Specificity of protein lysine methyltransferases and methods for detection of lysine methylation of non-histone proteins

Philipp Rathert,[a] Arunkumar Dhayalan,[a] Huimin Ma[ab] and Albert Jeltsch*[a]

DOI: 10.1039/b811673c

Post translational modification of histone proteins including lysine methylation is an important epigenetic mark, essential for gene regulation and development. Recently, several examples of lysine methylation of non-histone proteins have been discovered suggesting that this is a common post-translational modification for regulation of protein activity. Here, we review assays for the detection of protein methylation based on mass spectrometry, radiolabel and immunological approaches using protein and peptide substrates including application of SPOT peptide arrays. Candidates for new methylation targets of protein methyltransferases can be predicted using the specificity of the enzyme and protein interaction data.

Epigenetic regulation of gene expression by covalent modification of histone proteins and methylation of DNA plays an important role during development and disease processes (reviews 1–4). Post-translational modifications of histone proteins include acetylation, phosphorylation, methylation, ubiquitylation and sumoylation (reviews 5,6). Many of these modifications occur on the N-terminal tails of the histone proteins that protrude from the nucleosomes. Methylation of lysine residues in histone tails has been identified in histone H3 K4, K9, K27, and K36, histone H4 K20, and histone H1b K26 which all have different biological functions (reviews 6,7). For example, H3K9 methylation leads to condensation of the chromatin and inhibition of gene expression. Similarly, H3K27 methylation has been linked to gene silencing, whereas H3K4 methylation is marking active chromatin (review 8).

In 2000, Rea *et al.* identified Suv39H1 which was the first histone lysine methyltransferase (HKMT).[9] Today about 30 different enzymes of this class are known in different species (review 10). Most HKMTs contain a SET domain of approximately 130 amino acids as catalytic part and use the coenzyme *S*-adenosyl-L-methionine (AdoMet) as the donor for an activated methyl group (review 11). HKMTs not only differ in their substrate specificity but also in product pattern, because one, two or three methyl groups can be added to a single lysine residue, which can have different biological meanings (review 11).

The first non-histone substrate identified for an HKMT was the tumor suppressor p53 which is methylated by Set7/9, a SET domain containing enzyme initially identified as histone H3 K4 methyltransferase.[12] Methylation of p53 by Set7/9 at a specific residue within the C-terminal regulatory domain activates p53,[12] because it is required for the binding of the acetyltransferase Tip60 and for the subsequent acetylation of p53.[13] In contrast, methylation of p53 at two other lysine residues by SMYD2 or Set8 leads to the inactivation of p53.[14,15] The transcription factor TAF10 was the next non-histone protein found to be methylated by Set7/9.[16,17] Methylated TAF10 has an increased affinity for RNA polymerase II, pointing to a direct role of this modification in pre-initiation complex formation.[16] The *S. cerevisia* COMPASS complex methylates a kinetochore component called Dam1 during chromosome segregation.[18] In addition, amongst others, cytochrome C and Rubisco are methylated at lysine residues as well (reviews 19,20). Recently it was demonstrated, that the histone methyltransferase G9a (which was previously shown to methylate Histone H3 lysine 9 and 27[21]) catalyses automethylation and that it acts on a variety of different non-histone proteins.[22,23] The new targets include not only transcription factors, but also enzymes, which themselves are involved in transcriptional regulation by changing the epigenetic information.[23] The modification of non-histone proteins by the addition of methyl groups to lysines was demonstrated to have two different possible downstream signaling pathways. The first one has been shown for the G9a automethylation and the methylation of the transcription factors WIZ and ACINUS[22,23] where the post-translational modification of the protein attracts reading domains that specifically bind the methylated lysine similarly as established for reading the histone methylation code (review 24). Therefore, the methylation could direct the non-histone proteins to specific sites inside the cell. The second mechanism was demonstrated for the CDYL1 histone acetyltransferase:[23] the methylation of CDYL1 affects the binding of its chromodomain to H3K9 methylated peptides, because the methylated internal peptide competes with the external methylated H3K9 peptide for binding to the chromodomain.

[a] *Biochemistry Laboratory, School of Engineering and Science, Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany. E-mail: a.jeltsch@jacobs-university.de*
[b] *State Key Laboratory of Organic Geochemistry, Guangzhou Research Center of Mass Spectrometry, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou, 510640, China*

It is conceivable that, in addition to the HKMTs mentioned so far, other enzymes of this class might also have non-histone protein targets, which needs to be demonstrated, suggesting that protein lysine methyltransferase (PKMT) might be the more appropriate general name for enzymes of this family. In addition, a first example of a lysine demethylase acting on non-histone targets was recently discovered[25] indicating that regulation of proteins by lysine methylation is much more complex than it was believed so far.

Analysis of peptide methylation *in vitro*

Different assays exist to determine the activity of HKMTs using synthetic peptide substrates. The turnover of the coenzyme can be followed using a coupled fluorescent assay, where the accumulation of the methyl donor product *S*-adenosy-L-homocysteine during the methylation reaction is detected.[26] The *S*-adenosy-L-homocysteine is enzymatically hydrolyzed to homocysteine and adenosine and the homocysteine concentration is then determined by conjugation of its free sulfhydryl moiety to a fluorophore.

Alternative assays employ a coenzyme carrying a radioactively labeled methyl group and directly detect the transfer of methyl groups to peptide or protein substrates. These assays require the separation of methylated peptides and unreacted AdoMet which can be achieved by filtration through Hybond-C membranes[27] or spotting on P81 cellulose[28] and detection of the radioacitivity by scintillation counting. In the biotin/avidin microplate peptide methylation assay,[29] a synthetic peptide containing a biotin group at the N-terminus is incubated with a histone methyltransferase in the presence of radiolabeled AdoMet. For separation, avidin coated micro plates are used. After binding of the radiolabeled product the excess of the unreacted AdoMet is washed off. Subsequently, the peptide is released and the radioactivity is quantified by scintillation counting. This assay is robust, sensitive and yields quantitative results. Alternatively, streptavidin magnetic beads were used for the separation of the methylated peptide (carrying a biotin tag) from unreacted AdoMet.[28] Methylation of covalently bound peptide in microplates followed by antibody detection of the methylation reaction, recently has been applied in a high-troughput screen to identify novel HKMT inhibitors.[30]

Recently, a continuous variant of the radioactive microplate assay has been described,[31] which makes use of Streptavidin coated microplates (FlashPlate™, Perkin Elmer). In these plates, the interior of each well is permanently coated with a thin layer of polystyrene-based scintillant followed by covalent binding of streptavidin molecules. Using a biotin tag, the peptide is immobilized at the surface of the plate. After removing the unbound peptide, the enzyme is added in buffer containing radioactively labeled AdoMet. Due to the short range of the β-particles emitted by tritium, this does not lead to a strong scintillation signal, because most β-particles are quenched by solvent before they reach the scintillant bound to wall of the plate. After enzymatic transfer of the radioactive methyl groups to the peptide substrates, they closely approach the walls of the microplate which leads to a strong scintillation signal. Due to a small number of handling steps, this assay is well suited for high throughput applications. To demonstrate the principle feasibility of the assay, the Dim-5 histone lysine methyltransferase from *Neurospora crassa* was used. After adding enzyme, radioactive cofactor and buffer into FlashPlates coated with peptide, a rapid and very strong increase in scintillation signal was observed. No signal change was detectable in the absence of enzyme or peptide (Fig. 1A).

To study the sequence specificity of PKMTs in more detail, peptide arrays prepared by SPOT peptide synthesis were applied.[23,32] SPOT peptide synthesis (review 33) uses the standard Fmoc-chemistry based solid phase peptide synthesis with the only exception that a cellulose membrane is used as solid support instead of the generally employed resins. With this method, it is possible to synthesize large peptide arrays that

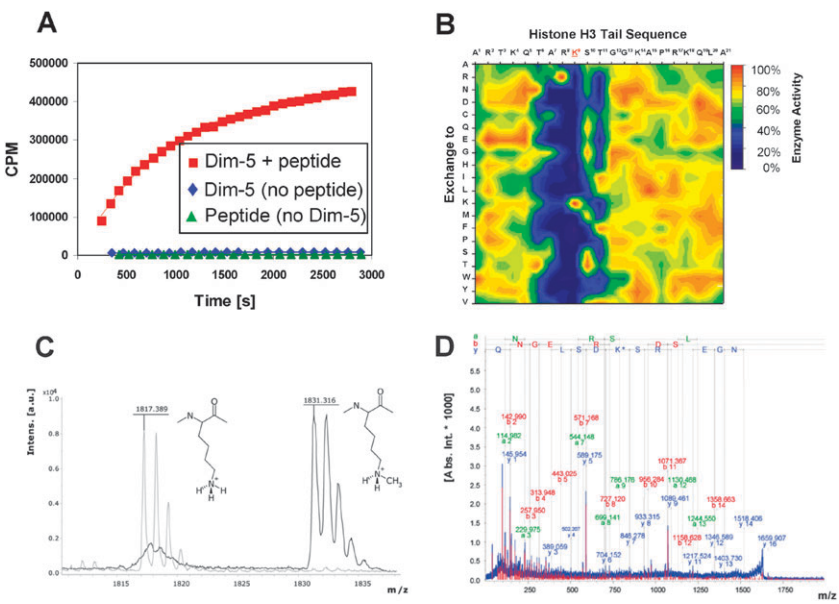

Fig. 1 (A) Methylation reaction kinetics using the Dim-5 H3K9 HKMT, an H3 tail peptide and radioactive labeled AdoMet carried out in Flashplates™.[31] A strong scintillation signal appears, which reflects the progress of the methylation reaction. If the enzyme is omitted or the reaction does not contain peptide, no signal was detected. (B) Specificity analysis of HKMTs using peptide arrays. Here, as an example an array comprising 420 peptides is shown.[23] The sequence of the H3 tail is given on the horizontal axis. Each residue was exchanged against all 20 natural amino acids (as indicated on the vertical axis) and the relative efficiency of methylation by G9a catalytic SET domain analyzed. (C) Detection of peptide methylation by mass spectrometry. The peptide was incubated with the catalytic domain of a PKMT in the presence of AdoMet and methylation of target peptide analyzed by MALDI-MS. Monomethylation of K10 in the peptide gave rise to the peak at 1831.3 Da with a mass shift of 14 Da. (D) Identification of monomethyl lysine by MALDI MS/MS. The spectrum of the peptide AANGEAGRSKDSLSRQ incubated with G9a catalytic domain identifies a monomethylation event at lysine 10.

afterwards can be used to screen for methylation activity of many modified substrates in parallel. The methylation of the respective substrates can be analyzed by following the enzymatic transfer of radioactively labeled methyl groups from AdoMet to the immobilized peptides. To investigate the influence of each amino acid of the histone H3 N-terminal peptide substrate on the activity of two PKMTs (Dim-5 and G9a), H3 tail arrays comprising 420 individual peptides were used, in which each peptide contained an exchange of one amino acid of the wild type H3 tail sequence against each of the 20 natural amino acids (Fig. 1B). Several independent membrane arrays were synthesized and incubated with the enzymes, before being analyzed quantitatively. After normalization differences were determined using statistical methods. In the case of G9a, for example, the reduced methylation of peptides carrying substitutions at positions 6-11 demonstrated an important role of T6, A7, R8, K9 (the target of methylation), S10 and T11 in the peptide recognition, while for Dim-5 recognition was seen for R8-G12.[23,32] In case of Dim-5, the results in general could be reconciled on the basis of the structure of Dim-5 with peptide. The specificity profiles of HKMTs can then be used to identify additional non-histone targets in an unbiased way.[23]

However, it is important to study not only the peptide sequence specificity of the various histone lysine methyltransferases, because HKMTs also differ in their product specificity in the sense that one, two or three methyl groups can be added to a single lysine. The modification state of peptides can be distinguished by mass spectrometric analysis.[34,35] After *in vitro* methylation with unlabelled AdoMet the peptide shows a mass shift of 14, 28 or 42 Da in the MALDI MS analysis which corresponds to the addition of one, two or three methyl group(s) (Fig. 1C). Further fragmentation of the peptide and MS/MS analysis can be performed to identify the modified lysine residue (Fig. 1D).

Analysis of protein methylation *in vitro*

Methylation of proteins can be analyzed *in vitro* after incubation of target protein and PKMT. Here, for detection, methyl specific antibodies can be used or radioactively labeled AdoMet is employed. In the first setup, the proteins can be methylated with unlabelled AdoMet, run on SDS-polyacrylamide gels and western blotted using methylation specific antibodies.[23] In the second setup, methylated proteins must be separated from unreacted AdoMet, which can be achieved by gel electrophoresis and detetion of the radioactivity in the gel. As an example, Fig. 2A shows a methylation experiment using a panel of purified non-histone protein domains incubated with radioactively labeled AdoMet and the catalytic domain of G9a. In comparison to the purified Histone H3 protein, three non-histone proteins, WIZ, ACINUS and CDYL1 as well as the N-terminal part of G9a HKMT show a much stronger G9a mediated methylation signal.[23] Alternatively, separation might be achieved by TCA precipitation of the methylated protein and scintillation counting.[36,37]

For identification of the positions of modifications, variants of the target protein can be used in which the potential target lysine residue(s) are exchanged by Ala or Arg. Alternatively, (and more directly) the site(s) of methylation can be determined by mass spectrometric analysis of the methylated proteins after digestion with proteases like ArgC, Chymotrypsin or Trypsin. As an example, Fig. 2B shows a peptide fragment of CDYL1 digested with ArgC protease after incubation of CDYL1 with the G9a PKMT. After digestion with the protease, the peptide fragments are examined using MALDI-MS analysis. Methylation of CDYL1 leads to a mass shift of 42 Da at the peptide fragment, residues 293–307, corresponding to trimethylation of one single lysine residue.

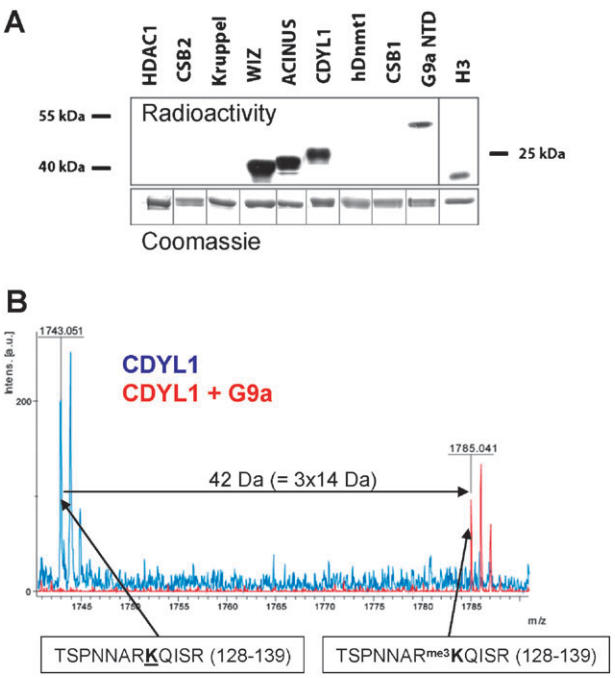

Fig. 2 (A) Detection of protein methylation by transfer of radioactively labeled methyl-groups.[23] Purified protein domains were incubated with the G9a catalytic SET domain in the presence of radioactively labeled [methyl-3H]-AdoMet, separated on an SDS-polyacrylamide gel and the methylation of the target proteins analyzed by autoradiography. The N-terminal domain of G9a is abbreviated as G9a NTD. A Coomassie stained loading control gel of the proteins is shown below. (B) Detection of protein methylation by mass spectrometry after *in vitro* methylation with purified G9a catalytic domain.[23] The GST-tagged non-histone G9a target protein CDYL1 was incubated for 5 h with the G9a catalytic SET domain in the presence of AdoMet and digested with ArgC protease. Partial digestion led to the fragment of 1743.04 Da (residues 125-139) and trimethylation of K135 in this fragment gave rise to the peak at 1785.0 Da (mass shift: 3 × 14 = 42 Da).

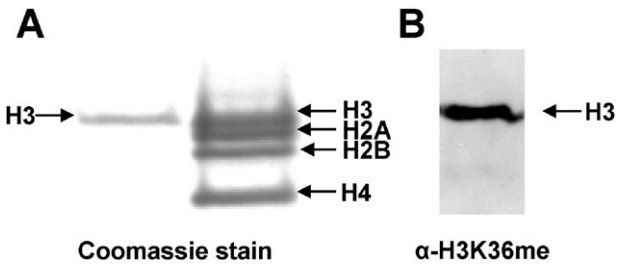

Fig. 3 (A) Coomassie stained 16% SDS PAGE gel of histones purified from HEK293 cells. Lane 1 shows commercially purchased H3 loaded on the gel as a control and lane 2 shows the core histone proteins after separation. (B) Western blot analysis of the purified histone proteins with anti-H3K36me3 antibody.

Analysis of protein methylation *in vivo*

For detection of lysine methylation in cells, purification of endogenous material is followed by either MALDI-MS analysis[34] or antibodies against different methylation states of lysines are used.[38] Fig. 3 shows an example of histone proteins isolated from human cells that were analyzed by western blot using anti-H3K36me3 antibody. In comparison to the coomassie stained control showing all four histone proteins, a specific signal appears at the size of the histone H3 protein in the western blot indicating the presence of K36 trimethylation on H3. Antibody staining can also be used in fixed cells to study the localization of particular modifications,[39–41] however the specificity of the antibody for the methylation site and state has to be verified and potential influences of other post translational modifications nearby must be considered.

Identification of novel non-histone protein methylation

Given the widespread importance of other covalent post-translational modifications (such as phosphorylation, acetylation and ubiquitylation) for function and stability of proteins, it was unlikely that lysine methylation is limited to histones. However, the detection of non-histone targets of PKMTs still is in its infancy. Apart from some candidate protein approaches, so far two different unbiased screens were applied to identify novel non-histone proteins that are subject to methylation. One approach is to isolate the whole proteome from PKMT knockout cells and incubate with recombinant PKMT and the radiolabeled cofactor AdoMet *in vitro*.[22] The appearance of radiolabeled bands then indicates the methylation targets of the PKMT. One limitation of this approach is in the detection and identification of the methylated targets, which is very difficult for low abundance proteins. In an alternative approach, a specificity profile of the PKMT is identified using peptide arrays. This motif is then used for screening protein databases to identify potential candidates taking into account additional information like cellular localization and known protein/protein interactions. The candidate sequences are then further investigated at peptide and protein level for methylation activity.[23,42] The advantage of this method is that it combines the strengths of an unbiased search with the candidate protein approach. However, it only works for enzymes that show a defined specificity profile.

Outlook

New results suggest that protein lysine methylation is a common post-translational modification of proteins, which regulates their biological activity in a complex manner. Future challenges in this field lie in the proteom-wide analysis of protein lysine methylation and in the identification of the specific biological role of the modification of each individual target.

References

1 E. Li, *Nat. Rev. Genet.*, 2002, **3**, 662–673.
2 P. A. Jones and S. B. Baylin, *Nat. Rev. Genet.*, 2002, **3**, 415–428.
3 G. Egger, G. Liang, A. Aparicio and P. A. Jones, *Nature*, 2004, **429**, 457–463.
4 A. P. Feinberg and B. Tycko, *Nat. Rev. Cancer*, 2004, **4**, 143–153.
5 R. Margueron, P. Trojer and D. Reinberg, *Curr. Opin. Genet. Dev.*, 2005, **15**, 163–176.
6 S. L. Berger, *Nature*, 2007, **447**, 407–412.
7 C. Martin and Y. Zhang, *Nat. Rev. Mol. Cell. Biol.*, 2005, **6**, 838–849.
8 B. Li, M. Carey and J. L. Workman, *Cell*, 2007, **128**, 707–719.
9 S. Rea, F. Eisenhaber, D. O'Carroll, B. D. Strahl, Z. W. Sun, M. Schmid, S. Opravil, K. Mechtler, C. P. Ponting, C. D. Allis and T. Jenuwein, *Nature*, 2000, **406**, 593–599.
10 T. Kouzarides, *Cell*, 2007, **128**, 693–705.
11 X. Cheng, R. E. Collins and X. Zhang, *Annu. Rev. Biophys. Biomol. Struct.*, 2005, **34**, 267–294.
12 S. Chuikov, J. K. Kurash, J. R. Wilson, B. Xiao, N. Justin, G. S. Ivanov, K. McKinney, P. Tempst, C. Prives, S. J. Gamblin, N. A. Barlev and D. Reinberg, *Nature*, 2004, **432**, 353–360.
13 J. K. Kurash, H. Lei, Q. Shen, W. L. Marston, B. W. Granda, H. Fan, D. Wall, E. Li and F. Gaudet, *Mol. Cell*, 2008, **29**, 392–400.
14 J. Huang, L. Perez-Burgos, B. J. Placek, R. Sengupta, M. Richter, J. A. Dorsey, S. Kubicek, S. Opravil, T. Jenuwein and S. L. Berger, *Nature*, 2006, **444**, 629–632.
15 X. Shi, I. Kachirskaia, H. Yamaguchi, L. E. West, H. Wen, E. W. Wang, S. Dutta, E. Appella and O. Gozani, *Mol. Cell*, 2007, **27**, 636–646.
16 A. Kouskouti, E. Scheer, A. Staub, L. Tora and I. Talianidis, *Mol. Cell*, 2004, **14**, 175–182.
17 J. F. Couture, E. Collazo, G. Hauk and R. C. Trievel, *Nat. Struct. Mol. Biol.*, 2006, **13**, 140–146.
18 K. Zhang, W. Lin, J. A. Latham, G. M. Riefler, J. M. Schumacher, C. Chan, K. Tatchell, D. H. Hawke, R. Kobayashi and S. Y. Dent, *Cell*, 2005, **122**, 723–734.
19 S. Clarke, *Curr. Opin. Cell. Biol.*, 1993, **5**, 977–983.
20 W. K. Paik, D. C. Paik and S. Kim, *Trends Biochem. Sci.*, 2007, **32**, 146–152.
21 M. Tachibana, K. Sugimoto, T. Fukushima and Y. Shinkai, *J. Biol. Chem.*, 2001, **276**, 25309–25317.
22 S. C. Sampath, I. Marazzi, K. L. Yap, S. C. Sampath, A. N. Krutchinsky, I. Mecklenbrauker, A. Viale, E. Rudensky, M. M. Zhou, B. T. Chait and A. Tarakhovsky, *Mol. Cell*, 2007, **27**, 596–608.
23 P. Rathert, A. Dhayalan, M. Murakami, X. Zhang, R. Tamas, R. Jurkowska, Y. Komatsu, Y. Shinkai, X. Cheng and A. Jeltsch, *Nat. Chem. Biol.*, 2008, **4**, 344–346.
24 S. D. Taverna, H. Li, A. J. Ruthenburg, C. D. Allis and D. J. Patel, *Nat. Struct. Mol. Biol.*, 2007, **14**, 1025–1040.
25 J. Huang, R. Sengupta, A. B. Espejo, M. G. Lee, J. A. Dorsey, M. Richter, S. Opravil, R. Shiekhattar, M. T. Bedford, T. Jenuwein and S. L. Berger, *Nature*, 2007, **449**, 105–108.
26 E. Collazo, J. F. Couture, S. Bulfer and R. C. Trievel, *Anal. Biochem.*, 2005, **342**, 86–92.
27 J. R. Wilson, C. Jing, P. A. Walker, S. R. Martin, S. A. Howell, G. M. Blackburn, S. J. Gamblin and B. Xiao, *Cell*, 2002, **111**, 105–115.

28 D. Patnaik, H. G. Chin, P. O. Esteve, J. Benner, S. E. Jacobsen and S. Pradhan, *J. Biol. Chem*., 2004, **279**, 53248–53258.
29 H. Gowher, X. Zhang, X. Cheng and A. Jeltsch, *Anal. Biochem*., 2005, **342**, 287–291.
30 S. Kubicek, R. J. O'Sullivan, E. M. August, E. R. Hickey, Q. Zhang, M. L. Teodoro, S. Rea, K. Mechtler, J. A. Kowalski, C. A. Homon, T. A. Kelly and T. Jenuwein, *Mol. Cell*, 2007, **25**, 473–481.
31 P. Rathert, X. Cheng and A. Jeltsch, *Biotechniques*, 2007, **43**, 602–608.
32 P. Rathert, X. Zhang, C. Freund, X. Cheng and A. Jeltsch, *Chem. Biol*., 2008, **15**, 5–11.
33 R. Frank, *J. Immunol. Methods*, 2002, **267**, 13–26.
34 T. Bonaldi, J. T. Regula and A. Imhof, *Methods Enzymo.l*, 2004, **377**, 111–130.
35 X. Zhang, Z. Yang, S. I. Khan, J. R. Horton, H. Tamaru, E. U. Selker and X. Cheng, *Mol. Cell*, 2003, **12**, 177–185.
36 X. Zhang, H. Tamaru, S. I. Khan, J. R. Horton, L. J. Keefe, E. U. Selker and X. Cheng, *Cell*, 2002, **111**, 117–127.
37 R. C. Trievel, B. M. Beach, L. M. Dirk, R. L. Houtz and J. H. Hurley, *Cell*, 2002, **111**, 91–103.
38 A. J. Bannister and T. Kouzarides, *Methods Enzymol*., 2004, **376**, 269–288.
39 A. H. Peters, S. Kubicek, K. Mechtler, R. J. O'Sullivan, A. A. Derijck, L. Perez-Burgos, A. Kohlmaier, S. Opravil, M. Tachibana, Y. Shinkai, J. H. Martens and T. Jenuwein, *Mol. Cell*, 2003, **12**, 1577–1589.
40 L. Perez-Burgos, A. H. Peters, S. Opravil, M. Kauer, K. Mechtler and T. Jenuwein, *Methods Enzymol*., 2004, **376**, 234–254.
41 G. Schotta, M. Lachner, K. Sarma, A. Ebert, R. Sengupta, G. Reuter, D. Reinberg and T. Jenuwein, *Genes. Dev*., 2004, **18**, 1251–1262.
42 P. Trojer and D. Reinberg, *Nat. Chem. Biol*., 2008, **4**, 332–334.

