Molecular recognition of histone lysine methylation by the Polycomb group repressor dSfmbt

Clemens Grimm1,4, Raquel Matos2, Nga Ly-Hartig2, Ulrich Steuerwald3, Doris Lindner3, Vladimir Rybin3, Jürg Müller2,*, Christoph W Müller3,*

1European Molecular Biology Laboratory, Grenoble Outstation, Grenoble, France, 2European Molecular Biology Laboratory, Gene Expression Unit, Heidelberg, Germany and 3European Molecular Biology Laboratory, Structural and Computational Biology Unit, Heidelberg, Germany

Polycomb group (PcG) proteins repress transcription by modifying chromatin structure in target genes. dSfmbt is a subunit of the Drosophila melanogaster PcG protein complex PhoRC and contains four malignant brain tumour (MBT) repeats involved in the recognition of various mono- and dimethylated histidine peptides. Here, we present the crystal structure of the four-MBT-repeat domain of dSfmbt in complex with a mono-methylated histone H4 peptide. Only a single histone peptide binds to the four-MBT-repeat domain. Mutational analyses show high-affinity binding with low peptide sequence selectivity through combinatorial interaction of the methyl-lysine with an aromatic cage and positively charged flanking residues with the surrounding negatively charged surface of the fourth MBT repeat. dSfmbt directly interacts with the PcG protein Scm, a related MBT-repeat protein with similar methyl-lysine binding activity. dSfmbt and Scm co-occupy Polycomb response elements of target genes in Drosophila and they strongly synergize in the repression of these target genes, suggesting that the combined action of these two MBT proteins is crucial for Polycomb silencing.

The EMBO Journal (2009) 28, 1965–1977. doi:10.1038/emboj.2009.147; Published online 4 June 2009

Subject Categories: chromatin & transcription; structural biology

Keywords: histone modifications; MBT repeat; methyl-lysine recognition; Polycomb group proteins

Introduction

Polycomb group (PcG) proteins are transcriptional regulators required for the repression of developmental control genes in animals and plants. PcG proteins exist in distinct multi-protein complexes that repress transcription by modifying the chromatin of target genes and thereby generating transcriptional off states that can be stably and heritably maintained (Francis and Kingston, 2001; Schwartz and Pirrotta, 2007). To date, three principal PcG multi-protein complexes have been identified and characterized: Pho repressive complex (PhoRC), PRC2 and the two related complexes PRC1 and dRAF (Schwartz and Pirrotta, 2007; Müller and Verrijzer, 2009). Among those, the PhoRC subunit Pho is the only sequence-specific DNA-binding PcG protein. Studies in Drosophila showed that PcG complexes assemble at specific cis-regulatory sequences in target genes, called Polycomb response elements (PRE), and that PhoRC has a central function in providing a PRE-binding platform that allows for the assembly of the chromatin-binding PRC1 and PRC2 complexes (Wang et al, 2004; Mohd-Sarip et al, 2005; Klymenko et al, 2006).

In addition to Pho, PhoRC contains dSfmbt (Klymenko et al, 2006). In Drosophila, dSfmbt, the PRC1 subunit Sex comb on midleg (Scm) and a third protein, called L(3)mbt, form a small protein family with a very similar and unique domain architecture. The central portion of each protein contains an MBT-repeat domain that consists of two (Scm), three (L(3)mbt) or four (dSfmbt) repeats, and each protein contains Zn-finger motifs in the N-terminus and a sterile alpha motif (SAM) domain at the very C-terminus. Studies on dSfmbt, first showed that MBT-repeat domains selectively bind to mono- and dimethylated lysine residues in histones, but that they show low specificity for any particular histone lysine (Klymenko et al, 2006). Recent studies reported the crystal structures of the MBT domains of Scm and L3MBTL1 in complex with methylated histone-tail peptides (Grimm et al, 2007; Li et al, 2007; Min et al, 2007; Santiveri et al, 2008). In both proteins, the mono- or dimethylated histone lysine residues bind to the second MBT repeat and the interactions between the methyl-lysine side chain and an aromatic pocket in this repeat contribute the major part of the binding energy, whereas histone residues adjacent to the methyl-lysine form few interactions (Grimm et al, 2007; Li et al, 2007; Min et al, 2007; Santiveri et al, 2008). Consistent with this mode of recognition, the MBT-repeat domain of Scm binds histone-tail peptides, mono-methylated at H3-K9 or H4-K20 with a low affinity of about 500–800 μM (Grimm et al, 2007; Santiveri et al, 2008), whereas for binding of L3MBTL1 to the same mono-methylated lysines in peptides,
two studies reported different affinities ranging from 140 to 400 μM (Min et al., 2007) or from 5 to 10 μM (Li et al., 2007).

Interestingly, two distinct MBT-repeat-containing proteins, Scm and dSfmbt, are both essential components of the PcG-repression system in Drosophila. Functional studies on Scm showed that mutations in the MBT-repeat domain that abolish methyl-lysine binding in vitro impede the Polycomb-repressor function of this protein in Drosophila (Grimm et al., 2007). Intriguingly, dSfmbt binds the same methylated lysines in histones bound by Scm but with about 100-fold higher affinity than Scm (Klymenko et al., 2006; Grimm et al., 2007; Santiveri et al., 2008). These observations, together with the lack of knowledge of sequence-specific methyl-lysine recognition by the L3MBT1 or Scm MBT-repeat domains prompted us to characterize the MBT-repeat domain of dSfmbt at the structural and functional level. Here, we report the crystal structure of the MBT-repeat domain of dSfmbt in complex with a histone H4 peptide, mono-methylated at lysine 20 (H4K20me1). Using isothermal calorimetry (ITC), we evaluate the binding specificity of dSfmbt for different histone-tail peptides methylated at particular lysine residues and assess the contribution of residues adjacent to the methyl-lysine residue by mutational analysis. Functional tests in Drosophila show that dSfmbt and Scm act in a highly synergistic manner to maintain repression at Polycomb target genes in vivo and suggest a role for the Scm–dSfmbt heterodimer in chromatin compaction.

Results and discussion

**Overall structure of the four-MBT-repeat domain of dSfmbt**

The structure of the four-MBT-repeat domain of D. melanogaster dSfmbt (dSfmbt-4MBT, Mr = 51 kDa, residues 535–977) was solved in complex with a histone H4 tail peptide centred onto H4K20me1 at 2.8 Å resolution (Table I). To favour crystallization, three point mutations (K715D, R886S and R900D) were introduced on the surface of the dSfmbt-4MBT construct; these mutations do not significantly affect H4K20me1 binding (Table II, Materials and Methods). The overall structure of the dSfmbt-4MBT–peptide complex is shown in Figure 1. As in Scm and L3MBT1, each MBT repeat consists of a central five-stranded β-core and an elongated N-terminal arm that contacts the neighbouring repeat. Repeat 2, 3 and 4 form a propeller-like structure with three-fold pseudo-symmetry similar to L3MBT1 (Wang et al., 2003). Repeat 1 is docked onto the outer rim of this propeller in the area of repeat 4 and forms most contacts with repeat 4 but also interacts with the adjacent repeat 2 through the N-terminal arm of this repeat. The arm of repeat 1 forms most of the contact surface to repeat 4 and its conformation is therefore less extended compared with the three other arms (Figure 1B).

Table I  Crystallographic data collection, phasing and model refinement statistics

| Data statistics | Native | Hg derivative<sup>a</sup> |
|-----------------|--------|--------------------------|
| Space group     | P22,2<sub>1</sub>,2<sub>c</sub> | C222 |
| Cell axes (Å)   | 70.8, 97.0, 214.1 | 100.4, 140.0, 275.2 |
| Wavelength (Å)  | 0.9330 | 1.0085 |
| Resolution range (Å)<sup>b</sup> | 30–2.8 (3.0–2.8) | 30–3.2 (3.37–3.2) |
| Reflections observed/unique | 226138/37129 | 137646/32417 |
| Completeness (%) | 99.8 (99.9) | 97.6 (91.4) |
| R<sub>sym</sub> (%)<sup>c</sup> | 12.3 (3.3) | 4.4 (1.7) |
| R<sub>free</sub> (%) | 13.7 (55.1) | 8.5 (49.0) |

Phasing statistics

| Resolution range (Å) | 30–2.8 (2.87–2.8) |
|----------------------|-------------------|
| R<sub>sym<sub>clus</sub></sub> (acentric) (%) | 0.82 |
| Phasing power (acentric) | 0.99 |
| Overall figure of merit | 0.50 |

Model refinement

| Resolution range (Å) | 30–2.8 |
|----------------------|--------|
| R<sub>cryst</sub>/R<sub>free</sub> (%)<sup>d</sup> | 22.3 (32.6)/25.4 (36.0) |
| Overall B-factor [Å<sup>2</sup>] | 28.9 |
| Protein main/side chain [Å<sup>2</sup>] | 28.0/29.9 |
| Peptide ligand [Å<sup>2</sup>] | 34.7 |
| r.m.s.d. bond lengths [Å] | 0.009 |
| r.m.s.d. angles (deg) | 1.35 |
| Residues in Ramachandran plot (%): | |
| Most favoured region | 89.4 |
| Allowed | 9.9 |
| Generously allowed | 1.1 |
| Disallowed | 0.0 |

<sup>a</sup>Hg derivative with three sites per molecule was obtained by soaking crystals for 1 h with 0.25 mM EMTS. Hg-λ.1 was used as reference data set in the two-wavelengths MAD experiment.

<sup>b</sup>Values in parentheses refer to the outer resolution shell.

<sup>c</sup>R<sub>sym</sub> = Σ|I<sub>o</sub>–I<sub>c</sub>|/ΣI<sub>c</sub>, where I<sub>c</sub> is the observed intensity of a given reflection.

<sup>d</sup>R<sub>cryst</sub> = Σ|F<sub>c</sub>–F<sub>o</sub>|/ΣF<sub>c</sub>, where F<sub>c</sub> and F<sub>o</sub> are the observed and calculated structure factor amplitudes, respectively. R<sub>free</sub> is equal to R<sub>cryst</sub> for a randomly selected 5% subset of reflections not used in the refinement.
The combination of these interactions between the four repeats thus results in a compact MBT-repeat domain. 

**H4K20me1 peptide binds to the fourth MBT repeat of dSfmbt**

In the complex structure, the H4K20me1 peptide (RHRKme1VLR) interacts with dSfmbt MBT repeat 4 (Figure 1A). Interactions between dSfmbt and the peptide are mediated by the central mono-methylated lysine, which points in the binding pocket on top of the β-barrel of the fourth MBT repeat (Figure 2) but also through a combination of polar and hydrophobic interactions of adjacent peptide residues with residues of repeat 4.

The methyl-lysine-binding pocket of the fourth repeat is formed by residues Phe941, Trp944 and Tyr948, whose aromatic planes are oriented perpendicular to each other, forming roughly the corner of a cube. The methyl-lysine side chain closely packs against the aromatic side chains of Tyr948 and Trp944. Compared with the ‘aromatic cage’ in Scm (Grimm et al, 2007), we observe a significant distortion of the ideal rectangular geometry, mainly because dSfmbt-residue Tyr948 is oriented at an angle of approximately 60° with respect to Trp944. On the other side of the binding pocket, Asp917 binds the ε-amino group of H4K20me1 through a direct hydrogen bond assisted by electrostatic interactions. Furthermore, the pocket is outlined by residue Cys925. In addition to the interactions with the mono-methylated lysine, a salt bridge connects dSfmbt Glu947 (corresponding to Scm Ala354) with Arg19 in histone H4, whereas the hydroxyl group of Tyr948 (corresponding to Scm Phe355) forms a

### Table II  Binding affinity of dSfmbt-4MBT for methyl-lysine-containing histone peptides

| Wild-type dSfmbt-4MBT | Mutant dSfmbt-4MBT protein | KD (μM) |
|-----------------------|---------------------------|---------|
| H4K20—KGGAKHRKVLRLDNIQ-CONH2 | D697A | >1000 |
| H4K20me1 | 1.1 ± 0.1 (11 ± 0.4) | 1.6 ± 0.7 | >1000 |
| H4K20me2 | 2.8 ± 0.2 | 2.9 ± 0.1 | >1000 |
| H4K20me3 | >1000 |
| H4K20—RHRKVLRLDNIQ-CONH2 | D808A | 1.5 ± 0.03 (12.8 ± 1.0) |
| H4K20—HRKVLRLDNIQ-CONH2 | D917A | 23 ± 1 |
| H4K20—RKVLRLDNIQ-CONH2 | E947A/Y948F | 40 ± 2.4 |
| H4K20—R19A12–27—KGGAKHRKVLRLDNIQ-CONH2 | 4.6 ± 0.3 (60 ± 4) |
| H4K20—(scrambled)—LNRQDIAGKVRLKRR-CONH2 | E947A-Y948F/D917A | 2.8 ± 0.1 (45 ± 3) |
| H3K4—ARTKQTARKSTGGKA-CONH2 | H3K4me1 | 2.8 ± 0.1 |
| H3K4me2 | 2.2 ± 0.1 |
| H3K9—ARTKQTARKSTGGKA-CONH2 | H3K9 | 85 ± 3 |
| H3K9me1 | 4.7 ± 0.3 (21 ± 0.4) | 1.9 ± 0.1 | 6.7 ± 0.4 | 340 ± 80 |
| H3K9me2 | 15.6 ± 0.9 (36 ± 4) |
| H3K9me3 | 85 ± 10 |
| H3K9—RKs-CONH2 | H3K9me1 | 134 ± 8 (125 ± 6) |
| H3K27—QLATKAARKVSAPATGGV-CONH2 | H3K27 | 40 ± 12 |
| H3K27me1 | 2.9 ± 0.4 |
| H3K27me2 | 12.0 ± 1.0 |
| H3K27me3 | >1000 |
| H3K36—SAPATGGVK35KPHRYRPG-CONH2 | H3K36me1 | 5.0 ± 0.1 |
| H3K36me2 | 3.4 ± 0.1 |
| H3K79—REIAQDFK79TDLRFQS-CONH2 | H3K79me1 | >1000 |
| H3K79me2 | >1000 |

*aValues in parentheses were measured at 150 mM NaCl.

b*KD* = 2.1 ± 0.1 μM for dSfmbt construct K715D/R886S/R900D (residue 535–977) used for co-crystallization.
hydrogen bond with the Nε atom of this arginine (Figure 2).
In the dsfmbt–peptide complex, electron density can be
unambiguously assigned for six of the seven peptide residues
(Figure 2). A peptide surface of 480 Å² contacts dsfmbt,
whereby 40% of the interaction surface is contributed by
the mono-methylated lysine residue.

Contributions of H4K20me1 and dsfmbt residues
to the peptide-binding affinity
We used ITC to evaluate binding of dsfmbt to methylated
histone-tail peptides. First, we tested the binding of dsfmbt-4MBT to 16-residue peptides that were either unmodified,
mono-, di- or tri-methylated at H4K20 (Table II, ITC profiles
are depicted in Supplementary Figure S1). Mono- and
dimethylated H4K20 peptides were bound with 1 and 3 µM
affinity, respectively, whereas unmethylated and tri-methyl-
ated H4K20 peptides were bound with approximately 500-
fold lower affinities (K_D ≈ 1000 µM, Table II). To probe the
contribution of residues flanking the methyl-lysine, we next
tested binding to shorter H4K20me1 peptides. The hepta-
meric peptide used for co-crystallization was bound with an
affinity comparable to the 16-residue peptide. However,
further shortening to a five-residue peptide reduced the
affinity to 23 µM (Table II). This suggests that contributions
provided by residues Arg17 and especially Arg23 that is well
ordered in the crystal structure (Figure 2) are responsible for
the approximately 15-fold higher affinity for the heptamer-
peptide. An even shorter three-residue peptide was bound
with a K_D value of 40 µM (Table II), indicating that His18 and
Leu22, both pointing away from the MBT surface, contribute
little to the binding affinity. The next residue Arg19 directly
adjacent to K20me1 is involved in polar interactions with
dsfmbt (Figure 2) and in the context of the 16-residue
H4K20me1 peptide, mutating Arg19 into alanine reduces
the binding affinity by about four-fold (Table II).

In a complementary set of experiments, we mutated
dsfmbt residues Glu947 and Tyr948 to generate a
dsfmbtE947A/Y948F protein (Figure 1C). Compared with wild-
type dsfmbt, the dsfmbtE947A/Y948F protein bound the 16-
residue H4K20me1 peptide with similar affinity (Table II),
previously because the change from Tyr948 to Phe948 still
permits the π–cation interaction with the guanidinium group
of Arg19. However, mutating the methyl-lysine-contacting
Asp917 into alanine in the single-mutant dsfmbtE947A or

Figure 1 Structure of the four MBT-repeat domain of dsfmbt. (A) Ribbon diagram of the four MBT repeats of dsfmbt coloured in blue
(repeat 1), green (repeat 2), yellow (repeat 3) and red (repeat 4). Histone H4K20me1 peptide is shown in grey. (B) Superposition of the core
folds (grey) of MBT repeats 1–4. For each repeat, helix a2 and the arm regions are coloured according to (A). (C) dsfmbt core domains of repeat
1–4 aligned with core domains of MBT repeats in Drosophila Scm and human L3MBTL1. Positions corresponding to cage-forming residues in
dsfmbt repeat 4, Scm repeat 2 and L3MBTL1 repeat 2 are marked with a red box and conserved cage-forming residues are depicted in red.

dsfmbt residues contacting the H4K20me1 peptide are indicated with an asterisk. dsfmbt, Scm and L3MBTL1 residues important for differential
peptide binding are drawn on red, blue and green background (compare text). N-terminal arm regions are less well conserved and are not
included in the alignment.
triple-mutant dSfmbt E947A/Y948F/D917A proteins completely abolished their ability to bind to H4K20me1 (Table II) without affecting the overall fold and thermal stability of the domain (data not shown). As control, we also tested alanine substitutions of the conserved Asp697 or Asp808 residues at the corresponding positions in the second or third repeat, respectively, (i.e. dSfmbt D697A and dSfmbt D808A) but found that these mutations did not significantly affect peptide binding (Table II).

In summary, these results suggest that dSfmbt binds H4K20me1 with high affinity through the combined interaction of the MBT-binding pocket with the mono-methylated lysine and multiple contacts on the MBT surface with histone residues flanking the methyl-lysine.

**Binding of dSfmbt to other methylated histone peptides**

Despite the high selectivity of dSfmbt in discriminating between different lysine methylation states, it is able to recognize mono- and dimethylated lysine in a broad range of sequence contexts: dSfmbt also binds histone peptides mono- or dimethylated at H3K4, H3K9, H3K27 or H3K36 with affinities ranging between 1 and 16 μM (Table II). Furthermore, a scrambled H4K20me1 peptide is bound with similar affinity as the native H4K20me1 peptide but more negatively charged peptides such as mono- or dimethylated H3K79 peptides (pI 4.4) are bound with an affinity below 1000 μM (Table II). It thus seems that charge complementarity between the positively charged amino acids in histone-tail peptides (pI values 11–12) and the overall negatively charged dSfmbt surface (Figure 3) rather than recognition of individual residues outside the methyl-lysine-binding pocket is important for the interaction. Given the low sequence specificity, we currently cannot exclude that dSfmbt recognizes methyl-lysine residues in other proteins, although so far only interactions between MBT-repeat proteins and mono- and di-methyl-lysine-containing histones have been reported (Kim et al., 2006; Trojer et al., 2007; Wu et al., 2007).

Previous binding studies using fluorescence polarization (FP) assays suggested more pronounced sequence selectivity for dSfmbt binding to H4K20me1/2 and H3K9me1/2 as opposed to binding to H3K4me1/2 or H3K27me1/2 (Klymenko et al., 2006). As our ITC measurements reported here provided little evidence for such binding selectivity, we repeated the binding assays with FP assays. To this end, we used a set of peptides that had been produced during the same synthesis reaction as those used for our ITC measurements but, in addition, had been modified by coupling fluorescent carboxylic acid to the N-terminus in the final synthesis step. In FP assays with these peptides, dSfmbt bound H4K20me1/2, H3K4me1/2, H3K9me1/2, H3K27me1/2 and H3K36me1/2 with comparably low micromolar affinities and the determined $K_D$ values were similar to those measured by ITC (Supplementary Table 1). The failure to detect high-affinity binding of dSfmbt to H3K4me1/2 or H3K27me1/2 by Klymenko et al. (2006) might be because of differences in the method of peptide labelling (i.e. post-synthetic labelling) used in the previous study (W Fischle, personal communication). Taken together, ITC and FP assays reported here both gave comparable results and suggest that mono- and dimethylated lysines in the N-termini of H3 and H4 are all bound with similar micromolar affinities, whereas unmethylated and tri-methylated peptides are bound with much reduced affinity.

**Comparison of the dSfmbt, L3MBTL1 and Scm MBT-repeat domains**

The three MBT repeats of L3MBTL1 can be superimposed onto dSfmbt repeats 2, 3 and 4 (r.m.s.d. $= 3.0$ Å, $Z$-score $= 22.6$) using programme DALI (Holm and Sander, 1993), which identifies repeat 1 as the additional repeat in dSfmbt (Figure 3A). Interestingly, the N-terminal ends of the superimposed L3MBTL1 and dSfmbt structures lie in close vicinity, supporting the hypothesis that repeat 1 of dSfmbt was inserted during evolution. Scm MBT-repeat 1 and 2 can be superimposed onto repeats 1 and 3 of L3MBTL1 (r.m.s.d. $= 2.0$ Å, $Z$-score $= 20.9$) and repeats 2 and 4 of dSfmbt (r.m.s.d. $= 3.4$ Å, $Z$-score $= 17.0$). Therefore, repeat 2 of L3MBTL1 and the homologous repeat 3 of dSfmbt seem as extra features inserted between the two flanking MBT repeats of Scm (Figure 3A). The high r.m.s.d. between dSfmbt repeats 2–4 and L3MBTL1 repeats 1–3 mainly results from the more open arrangement of the three L3MBTL1 repeats, which are arranged around a central channel running along their three-fold pseudo-symmetry axis (Figure 3B). In the crystal structure, this channel is filled with solvent and bound sucrose molecules used as cryoprotectant. However, it could also serve as additional ligand-binding site as it is lined with conserved residues (Figure 3C).

Scm binds mono-methyl-lysine-containing peptides with dissociation constants of approximately 500 μM (Grimm et al., 2007), whereas dSfmbt binds peptides with dissociation constants in the low micromolar range and up to 500 times better than Scm. These differences probably result from their differently charged surfaces (Figure 3B). In Scm, the methyl-lysine-binding pocket is lined by several basic residues (Lys326, Arg352 and His384, Figure 1C), which point towards the positively charged histone-tail peptide. In contrast, the
corresponding dSfmbt residues (Met919, Thr945 and Pro976) are uncharged and assist in peptide binding. In L3MBTL1, the corresponding residues (Met357, Asp383 and Asp415) can also assist in peptide binding, however, the negatively charged area around the methyl-lysine-binding pocket is less extended compared with dSfmbt (Figure 3B), which might explain the lower binding affinity.

Multiple binding sites in MBT-repeat proteins
Superposition of the fourth MBT repeat of dSfmbt with the three other repeats (Figure 4) shows that only the fourth repeat can accommodate methyl-lysine residues. In repeat 1, the crucial aspartate is substituted by an asparagine (Figure 1C), but more importantly the conformation of the loop bearing this residue is different. In the second repeat, two of the cage-forming aromatic residues are substituted by aspartate and serine, respectively, and in the third repeat, Tyr836 blocks the access of the methyl-lysine to the binding pocket. The MBT proteins, Scm and L3MBTL1, use their second MBT repeat for methyl-lysine binding and, indeed, the cage-forming residues, including Cys925 are well conserved in the second MBT repeat of L3MBTL1 and Scm. In contrast, in MBT repeats 1 and 3 of human L3MBTL1, Cys925 is substituted by bulkier residues that block the access to the binding pocket, whereas in Scm the cage-forming aromatic residues are substituted by smaller residues. In dSfmbt and Scm, conserved residues cluster around the methyl-lysine binding pocket, whereas the patch of strictly conserved residues is smaller in L3MBTL1 (Figure 3C).

**Figure 3** Comparison of the MBT-repeat-domain crystal structures of dSfmbt, L3MBTL1 and Scm. (A) Ribbon diagram of dSfmbt (left), L3MBTL1 (middle) and Scm (right). Equivalent MBT repeats as indicated by comparison of their tertiary structures are depicted with equivalent colours. (B) Electrostatic surface representation of dSfmbt, L3MBTL1 and Scm. The bound peptide ligands are depicted in yellow. In L3MBTL1, the methyl-lysine peptide is bound to MBT repeat 2. (C) Comparison of the surface conservation in dSfmbt, L3MBTL1 and Scm. Conserved regions with >50% sequence conservation are depicted in colour, dark green corresponds to strictly conserved residues. For surface comparison, orthologous sequences were aligned as depicted in Supplementary Figure S2. The following sequences were used for the alignments: dSfmbt: *Drosophila melanogaster*, Q9VK33, corresponds to the dSfmbt-4MBT crystal structure reported here; *Anopheles gambiae*, Q7Q0R1; *Xenopus laevis*, P59178; *Homo sapiens*, Q05BQ5; *Gallus gallus*, Q5ZLC2; *Tetraodon nigroviridis*, Q4T9NS. L3MBTL1: *Homo sapiens*, Q9Y468, corresponds to the L3MBTL1 crystal structure (PDB accession code 2RHI); *Bos taurus*, Q08DF3; *Gallus gallus*, XP_417302; *Danio rerio* XP_009604. Scm: *Drosophila melanogaster*, Q9VHA0, as present in the Scm crystal structure (PDB accession code 2R57); *Xenopus tropicalis*, Q0IHT6; *Homo sapiens*, SCML2; *Ciona intestinalis* Q4H2U6.
In conclusion, only a single MBT repeat in Scm, L3MBTL1 and dSfmbt can bind mono- and dimethylated lysine residues. It is possible that the other MBT repeats recognize other ligands. Indeed, in one of the crystal structures of L3MBTL1, the first MBT repeat binds a Pro–Ser-motif-containing peptide of a neighbouring molecule (Li et al., 2007), although the functional relevance of this interaction is not known.

Figure 4 Stereo view of superpositions of the MBT-repeat domains of dSfmbt, L3MBTL1 and Scm. Colour code corresponds to Figure 3 with dSfmbt repeats 1, 2, 3 and 4 depicted in blue, green, yellow, and red (top), L3MBTL1 repeats 1, 2 and 3 in green, yellow and red (middle) and Scm repeats 1 and 2 depicted in green and red (bottom).

dSfmbt and Scm interact functionally to maintain Polycomb repression

Previous structural/functional analyses of the MBT-repeat domain of Scm showed that a point mutation in the methyl-lysine-binding pocket that abolishes the methyl-lysine binding, or even complete deletion of the MBT-repeat domain, still permit these mutant Scm proteins to partially maintain PcG repression of target genes in a genetic-rescue
assay in *D. melanogaster* (Grimm et al., 2007). Similar observations were made with dSftmb; we found that not only the wild-type dSftmb protein but also the dSftmb<sup>E947A/Y948F/D917A</sup> protein (see above) is able to maintain PcG repression of target genes in a genetic-rescue assay in *dSftmb* null mutants (data not shown). One possible explanation for these findings would be that methyl-lysine binding by the MBT domains of dSftmb and Scm has only a minor function in PcG repression. However, as both the proteins have similar methyl-lysine-binding activities, an alternative possibility could be that the MBT-repeat domains in Scm and dSftmb function in a partially redundant manner to maintain PcG repression.

We therefore carried out a set of experiments to test whether and how dSftmb and Scm might interact. First, we analyzed the binding of dSftmb and Scm at PcG target genes in *in vivo*. We recently reported the genome-wide binding profile of dSftmb in developing *Drosophila* larvae (Oktaba et al., 2008). However, chromatin immunoprecipitation (ChiP) assays that monitor the binding of Scm have not yet been reported. We therefore carried out ChiP assays with antibodies against Scm and dSftmb in imaginal-disc tissues from *Drosophila* larvae. These analyses showed that both proteins are specifically bound at PREs of the PcG target genes *Ubx*, *Abd-B*, *en*, *ap*, *Dll*, *eve* and *pnr* (Figure 5). Scm and dSftmb are thus co-bound at PREs in *Drosophila*.

We next tested for the functional redundancy between dSftmb and Scm in the repression of these target genes. To this end, we removed dSftmb function in animals that lack wild-type Scm protein and instead express the MBT-mutant protein Scm<sup>D215N</sup>. Specifically, we induced clones of *dSftmb* null-mutant cells in Scm<sup>D215N</sup> mutant *Drosophila* larvae and analyzed the clones of *dSftmb Scm<sup>D215N</sup>* double-mutant cells for mis-expression of PcG target genes. In the wing imaginal disc, cell clones lacking dSftmb show widespread mis-expression of the PcG target gene *Ubx* (Klymenko et al., 2006), but they do not show mis-expression of *Abd-B* (Figure 6). Similarly, *Abd-B* is not mis-expressed in wing imaginal discs of Scm<sup>D215N</sup>-mutant animals (Figure 6). In striking contrast, *Abd-B* is strongly mis-expressed in clones of dSftmb Scm<sup>D215N</sup> double-mutant cells (Figure 6). A similar strong synergy between these two Polycomb repressor proteins is observed at the *en* gene. In imaginal discs with dSftmb single-mutant clones, *en* is only mis-expressed in a subset of clones in specific regions of the disc but remains repressed in other parts of the disc, and *en* is not mis-expressed in Scm<sup>D215N</sup> single mutants. In contrast, *en* is strongly mis-expressed in clones of dSftmb Scm<sup>D215N</sup> double-mutant cells (Figure 6). In addition, dSftmb Scm<sup>D215N</sup> double-mutant cell clones show a tumour-like phenotype that is characterized by unrestricted cell proliferation (Figure 6). This phenotype is not observed in either of the...
To test whether this strong genetic interaction between dSfmbt and Scm was specific, we used the same strategy to remove the function of the PcG gene calypso (Gaytán de Ayala Alonso et al., 2007) in ScmD215N-mutant larvae. Like in the case of dSfmbt, clones of calypso single-mutant cells in the wing imaginal disc show mis-expression of Ubx (Gaytán de Ayala Alonso et al., 2007) but maintain repression of Abd-B and en (Figure 6). In clones of calypso ScmD215N double-mutant cells, en remains fully repressed, and the clones do not show the tumour-like phenotype observed in dSfmbt ScmD215N double-mutant clones (Figure 6). Abd-B becomes mis-expressed in a fraction of calypso-ScmD215N clone cells but mis-expression is much less extensive than in dSfmbt ScmD215N double-mutant clones (Figure 6). Removal of dSfmbt function in ScmD215N-mutant animals therefore results in a much more severe Polycomb phenotypes compared with when calypso is removed in this genetic background. Taken together, these results suggest a particularly strong synergy between the PhoRC-component dSfmbt and the PcR1-component Scm in the repression of target genes and the control of cell proliferation.
Direct interaction between dSfmbt and Scm proteins

The strong genetic interaction between dSfmbt and Scm prompted us to test whether these two proteins might also physically interact with each other. To this end, we co-expressed Scm and dSfmbt in Sf9 cells using baculovirus and tested whether they form a stable complex, which can be purified from Sf9 cell extracts. As controls, we co-expressed Scm along with the PhoRC-component Pho or with Ph, the PRC1 component that had been reported to interact with Scm (Peterson et al., 1997, 2004). Flag-affinity purification from extracts of Sf9 cells that co-express Flag–Scm and untagged dSfmbt resulted in the isolation of a stable Scm–dSfmbt complex (Figure 7A). Ph also interacted weakly with Scm under the same assay condition but Pho did not form any complex with Scm (Figure 7A).

In the next step, we used C-terminal truncations of Scm and dSfmbt to define the interacting regions between the two proteins with a greater precision. N-terminal Flag-tagged dSfmbt constructs lacking the C-terminal SAM domain and the MBT repeats were still able to interact with Scm (Figure 7B, left panel), and the N-terminal Flag-tagged Scm constructs still interacted with the full-length and C-terminally truncated dSfmbt, also lacking the SAM domain and the MBT repeats (Figure 7B, middle and right panel). Our results identify the N-terminal moieties of dSfmbt and Scm containing Zn-finger motifs as the interacting regions (Figure 7C). Interestingly, interaction between Scm and dSfmbt does not seem to depend on the SAM domains. SAM domains of Scm and Ph form homo-polymeric structures, but are also thought to form Scm–Ph hetero-polymers (Kim et al., 2005). C-terminally truncated Scm lacking the SAM domain does no longer interact with Ph (Supplementary Figure S3), although it still binds to dSfmbt.

Our finding that Scm and dSfmbt can be isolated as a stable complex from Sf9 cells was somewhat unexpected because the biochemically purified PhoRC from Drosophila embryos does not seem to depend on the SAM domains. SAM domains of Scm and Ph form homo-polymeric structures, but are also thought to form Scm–Ph hetero-polymers (Kim et al., 2005). C-terminally truncated Scm lacking the SAM domain does no longer interact with Ph (Supplementary Figure S3), although it still binds to dSfmbt.
Methyl-lysine recognition by PcG protein
C Grimm et al

The EMBO Journal VOL 28 | NO 13 | 2009

1975

does not include Scm and similarly biochemically purified PRC1 contains substoichiometric quantities of Scm but no dSfmbt (Saurin et al, 2001; Klymenko et al, 2006). The failure to isolate dSfmbt–Scm complexes from Drosophila embryonic nuclear extracts may have different reasons. It could be that the dSfmbt–Scm interaction is weaker and becomes disrupted during complex purification. Alternatively, Scm and dSfmbt might interact only under certain conditions (i.e. once both are tethered to chromatin). Taken together, our genetic data, ChIP experiments and physical-interaction data show that dSfmbt and Scm interact directly and cooperate in a highly synergistic manner to maintain Polycomb repression.

Concluding remarks
Our results show how the MBT-repeat domain of dSfmbt binds mono- or dimethyl-lysine-containing histone-tail peptides. The binding affinity of dSfmbt for methylated lysines in the histone H3 and H4 N-termini is in the low micromolar range and is thus comparable to that of heterochromatin protein-1 or the double bromodomain of TAF250 that recognize modified histone lysines in specific sequence contexts (Ruthenburg et al, 2007). However, despite its high selectivity for different states of lysine methylation, dSfmbt-MBT recognizes mono- and dimethylated lysines in various sequence contexts. This broad binding specificity may be important for dSfmbt function within the PhoRC complex. Genome-wide-binding profiling showed that dSfmbt occupies 50% of its targets sites together with Pho, suggesting that dSfmbt is bound to those regions as a part of the PhoRC complex (Oktaba et al, 2008). Previous studies showed that dSfmbt binding at HOX genes crucially depends on Pho-protein-binding sites in PREs, and it is thus the DNA-binding activity of Pho that targets PhoRC to the genes it regulates (Klymenko et al, 2006). Similarly, L3MBTL1 is associated with an E2F–RBF complex (Lewis et al, 2004) and it thus seems likely that the association of L3MBTL1 with E2F target genes (Trojer et al, 2007) is mediated by the DNA-binding factor E2F. Histone methyl-lysine binding by these MBT-repeat proteins thus does not seem to be involved in targeting. Instead, the chromatin environment flanking Pho target sites may dictate which particular mono- and dimethylated lysines are recognized by dSfmbt in vivo.

What is the role of methyl-lysine binding of MBT-repeat proteins? It has been proposed that DNA-tethered MBT proteins use this binding activity for interactions with modified nucleosomes in the flanking chromatin to maintain a repressed-chromatin state (Klymenko et al, 2006; Trojer et al, 2007). The repeat structure of MBT-domain proteins also led to the suggestion that a single MBT-repeat domain could simultaneously recognize several methylation marks (Li et al, 2007; Trojer et al, 2007), which would provide a molecular mechanism for the observed chromatin compaction by L3MBTL1 in vitro (Trojer et al, 2007). However, the structure of the dSfmbt MBT-repeat domain bound to the H4K20me1 peptide and also the structures of L3MBTL1 and Scm bound to methyl-lysine-containing peptides (Grimm et al, 2007; Li et al, 2007; Min et al, 2007; Santiveri et al, 2008) argue against such a model. Only a single methyl-lysine-binding pocket is present in all MBT-repeat proteins, whereas the corresponding ‘pockets’ in the other repeats are shallower and less well conserved. Moreover, there is no biophysical evidence for simultaneous interaction with multiple methylated histone-tail peptides.

The physical and genetic interaction between dSfmbt and Scm suggests a close cooperation of these two proteins in Polycomb repression. Both proteins possess a similar methyl-lysine-binding capacity because of their MBT domains. It is therefore tempting to speculate that dSfmbt–Scm complexes may recognize methylated lysines in two different nucleosomes. Heterodimerization of dSfmbt and Scm with the MBT-repeat domain of each protein bound to a methylated-histone tail could provide a plausible mechanism for chromatin compaction.

Materials and methods
Protein expression and purification
Wild-type and mutant constructs of the four-MBT-repeat domain from D. melanogaster dSfmbt were generated using standard PCR and restriction-cloning techniques and the bacterial expression vector pETM11. The dSfmbt-4MBT protein and all variants were overexpressed in Escherichia coli strain BL21(DE3) as TEV-protease-cleavable N-terminal His6-fusion proteins at 18°C for 15 h. The cleared bacterial lysate in 50 mM Tris–HCl (pH 8.0), 150 mM NaCl, 10% (v/v) glycerol, 10 mM imidazole and 2 mM β-mercaptoethanol was incubated with Ni2+–NTA Sepharose (Qiagen) and the recombinant protein recovered by elution with imidazole followed by incubation with His-tagged TEV protease (0.01% w/w, overnight, 4°C). After dialysis to remove the imidazol, the protease was removed by incubation with Ni2+–NTA Sepharose. The final purification step comprised a gel-filtration step using a Superdex-200 column (GE Healthcare) in a buffer containing 10 mM Tris–HCl (pH 8.0), 150 mM NaCl and 5 mM DTT and protein at concentration of 30 mg/ml.

Crystallographic and data collection
Wild-type dSfmbt-4MBT protein (residues 532–980) was crystalized using the hanging-drop method by mixing 5 μl of protein solution at 50 mg/ml with 5 μl of reservoir solution (0.8 M sodium acetate, 100 mM imidazole, pH 6.5). Crystals were cooled for data collection to 100K in the mother liquor containing 25% (v/v) glycerol as cryoprotectant. Crystals diffraction to 3.2 Å resolution at the ESRF synchrotron and belonged to space-group C222 with three molecules in the asymmetric unit. Three point mutations (K715D, R886S and R900D) were introduced on the surface of a slightly shorter dSfmbt-4MBT construct (residues 535–977). Co-crystals of this mutant construct with peptide RHRKme1VLR were obtained by mixing protein solution at 53 mg/ml in presence of 3 mg/ml peptide with 3.7 M NaCl as the precipitant. These crystals diffraction to 2.8 Å at 100K in the mother liquor containing 35% (w/v) sucrose as cryoprotectant and belonged to space group P2221 with two molecules in the asymmetric unit.

Phase determination and refinement
The structure of wild-type dSfmbt-4MBT protein (residues 532–980) was solved by a two-wavelength MAD experiment in crystal form C222 using a mercury derivative. Four heavy-atom sites were identified using program SOLVE (Terwilliger and Berendzen, 1999). Coordinates for these sites were refined and five more sites were identified using program SHARP (de la Fortelle and Bricogne, 1997). The resulting experimental phases were further improved by solvent flattening and averaging using program DM (Cowtan and Zhang, 1999). In the resulting electron density, the MBT core fold could be located and a partial model could be built. However, the remaining parts of the molecule were disordered and the poor quality of the electron density in these regions prevented us from building a complete model. Molecular replacement was carried out using this partial model and a dataset from the P22 121 crystals at 2.8 Å resolution using program Phaser (McCoy et al, 2008), which yielded a solution for two molecules. The resulting electron density maps allowed us to complete the missing parts of the model and to locate and to build the bound peptide. Several rounds of manual building using program O (Jones and Kjeldgaard, 1997) and
automated refinement using program REFMAC, including TLS refinement (Murshudov et al., 1997) led to a final model with excellent geometry (Table 1).

**ITC and FP measurements**

ITC was carried out using a VP-ITC Microcalorimeter (Microcal, Northampton, MA, USA). Peptides were purified by reverse-phase HPLC in the presence of trifluoroacetic acid. To remove traces of trifluoroacetic acid, all titrations, proteins were dialysed extensively against ITC buffer [20 mM Tris–HCl (pH 8.0), 20 mM or 150 mM NaCl, 2 mM β-mercaptoethanol]. The experiments were carried out at 25°C. A typical titration consisted of injecting 5–10 µl aliquots of 1–5 mM peptide into a solution of 50–200 mM dsfmbt–4MBT protein at time intervals of 5 min to ensure that the titration peak returned to the baseline. The ITC data were analyzed and corrected for the heat of dilution of peptides in the absence of protein using program Origin version 5.0 provided by the manufacturer.

Fluorescein-labelled peptides were synthesized at Protein Specialty Laboratories, Heidelberg. FP assays were carried out at 20 mM Tris–HCl (pH 8.0), 20 mM NaCl, 2 mM β-mercaptoethanol using fluorescein-labelled peptides at an 80 mM concentration on a Synergy 4 instrument (BioTek Instruments). To calculate the K_d values the experimental data were imported and analyzed by program Origin 7.5 as previously described (Jacobs et al., 2004).

**Flag-affinity purification of Scm–dsfmbt complexes**

Baculoviruses expressing full-length Ph, Pho and dsfmbt have been described earlier (Francis et al., 2001; Klymenko et al., 2006). Flag-Scm 877 was a gift from Jeff Simon. The detailed plasmid maps of Scm and dsfmbt constructs used in this study are available on request.

Sf9 cells were co-infected for 48 h with untagged dsfmbt and with different Flag–Scm constructs or with untagged Scm and Flag–dsfmbt construct. The whole-cell extracts were prepared according to Klymenko et al. (2006). 0.2 ml anti-Flag beads (Sigma) were used for 10 ml of extracts. Binding was carried out overnight at 4°C in extraction buffer A (20 mM Tris–HCl (pH 8.0), 300 mM NaCl, 20% (v/v) glycerol, 4 mM MgCl_2, 0.4 mM EDTA and 2 mM DTT) with 0.05% NP40, 0.2 mM protease inhibitors and 0.5 mM DTT) with 0.05% NP40, 0.2 mM protease inhibitors and 0.5 mM DTT. Beads were eluted at 4°C with 0.4 mg/ml Flag peptide in buffer B, containing 300 mM KCl. The supernatant was analyzed by SDS–PAGE followed by Coomassie staining.

**Functional analysis of dsfmbt and Scm in imaginal discs**

Imaginal discs were dissected from third instar larvae that were produced by crossing the appropriate mutant by strains listed below:

| Strain              | Description                          |
|---------------------|--------------------------------------|
| yw hs-flp; hs–nGFP FRT40 |                                      |
| yw hs-flp; [hs–nGFP FRT40; Scm 877/3102]/SM5-TM6B |                                      |
| w; dsfmbt FRT40/SM6B  |                                      |
| w; FRT32 Scm 877/TM6B  |                                      |
| w; [dsfmbt FRT40; FRT2 Scm 877/3102]/SM5-TM6B |                                      |
| yw; FRT40 FRT42D P(lv+) calypso 3/SM6B  |                                      |
| yw hs-flp; [FRT40 FRT42D P(lv+) calypso 3/SM6B 3102]/SM5-TM6B |                                      |

Note, the Scm 877/3102 allele encodes Scm 877DIN.

Clone induction and staining of imaginal discs with antibodies against Abd-B (Celinker et al., 1989) or En (mouse monoclonal 4D9) was done as described earlier (Beuchle et al., 2001).

**Supplementary data**

Supplementary data are available at The EMBO Journal Online (http://www.embojournal.org).

**Acknowledgements**

We thank J Simon for the gift of the Flag–Scm baculovirus expression vector. RM is supported by a grant from the Deutsche Forschungsgemeinschaft. We thank the EMBL–ESRF Joint Structural Biology Group for access and support at the ESRF beamlines. We also acknowledge the support of the crystallization facility of the Partnership for Structural Biology, Grenoble and the proteomic core facility at EMBL, Heidelberg.

**References**

Beuchle D, Struhl G, Muller J (2001) Polycomb group proteins and heritable silencing of Drosophila Hox genes. *Development* **128**: 993–1004

Bornemann D, Miller E, Simon J (1998) Expression and properties of wild-type and mutant forms of the Drosophila sex comb on midleg (SCM) repressor protein. *Genetics* **150**: 675–686

Celinker SE, Reelan DJ, Lewis EB (1989) The molecular genetics of the bithorax complex of Drosophila: characterisation of the products of the Abdominal-B domain. *Genes Dev* **3**: 1424–1436

Cowtan KD, Zhang KY (1999) Density modification for macromolecular phase improvement. *Prog Biophys Mol Biol* **72**: 245–270

Francis NJ, Kingston RE (2001) Mechanisms of transcriptional memory. *Nat Rev Mol Cell Biol* **2**: 409–421

Francis NJ, Saurin AJ, Shao Z, Kingston RE (2001) Reconstitution of a functional core Polycomb repressive complex. *Mol Cell* **8**: 545–556

Gaytán de Ayala Alonso A, Gutiérrez L, Fritsch C, Papp B, Beuchle D, Müller J (2007) A genetic screen identifies novel Polycomb group genes in Drosophila. *Genetics* **176**: 2099–2108

Grimm C, de Ayala Alonso AC, Rybin V, Steuerwald U, Ly-Hartig N, Fischle W, Muller J, Muller CW (2007) Structural and functional analyses of methyl-lysine binding by the malignant brain tumour repeat protein Sex comb on midleg. *EMBO Rep* **8**: 1031–1037

Holm L, Sander C (1993) Protein structure comparison by alignment of distance matrices. *J Mol Biol* **233**: 123–138

Jacob SA, Fischle W, Khorasanizadeh S (2004) Assays for the determination of structure and dynamics of the interaction of the chromodomain with histone peptides. *Methods Enzymol* **376**: 131–148

Jones TA, Kjeldgaard M (1997) Electron-density map interpretation. *Methods Enzymol* **277**: 173–208

Kim CA, Sawaya MR, Cascio D, Kim W, Bowie JU (2005) Structural organization of a Sex-comb-on-midleg/Polyhomeotic cooperator. *J Biol Chem* **280**: 27769–27775

Kim J, Daniel J, Espejo A, Lake A, Krishna M, Xia L, Zhang Y, Bedford MT (2006) Tudor, MRT and chromodomains gauge the degree of lysine methylation. *EMBO Rep* **7**: 397–403

Klymenko T, Papp B, Fischle W, Kocher T, Schelder M, Fritsch C, Wild B, Wilm M, Muller J (2006) A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. *Genes Dev* **20**: 1110–1122

Lewis PW, Beall EL, Fleischer TC, Georlette D, Link AJ, Botchan MR (2004) Identification of a Drosophila Mvb-EZF2/RFB transcriptional repressor complex. *Genes Dev* **18**: 2929–2940

Li H, Fischle W, Wang W, Duncan EM, Liang L, Murakami-Ishibe S, Allis CD, Patel DJ (2007) Structural basis for lower lysine methylation state-specific readout by MRT repeats of L3MBTL1 and an engineered PHD finger. *Mol Cell* **28**: 677–691

McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* **61**: 458–464
Min J, Allali-Hassani A, Nady N, Qi C, Ouyang H, Liu Y, MacKenzie F, Vedadi M, Arrowsmith CH (2007) L3MBTL1 recognition of mono- and dimethylated histones. Nat Struct Mol Biol 14: 1229–1230

Mohd-Sarip A, Cleard F, Mishra RK, Karch F, Verrijzer CP (2005) Synergistic recognition of an epigenetic DNA element by Pleiohomeotic and a Polycomb core complex. Genes Dev 19: 1755–1760

Muller J, Verrijzer CP (2009) Biochemical mechanisms of gene regulation by Polycomb group protein complexes. Curr Opin Genet Dev 19: 150–158

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53: 240–255

Oktaba K, Gutierrez L, Gagneur J, Girardot C, Sengupta AK, Furlong EEM, Muller J (2008) Dynamic regulation by Polycomb group protein complexes controls pattern formation and the cell cycle in Drosophila. Dev Cell 25: 877–889

Peterson AJ, Kyba M, Bornemann D, Morgan K, Brock HW, Simon J (1997) A domain shared by the Polycomb group proteins Scm and ph mediates heterotypic and homotypic interactions. Mol Cell Biol 17: 6683–6692

Peterson AJ, Mallin DR, Francis NJ, Ketel CS, Stamm J, Voeller RK, Kingston RE, Simon JA (2004) Requirement for sex comb on midleg protein interactions in Drosophila Polycomb group repression. Genetics 167: 1225–1239

Ruthenburg AJ, Li H, Patel DJ, Allis CD (2007) Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol 8: 983–994

Santiveri CM, Lechtenberg BC, Allen MD, Sathyamurthy A, Jaulent AM, Freund SM, Bycroft M (2008) The Malignant brain tumor repeats of human SCML2 bind to peptides containing monomethylated lysine. J Mol Biol 382: 1107–1112

Saurin AJ, Shao Z, Erdjument-Bromage H, Tempst P, Kingston RE (2001) A Drosophila Polycomb group complex includes Zeste and dTAFII proteins. Nature 412: 655–660

Schwartz YB, Pirrotta V (2007) Polycomb silencing mechanisms and the management of genomic programmes. Nat Rev Genet 8: 9–22

Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. Acta Crystallogr D Biol Crystallogr 55: 849–861

Trojer P, Li G, Sims III RJ, Vaquero A, Kalakonda N, Bocconi P, Lee D, Erdjument-Bromage H, Tempst P, Nimer SD, Wang YH, Reinberg D (2007) L3MBTL1, a histone-methylation-dependent chromatin lock. Cell 129: 915–928

Wang L, Brown JL, Cao R, Zhang Y, Kassis JA, Jones RS (2004) Hierarchical recruitment of Polycomb group silencing complexes. Mol Cell 14: 637–646

Wang WK, Tereshko V, Bocconi P, MacGrogan D, Nimer SD, Patel DJ (2003) Malignant brain tumor repeats: a three-leaved propeller architecture with ligand/peptide binding pockets. Structure 11: 775–789

Wu S, Trievel RC, Rice JC (2007) Human SFMBT is a transcriptional repressor protein that selectively binds the N-terminal tail of histone H3. FEBS Lett 581: 3289–3296

The EMBO Journal is published by Nature Publishing Group on behalf of European Molecular Biology Organization. This article is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 Licence. [http://creativecommons.org/licenses/by-nc-nd/3.0]