

Communication

# Quantitative Chemical Proteomics Approach to Identify Post-translational Modification-mediated Protein-protein Interactions

Xiang Li, Emily A. Foley, Kelly R. Molloy, Yinyin Li, Brian Trevor Chait, and Tarun M. Kapoor

*J. Am. Chem. Soc.*, **Just Accepted Manuscript** • DOI: 10.1021/ja210528v • Publication Date (Web): 09 Jan 2012

Downloaded from <http://pubs.acs.org> on January 17, 2012

## Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.



**ACS Publications**  
High quality. High impact.

# Quantitative Chemical Proteomics Approach to Identify Posttranslational Modification-mediated Protein-protein Interactions

Xiang Li<sup>†‡</sup>, Emily A. Foley<sup>†</sup>, Kelly R. Molloy<sup>§</sup>, Yinyin Li<sup>§</sup>, Brian T. Chait<sup>§</sup>, Tarun M. Kapoor<sup>†\*</sup>

<sup>†</sup>Laboratory of Chemistry and Cell Biology, <sup>§</sup>Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, the Rockefeller University, New York, NY, 10065. <sup>‡</sup>Current address: Department of Chemistry, The University of Hong Kong, Pokfulam Road, Hong Kong, China.

[kapoor@rockefeller.edu](mailto:kapoor@rockefeller.edu)

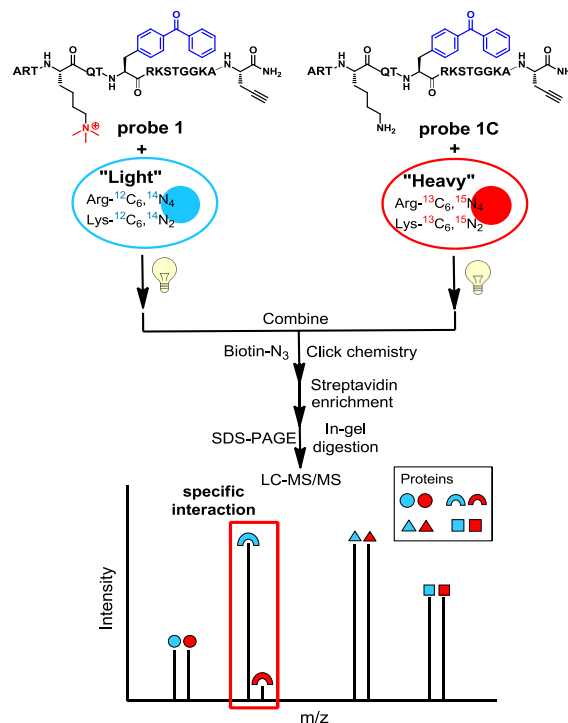
**ABSTRACT:** Post-translational modifications (PTMs) (e.g. acetylation, methylation, and phosphorylation) play crucial roles in regulating the diverse protein-protein interactions involved in essentially every cellular process. While significant progress has been made to detect PTMs, profiling protein-protein interactions mediated by these PTMs remains a challenge. Here, we report a method that combines a photo-crosslinking strategy with stable isotope labeling in cell culture (SILAC)-based quantitative mass spectrometry to identify PTM-dependent protein-protein interactions. To develop and apply this approach, we focused on trimethylated lysine-4 at the histone H3 ‘tail’ (H3K4Me<sub>3</sub>), a PTM linked to actively transcribed regions on chromosomes. Our approach identified proteins previously known to recognize this modification and MORC3 as a new H3K4Me<sub>3</sub> ‘reader’. This study indicates that our cross-linking-assisted and SILAC-based protein identification (CLASPI) approach can be used to profile protein-protein interactions mediated by PTMs, such as lysine methylation.

To fully understand a protein’s function in cellular processes, it is critical that its binding partners, including those that recognize different PTMs<sup>1</sup>, are identified. Currently, methods such as affinity chromatography<sup>2</sup>, immunoprecipitation<sup>3,4</sup>, and yeast two-hybrid<sup>5</sup> approaches are used to identify protein-protein interactions. However, we lack reliable methods to find proteins that interact exclusively with a PTM-carrying form of another protein. There are several reasons why developing such methods has been challenging, including the following. First, amino acids carrying the PTMs cannot be readily incorporated into proteins in cellular contexts, limiting the applicability of two-hybrid-type approaches. Second, PTMs can be dynamic, present at sub-stoichiometric levels, and mediate relatively weak (micro molar) interactions<sup>1</sup>, thereby reducing yields in ‘pull-down’ experiments.

To develop a robust method that can address difficulties in analyzing PTM-dependent protein-protein interactions, we focused on histones. These widely conserved proteins assemble into nucleosomes, around which DNA is packaged to form chromatin<sup>6</sup>. Numerous histone PTMs, such as methylation, phosphorylation, acetylation and ubiquitylation, have been identified<sup>7</sup>. Histone PTMs recruit binding proteins, or ‘readers’ of these ‘marks’ to execute cellular processes, such as gene transcription, DNA replication and chromosome segregation<sup>8-10</sup>. The important roles of histone PTMs in cell proliferation and the available data indicating that histone PTMs are hubs for protein-protein interactions made them an attractive starting point for our studies.

We recently reported a photo-crosslinking-based peptide probe (probe **1**, Fig. 1) to covalently capture proteins that recognize trimethylation of histone H3 at lysine 4 (H3K4Me<sub>3</sub>)<sup>11</sup>, a ‘mark’ associated with actively transcribed gene promoters. This probe is based on the unstructured N-terminal ‘tail’ of histone H3, with lysine-4 trimethylated, a photo-crosslinker (benzophenone) appended to alanine-7 and an alkyne incorporated at the C-terminal

residue. Probe **1** was designed to address two difficulties in identifying PTM-dependent protein-protein interactions. First, the probe contained stoichiometric levels of tri-methylated lysine-4. Second, photo-crosslinking converted weak interactions into covalent bonds. In *in vitro* studies, we showed that this probe selectively labeled a subset of known H3K4Me<sub>3</sub>-binding proteins<sup>11</sup>. However, the utility of this probe to profile proteins that ‘read’ this histone ‘mark’ had not been tested. Here, we combine our photo-crosslinking-based strategy with quantitative SILAC technology<sup>12</sup> to develop a robust chemical proteomics approach (named CLASPI) for profiling proteins that recognize methylated histone ‘tails’.

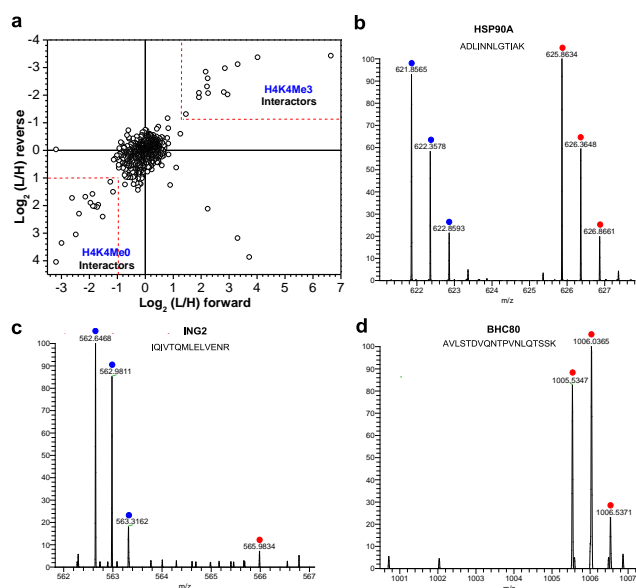


**Figure 1.** Schematic for the CLASPI strategy to profile H3K4Me<sub>3</sub> binding partners in whole-cell proteomes.

We first examined whether probe **1** could be used to profile H3K4Me<sub>3</sub>-binding proteins in whole-cell proteomes. Probe **1** or a control probe (probe **1C**) was incubated with HeLa cell lysate, and exposed to UV light for photo-crosslinking. The crosslinked proteins were conjugated to biotin through azide-alkyne click chemistry and isolated by streptavidin-coated beads. Western blot analysis revealed that, as expected, endogenous ING2<sup>13</sup>, a known H3K4Me<sub>3</sub> ‘reader’, was selectively captured by probe **1** but not by probe **1C** (Fig. S1). However, silver staining of the proteins captured by probe **1**- and probe **1C** revealed few visible differences indicating that detection of H3K4Me<sub>3</sub>-specific binding proteins would require a quantitative and sensitive method (Fig. S1).

We therefore combined our photo-crosslinking approach with SILAC technology, which has been widely used for quantitative mass spectrometry.

As illustrated in Fig. 1, whole-cell lysates derived from HeLa cells grown in medium containing either ‘light’ (natural isotope abundance forms) and ‘heavy’ ( $^{13}\text{C}$ ,  $^{15}\text{N}$ -substituted arginine and lysine) were incubated with probe **1** and probe **1C**, respectively. After photo-crosslinking, the ‘light’ and ‘heavy’ lysates were pooled. Then the crosslinked proteins were biotinylated using click chemistry, followed by affinity purification and gel electrophoresis. Following in-gel trypsin digestion, the peptide mixtures were separated by HPLC and analyzed with a LTQ-Orbitrap mass spectrometer. By this method, proteins that were enriched in the ‘light’ fraction are likely H3K4Me<sub>3</sub>-binders, while proteins that were recovered at similar levels by both probes are non-selective binders (Fig. 1). To obtain a robust analysis of specific ‘readers’ of H3K4Me<sub>3</sub>, we performed two independent CLASPI experiments. In a ‘forward’ experiment, the ‘light’ lysate was cross-linked with probe **1**, and the ‘heavy’ lysate was crosslinked with probe **1C**. In a ‘reverse’ experiment, the probes were switched. Only H3K4Me<sub>3</sub>-specific proteins that were identified in both experiments were further analyzed (Table S1).



**Figure 2.** Profiling of H3K4Me<sub>3</sub> readers by CLASPI. (a) A 2D plot showing the  $\text{Log}_2$  values of the SILAC ratios (L/H) of each identified protein for the ‘forward’ (x axis) and ‘reverse’ (y axis) experiments. The H3K4Me<sub>3</sub>-binding and unmodified H3 (H3K4Me<sub>0</sub>)-interacting partners are indicated in the top right and bottom left quadrants, respectively. Representative MS spectra of peptide from (b) non-specific binders (HSP90A), (c) H3K4Me<sub>3</sub> binders (ING2) and (d) H3K4Me<sub>0</sub> binders (BHC80). The ‘light’ and ‘heavy’ peptide isotopes are indicated by blue and red dots, respectively.

A logarithmic ( $\text{Log}_2$ ) SILAC ratio (L/H) of the identified proteins was then plotted with data from the ‘forward’ experiment along the x-axis and ‘reverse’ experiment along the y-axis (Fig. 2a). As expected, most proteins showed a SILAC ratio (L/H) around 1:1 (e.g., Fig. 2b), clustering at the intersection of the x and y axes in Fig. 1b, indicating that they are not likely to be involved in methylation-sensitive interactions. A small subset of proteins, which showed a high SILAC ratio (L/H) in the ‘forward’ experiment (e.g., Fig. 2c) and a low ratio (L/H) in the ‘reverse’ experiment, appeared in the top right quadrant of the plot (Fig. 2a), indicating that they bind preferentially to the K4-trimethylated histone H3 ‘tail’. Many biochemically or structurally characterized H3K4Me<sub>3</sub> binders were identified in this region,

including ING1<sup>13</sup>, ING2<sup>13</sup>, PYGO2<sup>14</sup>, PHF23<sup>15</sup>, SGF29<sup>16</sup> and JMJD2A<sup>17</sup>, validating our approach (Table 1). In addition, we also identified a number of potential H3K4Me<sub>3</sub> binders whose ability to recognize this modification had not yet been well-characterized.

**Table 1.** CLASPI-identified proteins that were enriched or excluded by the H3K4Me<sub>3</sub> probe (**1**), relative to the unmodified probe (**1C**).

H3K4Me <sub>3</sub> interactors	Ratio (L/H)	H3K4Me <sub>0</sub> interactors	Ratio (L/H)
SGF29 <sup>a</sup>	99.60	BHC80 <sup>b</sup>	0.11
ING1 <sup>b</sup>	16.21	RPL4	0.13
PHF23 <sup>b</sup>	9.89	HAT1	0.16
ING2 <sup>b</sup>	7.69	RBBP7	0.18
JMJD2A <sup>a</sup>	7.19	TRIM33 <sup>b</sup>	0.19
CXXC1 <sup>b</sup>	7.01	JADE2 <sup>b</sup>	0.23
MORC3	4.75	JADE1 <sup>b</sup>	0.26
PYGO2 <sup>b</sup>	4.70	RBBP4	0.27
JADE3 <sup>b</sup>	4.63	DPF2 <sup>b</sup>	0.27
UBR7 <sup>b</sup>	4.47	MTA2	0.28
SPIN1 <sup>a</sup>	3.79	MTA1	0.30
DIDO1 <sup>b</sup>	3.78	UHRF2 <sup>b</sup>	0.31
SIN3A	2.47	RPL17	0.34
		UHRF1 <sup>b</sup>	0.45

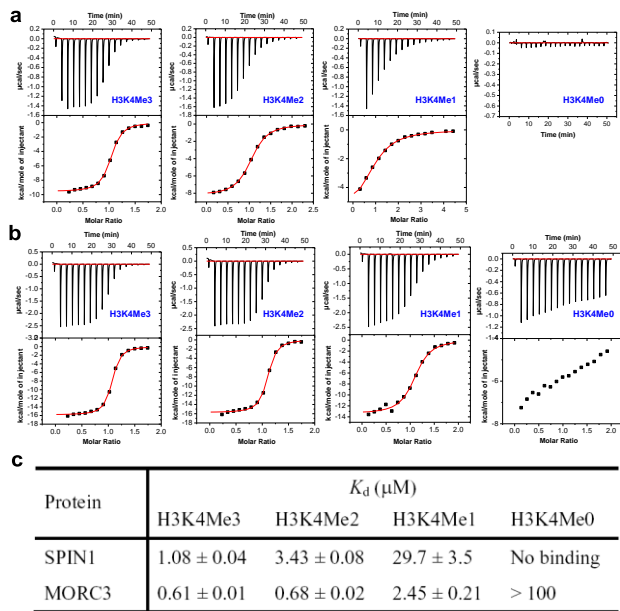
The criteria used to select these proteins was >1.5-fold change in both the ‘forward’ and ‘reverse’ experiments and statistically significant with the product of outlier probabilities of both the experiments ( $P = P^{\text{(for)}} \times P^{\text{(rev)}} < 0.0004$ ). Proteins are sorted according to their SILAC ratios (L/H) in the ‘forward’ experiment (see Table S1 for their SILAC ratios (L/H) in the ‘reverse’ experiment). <sup>a</sup> Tudor domain-containing protein; <sup>b</sup> PHD finger-containing protein.

Interestingly, we also identified a subset of proteins that mapped to the bottom-left quadrant of the 2D plot. These proteins showed a low ratio in the forward experiment and a high ratio in the reverse experiment, indicating that they were likely excluded by H3K4Me<sub>3</sub>. This data set includes proteins that are known to bind preferentially unmodified histone H3 tail such as BHC80<sup>18</sup> (Fig. 2d), TRIM33<sup>19</sup> and JADE1<sup>20</sup> (Table 1). Furthermore, non-labeled exogenous contaminants such as keratin and proteins derived from the medium did not change the ratio of L/H from the ‘forward’ to the ‘reverse’ experiment and therefore appeared in the lower right-hand quadrant. Together, this analysis allowed the identification of potential interacting partners of both K4-trimethylated and unmodified histone H3 tails.

A common feature of the structurally characterized histone lysine methylation ‘readers’ is the presence of an ‘aromatic cage’ that mediates the recognition of trimethyl lysine through  $\pi$ -cation interactions<sup>9</sup>. We therefore examined, at the sequence level, whether the proteins identified by CLASPI had this feature. For this analysis, we focused on the proteins that contained plant homeodomain (PHD) fingers (Table 1), a fold known to recognize histone lysine methylations. As shown in Fig. S2 (top), primary sequence alignment revealed that all putative H3K4Me<sub>3</sub>-binding proteins except JADE, contain key residues that constitute the aromatic cage for trimethyl lysine recognition. In contrast, these residues were not conserved in the PHD fingers of the proteins that preferentially bound unmodified histone H3 (Fig. S2, bottom). These analyses suggest that these CLASPI-identified proteins that have conserved aromatic cage residues are likely to recognize the H3K4Me<sub>3</sub> ‘mark’.

We next examined whether H3K4Me<sub>3</sub>-binding proteins identified by CLASPI can directly and selectively bind to this histone mark *in vitro*. We first focused on SPIN1, a protein that has re-

cently been reported as a potential H3K4Me<sub>3</sub> binder in a SILAC-based proteomics study<sup>21</sup>. However the interaction between SPIN1 and H3K4Me<sub>3</sub> has not yet been verified or biophysically characterized. Therefore, we purified recombinant SPIN1 and measured its binding affinities toward various histone H3 peptides at different methylation states using isothermal titration calorimetry (ITC). Consistent with our CLASPI analysis, binding between SPIN1 and unmodified H3 peptide was undetectable, whereas SPIN1 bound to an H3K4Me<sub>3</sub> peptide with  $K_d = 1.1 \mu\text{M}$ . Additionally, we found that SPIN1 bound more tightly to a trimethylated H3K4 peptide, compared to mono- or dimethylated peptides (Figs. 3a and c).



**Figure 3.** Characterization of CLASPI-identified histone methylation ‘readers’. Isothermal titration calorimetry measurements for the binding affinities of (a) SPIN1 and (b) MORC3 for the indicated histone H3 peptides. (c) A summary of dissociation constants ( $K_d$ ) of SPIN1 and MORC3 for histone peptides. The values are the average of two independent measurements  $\pm$  standard deviation.

We next characterized MORC3, which has not previously been linked to H3K4Me<sub>3</sub> binding. Sequence analysis of MORC3 revealed a zinc finger CW domain, which, like PHD fingers and tudor domains, is known to bind methyl lysines on histones<sup>22,23</sup>. We therefore expressed in bacteria and purified the MORC3 CW domain and analyzed histone H3 ‘tail’ interactions. ITC revealed that the MORC3 CW domain bound to H3K4Me<sub>3</sub> peptide with  $K_d = 0.6 \mu\text{M}$ , Figs. 3b and c). Interestingly, while this domain can distinguish methylated H3K4 peptides from the unmodified H3 peptide, it only modestly discriminates between the different methylation states of H3 peptides (Fig. 3c). Together, these results validate the specific interactions between histone H3K4Me<sub>3</sub> and two proteins we identified by the CLASPI experiments.

In summary, we have developed and used the CLASPI strategy to profile proteins that recognize histone H3K4Me<sub>3</sub> in whole-cell proteomes. We believe this approach complements conventional ‘pull down’ approaches and offers the following advantages. First, relatively weak ( $\mu\text{M}$ ) interactions can be identified. Second, when compared with non-specific chemical cross-linking (e.g., using formaldehyde), the incorporation of a covalent photo-cross linker proximal to the PTM can ensure that the cross-linked protein directly binds the probe. In contrast, the gentle wash conditions necessary to preserve weak binding interactions in traditional pull-down methods can inadvertently result in isolation of multi-protein complexes. For example, eight subunits of the SAGA

complex were detected in association with H3K4Me<sub>3</sub> peptide using a pull-down approach<sup>16</sup>, whereas our CLASPI approach captured the one SAGA subunit responsible for H3K4Me<sub>3</sub> binding, SGF29. On the other hand, while we identified several known H3K4Me<sub>3</sub> binders by CLASPI, others such as TAF3 and CHD1 were missed. One likely explanation is that the bulky benzophenone moiety, which replaced the original small methyl side chain of Ala-7 in probe 1, may ‘bump’ into the interaction surface of some ‘readers’, such as TAF3<sup>24</sup>. To address this limitation in future studies, the benzophenone group could be substituted with other photoreactive groups (e.g. diazirine), and/or repositioned relative to the PTM. We anticipate that CLASPI can be extended to analyze protein-protein interactions in complex proteomes that depend on other PTMs (e.g. phosphorylation). These studies will be reported in due course.

**Acknowledgement** TMK acknowledges support by the NIH (GM98579).

**Supporting Information Available:** The experimental method and the supplementary figures and table. This material is available free of charge via the Internet at <http://pubs.acs.org>.

**References**

- (1) Walsh, C. T. *Posttranslational Modifications of Proteins: Expanding Nature's Inventory*; Roberts and Co. Publishers: Greenwood Village, CO, 2006.
- (2) Beeckmans, S. *Methods* **1999**, *19*, 278.
- (3) Burgess, R. R.; Thompson, N. E. *Curr Opin Biotechnol* **2002**, *13*, 304.
- (4) Kellogg, D. R.; Moazed, D. *Methods Enzymol* **2002**, *351*, 172.
- (5) Fields, S.; Song, O. *Nature* **1989**, *340*, 245.
- (6) Kornberg, R. D.; Lorch, Y. *Cell* **1999**, *98*, 285.
- (7) Kouzarides, T. *Cell* **2007**, *128*, 693.
- (8) Jenuwein, T.; Allis, C. D. *Science* **2001**, *293*, 1074.
- (9) Taverna, S. D.; Li, H.; Ruthenburg, A. J.; Allis, C. D.; Patel, D. J. *Nat Struct Mol Biol* **2007**, *14*, 1025.
- (10) Seet, B. T.; Dikic, I.; Zhou, M. M.; Pawson, T. *Nat Rev Mol Cell Biol* **2006**, *7*, 473.
- (11) Li, X.; Kapoor, T. M. *J Am Chem Soc* **2010**, *132*, 2504.
- (12) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. *Mol Cell Proteomics* **2002**, *1*, 376.
- (13) Shi, X.; Hong, T.; Walter, K. L.; Ewalt, M.; Michishita, E.; Hung, T.; Carney, D.; Pena, P.; Lan, F.; Kaadige, M. R.; Lacoste, N.; Cayrou, C.; Davrazou, F.; Saha, A.; Cairns, B. R.; Ayer, D. E.; Kutateladze, T. G.; Shi, Y.; Cote, J.; Chua, K. F.; Gozani, O. *Nature* **2006**, *442*, 96.
- (14) Fiedler, M.; Sanchez-Barrena, M. J.; Nekrasov, M.; Mieszczanek, J.; Rybin, V.; Muller, J.; Evans, P.; Bienz, M. *Mol Cell* **2008**, *30*, 507.
- (15) Wang, G. G.; Song, J.; Wang, Z.; Dormann, H. L.; Casadio, F.; Li, H.; Luo, J. L.; Patel, D. J.; Allis, C. D. *Nature* **2009**, *459*, 847.
- (16) Vermeulen, M.; Eberl, H. C.; Matarese, F.; Marks, H.; Denissov, S.; Butter, F.; Lee, K. K.; Olsen, J. V.; Hyman, A. A.; Stunnenberg, H. G.; Mann, M. *Cell* **2010**, *142*, 967.
- (17) Huang, Y.; Fang, J.; Bedford, M. T.; Zhang, Y.; Xu, R. M. *Science* **2006**, *312*, 748.
- (18) Lan, F.; Collins, R. E.; De Cegli, R.; Alpatov, R.; Horton, J. R.; Shi, X.; Gozani, O.; Cheng, X.; Shi, Y. *Nature* **2007**, *448*, 718.
- (19) Agricola, E.; Randall, R. A.; Gaarenstroom, T.; Dupont, S.; Hill, C. S. *Mol Cell* **2011**, *43*, 85.
- (20) Saksouk, N.; Avvakumov, N.; Champagne, K. S.; Hung, T.; Doyon, Y.; Cayrou, C.; Paquet, E.; Ullah, M.; Landry, A. J.; Cote, V.; Yang, X. J.; Gozani, O.; Kutateladze, T. G.; Cote, J. *Mol Cell* **2009**, *33*, 257.
- (21) Bartke, T.; Vermeulen, M.; Xhemalce, B.; Robson, S. C.; Mann, M.; Kouzarides, T. *Cell* **2010**, *143*, 470.
- (22) He, F.; Umehara, T.; Saito, K.; Harada, T.; Watanabe, S.; Yabuki, T.; Kigawa, T.; Takahashi, M.; Kuwasako, K.

1 Tsuda, K.; Matsuda, T.; Aoki, M.; Seki, E.; Kobayashi, N.; Guntert, P.;  
2 Yokoyama, S.; Muto, Y. *Structure* **2010**, *18*, 1127.  
3 (23) Hoppmann, V.; Thorstensen, T.; Kristiansen, P. E.;  
4 Veiseth, S. V.; Rahman, M. A.; Finne, K.; Aalen, R. B.; Aasland, R.  
5 *EMBO J* **2011**, *30*, 1939.

(24) van Ingen, H.; van Schaik, F. M.; Wienk, H.;  
6 Ballering, J.; Rehmann, H.; Dechesne, A. C.; Kruijzer, J. A.; Liskamp, R.  
7 M.; Timmers, H. T.; Boelens, R. *Structure* **2008**, *16*, 1245.

