H3K9me2/K14ac was reduced 18-fold and binding to H3K9me1/K14ac was completely lost, but H3K9me3/K14ac binding was reduced only mildly (2-fold) (Fig. 2 and Supplementary Fig. 6c). These data suggest that TD3 is mainly involved in H3K9me1/2/

K14ac binding. The weaker binding of H3K9me1/2/K14ac to W358A could be explained by the loss of the H-bond between the W358 side chain and the carbonyl-oxygen of S10, leading to a general reduction in peptide binding. We conclude that both TD2 and TD3 contribute to K9me binding, with TD2 showing preference for H3K9me3 and TD3 preferring H3K9me1/2.

We further solved the structures of the W358A mutant in complex with H3K9me3/K14ac or H3K9me2/K14ac peptides (Fig. 4 and Supplementary Table 1). In the W358A H3K9me3/K14ac structure, K9me3 indeed was bound to TD2, while the peptide residues following S10 (including K14ac) adopted the same conformation as in the wildtype structures (Fig. 4a–c). The K9me3 was bound to an open aromatic cage in TD2 formed by Y268, W275, and Y277, where W275 was rearranged compared to the apo-structure to accept the trimethylated K9 (Fig. 4d, e). In the W358A–H3K9me2/K14ac structure, electron density for the K9me2 residue did not permit confident placement of the dimethylated lysine, suggesting that it is partially disordered (Supplementary Fig. 3). In summary, the Tudor domains of SETDB1 collaborate to recognize the combinatorial histone marks H3K9me/K14ac, with distinct Tudor domains showing different methylation state selectivity.

**SETDB1 is active on K14ac substrates on peptide arrays.** To investigate the peptide methylation activity of SETDB1, peptide arrays were methylated by the N-terminal-deleted SETDB1 protein (190–1291) (Supplementary Fig. 7). The results showed that SETDB1 methylates peptides containing K9 in un-, mono-, or dimethylated forms, indicating that SETDB1 could methylate



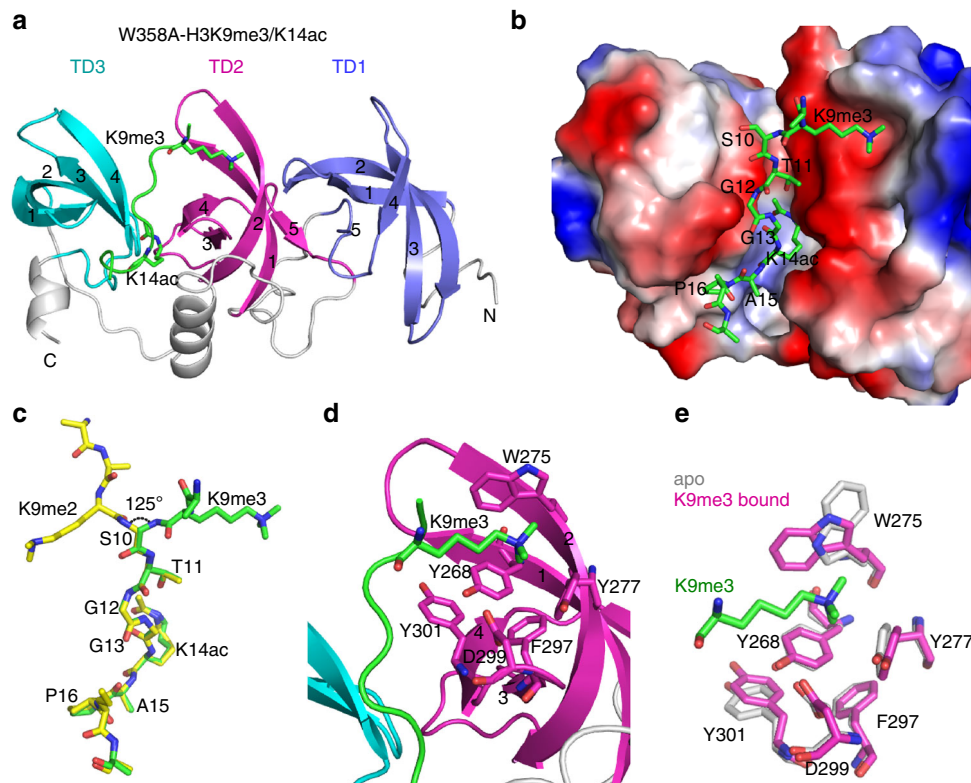


**Fig. 3** Details of the 3TD structure bound with a doubly modified peptide. **a** Structure of wildtype 3TD bound to H3K9me2/K14ac. 3TD is shown in ribbon representation with the peptide shown in yellow. **b** Surface representation of the protein in the same orientation as in **(a)** with the peptide shown in stick mode. **c** Details of the K9me2/K14ac peptide interaction in the cleft between TD2 and TD3. **d** Details of the K14ac recognition at the interface of TD2 and TD3. **e** Details of the K9me2-binding aromatic pocket in TD3. **f** Comparison of the aromatic pocket in TD3 in the apo form and **g** after binding of different H3K9me/K14ac peptides

histone H3K9 up to trimethylation. SETDB1 activity was observed on K14 acetylated and unacetylated substrates, and our data revealed a strong inhibition of the SETDB1 activity by phosphorylation of H3 peptides at S10 (as observed before<sup>9</sup>) or T11 (not yet reported). Previously, methylation experiments with peptides in solution showed that the K14ac modification inhibits the enzymatic activity of SETDB1<sup>9</sup>. The discrepancy of these results is possibly due to the different substrates used, because in solution, the SET and 3TD domains would compete for binding of substrate peptides, and acetylation of K14, which increases 3TD binding, would reduce catalytic activity. In contrast, in peptide array methylation experiments, the substrate peptides are present on the array at high density, and the binding of the 3TD to one peptide in a spot would allow the SET domain to methylate adjacent peptides in the same spot. Moreover, it cannot be ruled out that the different enzyme sources in both studies affected the results, because previously FLAG-tagged full-length SETDB1 purified from HEK293 cells was used, whereas we used an N-terminal-truncated SETDB1 purified from Sf9 cells. In the peptide array methylation experiments, peptides containing

K14ac were methylated more strongly by SETDB1 (Supplementary Fig. 7), which would be in agreement with the recruitment of this enzyme to these peptide spots via the 3TD domain. However, given the limitations of the peptide array-based activity assay, this result needs further validation.

**H3 and nucleosomes binding of 3TD is K9me/K14ac dependent.** Having shown the specific binding of 3TD to H3 peptides modified with K14ac and K9me, we next aimed to investigate the role of H3 modifications in 3TD binding to native histones and nucleosomes. Native histones purified from human HEK293 cells and recombinant unmodified H3 were separated on SDS gels, and the binding of 3TD to the blotted histone proteins was investigated by far-western assays. As shown in Fig. 5a, 3TD specifically binds to the native histone H3 (containing endogenous modifications), but not to the unmodified recombinant H3. To investigate the role of lysine acetylation in 3TD binding to H3, we next used histones isolated from HEK293 cells treated with HDAC inhibitors (trichostatin A or sodium butyrate). Histones



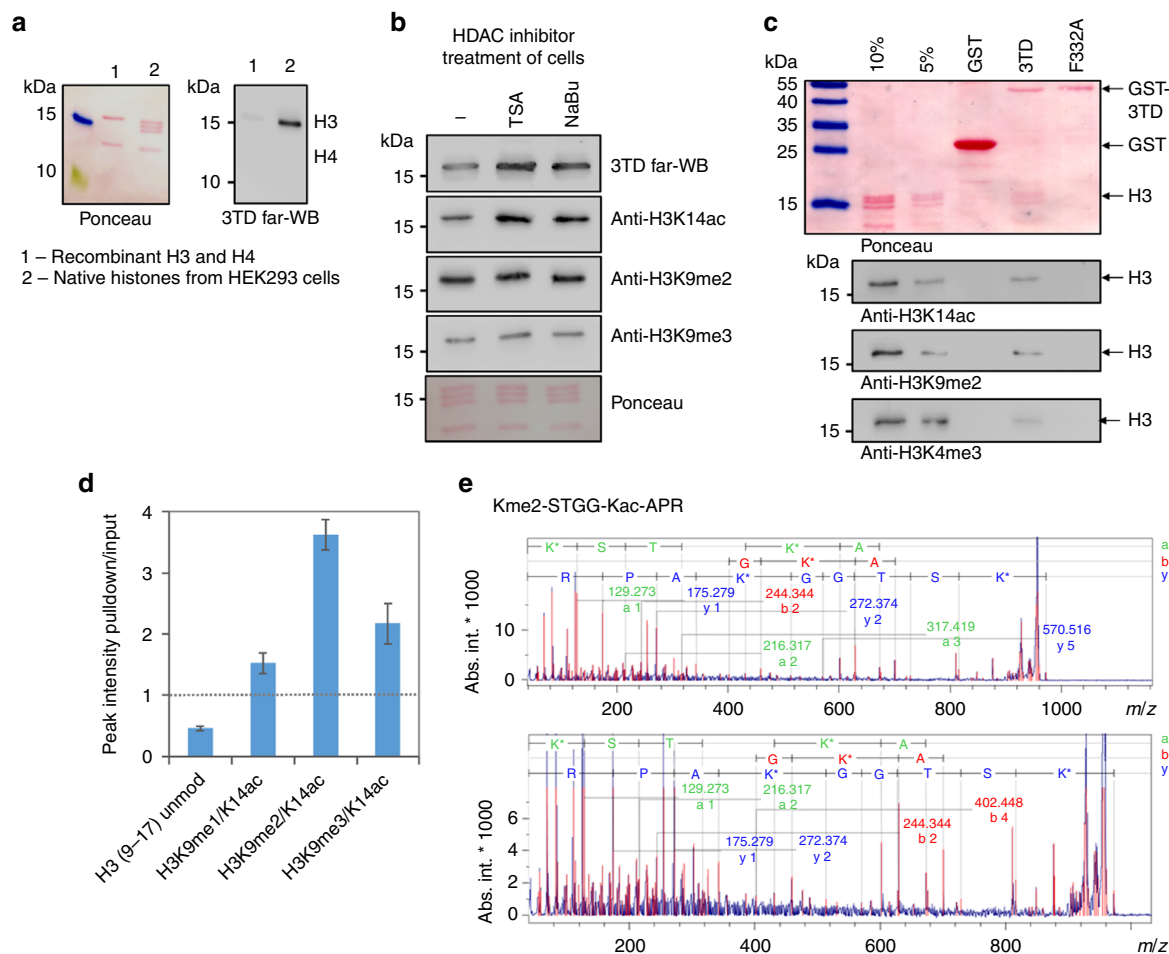
**Fig. 4** Details of the 3TD W358A structure bound with a doubly modified peptide. **a** Structure of 3TD W358A mutant bound to H3K9me3/K14ac shown in ribbon representation with the peptide shown in green. **b** Surface representation of the mutant in the same orientation as in (a). **c** Comparison of the conformations of the H3K9me2/K14ac (in yellow) and H3K9me3/K14ac (in green) peptides bound to wildtype 3TD and 3TD W358A mutant, respectively. The side chain of H3S10 is not shown for clarity. **d** Details of the K9me3-binding aromatic pocket in TD2. **e** Comparison of the aromatic pocket in TD2 in the apo form and after binding of the H3K9me3/K14ac peptides

were separated on SDS gels and their modification state was analyzed by western blot with anti-K9me and anti-K14ac antibodies. As expected, H3 proteins isolated from cells treated with HDAC inhibitors showed an increased H3K14ac signal, while the content of K9 methylation did not change (Fig. 5b). Far-western blots were conducted to study 3TD binding to both histone preparations. Consistent with the gain in the H3K14ac level, an increased 3TD binding was observed, confirming that the 3TD binding to native histones is conferred by histone acetylation. Next, we used the GST-tagged 3TD to pull-down mononucleosomes in chromatin domain immunoprecipitation (CIDOP) experiments<sup>29,30</sup>. Wildtype 3TD, but not the F332A mutant (with disrupted acetyl-lysine binding pocket), was able to pull-down H3 containing K14ac and K9me2 (Fig. 5c). H3K4me3 was analyzed as unrelated active mark and found not to be enriched in the pull-down. The pull-down was analyzed by mass spectrometry and showed strong enrichment of doubly modified H3K9me/K14ac peptides (1.5 to 2-fold for K9me1/K14ac and K9me3/K14ac, up to 4-fold for K9me2/K14ac) (Fig. 5d and Supplementary Fig. 8). The identity of the K9me2–K14ac peptide was confirmed by MS/MS (Fig. 5e). Unmodified H3 (9–17) and unrelated peptides detected by mass spectrometry analysis were depleted or not enriched confirming the specificity of the pull-down reaction (Supplementary Fig. 8).

**SETDB1 co-localizes with H3K14ac/K9me3 chromatin.** Next, we analyzed available ChIP-seq datasets from mESC cells for SETDB1<sup>31</sup>, H3K9me3<sup>32</sup>, and H3K14ac<sup>33</sup> to investigate the in vivo genome binding of SETDB1. As shown in Fig. 6a, we observed co-occurrence of these signals at many sites. Peak calling and

analysis of the overlap of peaks and peak regions confirmed a strong enrichment of SETDB1 at H3K9me3 peaks (Fig. 6b), which is in agreement with the literature findings<sup>31,32</sup>. This correlation was expected, because H3K9me3 is the methylation product of SETDB1. However, we also found that 14% of all SETDB1 peaks overlap with H3K14ac peaks and 28% of all H3K14ac peaks overlap with H3K9me3 peaks (Fig. 6b). Strikingly, we observed that 9.7% of all SETDB1 peaks (2231) overlap with H3K9me3 and H3K14ac peaks, which is a strong enrichment, because only 3.9% were expected by chance (=28% of 14%). Further evaluation of the significance of all binary and ternary overlaps was performed by Fisher's exact test, considering the size and number of regions in the data sets. This resulted in highly significant two-sided *p*-values in all cases (always *p* < 1E-300). In addition, the high significance of the overlap of H3K14ac peaks with the intersection of SETDB1 and H3K9me3 peaks was confirmed by randomization of the H3K14ac regions (Supplementary Note 2).

Clustering of the data on SETDB1 peaks showed that H3K9me3 signals were enriched at or near the vast majority of the SETDB1 binding sites (Fig. 6c, clusters 1–7), and two clusters showed co-occurrence of SETDB1 and H3K9me3 with H3K14ac: cluster 6 (strongly) and cluster 7 (weakly). Cluster 6 contains 12.2% of all SETDB1 peaks, which is in line with the analysis presented above. We conclude that K14ac is present at a subset of SETDB1 binding sites. Prompted by the documented function of SETDB1 to introduce H3K9 methylation at repeats, we analyzed the overlap of the SETDB1 peaks in the different clusters with repeats (Fig. 6d). We observed an enrichment of LTRs in clusters 2–4 (strongly), 7 and 8 (moderately). Moreover, cluster 6,



**Fig. 5** Binding of 3TD to modified H3 and nucleosomes. **a** Far-western blot analysis of 3TD binding to recombinant histones H3 and H4 (lane 1) and native histones (lane 2). Positions of the H3 and H4 proteins are indicated. The Ponceau-stained membrane is shown as a loading control. **b** Far-western blot analysis of 3TD binding to native histones isolated from HEK293 cells after treatment with histone deacetylase inhibitors trichostatin A (TSA) or sodium butyrate (NaBu). Untreated (–) cells are used as a control. The abundance of histone PTMs in the histone preparations was analyzed by western blot. The Ponceau-stained membrane is shown as a loading control. **c** Western blot analysis of the GST-3TD wildtype and F332A mutant mononucleosomal pull-down. Mononucleosomes were isolated from HEK293 cells. GST was used as a negative control. **d** MALDI MS analysis of the 3TD mononucleosomal pull-down. Quantification of the MS data (Supplementary Fig. 8) shows an enrichment of the H3K9me/K14ac peptides in 3TD pull-down and depletion of the unmodified H3(9–17) peptide when compared to input. Internal normalization was done using peaks of unrelated, unmodified H3 peptides. The error bars represent the SEM of two biological repeats. MS/MS for H3(9–17) and H3(41–49) are shown in Supplementary Fig. 8c,d. **e** MS/MS analysis of the H3 (9–17) Kme2-STGG-Kac-APR peptide fragment (theoretical mass  $MH^+$ : 971.563 Da; experimental mass: 971.482 Da). Fragment ions of the a, b, and y series are indicated in green, red, and blue, respectively. The two panels show the same analysis with different x- and y-axis scales

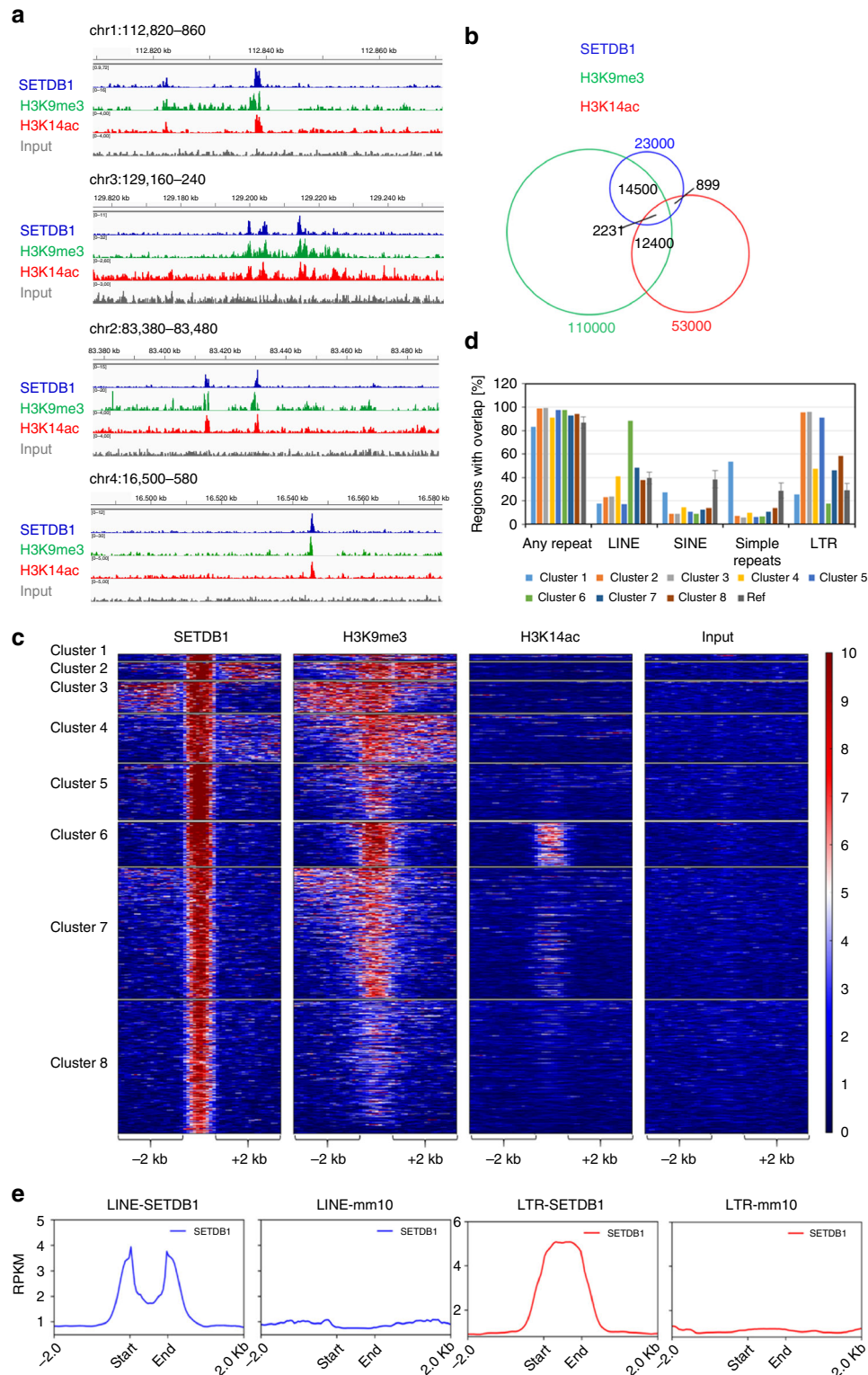
containing regions with SETDB1, H3K9me3, and H3K14ac signals, showed a very strong enrichment of LINE elements. We conclude that a subset of SETDB1 binding sites contains the H3K9me3/H3K14ac double modification and these sites are enriched in LINE elements. Next, we calculated meta-profiles of SETDB1 read density using LTR and LINE elements overlapping with SETDB1 peaks (Fig. 6e). These profiles clearly show that the binding of SETDB1 to LTRs and LINES differs. While SETDB1 binds to LTRs in the center of the elements, LINES are bound at the edges. Further analyses showed that H3K9me3 density mirrors SETDB1 density on repeat elements overlapping with SETDB1 peaks (Supplementary Fig. 9a). On LINE elements, a weak H3K14ac signal was detected as well, which (similar to the SETDB1 signal) was centered at the edges of the LINE elements. In agreement with our analysis, the H3K14ac signal was strongly increased on LINE elements from cluster 6 (Supplementary Fig. 9b). The clear H3K14ac enrichment at the flanks of SETDB1 overlapping LINE elements (in particular from cluster 6) strongly

supports our conclusion that H3K14ac is involved in the targeting of SETDB1.

#### Sub-nuclear localization of 3TD-binding pocket mutants.

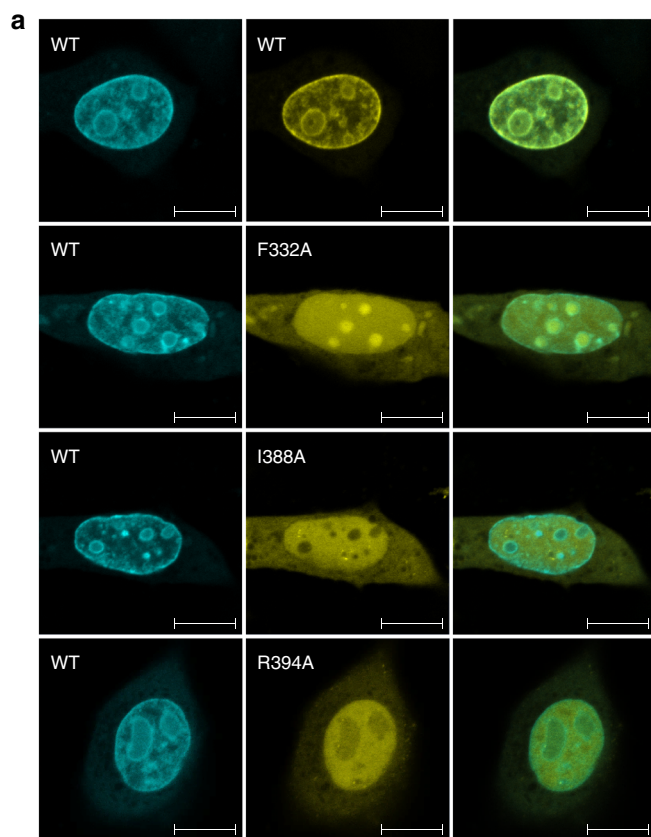
Finally, we investigated if the sub-nuclear localization of 3TD in mouse NIH3T3 cells depends on its H3 tail interaction (Fig. 7 and Supplementary Figs. 10 and 11). 3TD and 3TD mutants were expressed as fusion constructs with mCerulean or mVenus in mouse NIH3T3 cells. Wildtype 3TD showed a granular distribution with perinucleolar enrichment. As expected, co-expression of two wildtype constructs resulted in a perfect overlay of the two fluorescence signals. In contrast, all mutants showed clear differences in their sub-nuclear localization. F332A displayed a diffused localization with enrichment in nucleoli, while I388A and R394A showed diffused localization with nucleolar exclusion. These data show that substitutions of single residues important for peptide binding (F332, I388, or R394) changed the sub-nuclear localization of the mutants, indicating





**Fig. 6** Genomic binding of SETDB1. **a** Example browser views of the co-occurrence of SETDB1, H3K9me3, and H3K14ac peaks. **b** Venn diagram of the overlap of annotated SETDB1, H3K9me3, and H3K14ac peaks. **c** Clustering analysis of RPKM read density centered on SETDB1 peaks ( $\pm 2$  kb) showing the strong overlap of SETDB1 with H3K9me3 in cluster 1–7 and the overlap of SETDB1, H3K9me3, and H3K14ac peaks in cluster 6. **d** Overlap of the SETDB1 binding regions from cluster 1–8 with different classes of repeat elements. “Ref” refers to the average of 10 control distributions calculated by EpiExplorer<sup>61</sup> for the different clusters. The standard errors of the average repeat contents in these control distributions are indicated by the error bars. More details and statistics are provided in Supplementary Note 3. Note the strong enrichment of LTRs in clusters 2, 3, and 5 and of LINES in cluster 6. **e** Meta-profiles of SETDB1 density on SETDB1 overlapping LINE (LINE-SETDB1), all LINE (LINE-mm10), SETDB1 overlapping LTR (LTR-SETDB1), and all LTR elements (LTR-mm10)





**b**

| Pattern                | A    | B     | C     | D    | E    | F    | Total number of counted cells |
|------------------------|------|-------|-------|------|------|------|-------------------------------|
| Schematic illustration |      |       |       |      |      |      |                               |
| WT                     | 100% | 0     | 0     | 0    | 0    | 0    | 82                            |
| F332A                  | 0    | 0     | 96.6% | 2.3% | 0    | 1.1% | 87                            |
| I388A                  | 0    | 97.5% | 0     | 1.3% | 1.3% | 0    | 79                            |
| R394A                  | 0    | 96.4% | 0     | 2.4% | 0    | 1.2% | 83                            |

**Fig. 7** Subnuclear localization of 3TD and 3TD mutants. Pairs of 3TD proteins were co-expressed as fusion constructs with mCerulean (shown in cyan) or mVenus (shown in yellow) in mouse NIH3T3 cells. An overlay of both signals is shown in the third column. **a** Exemplary images of wildtype/wildtype controls (top row) showing a perfect overlay of the two fluorescence signals. In contrast, images of mutant/wildtype combinations clearly show differences in the sub-nuclear localization of the mutants. Scale bar: 10  $\mu$ m. Additional images are shown in Supplementary Figs. 10 and 11. **b** Summary of the localization data. Around 80 cells were inspected per experiment, classified to different phenotypes and counted. The following phenotypes were used: A (wildtype like)-granular, perinucleolar with nucleolar exclusion, B-diffused with nucleolar exclusion, C-diffused with nucleolar enrichment, D-diffused, E-diffused, perinucleolar with nucleolar exclusion, F-diffused, perinucleolar with nucleolar enrichment. Phenotypes D-F were observed very rarely

that H3 tail binding affects the sub-nuclear localization of 3TD. We also investigated the contribution of 3TD to the localization of the full-length SETDB1, but quick and efficient export of SETDB1 from the nucleus, consistent with previous reports<sup>34,35</sup>, prevented us from drawing any conclusions.

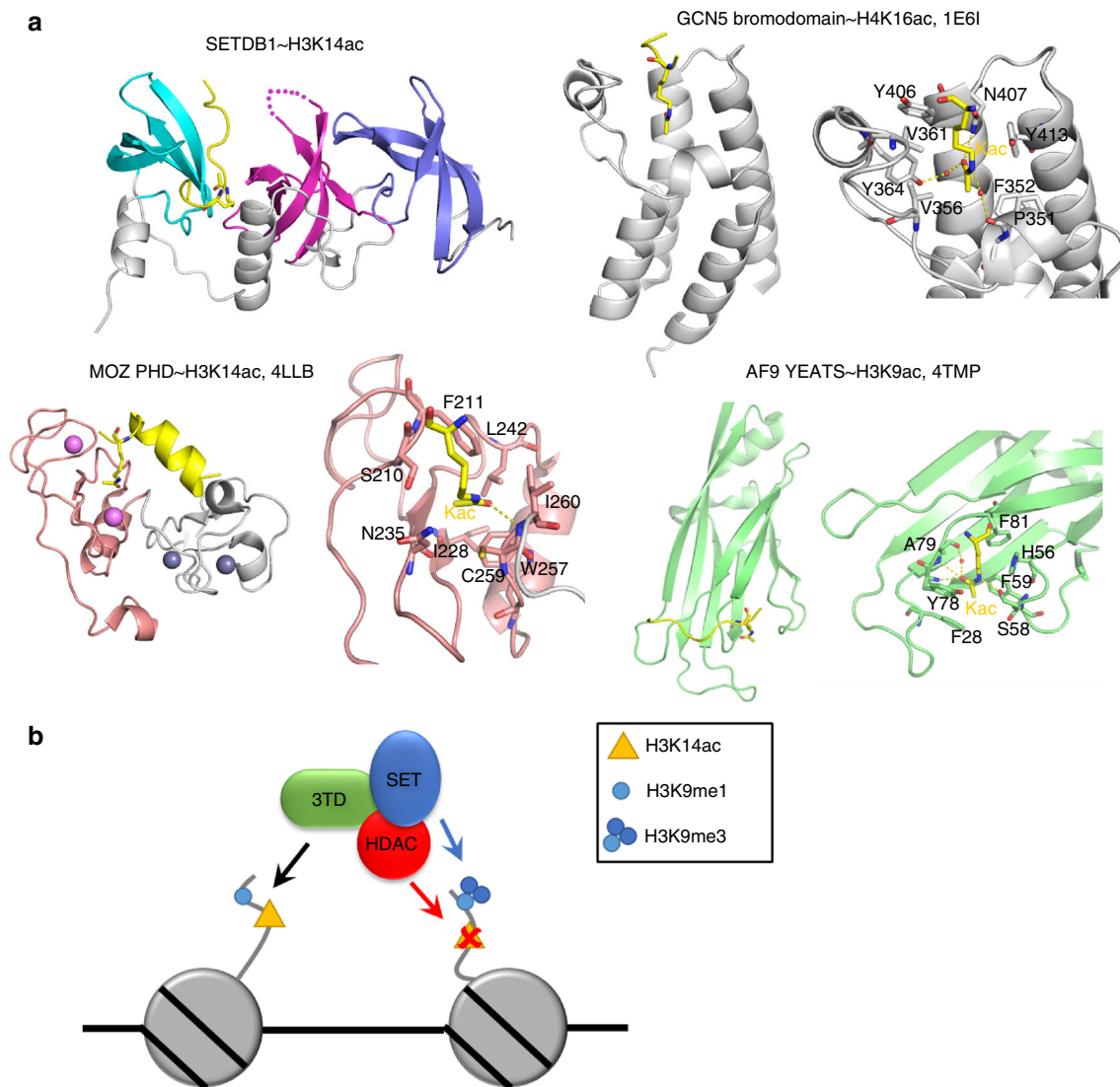
## Discussion

It is well established that epigenetic writers frequently also contain reading functions<sup>36</sup>. Often the modification introduced by the writer is read by the reader, as exemplified in the cases of the

SUV39H1<sup>37</sup> and G9a H3K9 PKMTs<sup>38</sup>, a process that can contribute to propagation and spreading of the mark and by doing so to ensure stability of the epigenome network. However, under specific conditions, efficient switching mechanisms are required to change functional chromatin states. Under these circumstances, recruitment of the required writers would depend on the reading of an “opposing” signal, as discovered in this work for the SETDB1 H3K9 trimethylase. We show here that the 3TD domain of SETDB1 specifically binds to histone H3 tails containing a combinatorial modification at K9 and K14. The activating modification, H3K14ac, is recognized as a primary mark responsible for binding, and must be accompanied by methylation of H3K9, which is associated with active or repressed chromatin depending on the methylation state. Co-occurrence of K9 methylation and K14 acetylation had been shown previously by mass spectrometry in histone H3 isolated from different human cell types (Supplementary Fig. 12)<sup>39,40</sup>, but the functions of these double marks are not known. We show activity of SETDB1 on H3K14-acetylated substrates in peptide array methylation experiments, suggesting that recruitment of SETDB1 to H3K9me/K14ac chromatin can result in silencing of the target regions, by introduction of H3K9me3 via the SET domain of SETDB1 and by deacetylation of K14ac via SETDB1-associated HDAC activity. This finding is in agreement with the observation that H3K9me3 levels decrease on H3 tails with acetylated K14 after the combined knock-down of the three H3K9 methyltransferases EHMT1, EHMT2, and SETDB1<sup>40</sup>, and suggests that SETDB1 is responsible for this effect.

The three Tudor modules of SETDB1 fold as a tightly connected protein domain. Binding of K14ac occurs at the interface between TD2 and TD3 and is mediated by hydrogen bonding and hydrophobic interactions. Previously, the bromodomain has been characterized as an acetyl-lysine binding module<sup>41</sup>, and tandem PHD fingers of DPF3b<sup>42</sup> and MOZ<sup>43–45</sup> have also been reported to recognize H3K14ac. The MOZ tandem PHD finger in addition interacts with acetylated lysine<sup>46</sup>. Moreover, the YEATS domains of AF9 and Taf14 were reported to recognize H3K9ac<sup>47,48</sup> as well as lysine crotonylation<sup>49,50</sup>. In contrast, Tudor domains have been identified as methyl-lysine or methyl-arginine reading domains<sup>26,27</sup>. Our biochemical and structural analyses demonstrate that 3TD of SETDB1 functions as a recognition domain for acetyl-lysine (Fig. 8). Previously, binding to a similar bivalent H3K9me3/H3K18ac modification has been observed for TRIM33<sup>51</sup>. However, in that case the binding is mediated by two separate reading domains (one PHD domain and one bromodomain), which need to reside in a greater distance to one another to allow for specific binding of H3K9me3/H3K18ac to happen. The observation that TRIM33 binding to H3K9me3/H3K18ac activates gene expression, while SETDB1 binding to H3K9me/K14ac likely leads to gene silencing (because H3K9me3 is introduced and K14ac could be removed), suggests that these combinatorial K9me/Kac modification states have distinct roles. Interestingly, the conformation of the peptide bound to 3TD depends on the modification state of K9. While K9me3 binding occurs at the TD2 domain with higher affinity, TD3 preferentially recognizes K9me1 and K9me2 substrates—illustrating a pocket switch mechanism in a histone reading domain. It is conceivable that this conformational change of the bound peptide triggers downstream effects in SETDB1, for example allosterically modulating its activity or controlling the interaction with its complex partners, like KAP-1<sup>9</sup>, the NuRD complex<sup>19–21</sup>, DNMT3A<sup>17</sup>, other H3K9 PKMTs<sup>18</sup>, or the HUSH complex<sup>16</sup>. Future work will be needed to explore these effects in more detail.

Our data suggest that H3K9me and K14ac define a bivalent chromatin state, which is bound by 3TD of SETDB1, in agreement with other examples documenting roles of SETDB1 at



**Fig. 8** Overview of Kac binding domains and model of the function of SETDB1. **a** Comparison of the 3TD Kac binding site with Kac binding to the GCN5 bromodomain, MOZ PHD domain, and AF9 YEATS domain. **b** Model illustrating the putative role of 3TD in the function of SETDB1. Binding of 3TD to H3 tails containing K9me-K14ac recruits the SETDB1 complex to these regions. Afterward, the SET domain of SETDB1 introduces H3K9me3, and depending on the complex composition, the HDAC activity of NuRD may remove H3K14 acetylation, finally leading to the silencing of the genomic region

bivalent chromatin<sup>32,52</sup>. By this process, SETDB1-mediated silencing of endogenous retroviral elements might be connected to the activating H3K14ac modification, when it co-occurs with methylated H3K9. After recruitment of SETDB1 to H3K9me1/K14ac or H3K9me2/K14ac sites, SETDB1 is expected to methylate them to the K9me3 state (Fig. 8b), which is in agreement with the binding of SETDB1 at LINE elements containing K9me3/K14ac. Depending on the exact composition of the SETDB1 complex bound, deacetylation of K14ac could also be catalyzed, suggesting that SETDB1 bound to K9me3 marked LTRs initially could have been recruited by K14ac, but this modification has already been removed. Recruitment by K14ac might help to focus SETDB1 activity on the most active LTR and LINE elements, supporting the silencing of active and hence most dangerous repeat elements. In agreement with this model, the SETDB1-mediated silencing of endogenous retroviral elements has been also linked to histone variant H3.3, which is another signal often associated with gene activation<sup>53</sup>. Hence, the SETDB1 silencing system might be specifically attracted to repeat elements decorated with activating marks, such as H3.3 and H3K14ac. In

summary, our data suggest that recruitment of SETDB1 to H3K9me/H3K14ac might allow to efficiently counteract acetylation of these elements, introduce K9 trimethylation, and ensure silencing of retrotransposons containing internal activating elements. Future work may include the generation of endogenous SETDB1 mutants containing amino acid exchanges in 3TD by CRISPR/Cas9-mediated homologous recombination and subsequent studies of the biological consequences of these exchanges.

## Methods

**Protein expression and purification.** Wildtype SETDB1-3TD was cloned into the pET28a-MHL (residues 190–410), pGEX6P2, and pGEX-4T-1 (residues 197–403) vectors. Point mutants were generated by site-directed mutagenesis. Wildtype and mutant proteins were expressed in an SGC-generated derivative strain of BL21 *Escherichia coli* with the pRARE plasmid for overcoming the codon bias. Cells were grown in TB medium or in M9 minimal medium supplemented with <sup>15</sup>NH<sub>4</sub>Cl for the NMR analyses. Expression was induced with 0.5 mM IPTG, and the bacteria were harvested by centrifugation and lysed by sonication. The GST-fusion proteins were purified using Glutathione-Sepharose 4B beads (GE Healthcare). The GST-tag was either cleaved with PreScission protease, or left for the purposes of western blot analysis or peptide array probing, in which case the GST-fusion protein was eluted off the Glutathione-Sepharose beads using 50 mM reduced L-glutathione (Sigma-

Aldrich). The 6×His-tagged protein was purified by an Ni-NTA column (Qiagen) followed by a gel filtration column (Superdex 75, GE Healthcare). For SETDB1 methylation kinetics, an N-terminus-deleted construct was cloned into the pFBOH-SUMOstar-TEV vector as 6×His-SUMO-TEV-SETDB1 (190–1291) fusion protein. The resulting virus was transformed into Sf9 insect cell for protein expression, and purified using Ni-NTA column (Qiagen) followed by ion exchange chromatography (Hitrap Q HP column).

**CD spectroscopy.** Folding of purified proteins was analyzed by circular dichroism (CD) spectroscopy. CD measurements were performed using a J-815 circular dichroism spectrophotometer (JASCO Corporation, Tokyo, Japan) using GST-tagged 3TD domains in 100 mM KCl, 5 mM HEPES pH 7.5, and 5% glycerol. Depending on the protein preparation, spectra were measured at 10 μM or 2 μM protein. The spectra were collected at 22 °C using a 0.1-mm cuvette in a wavelength range between 195 and 240 nm. For each sample, at least 60 scans were collected and averaged.

**Protein crystallization.** Purified proteins mixed with a 3- to 5-fold stoichiometric excess of the respective peptides were crystallized using the sitting drop vapor diffusion method at 18 °C by mixing 0.5 μl of the protein with 0.5 μl of the reservoir solution. The 3TD bound with H3K9me2/K14ac or H3K9me3/K14ac complex was crystallized in 25% PEG 3350, 0.2 M NaCl, 0.1 M HEPES pH 7.5, and 5% glycerol. The 3TD–H3K9me0/K14ac complex was crystallized in 20% PEG 5000 mono-methyl ether, 0.1 M Bis-Tris pH 6.5. The 3TD W358A mutant bound with H3K9me2/K14ac or H3K9me3/K14ac complex was crystallized in 25% PEG 3350, 0.2 M Li<sub>2</sub>SO<sub>4</sub>, and 0.1 M HEPES pH 7.5. Further details of crystallization and structure determination are given in Supplementary Note 1.

**Nuclear magnetic resonance spectroscopy.** NMR experiments were collected on a Bruker Avance 500-MHz spectrometer equipped with a cryogenic probe. Chemical shift perturbation experiments were carried out using uniformly <sup>15</sup>N-labeled wildtype 3TD protein in 20 mM Tris/HCl pH 7.0 and 150 mM NaCl. <sup>1</sup>H, <sup>15</sup>N HSQC spectra were recorded at 25 °C in the presence of increasing concentrations of histone tail peptides. *K<sub>d</sub>* values were calculated by a nonlinear least-squares analysis in Kaleidagraph using the equation

$$\Delta\delta = \Delta\delta_{\max} \left( \frac{c_L + c_P + K_d}{c_L + c_P + K_d} - \sqrt{\left( \frac{c_L + c_P + K_d}{c_L + c_P + K_d} \right)^2 - 4c_L c_P} \right) / 2c_P,$$

where *c<sub>L</sub>* is the concentration of the peptide, *c<sub>P</sub>* is the concentration of the protein,  $\Delta\delta$  is the observed normalized chemical shift change, and  $\Delta\delta_{\max}$  is the normalized chemical shift change at saturation, calculated as

$$\Delta\delta = \sqrt{(\Delta\delta_H)^2 + (\Delta\delta_N/5)^2},$$

where  $\delta$  is the chemical shift in p.p.m.

**Isothermal titration calorimetry.** ITC experiments were carried out at 25 °C on a VP-ITC calorimeter (MicroCal). Protein and peptide were kept in an identical buffer of 20 mM Tris/HCl pH 7.0 and 150 mM NaCl. ITC measurements were carried out with protein and ligand concentrations ranging from 50 to 100 μM and 1–2 mM, respectively. Binding isotherms were analyzed by nonlinear least-squares fitting of the data using Microcal ORIGIN software (Microcal) using a one-site binding model.

**Mass spectrometry.** 3TD pull-down and input control mononucleosomal chromatin isolated from HEK293 cells were resolved on 15% SDS-PAGE gels. The bands corresponding to histone H3 were excised from the gels and shredded into small 1-mm cubes. Coomassie, detergents, and other chemicals that could interfere with trypsin digestion or MALDI analysis were removed in three subsequent 15-min incubation steps with 100 μl of 1:1 (v/v) mixture of 50 mM ammonium bicarbonate and acetonitrile. After each incubation, the gel pieces were dehydrated by addition of 100 μl pure acetonitrile and the liquid was discarded after each wash. For in-gel digestion, the gel pieces were rehydrated in 50 μl 25 mM ammonium bicarbonate containing 5 ng/μl sequencing-grade modified trypsin (Promega) and incubated at 37 °C for 30 min. Afterward, the supernatant containing the digested peptides was transferred into a new tube, flash frozen in liquid nitrogen for later use, or directly spotted on an AnchorChip (Bruker) target plate, dried, and covered with alpha-cyano-4-hydroxycinnamic acid (HCCA). Peptide mass fingerprinting was done on an AutoFlex Speed MALDI TOF/TOF mass spectrometer (Bruker) operated in reflected positive mode. The spectra were recorded between 500 and 4000 *m/z*. Peptide calibration standard (Bruker) was used for mass calibration. The obtained spectra were analyzed with the BioTools package (Bruker). For MS/MS analysis, selected peaks were fragmented and measured in LIFT mode, the mass spectra were matched against possible peptides found in nonredundant (nr) database using MASCOT software, or matched to the theoretical sequence of the peptide identified in the first PMF analysis.

**Cell lines.** Human HEK293 and mouse NIH3T3 cells were obtained from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures. The mouse ES cell line E14 (129/Ola) was obtained from Wolf Reik (Cambridge, UK). All cells were free of mycoplasma.

**Binding and methylation of peptide arrays.** Experiments to determine the binding specificity of 3TD wildtype and mutants were performed using CelluSpots peptide arrays (Modified Histone Peptide Array, Active Motif), as described previously, in 20 mM HEPES pH 7.5, 100 mM KCl, 1 mM EDTA, 0.2 mM DTT, and 10% glycerol<sup>54</sup>. All peptide array binding studies were conducted at least in two technical repeats. Annotations of all spots are given in Supplementary Table 2. The amounts of GST-fused Tudor proteins used in peptide array binding and western blot analyses ranged from 0.5 to 100 nM. For methylation analysis, Modified Histone Peptide arrays (Active Motif) were incubated with 20 nM SETDB1 in buffer (50 mM Tris/HCl pH 9.0, 4 mM DTT, 5 mM MgCl<sub>2</sub>) supplemented with 0.38 μM radioactively labeled [methyl-<sup>3</sup>H]-AdoMet (Perkin Elmer) for 2 h at room temperature. The arrays were then washed 5 times with wash buffer (100 mM NH<sub>4</sub>HCO<sub>3</sub> and 1% SDS), incubated in Amplify NAMP100 (GE Healthcare) for 5 min, and then exposed to Hyperfilm TM ECL high-performance chemiluminescence film (GE Healthcare) in the dark at –80 °C for 1–7 days.

**Binding to histones and histone tails.** Native histones were isolated by acid extraction<sup>55</sup> from HEK293 cells, and recombinant histone H3 was purchased from New England Biolabs. For the experiment with HDAC inhibitors, cells were treated for 48 h with 0.4 mM TSA or 5 mM sodium butyrate. Five micrograms of native histones and matching amount of recombinant H3 were electrophoresed on 18% SDS-PAGE and transferred onto nitrocellulose membranes. The membrane was then blocked in 5% milk in TBST and incubated with purified GST-tagged 3TD, followed by the anti-GST (GE Healthcare 27457701A, dilution 1:5000) and anti-goat antibody (Sigma A4174, dilution 1:10000). The amount of GST-fused Tudor protein used in western blot analyses was 100 nM. All experiments were performed in duplicates or triplicates. In the western blots, following antibodies were used: anti-K14ac (Millipore MABE351, dilution for checking specificity: 1:2000 or 1:1000; for western blot: 1:500), anti-H3K9me3 (Abcam ab8898, dilution 1:2000), anti-H3K9me2 (Abcam ab1220, dilution 1:1000), and anti-H3K4me3 (Abcam ab8580, dilution 1:2000). For specificity analysis of antibodies, see Supplementary Fig. 13. Secondary antibodies were anti-rabbit IgG HRP (GE Healthcare NA934, dilution 1:5000) and anti-mouse IgG HRP (GE Healthcare NXA931, dilution 1:5000). Native histone and nucleosome pull-down experiments were done at least in two biological repeats. All antibodies were used following the manufacturer's recommendations. Uncropped gel images are shown in Supplementary Fig. 14.

**Analysis of SETDB1 genomic binding.** SETDB1, H3K9me3, and K14ac ChIP data from murine embryonic stem cells were from GSE17642<sup>31</sup>, GSE18371<sup>32</sup>, and GSE31284<sup>33</sup>, respectively. The raw reads were mapped to mm10 with Bowtie2, using default settings<sup>56</sup> and RPKM normalized. Peaks were called with MACS<sup>57</sup> in the Galaxy environment<sup>58</sup>. The quantitative analyses of peak overlap were carried out in Galaxy using the Intersect tools. ComputeMatrix and PlotHeatmap in DeepTools was used for generation of heatmaps, meta-profiles, and K-means clustering<sup>59</sup>. Statistics were determined using the Fisher's Test and ShuffleBED tools in DeepTools. Overlap regions were checked not to include blacklisted regions for mm10 downloaded from ENCODE. LINE and LTR elements were downloaded from NCBI table browser. The Integrative Genomics Viewer was used for visualization of the data and preparation of browser shots<sup>60</sup>. For annotation of repeats and known chromatin modifications to genomic regions, EpiExplorer was used<sup>61</sup>.

**Fluorescence microscopy.** For sub-nuclear localization studies, SETDB1 3TD wildtype and mutant domains were cloned in expression constructs containing an NLS and fused to mCerulean and mVenus. NIH3T3 cells were seeded on tissue culture dishes with glass bottom and 1 day later co-transfected with plasmids expressing one mCerulean and one mVenus tagged 3TD wildtype or mutant domain using FuGENE HD (Promega) according to the manufacturer's instructions. After 24 h, images were collected using a Zeiss LSM 710 confocal microscope and Plan-Apochromat 63×/1.40 Oil DIC M27 objective.

**Data availability.** Structural data have been submitted to the PDB database under accession numbers 6BHD, 6BHE, 6BHG, 6BHH, and 6BHI. All other data supporting the findings of this study are available within the paper and its supplementary information files. Additional primary data are available from the corresponding authors upon reasonable request.

Received: 16 November 2016 Accepted: 14 November 2017

Published online: 12 December 2017



## References

- Chi, P., Allis, C. D. & Wang, G. G. Covalent histone modifications—miswritten, misinterpreted and mis-erased in human cancers. *Nat. Rev. Cancer* **10**, 457–469 (2010).
- Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381–395 (2011).
- Martin, C. & Zhang, Y. The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* **6**, 838–849 (2005).
- Krishnan, S., Horowitz, S. & Trievel, R. C. Structure and function of histone H3 lysine 9 methyltransferases and demethylases. *ChemBiochem* **12**, 254–263 (2011).
- Becker, J. S., Nicetto, D. & Zaret, K. S. H3K9me3-dependent heterochromatin: barrier to cell fate changes. *Trends Genet.* **32**, 29–41 (2016).
- Mozzetta, C., Boyarchuk, E., Pontis, J. & Ait-Si-Ali, S. Sound of silence: the properties and functions of repressive Lys methyltransferases. *Nat. Rev. Mol. Cell Biol.* **16**, 499–513 (2015).
- Barski, A. et al. High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
- Shahbazian, M. D. & Grunstein, M. Functions of site-specific histone acetylation and deacetylation. *Annu. Rev. Biochem.* **76**, 75–100 (2007).
- Schultz, D. C. et al. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* **16**, 919–932 (2002).
- Dodge, J. E., Kang, Y. K., Beppu, H., Lei, H. & Li, E. Histone H3-K9 methyltransferase ESET is essential for early development. *Mol. Cell Biol.* **24**, 2478–2486 (2004).
- Karimi, M. M. et al. DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell* **8**, 676–687 (2011).
- Matsui, T. et al. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* **464**, 927–931 (2010).
- Rowe, H. M. et al. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **463**, 237–240 (2010).
- Liu, S. et al. Setdb1 is required for germline development and silencing of H3K9me3-marked endogenous retroviruses in primordial germ cells. *Genes Dev.* **28**, 2041–2055 (2014).
- Collins, P. L., Kyle, K. E., Egawa, T., Shinkai, Y. & Oltz, E. M. The histone methyltransferase SETDB1 represses endogenous and exogenous retroviruses in B lymphocytes. *Proc. Natl Acad. Sci. USA* **112**, 8367–8372 (2015).
- Tchasovnikarova, I. A. et al. GENE SILENCING. Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells. *Science* **348**, 1481–1485 (2015).
- Li, H. et al. The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells. *J. Biol. Chem.* **281**, 19489–19500 (2006).
- Fritsch, L. et al. A subset of the histone H3 lysine 9 methyltransferases Suv39h1, G9a, GLP, and SETDB1 participate in a multimeric complex. *Mol. Cell* **37**, 46–56 (2010).
- Capili, A. D., Schultz, D. C., Rauscher, I. F. & Borden, K. L. Solution structure of the PHD domain from the KAP-1 corepressor: structural determinants for PHD, RING and LIM zinc-binding domains. *EMBO J.* **20**, 165–177 (2001).
- Schultz, D. C., Friedman, J. R. & Rauscher, F. J. 3rd Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2 $\alpha$  subunit of NuRD. *Genes Dev.* **15**, 428–443 (2001).
- Yang, L. et al. An ERG (ets-related gene)-associated histone methyltransferase interacts with histone deacetylases 1/2 and transcription co-repressors mSin3A/B. *Biochem. J.* **369**, 651–657 (2003).
- Castro-Diaz, N. et al. Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes Dev.* **28**, 1397–1409 (2014).
- Walter, M., Teissandier, A., Perez-Palacios, R. & Bourc'his, D. An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells. *eLife* **5**, e11418 (2016).
- Kim, J. et al. Maternal Setdb1 is required for meiotic progression and preimplantation development in mouse. *PLoS Genet.* **12**, e1005970 (2016).
- Patel, D. J. & Wang, Z. Readout of epigenetic modifications. *Annu. Rev. Biochem.* **82**, 81–118 (2013).
- Lu, R. & Wang, G. G. Tudor: a versatile family of histone methylation 'readers'. *Trends Biochem. Sci.* **38**, 546–555 (2013).
- Chen, C., Nott, T. J., Jin, J. & Pawson, T. Deciphering arginine methylation: Tudor tells the tale. *Nat. Rev. Mol. Cell Biol.* **12**, 629–642 (2011).
- Kim, J. et al. Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep.* **7**, 397–403 (2006).
- Kungulovski, G. et al. Application of histone modification-specific interaction domains as an alternative to antibodies. *Genome Res.* **24**, 1842–1853 (2014).
- Kungulovski, G., Mauser, R., Reinhardt, R. & Jeltsch, A. Application of recombinant TAF3 PHD domain instead of anti-H3K4me3 antibody. *Epigenetics Chromatin* **9**, 11 (2016).
- Yuan, P. et al. Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. *Genes Dev.* **23**, 2507–2520 (2009).
- Bilodeau, S., Kagey, M. H., Frampton, G. M., Rahl, P. B. & Young, R. A. SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev.* **23**, 2484–2489 (2009).
- Karmodiya, K., Krebs, A. R., Oulad-Abdelghani, M., Kimura, H. & Tora, L. H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genom.* **13**, 424 (2012).
- Tachibana, K. et al. Analysis of the subcellular localization of the human histone methyltransferase SETDB1. *Biochem. Biophys. Res. Commun.* **465**, 725–731 (2015).
- Cho, S., Park, J. S. & Kang, Y. K. Regulated nuclear entry of over-expressed Setdb1. *Genes Cells* **18**, 694–703 (2013).
- Liu, N. et al. Recognition of H3K9 methylation by GLP is required for efficient establishment of H3K9 methylation, rapid target gene repression, and mouse viability. *Genes Dev.* **29**, 379–393 (2015).
- Muller, M. M., Fierz, B., Bittova, L., Liszczak, G. & Muir, T. W. A two-state activation mechanism controls the histone methyltransferase Suv39h1. *Nat. Chem. Biol.* **12**, 188–193 (2016).
- Collins, R. E. et al. The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. *Nat. Struct. Mol. Biol.* **15**, 245–250 (2008).
- Young, N. L. et al. High throughput characterization of combinatorial histone codes. *Mol. Cell. Proteom.* **8**, 2266–2284 (2009).
- Sridharan, R. et al. Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1 $\gamma$  in reprogramming to pluripotency. *Nat. Cell Biol.* **15**, 872–882 (2013).
- Sanchez, R. & Zhou, M. M. The role of human bromodomains in chromatin biology and gene transcription. *Curr. Opin. Drug Discov. Dev.* **12**, 659–665 (2009).
- Zeng, L. et al. Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature* **466**, 258–262 (2010).
- Ali, M. et al. Tandem PHD fingers of MORF/MOZ acetyltransferases display selectivity for acetylated histone H3 and are required for the association with chromatin. *J. Mol. Biol.* **424**, 328–338 (2012).
- Qiu, Y. et al. Combinatorial readout of unmodified H3R2 and acetylated H3K14 by the tandem PHD finger of MOZ reveals a regulatory mechanism for HOXA9 transcription. *Genes Dev.* **26**, 1376–1391 (2012).
- Dreveny, I. et al. The double PHD finger domain of MOZ/MYST3 induces alpha-helical structure of the histone H3 tail to facilitate acetylation and methylation sampling and modification. *Nucleic Acids Res.* **42**, 822–835 (2014).
- Klein, B. J. et al. Recognition of histone H3K14 acylation by MORF. *Structure* **25**, 650–654 (2017). e652.
- Li, Y. et al. AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation. *Cell* **159**, 558–571 (2014).
- Shanle, E. K. et al. Association of Taf14 with acetylated histone H3 directs gene transcription and the DNA damage response. *Genes Dev.* **29**, 1795–1800 (2015).
- Andrews, F. H. et al. The Taf14 YEATS domain is a reader of histone crotonylation. *Nat. Chem. Biol.* **12**, 396–398 (2016).
- Li, Y. et al. Molecular coupling of histone crotonylation and active transcription by AF9 YEATS domain. *Mol. Cell* **62**, 181–193 (2016).
- Xi, Q. et al. A poised chromatin platform for TGF-beta access to master regulators. *Cell* **147**, 1511–1524 (2011).
- Matsumura, Y. et al. H3K4/H3K9me3 bivalent chromatin domains targeted by Lineage-Specific DNA methylation pauses adipocyte differentiation. *Mol. Cell* **60**, 584–596 (2015).
- Elsasser, S. J., Noh, K. M., Diaz, N., Allis, C. D. & Banaszynski, L. A. Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature* **522**, 240–244 (2015).
- Bock, I. et al. Application of Celluspot peptide arrays for the analysis of the binding specificity of epigenetic reading domains to modified histone tails. *BMC Biochem.* **12**, 48 (2011).
- Shechter, D., Dormann, H. L., Allis, C. D. & Hake, S. B. Extraction, purification and analysis of histones. *Nat. Protoc.* **2**, 1445–1457 (2007).
- Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
- Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
- Afgan, E. et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* **44**, W3–W10 (2016).
- Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).



60. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
61. Halachev, K., Bast, H., Albrecht, F., Lengauer, T. & Bock, C. EpiExplorer: live exploration and global analysis of large epigenomic datasets. *Genome Biol.* **13**, R96 (2012).

### Acknowledgements

We thank R. Scott Houlston for technical assistance in NMR experiments, Maria Amaya for determining the apostructure of SETDB1 (3DLM), and Aiping Dong for collecting some of the diffraction data. We gratefully acknowledge provision of mES E14 cells by Ferdinand v. Meyen and Wof Reik (Babraham Institute, UK). This work has been supported by the Deutsche Forschungsgemeinschaft (JE 252/26-1). It was also partially supported by National Natural Science Foundation of China (31500615). The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada through Ontario Genomics Institute, Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, Merck & Co., Novartis Pharma AG, Ontario Ministry of Economic Development and Innovation, Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and the Wellcome Trust. Some of the work described here was performed at Structural Biology Center and GM/CA beamlines. GM/CA@APS has been funded in whole or in part with Federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

### Author contributions

R.Z.J., S.Q., G.K., J.M., and A.J. designed the experiments. R.Z.J. with help of J.S., S.K., S.W., T.P.J., P.B., A.J., and R.T. conducted the array binding and methylation, histone and nucleosome binding, microscopy, and mass spectrometry experiments. S.Q. performed the ITC experiments and NMR studies. R.Z.J., H.W., and P.L. cloned the constructs. S.Q., Y.L., and L.D. purified and crystallized the protein. W.T. prepared crystallographic models of 3TD-K9meX/K14ac complexes. G.K. with help of R.Z.J. and A.J. conducted the

genomics experiments. All authors contributed to data analysis and interpretation. A.J., R.Z.J., S.Q., and J.M. wrote the manuscript. All authors contributed to the editing of the manuscript and approved its final version. R.Z.J. and S.Q. contributed equally to this work.

### Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-017-02259-9>.

**Competing interests:** The authors declare that they have no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017