Supporting Information Appendix:

Antigen clasping by two antigen-binding sites of an exceptionally specific antibody for histone methylation

Takamitsu Hattori, Darson Lai, Irina S. Dementieva, Sherwin P. Montaño, Kohei Kurosawa, Yupeng Zheng, Louesa R. Akin, Kalina M. Świst-Rosowska, Adrian T. Grzybowski, Akiko Koide, Krzysztof Krajewski, Brian D. Strahl, Neil L. Kelleher, Alexander J. Ruthenburg and Shohei Koide

Contents

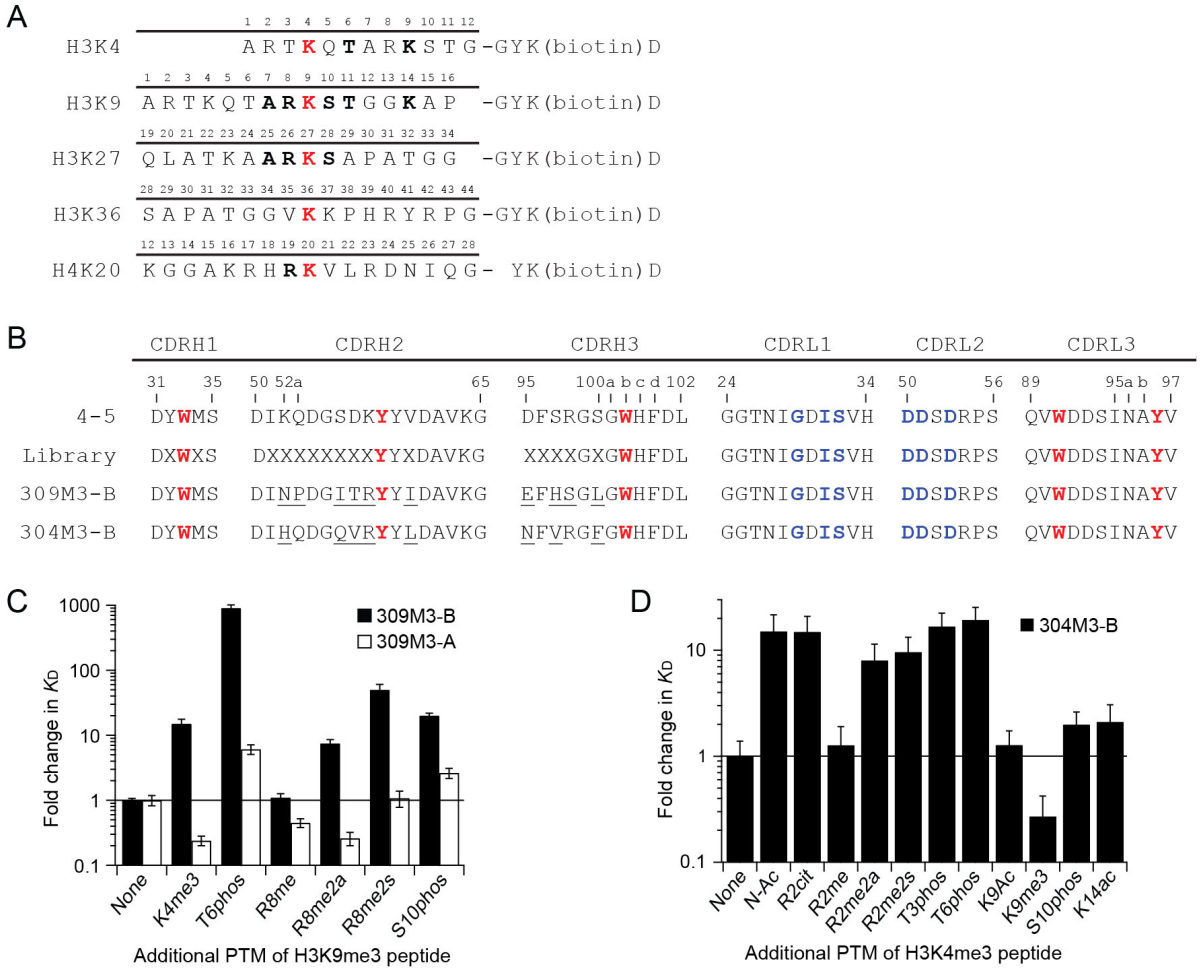
Supplementary figures

Supplementary Tables

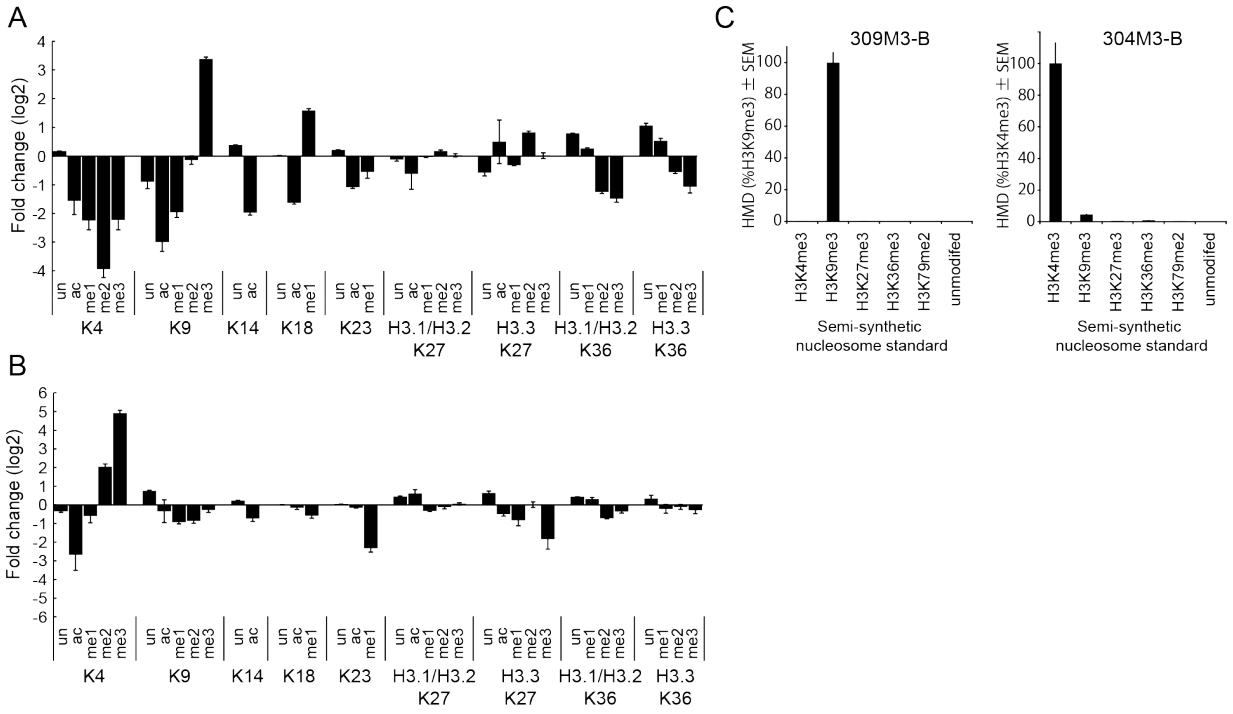
1. Supplementary Table 1. IP followed by mass spectrometry analysis.
2. Supplementary Table 2. Data collection and refinement statistics for the x-ray crystal structures.

Supplementary materials and methods

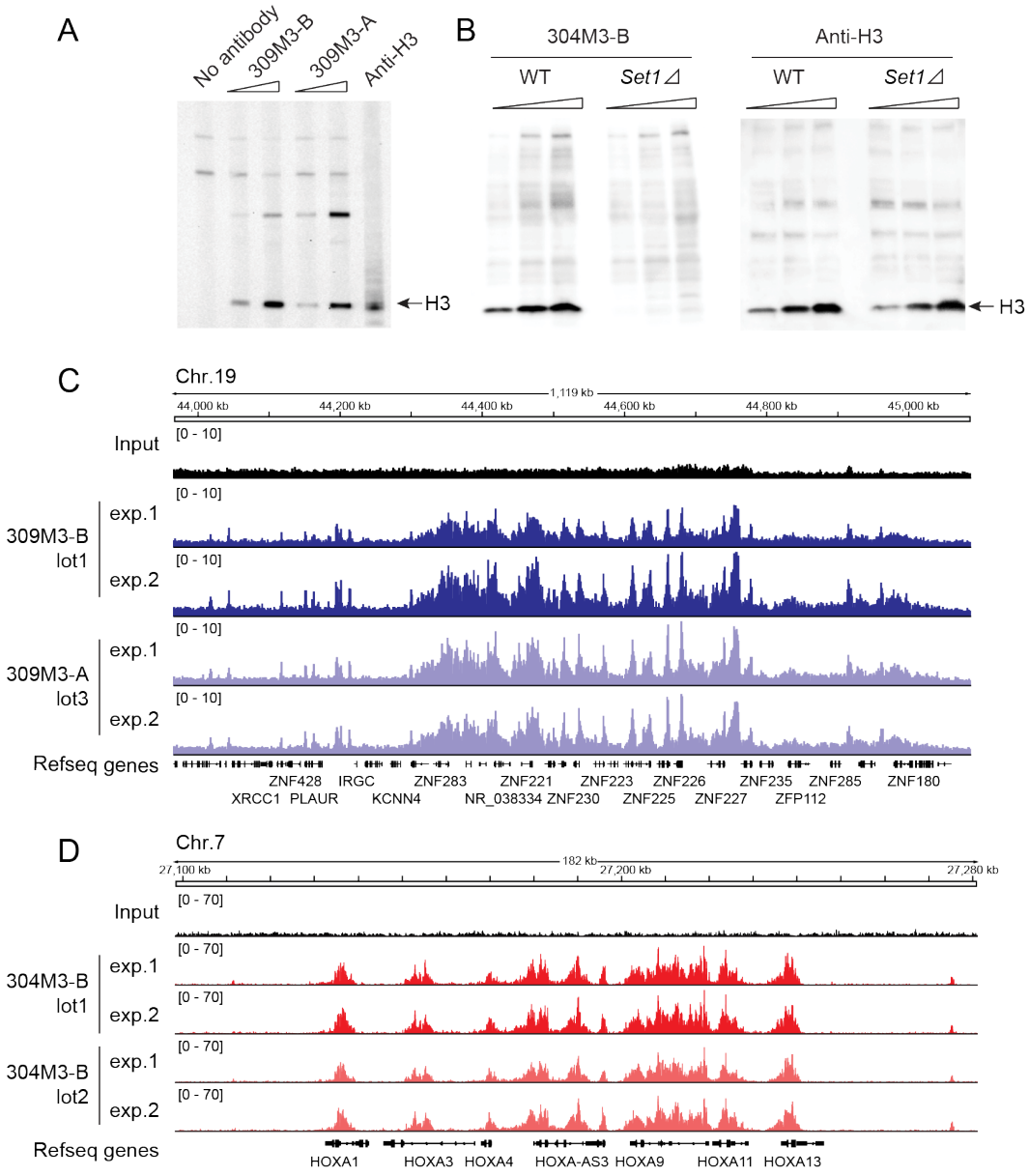
References for supporting information



Supplementary Figure 1. Amino acid sequences of antigen peptides, and CDR sequences and characterization of 309M3-B and 304M3-B. (A) The amino acid sequences of peptides used in this study. The residue numbering in the histone proteins is also shown. The lysine residues harboring PTMs are labeled in red. Conserved residues with respect to those adjacent to H3K9 are in bold. The lysine residue at the GYKD tag is biotinylated and this tag is used for quantification. (B) The CDR sequences of recombinant antibodies. The CDRs were defined and the residues are numbered according to the Kabat scheme (1). The residues that form the aromatic cage and the A_1R_2-binding pocket are labeled in red and blue, respectively. The residues that are different from clone 4-5 are underlined. "X" in the library denotes the randomized amino acid. (C and D) Influence of neighboring modifications on binding affinity. The fold changes in K_D values of 309M3-B to H3K9me3 peptides containing an additional modification (C) and 304M3-B to H3K4me3 peptides (D) are shown. Abbreviations used are: N-Ac, N-terminal acetylation; me2a, asymmetric dimethylation; me2s, symmetric dimethylation; phos, phosphorylation; ac, acetylation; and cit, citrulline substitution. Data for 309M3-A are taken from Hattori *et al* (2013)(2).

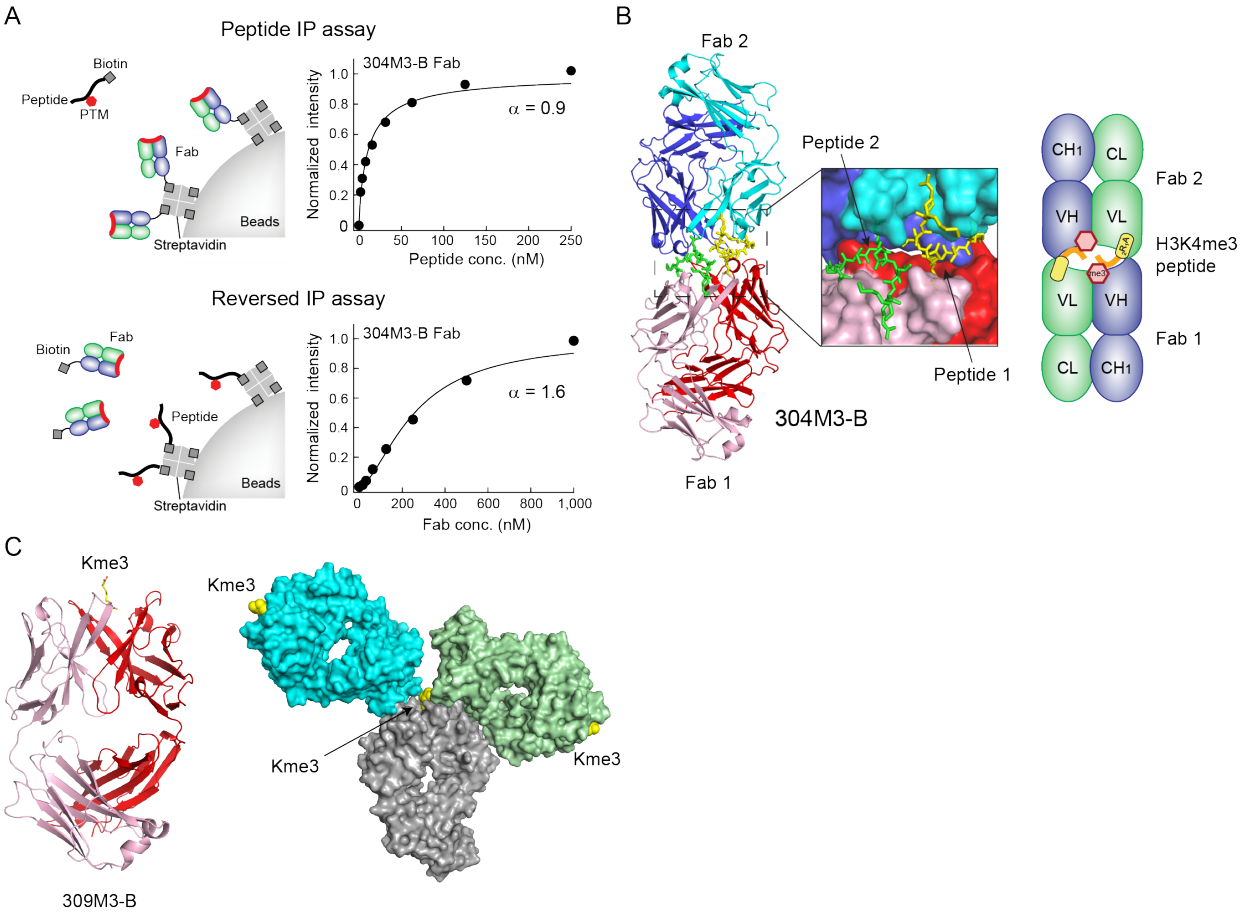


Supplementary Figure 2. Specificity analysis of 309M3-B and 304M3-B by IP followed by mass spectrometry (IP-MS) and internal standard calibrated ChIP (ICeChIP). (A) Changes in the amounts of the indicated modifications in the GluC digested histone mixture from HeLa cells after IP with 309M3-B are shown. Data revealed substantial enrichment of H3K9me3 and no enrichment of unmodified K9, K9ac, K9me1 and K9me2 and the other Kme3 marks on H3. Acetylation at K4, K14, K18 and K23, K4me1/me2/me3, and K36me2/me3 on all H3 histone variants (i.e., H3.1, H3.2 and H3.3), were strongly depleted, suggesting a negative correlation between these marks and H3K9me3 in HeLa H3. Interestingly, H3K18me1 was enriched 3-fold after H3K9me3 IP, suggesting the co-existence of these modifications on the same histone tail. See also Supplementary Table 1. (B) Specificity of 304M3-B tested by IP-MS using proteolytically digested histone mixture. Data are shown in the same manner as (A). IP with 304M3-B resulted in a 30-fold enrichment of H3K4me3 and no enrichment of unmodified K4, K4ac or K4me1. Importantly, no other marks with H3K4me3, including those on different histone proteins, were enriched. We observed enrichment in H3K4me2, but the difference in enrichment between H3K4me3 and H3K4me2 was >7-fold. Interestingly, K23me1 and K27me3 on histone H3.3 variant were substantially depleted after 304M3-B IP, suggesting these PTMs are negatively correlated with H3K4me3 in HeLa cells. See also Supplementary Table 1. All data shown here are from triplicate measurements. (**C**) Native ICeChIP-qPCR analysis with biotinylated 309M3-B (left panel) and 304M3-B (right panel) as described in Grzybowski *et al*. (2015)(3). Relative IP efficiency (%) for each barcoded semi-synthetic nucleosome indicated is normalized to on-target ladder capture; error bars are S.E.M. for 2 and 3 technical replicates, respectively.



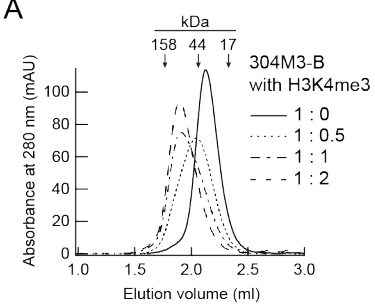
Supplementary Figure 3. Validation of recombinant antibodies in epigenetics applications. (A and B) Western blot analyses of 309M3-B (A) and 304M3-B (B). (A) Whole cell lysates of K562 cells were blotted with 309M3-B (100 and 400 nM) and 309M3-A (25 and 100nM) and the anti-histone H3 pAb. The arrows indicate the location of histone H3. 309M3-B recognized full-length native histone H3. The previously validated 309M3-A antibody is included for comparison. In (A) the upper two bands are due to the secondary reagent, and the central band is due to cross-reactivity to an unidentified cytoplasmic protein (2). (B) Ten, twenty and thirty micro-grams of yeast whole cell extracts from wild-type and *Set*1-deleted strains (lacking H3K4me3) (4) were probed with 304M3-B (left two panels) and anti-histone H3 polyclonal antibody (right two panels). Note that the blots were intentionally over-developed to visualize weak background signals. 304M3-B recognized histone H3 in the context of wild type cells, no signal was detected in cells deleted for *Set*1 – the sole H3K4 methyltransferase in this organism,

strongly suggesting that 304M3-B specifically recognizes native histone H3 harboring K4me3. (C and D) Native ChIP followed by sequencing (ChIP-seq) of HEK293 cells with 309M3-B and 309M3-A (C), and with two different lots of 304M3-B (D). Duplicate experiments with each antibody are shown. The number of reads for Zinc finger regions (*ZNFs*) at chromosome 19 (C) and *HOXA* regions at chromosome 7 (D) are plotted. Native ChIP with 309M3-B showed enrichment for H3K9me3 in specific loci including the 3' ends of a cluster of zinc finger genes on chromosome 19, peak patterns similar to those obtained with 309M3-A that has already been validated in ChIP (2) (C). ChIP-seq of HEK293 cells using 304M3-B showed substantial enrichment in loci including the *HOXA* regions that are marked with H3K4me3(5, 6) (D).



Supplementary Figure 4. Dimerization of 304M3-B triggered by antigen binding and the absence of dimerization of 309M3-B bound to trimethylated Lys. (A) The 304M3-B antibody showed cooperative binding when the orientation of the peptide IP assay was reversed. The curves show the best fit of the Hill equation with the values of the Hill coefficient, α, indicated. Note different concentration ranges on the horizontal axis. Data shown are from triplicate measurements. (B) The overall structures of 304M3-B in complex with the H3K4me3 peptide. 304M3-B forms a head-to-head dimer with the H3K4me3 peptides. The structure is shown in the same manner as Fig. 2B. The H3K4me3 peptide 1 and peptide 2 are shown in yellow and green, respectively. (C) The structure of 309M3-B in complex with trimethylated Lys (Kme3). The left panel shows the overall structure of the 309M3-B//Kme3 complex, in which the heavy and light chains are shown in red and pink, respectively. The right panel shows the crystal packing of the Fab//Kme3 complex, illustrating that this complex does not form a head-to-head dimer in the crystal.

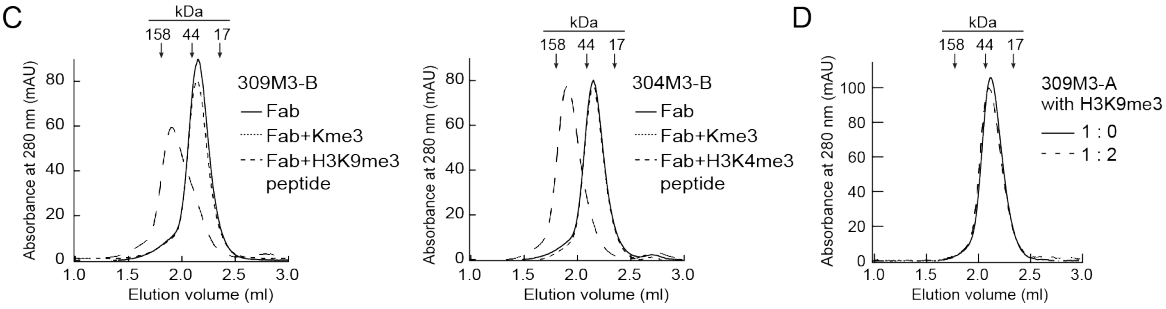
A



B

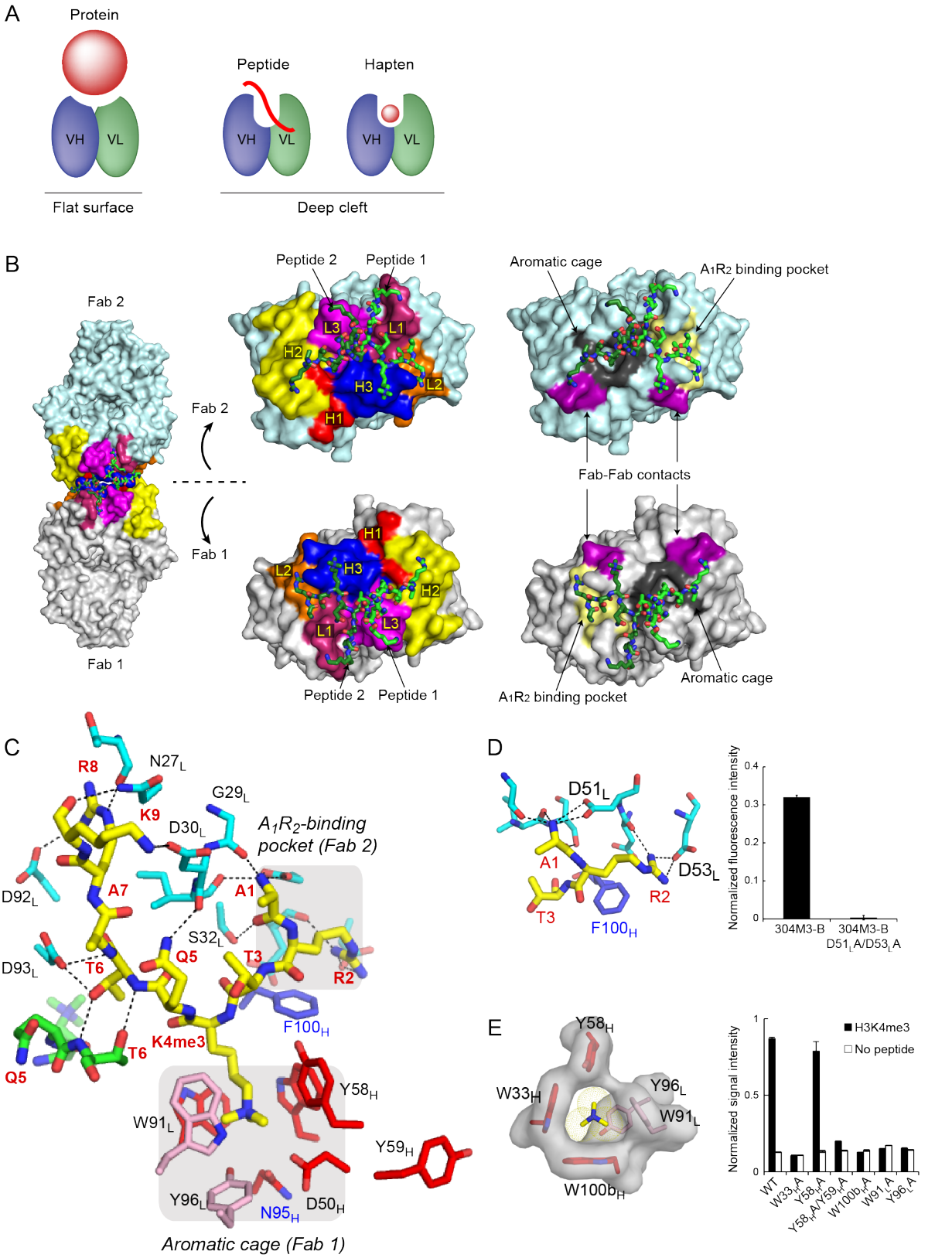
Fab	Peptide	Ratio	Radius (nm)	%Pd	Mw-R (kDa)	%Intensity	%Mass
309M3-B	H3K9me3	1 : 0	3.0 ± 0.02	3.4 ± 2	45 ± 1	99 ± 1	99.9 ± 0.1
309M3-B	H3K9me3	1 : 0.5	4.3 ± 0.1	12 ± 0.1	104 ± 4	98 ± 2	99.8 ± 0.2
309M3-B	H3K9me3	1 : 1	4.4 ± 0.01	12 ± 0.1	110 ± 1	100	100
309M3-B	H3K9me3	1 : 2	4.4 ± 0.02	12 ± 0.1	110 ± 1	97 ± 0.3	100
304M3-B	H3K4me3	1 : 0	3.0 ± 0.1	6.7 ± 1	44 ± 3	98 ± 2	100
304M3-B	H3K4me3	1 : 0.5	4.0 ± 0.1	9.7 ± 2	89 ± 5	98 ± 2	99.8 ± 0.2
304M3-B	H3K4me3	1 : 1	4.4 ± 0.1	12 ± 0.2	110 ± 4	99 ± 1	99.9 ± 0.1
304M3-B	H3K4me3	1 : 2	4.5 ± 0.01	11 ± 0.1	110 ± 1	95 ± 2	99.9 ± 0.1
309M3-A	H3K9me3	1 : 0	3.1 ± 0.1	5.6 ± 3	47 ± 2	95 ± 5	99.9 ± 0.1
309M3-A	H3K9me3	1 : 2	3.2 ± 0.1	11 ± 1	52 ± 3	95 ± 2	100

C

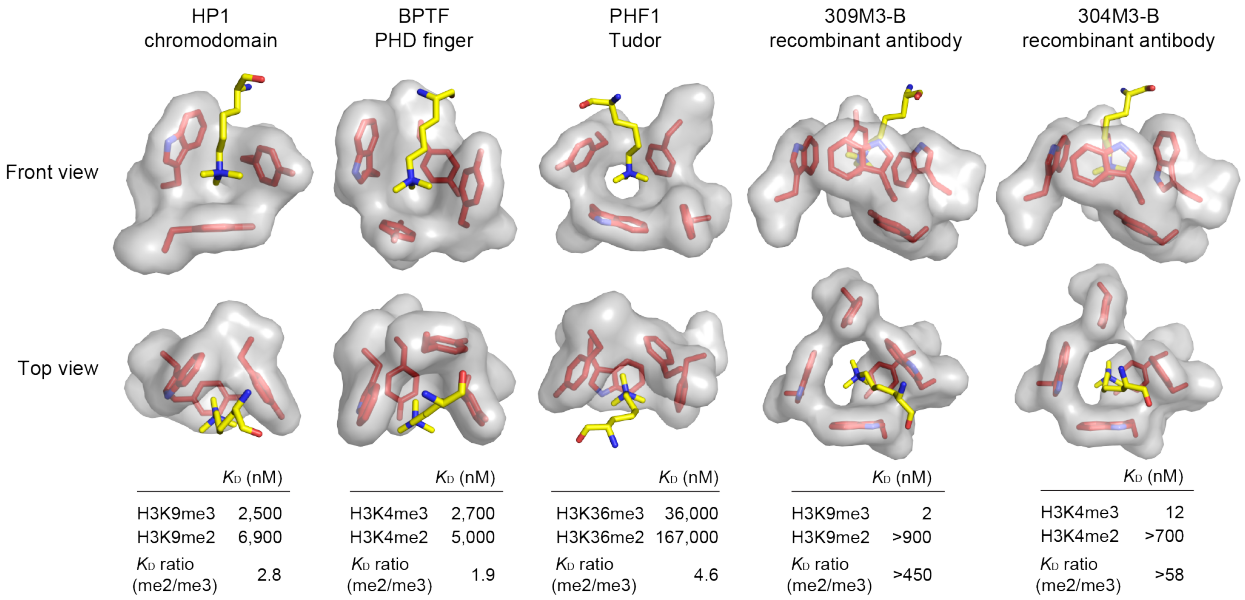


D

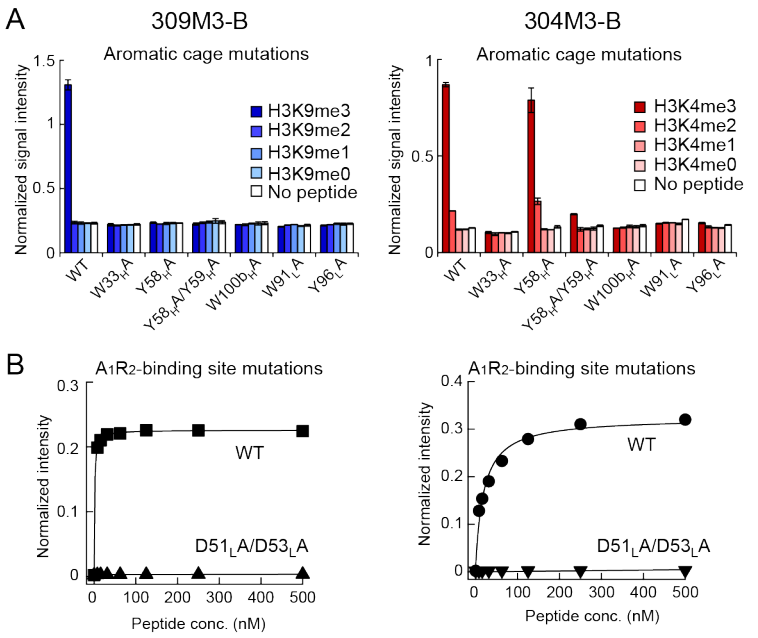
Supplementary Figure 5. Analysis of dimerization of recombinant antibodies. (A**)** Dimerization of 304M3-B in solution analyzed using gel filtration chromatography. The Fab samples were mixed with the peptides with the indicated ratios, and analyzed. Allows indicate the peak positions of calibration proteins (γ-globulin, ovalbumin and myoglobin from left to right). (B) The complete data set for the dynamic light scattering analysis. Pd is the polydispersity and Mw-R is the molecular mass estimated from the measurements. Data shown here are the averages of triplicate measurements with errors indicating S.D. Interestingly, although 309M3-A, another recombinant antibody to H3K9me3, was also generated from the same phage-display library (2), it did not form a dimer with the H3K9me3 peptide. (C) Dimerization of antibody requires peptide antigen. 309M3-B (left panel) and 304M3-B (right panel) were mixed with trimethyl Lys (Kme3) or respective peptide with a molar ratio of 1:2 and analyzed by gel filtration. (D) Gel filtration analysis of 309M3-A in the complex with H3K9me3 peptide, demonstrating that this Fab does not form a dimer upon binding to its target peptide.


A
Protein
VH
VL
Peptide
Hapten
Flat surface
Deep cleft
B
Fab 2
Fab 1
Peptide 2
Peptide 1
L3
L1
H2
H3
L2
H1
Aromatic cage
A1R2 binding pocket
Fab-Fab contacts
C
R8
N27L
G29L
K9
D30L
A1R2-binding pocket (Fab 2)
D92L
A7
A1
S32L
D93L
T6
Q5
T3
R2
F100H
K4me3
Y58H
W91L
Y59H
Y96L
N95H
D50H
Aromatic cage (Fab 1)
D
D51L
A1
D53L
R2
T3
F100H
Normalized fluorescence intensity
304M3-B
304M3-B D51LA/D53LA
E
Y58H
W33H
Y96L
W91L
W100bH
Normalized signal intensity
H3K4me3
No peptide
WT
W33LA
Y58HA
Y58HA/Y59HA
W100bHA
W91LA
Y96LA

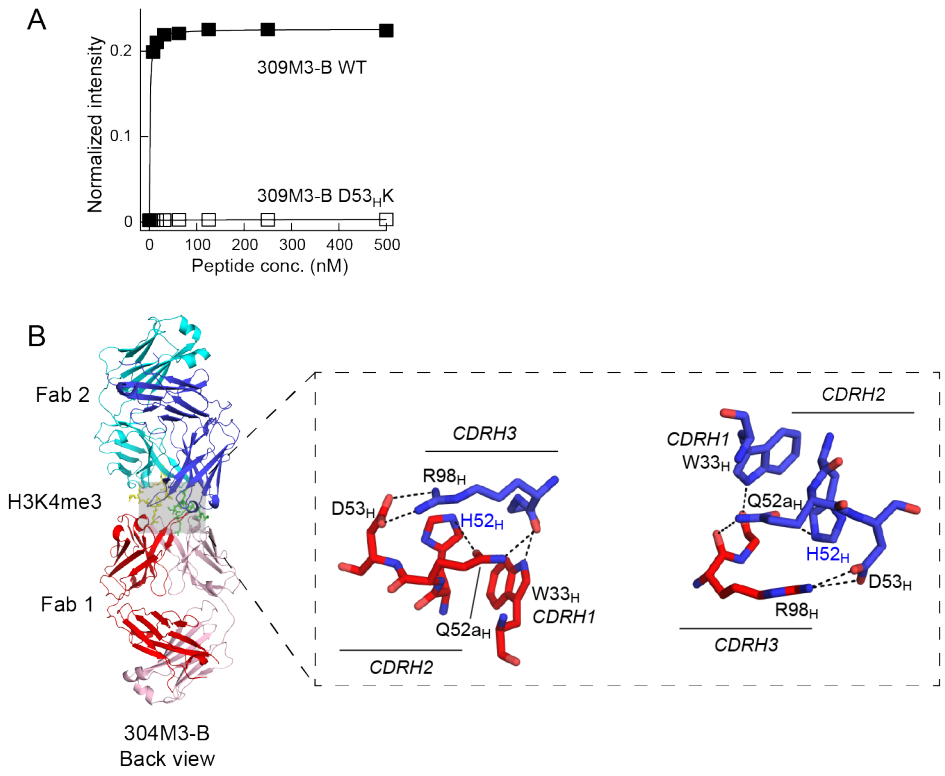
Supplementary Figure 6. The new antibodies form antigen-binding sites with unconventional topography for anti-peptide antibodies. (A) Schematic drawing showing different topographies of the conventional antigen-binding site for recognizing different types of antigens. Antibodies directed to large spherical antigens such as structured proteins generally have flat antigen-binding surfaces (left panel), whereas anti-peptide or anti-hapten antibodies have long CDRs that create a deep cleft (right panels). (B) The structures of 304M3-B in complex with H3K4me3 peptides. The structure is shown in the same manner as Fig. 3A. Total 12 CDRs are involved in the peptide recognition. (C-E) The antibody-peptide interaction interface in the 304M3-B//H3K4me3 peptide complex. The structures and binding analyses are shown in the same manner as Fig. 3B, C and D. In (C), one molecule of the two H3K4me3 peptides bound to the Fab dimer is focused for clarity. Water molecules are omitted here for clarity. All binding data shown here are from triplicate measurements. See also Supplementary Fig. 8 for additional data. In (E), alanine mutations of residues constituting the aromatic cage are shown. Mutations of these residues strongly disrupted binding of the antibodies to trimethylated histone peptides except that $Y58_HA$ of 304M3-B retained binding. $Y58_H$ is the first residue of a YY sequence, and a double mutant, $Y58_HA/Y59_HA$, abolished binding, suggesting that the $Y59_H$ residue somehow substituted the role of $Y58_H$ in the $Y58_HA$ mutant, although $Y59_H$ was not a part of the aromatic cage in wild-type 304M3-B as shown in (C). These results indicate that all of the five aromatic residues were essential for recognizing trimethylated Lys.



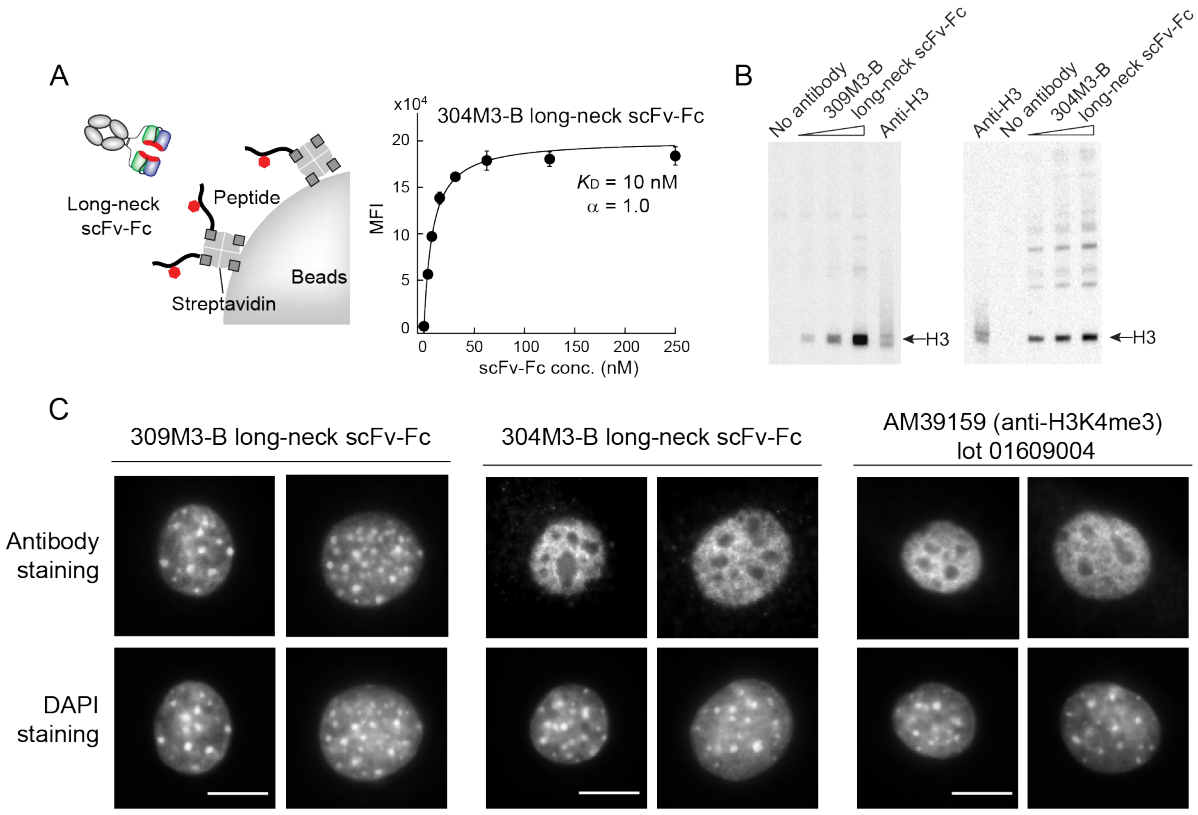
Supplementary Figure 7. Structural comparisons of the aromatic cages in histone reader proteins and recombinant antibodies. Trimethylated Lys residues are shown as stick modes with carbon atoms in yellow. The K_D values of histone reader proteins to indicated peptides were taken from Jacobs *et al* (2002)(7), Li *et al* (2006)(8) and Musselman *et al* (2012)(9). The ratio of K_D values between peptides harboring me3 and me2 is 2~5 folds among natural proteins, whereas 60~450 folds among recombinant antibodies, indicating methylation-state specificity of recombinant antibodies is considerably higher than native proteins. PDB codes: HP1 chromodomain, 1KNE; BPTF PHD finger, 2F6J; PHF1 Tudor, 4HCZ.



Supplementary Figure 8. The aromatic cage and A1R2-inding pocket are essential for antigen binding. (A) The aromatic cage is essential for antigen binding of 309M3-B (left panel) and 304M3-B (right panel), characterized by phage ELISA. The peptides corresponding to 5 nM in the volume of the binding reaction were immobilized in the wells of a microtiter plate coated with Neutravidin for binding analysis, and anti-V5tag monoclonal antibody (SIGMA, lot PC196809) was directly coated on the wells for detecting antibody expression on phages. The binding signals were normalized with respect to the signal for antibody expression. (B) The importance of the A_1R_2-binding site. Effects of mutations in 309M3-B (left panel) and 304M3-B (right panel) were assessed using the peptide IP assay. The binding signal was normalized with respect to the amount of Fab immobilized on beads.



Supplementary Figure 9. Antibody-antibody contacts are obligatory for antigen binding. (A) Mutation in antibody-antibody contacts in 309M3-B abolished the peptide binding, assessed by the peptide IP assay. The binding signal was normalized with respect to the amount of Fab immobilized on beads. Data shown here are from triplicates measurements. Data for 309M3-B WT are the same as Supplementary Fig. 8B, because these data were from the same experiments. (B) Antibody-antibody contacts between two 304M3-B molecules. Polar interactions are marked as dash lines. The color scheme is the same as in Fig. 2B. The residues that are different from clone 4-5 are labeled in blue.



Supplementary Figure 10. The long-neck scFv-Fc antibodies performed well in epigenetics applications. (A) The reversed IP assay of the long-neck scFv-Fc version of 304M3-B. The long-neck scFv-Fc did not show cooperative binding. Compare this result with Supplementary Fig. 4A. Data shown here are from triplicate measurements. (B) The long-neck scFv-Fc versions of 309M3-B and 304M3-B recognized native histone H3 in Western blot analysis. Whole cell lysates of K562 cells were blotted with 309M3-B (0.05, 0.2 and 0.8 nM; left panel) and 304M3-B (10, 20 and 40 nM; right panel) in the long-neck scFv-Fc format, and an anti-histone H3 pAb. The arrows indicate the location of histone H3. (C) Immunofluorescence analysis. NIH 3T3 cells were stained with 50 nM of 309M3-B long-neck scFv-Fc (top panels), 200 nM of 304M3-B long-neck scFv-Fc (bottom left panels) and 1:500 diluted anti-H3K4me3 rabbit polyclonal serum (bottom right panels). Scale bar, 10 μm. Cells were co-stained with fluorochrome 4',6'-diamidino-2-phenylindole (DAPI). Immunofluorescence staining with the 309M3-B long neck scFv-Fc antibody showed a punctate pattern where the foci overlapped with those of the DAPI staining, indicating that 309M3-B long-neck scFv-Fc antibody was concentrated in the pericentric heterochromatin region where H3K9me3 is enriched (2), The staining pattern of 304M3-B long-neck scFv-Fc was highly similar with its from AM39159 (lot 01609004) that has previously been validated in the peptide IP assay (10), suggesting that 304M3-B long-neck scFv-Fc localized in regions where H3K4me3 is enriched.

Supplementary Table 1. IP followed by mass spectrometry analysis, related to Supplementary Fig. 2. Changes in the amounts of the indicated peptides in the Glu-C digested histone mixture from HeLa cells after IP with 309M3-B and 304M3-B are shown. Data shown here are the averages of triplicate measurements with errors indicating S.D.

Peptide	309M3-B Fold change		304M3-B Fold change	
	Average	S.D.	Average	S.D.
H3K4un	1.1	0.003	0.8	0.03
H3K4Ac	0.4	0.1	0.2	0.1
H3K4me1	0.2	0.1	0.8	0.2
H3K4me2	0.1	0.02	4.3	0.5
H3K4me3	0.2	0.1	31	3.6
H3K9unK14Ac	0.3	0.03	0.9	0.1
H3K9unK14un	0.7	0.1	2.0	0.1
H3K9AcK14Ac	0.1	0.01	0.6	0.3
H3K9AcK14un	0.2	0.1	1.1	0.5
H3K9me1K14Ac	0.3	0.04	0.6	0.2
H3K9me1K14un	0.3	0.03	0.5	0.1
H3K9me2K14Ac	0.2	0.03	0.3	0.03
H3K9me2K14un	1.2	0.1	0.7	0.1
H3K9me3K14Ac	1.4	0.2	0.1	0.03
H3K9me3K14un	11	0.5	0.9	0.1
H3K18unK23un	1.2	0.01	1.0	0.004
H3K18me1K23un	3.0	0.2	0.7	0.1
H3K18AcK23un	0.4	0.01	0.8	0.03
H3K18unK23Ac	0.5	0.02	0.9	0.004
H3K18unK23me1	0.7	0.1	0.2	0.03
H3K18AcK23Ac	0.2	0.1	1.2	0.2
H3.1/H3.2 K27unK36un	2.4	0.04	2.3	0.1
H3.1/H3.2 K27me1K36un	3.2	0.003	1.7	0.1
H3.1/H3.2 K27me2K36un	1.2	0.03	1.0	0.03
H3.1/H3.2 K27me3K36un	1.0	0.04	1.0	0.04
H3.1/H3.2 K27AcK36un	2.1	0.8	3.7	1.0
H3.1/H3.2 K27unK36me1	1.4	0.04	2.0	0.2
H3.1/H3.2 K27me1K36me1	1.4	0.01	1.4	0.03
H3.1/H3.2 K27me2K36me1	1.0	0.06	1.0	0.1
H3.1/H3.2 K27me3K36me1	1.3	0.03	1.0	0.1
H3.1/H3.2 K27AcK36me1	0.7	0.3	2.6	1.0
H3.1/H3.2 K27unK36me2	0.4	0.03	1.0	0.03
H3.1/H3.2 K27me1K36me2	0.4	0.02	0.5	0.04
H3.1/H3.2 K27me2K36me2	0.7	0.02	0.3	0.2
H3.1/H3.2 K27me3K36me2	0.6	0.2	0.2	0.1

(Supplementary Table 1, Continued)

Peptide	309M3-B Fold change		304M3-B Fold change	
	Average	S.D.	Average	S.D.
H3.1/H3.2 K27AcK36me2	0.3	0.1	0.5	0.2
H3.1/H3.2 K27unK36me3	0.4	0.05	1.1	0.1
H3.1/H3.2 K27me1K36me3	0.4	0.03	0.6	0.1
H3.1/H3.2 K27me2K36me3	0.5	0.06	0.3	0.1
H3.1/H3.2 K27me3K36me3	0.2	0.1	0.8	0.6
H3.1/H3.2 K27AcK36me3	0.3	0.06	1.0	0.2
H3.3 K27unK36un	3.2	0.3	2.9	0.8
H3.3 K27me1K36un	4.7	1.1	1.5	0.6
H3.3 K27me2K36un	1.8	0.1	1.2	0.1
H3.3 K27me3K36un	1.0	0.1	0.2	0.05
H3.3 K27AcK36un	16	8.8	8.7	0.7
H3.3 K27unK36me1	0.6	0.2	3.1	0.6
H3.3 K27me1K36me1	1.0	0.1	0.5	0.2
H3.3 K27me2K36me1	1.8	0.1	0.6	0.1
H3.3 K27me3K36me1	1.2	0.1	1.2	1.2
H3.3 K27AcK36me1	1.2	0.8	2.1	0.9
H3.3 K27unK36me2	0.7	0.02	1.5	0.2
H3.3 K27me1K36me2	0.7	0.03	0.6	0.1
H3.3 K27me2K36me2	1.2	0.2	0.1	0.02
H3.3 K27me3K36me2	2.2	0.6	0.9	0.7
H3.3 K27AcK36me2	1.1	0.02	0.3	0.1
H3.3 K27unK36me3	0.4	0.1	1.4	0.2
H3.3 K27me1K36me3	0.6	0.1	0.4	0.2
H3.3 K27me2K36me3	0.7	0.3	0.9	0.7
H3.3 K27me3K36me3	0.1	0.04	0.3	0.3
H3.3 K27AcK36me3	1.1	0.4	0.5	0.3
H4K20un	1.3	0.04	1.6	0.1
H4K20me1	0.3	0.2	0.3	0.1
H4K20me2	0.9	0.02	0.7	0.04
H4K20me3	0.4	0.2	0.3	0.1

Supplementary Table 2. Data collection and refinement statistics

	309M3-B//H3K9me3 peptide (4YHP)	309M3-B//Kme3 (4YHY)	304M3-B//H3K4me3 peptide (4YHZ)
Data collection			
Space group	*C*121	$P2_12_12_1$	$P3_121$
Cell dimensions			
a, *b*, *c* (Å)	162.77, 159.89, 128.76	74.98, 100.34, 135.29	130.38, 130.38, 96.37
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 120
Resolution (Å)	50-2.5 (2.54-2.50)[a,b]	50-1.9 (1.93-1.90)[a,b]	50-2.3 (2.34-2.30)[a,b]
R_{merge}	0.065 (0.611)	0.089 (0.587)	0.098 (-)
I / σI	20.6 (2.4)	25.4 (2.5)	20.3 (2.4)
Completeness (%)	99.8 (100)	98.6 (84.8)	100 (100)
Redundancy	3.8 (3.8)	7.2 (5.5)	9.1 (8.7)
Refinement			
Resolution (Å)	32.7-2.51	39.9-1.90	48.7-2.30
No. reflections	111,976	79,566	42,022
R_{work} / R_{free}	0.20/0.24	0.21/0.24	0.18/0.21
No. atoms			
Protein	13,291	6,524	3,361
Ligand/ion	-	-	47
Water	486	604	174
B-factors			
Protein	40.6	26.1	50.4
Ligand/ion	-	-	59.8
Water	39.4	33.4	49.3
R.m.s. deviations			
Bond lengths (Å)	0.005	0.007	0.007
Bond angles (°)	0.87	1.08	1.09
Interface analysis			
Interface area (Å^2)[c]			
Peptide 1-Fab 1	462	-	338
Peptide 1-Fab 2	715	-	649
Peptide 1-Fab (1+2)	1,177	-	987
Fab 1-Fab 2	382	-	529

[a]Values in parentheses are for highest-resolution shell. [b]One crystal for each structure was used for data collection and structure determination. [c]Interface area calculation was performed using PDBePISA server (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html). Interface areas were calculated using peptide 1, Fab 1 and Fab 2 (symmetric unit) in each complex. Ramachandran statistics are described in the X-ray crystallography and structural determination section of Materials and Methods.

Supplementary Materials and Methods

Synthetic peptides

Histone peptides were purchased from Abgent, Anaspec or Genemed Synthesis, or synthesized in-house as described previously (11). Their purity was 95% or higher as judged by reversed-phase chromatography. The design of the peptides is shown in Supplementary Fig. 1A.

In vitro selection, expression and purification of recombinant antibodies

Selection of recombinant antibodies in the single-chain antibody format (scFv) from phage-display libraries was performed as previously described (2). Recombinant antibodies were selected from a custom-made phage-display library, designed based on clone 4-5, by the combined use of positive selection with the target (i.e. the H3K9me3 or H3K4me3 peptide) and negative selection with off-target peptides (i.e. peptides corresponding to other methylation states of the same site and those corresponding to Kme3 at different sites). Clones specifically bind to H3K9me3 or H3K4me3 peptides were identified by phage ELISA.

The identified clones were converted into the Fab format, expressed and purified as described previously (2). The genes for 309M3-B and 304M3-B were reformatted into the Fab format that contains a biotinylation acceptor peptide at the C-terminus of the heavy chain. The Fab clones were expressed in *E. coli* 55244 (ATCC) that co-expressed BirA biotin ligase in culture media containing biotin, and then purified using protein G affinity chromatography followed by cation exchange chromatography. The mutants of 309M-B and 304M3-B were expressed and purified in the same manner.

Cloning, expression and purification of long-neck scFv-Fc and IgG antibodies

The long-neck scFv-Fc antibody was generated by linking the scFv version of 309M3-B and 304M3-B with the Fc region of mouse IgG1 with a 17 amino acid linker, RSSSSSGSSSSGSSGRS. The DNA fragments encoding 309M3-B and 304M3-B scFv connected with Fc region by the linker were cloned into pFUSE-mIgG1-Fc2 vector (InvivoGen). The IgG antibody was generated by fusing mouse Fc region to 309M3-B Fab. The DNA fragment encoding the heavy chain of 309M3-B Fab (VH-CH1) was cloned into pFUSE-mIgG1-Fc2 vector (InvivoGen) and the light chain (VL-CL) was cloned into pFUSE2ss-CLIg-mk vector

(InvivoGen). A His_6-tag was attached at the C-terminus of long-neck scFv-Fc and the C-terminus of CH3 in IgG to facilitate protein purification.

HEK293T cells (ATCC) were grown in DMEM supplied with 10% FBS and antibiotics. Cells were transfected with the plasmid DNA using polyethylenimine (PEI) (SIGMA), and the supernatant was collected after 48 hour incubation. The scFv-Fc and IgG proteins were purified from supernatant using Ni Sepharose 6 Fast Flow (GE Healthcare) according to manufacturer's protocol. The purity was confirmed by SDS-PAGE.

Peptide IP assay

Peptide IP assay was performed essentially as described previously (2, 10). For biotinylated Fab antibodies, antibody-coated beads were prepared by incubating 24 μL of Dynabeads M280 streptavidin (Life Technology) and 0.6 μg of a biotinylated Fab (molar equivalent to 1 μg of IgG antibody) for 1hr at 4 ˚C. Excess biotin-binding sites of streptavidin were blocked with biotin. The antibody-coated beads were used for the peptide IP assay as described previously (2, 10), except that free biotin was added to the binding reaction buffer to prevent undesired binding of the biotinylated peptide to the Fab-immobilized beads. For analysis of mutants of 309M3-B and 304M3-B, equal loading of different Fab samples onto beads were confirmed by staining with Alexa Fluor647-conjugated goat anti-$F(ab')_2$ fragment antibody (Jackson ImmunoResearch, lot 110126), as many of these mutants lost binding to histone peptides. The binding signal was normalized with respect to the amount of Fab immobilized on beads. For the peptide IP assay in the reversed orientation, 3.3 pmol of biotinylated peptide was incubated with 5 μL of Dynabeads M280 streptavidin, and then excess biotin-binding sites of streptavidin were blocked with biotin. The biotinylated Fab or scFv-Fc were titrated into the solution containing peptide-immobilized beads in the presence of excess biotin, and washing and detection were performed as described previously (2, 10). The biotinylated Fab antibodies were detected with Dylight 650 conjugated Streptavidin (Pierce), and scFv-Fc antibodies were detected with Dylight 650 conjugated goat anti-mouse IgG (Pierce, lot NH171405B). The Hill coefficient values were determined for the titrations by nonlinear least-squares fitting in the Igor program (WaveMetrics) of the equation

$$MFI(x) = MFI_0 + \Delta MFI \frac{x^a}{K_D + x^a}$$

where MFI(x) is the median fluorescent intensity (MFI) observed at the peptide (or antibody) concentration x, MFI_0 is the MFI in the absence of a peptide (or antibody), ΔMFI is the difference in the MFI in the absence and presence of the saturating concentration of a peptide (or antibody), α is Hill coefficient and K_D is the apparent dissociation constant (10, 12).

Binding analysis using yeast surface display

Yeast transformation and antibody expression on yeast surface were performed essentially as described previously (13). The DNA encoding aromatic cage mutants of 309M3-B and 304M3-B scFv antibodies were cloned into yeast-display vector, pGalAgaCamR using *Nco*I and *Xho*I sites. The EBY100 yeast strain was transformed with the plasmid DNA and individual clones were confirmed by sequencing. Expression of the scFv antibody was induced by growing the cells in SG-CAA media at room temperature for 48hr.

Binding reaction, washing and detection were performed as described above under "Peptide IP assay". For detecting soluble Fab bound to the displayed scFv antibody (by forming a heterotrimer), the biotinylated $D51_LA/D53_LA$ double-alanine mutant Fab antibodies were mixed with Dylight 650 conjugated Streptavidin (Thermo) at a molar ratio of 4:1 to make Fab-streptavidin complex, and mixed with antibody-expressed yeast cells in the presence or absence of non-biotinylated peptide. Four micro-molar of peptide and 100 nM of Fab-streptavidin complex were used for these assays.

IP followed by MS analysis

IP-MS of synthetic peptide mixtures was performed as follows. Equal molar amounts of the synthetic histone H3 peptides (residue 1-21 for K4 and K9 peptides, residue 21-44 for K27 and K36 peptides) spanning unmodified, Kme1, Kme2 and Kme3 (Anaspec) were mixed to give the total amount of peptide of 9 μg. Acid extracted histones from HeLa S3 cells were prepared and digested with endoproteinase GluC (Roche) as described previously (2, 14). For analysis with natural histone proteins, we proteolytically cleaved acid extracted histone proteins with endoproteinase GluC to produce a mixture of peptides including the histone H3 peptide spanning residues 1-50. We used the cleaved peptide mixture rather than the full-length, denatured histone proteins to minimize nonspecific interactions between histone proteins with the immobilized antibody. The cleaved histone mixture was used as the input for IP without further purification so

that we could detect cross-reactivity to PTMs that occur outside histone H3 residues 1-50, e.g. histone H4. Fab-coated M280 beads were prepared as described above under "Peptide IP assay". Ten micro-gram of antibody-coated beads were incubated with 9 μg of the mixture of synthetic peptides, or with 25 μg of the GluC digested histone mixture for 2 hr at 4 ˚C in the immunoprecipitation buffer (83 mM sodium phosphate (pH 7.2) buffer containing 100 mM KCl, 2 mM $MgCl_2$, 10% v/v glycerol, 0.1% v/v NP-40, 200 μM PMSF and protease inhibitor cocktail (Roche)). The beads were washed as described previously (2), and bound synthetic peptides or GluC digested histones were eluted in 0.2% TFA. Input and speed vacuum concentrated IP products were derivatized using propionic anhydride (SIGMA) and digested with trypsin (Promega) as described previously (15). The processed peptides were analyzed by nano-capillary liquid chromatography (nLC, Dinoex UltiMate 3000) coupled triple quadrupole mass spectrometer (QqQ MS, ThermoFisher Scientific TSQ Quantum) using nLC setup and QqQ instrument parameters reported previously (15). Selective reaction monitoring (SRM) method was used to measure histone peptide carrying various PTMs (15). Acquired MS raw files were analyzed using Skyline software (MacCoss Lab) (16). Integrated chromatogram peak areas from 2 or 3 transitions were used for the quantitation of each peptide. The ratio of quantified peak area from IP experiment against input was reported as % input. Relative level of each modification form was normalized by the sum of all the monitored modification forms from the same peptide sequence. Fold enrichment was then determined by the ratio of relative levels of each modification from input and IP experiment. Data was collected in biological triplicates (IP) and technical triplicates (MS assay).

Internal Standard Calibrated ChIP (ICeChIP) analysis

ICeChIP-qPCR analysis was performed as previously described (3). Native ICeChIP-qPCR was performed using 1.8 μg of biotinylated Fab immobilized on 73 μL of M-280 Streptavidin Dynabeads on a sample of 250 ng of purified human chromatin, from approximately 30,000 cells, in the presence of uniquely barcoded semisynthetic nucleosome standards representing H3K4me3, H3K9me3, H3K27me3, H3K36me3, H3K79me2 and unmodified nucleosome core particles. IP efficiency (%), defined as (IP/input*100%), was determined by qPCR for each barcoded ladder normalized to on-target ladder capture.

Western blot analysis

The W303 (WT) and *Set1Δ* yeast strains were grown in YPD media. Preparation of whole cell extracts (WCE) and western blot analysis were performed as previously described (17). The total protein amount of WCE was determined by Coomassie Plus Protein Assay Reagent (Thermo). Next, 10, 20 and 40 μg of WCE were separated in SDS-PAGE and transferred to a PVDF membrane. The membrane was blocked with PBS (9.57 mM phosphate, 137 mM NaCl and 2.68 mM KCl (pH7.4)) containing 5% skim milk and rinsed with PBS containing 0.1% Tween20 (PBST). The biotinylated 304M3-B was incubated with horseradish peroxidase (HRP) conjugated neutravidin (Thermo) at a molar ratio of 4:1 to make antibody-neutravidin HRP complex prior to blot. Biotin was added into antibody-neutravidin HRP complex to block remaining biotin-binding sites. The membrane was probed with 5 nM of the antibody-neutravidin HRP complex in PBST containing 1% BSA, washed with PBST and detected, or probed with anti-histone H3 polyclonal antibody (Abcam Ab1791, lot GR64775-1) and detected with goat anti-rabbit IgG-HRP (Thermo).

The Western blot analysis of the whole cell lysate of K562 cells with the 309M3-B Fab or the long-neck scFv-Fc antibodies were performed essentially as described previously (2). For detection of the long-neck scFv-Fc antibodies, HRP conjugated anti-mouse IgG Fc (Pierce, lot PI19091913) was used.

Immunofluorescence analysis

The NIH 3T3 cells were grown in DMEM (Cellgro) media supplied with 10% FBS, 2 mM L-glutamine and antibiotics. Immunofluorescence analysis was performed as described previously (2). Anti-H3K4me3 rabbit polyclonal serum (AM39159, lot 01609004), which was validated in peptide IP assay previously (10), was used for a positive control for H3K4me3 staining. The long-neck scFv-Fc antibodies were detected with Dylight 650 conjugated goat anti-mouse IgG (Pierce, lot NH171405B), and the rabbit polyclonal antibody was detected with Dylight650 conjugated anti-rabbit polyclonal antibody (Pierce, lot ND170165).

Native ChIP followed by sequencing (ChIP-Seq)

The HEK293 cells were grown in the DMEM media (Thermo) with 10% FBS and antibiotics. Native ChIP experiments using HEK293 cells were performed as previously

described (2, 18), except that antibody-coated beads were prepared by incubating Dynabeads M280 streptavidin (Life technology) with a recombinant antibody. Briefly, 3.3 μg of Fab (equivalent to 5 μg of IgG) was immobilized on the M280 beads as described above under "peptide IP assay", and the beads were incubated with 10 μg of purified nucleosomes at 4˚C for overnight. Washing, elution and DNA purification were performed as previously described (2).

Library preparation for sequencing was performed using TruSeq ChIP sample preparation Kit, Set A (Illumina) according to the manufacturer's protocol. Four DNA libraries were multiplexed and analyzed by Illumina HiSeq 2500 according to the manufacturer's standard protocol.

Data analysis was facilitated using Galaxy platform (19, 20). Integrity of the data was verified with FASTQ Groomer (21). Sequences were aligned to the UCSC human genome assembly HG19 using Bowtie2 at 'sensitive' setting (22). Subsequently unmapped reads, as well as reads with mapQ score lower than 13 (p-value 0.05) were removed from analysis. Genome coverage was calculated using BEDTools (23). Peaks were called using MACS14 with following settings: (Effective genome size 2.7e9; tag size 50; band width 150; p-value cutoff 1e-7; MFOLD range 0,30). Peaks for dataset to be compared were concatenated, merged and clustered (maximum distance between intervals 150; minimum cluster per interval 1). Subsequently mean genomic coverage over each peak for compared datasets were calculated using Java Genomics Toolkit (www.palpant.us/java-genomics-toolkit/). These values were corrected for 1M reads of sequencing for a given dataset. Outliers with values above three standard deviations (3SD) were removed from analysis. Correlation was tested using Pearson product-moment correlation coefficient. Figures were generated using Integrative Genomics Viewer (IGV) (24).

Dimerization analysis by gel filtration and dynamic light scattering

The Fab antibodies at the final concentration of 10 μM was mixed with the synthetic peptides in PBS, and incubated for 30 min at room temperature. The antibody-peptide mixture was analyzed using a Superdex 200 5/150 GL column (GE Healthcare) in PBS at a flow rate of 0.3 ml/min on an AKTA Pure system (GE Healthcare). Dynamic light scattering measurements were performed using DynaPro NanoStar (Wyatt Technology) according to the manufacturer's protocol, and data were analyzed using the DYNAMICS software (Wyatt Technology).

X-ray crystallography and structural determination

For crystallization, the antibodies were produced as the Fab format without the biotinylation acceptor peptide. Fab was expressed and purified using a Protein G column as described above, followed by purification using a HiLoad 16/160 Superdex 75 pg column (GE Healthcare) in 20 mM HEPES (pH 7.8) buffer containing 150 mM NaCl. 309M3-B was concentrated to ~8 mg/ml and mixed with Nε, Nε, Nε-Trimethyl lysine hydrochloride (SIGMA) in a 1:5 molar ratio, or a synthetic H3K9me3 peptide corresponding to residues 1-16, (Abgent) in a 1:2 molar ratio. The complex of 309M3-B with Kme3 was crystalized in 24% PEG 3350, 10% Tacsimate (pH7.0), 10 mM EDTA, 1% MPD and 10 mM Sarcosine. The complex of 309M3-B with H3K9me3 peptide was crystallized in 22% PEG 3350, 8% Tacsimate (pH8.0) and 3% Xylitol. Both complexes were crystalized at 19 ˚C using hanging-drop vapor-diffusion method. These crystals were cryoprotected in the presence of 17% ethylene glycol and flash-frozen in liquid nitrogen. 304M3-B was concentrated to ~14 mg/ml and mixed with a synthetic H3K4me3 peptide corresponding to residues 1-12 (Abgnet) in a 1:2 molar ratio, and crystallized in 20% glycerol, 1.8 M ammonium sulfate and 0.1 M HEPES (pH 7.5) at 19 ˚C using the hanging-drop vapor-diffusion method. The crystals were flash-frozen in liquid nitrogen without using additional cryoprotectant.

X-ray diffraction data were collected at 100K using the 19ID-D (for the 309M3-B//Kme3 complex (wavelength 0.97899 Å) and the 304M3-B//H3K4me3 peptide complex (wavelength 0.97915 Å)) and 19BM (for the 309M3-B//H3K9me3 peptide complex (wavelength 0.97933 Å)) beam lines of the Advanced Photon Source (Argonne National Laboratory). Data were processed with HKL2000 or HKL3000 (25, 26). The structures were determined by molecular replacement using MOLREP in the CCP4 program suite (27). To determine the structure of 309M3-B, we first built a chimera Fab model using 3 different coordinates (PDB ID: 2R8S for CH and CL, 2DD8 for VL and 1MHP for VH) (28-30), and then used it as a search model. The structure of 309M3-B was next used as a search model for the determination of the 304M3-B structure. The coordinate of trimethylated Lys was taken from the PDB entry, 2H23 (31). The H3K9me3 and H3K4me3 peptides were manually built into electron density maps using Coot (32). Manual model fitting, solvent addition and refinement of the structures were performed iteratively using Coot and Phenix.refine (33). Data collection and refinement statistics are listed in **Supplementary Table 2**. In the final models of 309M3-B//H3K9me3 peptide complex , 309M3-

B//Kme3 complex and 304M3-B//H3K4me3 peptide complex, 96.3% , 98.7% and 96.1% of the resides were in favored regions of the Ramachandran plot with 3.3% , 1.3% and 3.9% in additional allowed regions, respectively. Molecular graphics were generated using Pymol (www.pymol.org).

References in supporting information

1. Wu TT & Kabat EA (1970) An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J Exp Med* 132:211-250.
2. Hattori T*, et al.* (2013) Recombinant antibodies to histone post-translational modifications. *Nat Methods* 10:992-995.
3. Grzybowski AT, Chen Z, & Ruthenburg AJ (2015) Calibrating ChIP-Seq with Nucleosomal Internal Standards to Measure Histone Modification Density Genome Wide. *Mol Cell* 58:886-899.
4. Briggs SD*, et al.* (2001) Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. *Genes Dev* 15:3286-3295.
5. Wysocka J*, et al.* (2005) WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* 121:859-872.
6. Wysocka J*, et al.* (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* 442:86-90.
7. Jacobs SA & Khorasanizadeh S (2002) Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* 295:2080-2083.
8. Li H*, et al.* (2006) Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* 442:91-95.
9. Musselman CA*, et al.* (2012) Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. *Nat Struct Mol Biol* 19:1266-1272.
10. Nishikori S*, et al.* (2012) Broad ranges of affinity and specificity of anti-histone antibodies revealed by a quantitative Peptide immunoprecipitation assay. *J Mol Biol* 424:391-399.
11. Fuchs SM, Krajewski K, Baker RW, Miller VL, & Strahl BD (2011) Influence of combinatorial histone modifications on antibody and effector protein recognition. *Curr Biol* 21:53-58.
12. Goutelle S*, et al.* (2008) The Hill equation: a review of its capabilities in pharmacological modelling. *Fundam Clin Pharmacol* 22:633-648.
13. Koide A, Wojcik J, Gilbreth RN, Hoey RJ, & Koide S (2012) Teaching an old scaffold new tricks: monobodies constructed using alternative surfaces of the FN3 scaffold. *J Mol Biol* 415:393-405.
14. Zheng Y, Tipton JD, Thomas PM, Kelleher NL, & Sweet SM (2014) Site-specific human histone H3 methylation stability: fast K4me3 turnover. *Proteomics* 14:2190-2199.

15. Zheng Y, Thomas PM, & Kelleher NL (2013) Measurement of acetylation turnover at distinct lysines in human histones identifies long-lived acetylation sites. *Nat Commun* 4:2203.
16. MacLean B*, et al.* (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26:966-968.
17. Kizer KO, Xiao T, & Strahl BD (2006) Accelerated nuclei preparation and methods for analysis of histone modifications in yeast. *Methods* 40:296-302.
18. Brand M, Rampalli S, Chaturvedi CP, & Dilworth FJ (2008) Analysis of epigenetic modifications of chromatin at specific gene loci by native chromatin immunoprecipitation of nucleosomes isolated using hydroxyapatite chromatography. *Nat Protoc* 3:398-409.
19. Giardine B*, et al.* (2005) Galaxy: a platform for interactive large-scale genome analysis. *Genome Res* 15:1451-1455.
20. Goecks J, Nekrutenko A, Taylor J, & Galaxy T (2010) Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* 11:R86.
21. Blankenberg D*, et al.* (2010) Manipulation of FASTQ data with Galaxy. *Bioinformatics* 26:1783-1785.
22. Langmead B, Trapnell C, Pop M, & Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25.
23. Quinlan AR & Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841-842.
24. Robinson JT*, et al.* (2011) Integrative genomics viewer. *Nat Biotechnol* 29:24-26.
25. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A* 276:307-326.
26. Minor W, Cymborowski M, Otwinowski Z, & Chruszcz M (2006) HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr* 62:859-866.
27. Vagin A & Teplyakov A (2000) An approach to multi-copy search in molecular replacement. *Acta Crystallographica Section D-Biological Crystallography* 56:1622-1624.
28. Ye JD*, et al.* (2008) Synthetic antibodies for specific recognition and crystallization of structured RNA. *Proc Natl Acad Sci U S A* 105:82-87.
29. Prabakaran P*, et al.* (2006) Structure of severe acute respiratory syndrome coronavirus receptor-binding domain complexed with neutralizing antibody. *J Biol Chem* 281:15829-15836.
30. Karpusas M*, et al.* (2003) Crystal structure of the alpha1beta1 integrin I domain in complex with an antibody Fab fragment. *J Mol Biol* 327:1031-1041.
31. Couture JF, Hauk G, Thompson MJ, Blackburn GM, & Trievel RC (2006) Catalytic roles for carbon-oxygen hydrogen bonding in SET domain lysine methyltransferases. *J Biol Chem* 281:19280-19287.
32. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126-2132.
33. Adams PD*, et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213-221.