JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Quantitative Chemical Proteomics Approach To Identify Post-translational Modification-Mediated Protein−Protein Interactions

Xiang Li,[†,‡] Emily A. Foley,[†] Kelly R. Molloy,[§] Yinyin Li,[§] Brian T. Chait,[§] and Tarun M. Kapoor*[,†]

[†]Laboratory of Chemistry and Cell Biology and [§]Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, New York 10065, United States

Supporting Information

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-cross-linking 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 *N*-terminus ($H3K4Me_3$), a PTM linked to actively transcribed gene promoters. Our approach identified proteins previously known to recognize this modification and MORC3 as a new protein that binds $H3M4Me_3$. 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,[1] are identified. Currently, methods such as affinity chromatography,[2] immunoprecipitation,[3,4] and yeast two-hybrid[5] 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 substoichiometric levels, and mediate relatively weak (micromolar) interactions,[1] 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.[6] Numerous histone PTMs, such as methylation, phosphorylation, acetylation, and ubiquitylation, have been identified.[7] Histone PTMs recruit binding proteins, or "readers" of these "marks" to execute cellular processes, such as gene transcription, DNA replication and chromosome segregation.[8−10] 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-cross-linking-based peptide probe (**1**, Figure 1) to covalently capture proteins that recognize trimethylation of histone H3 at lysine 4 ($H3K4Me_3$),[11] 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-cross-linker (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

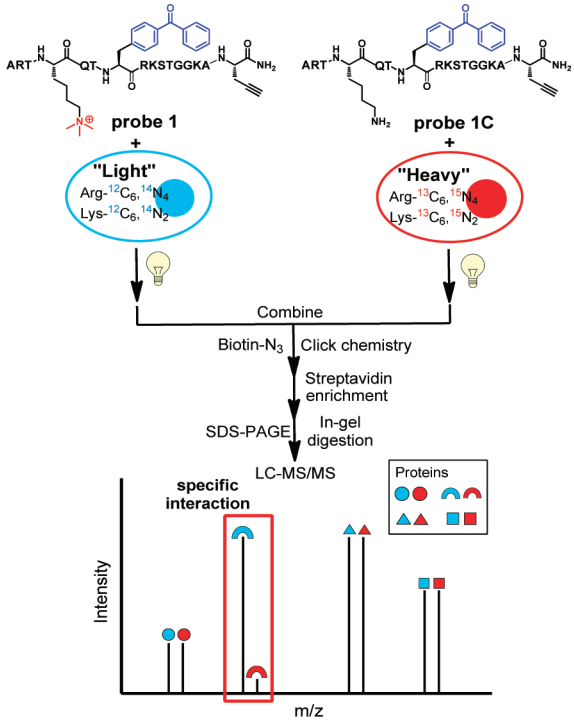


Figure 1. Schematic for the CLASPI strategy to profile $H3K4Me_3$ binding partners in whole-cell proteomes.

Received: November 16, 2011
Published: January 9, 2012

contained stoichiometric levels of trimethylated lysine-4. Second, photo-cross-linking converted weak interactions into covalent bonds. In *in vitro* studies, we showed that this probe selectively labeled a subset of known $H3K4Me_3$-binding proteins.[11] However, the utility of this probe to profile proteins that read this histone "mark" from whole cell proteomes had not been tested. Here, we combine our photo-cross-linking-based strategy with quantitative SILAC technology[12] to develop a robust chemical proteomics approach (named CLASPI) for profiling proteins that recognize methylated histone "tails."

We first examined whether probe **1** could be used to profile $H3K4Me_3$-binding proteins in whole-cell proteomes. Probe **1** or a control probe (**1C**) was incubated with HeLa cell lysate, and exposed to UV "light" for photo-cross-linking. The cross-linked 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,[13] a known $H3K4Me_3$ "reader," was selectively captured by probe **1** but not by probe **1C** (Figure S1). However, silver staining of the proteins captured by probe **1** and probe **1C** revealed few visible differences, indicating that detection of $H3K4Me_3$-specific binding proteins would require a quantitative and sensitive method (Figure S1). We therefore combined our photo-cross-linking approach with SILAC technology, which has been widely used for quantitative mass spectrometry.

As illustrated in Figure 1, whole-cell lysates derived from HeLa cells grown in medium containing either "light" (natural isotope abundance forms) or "heavy" lysates (^{13}C,^{15}N-substituted arginine and lysine) were incubated with probe **1** and probe **1C**, respectively. After photo-cross-linking, the "light" and "heavy" lysates were pooled. The cross-linked proteins were then 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 $H3K4Me_3$-binders, while proteins that were recovered at similar levels by both probes are nonselective binders (Figure 1). To obtain a robust analysis of specific "readers" of $H3K4Me_3$, we performed two independent CLASPI experiments. In a "forward" experiment, the "light" lysate was cross-linked with probe **1**, and the "heavy" lysate was cross-linked with probe **1C**. In a "reverse" experiment, the probes were switched. Only $H3K4Me_3$-specific proteins that were identified in both experiments were further analyzed (Table S1).

A logarithmic (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 (Figure 2a). As expected, most proteins showed a SILAC ratio (L/H) around 1 (e.g., Figure 2b), clustering at the intersection of the x and y axes, 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., Figure 2c) and a low ratio (L/H) in the "reverse" experiment, appeared in the top right quadrant of the plot (Figure 2a), indicating that they bind preferentially to the K4-trimethylated histone H3 "tail." Many biochemically or structurally characterized $H3K4Me_3$ binders were identified in this region, including ING1,[13] ING2,[13] PYGO2,[14] PHF23,[15] SGF29[16] and JMJD2A,[17] validating our approach (Table 1). In addition, we also identified a number of potential $H3K4Me_3$ binders whose ability to recognize this modification had not yet been well-characterized.

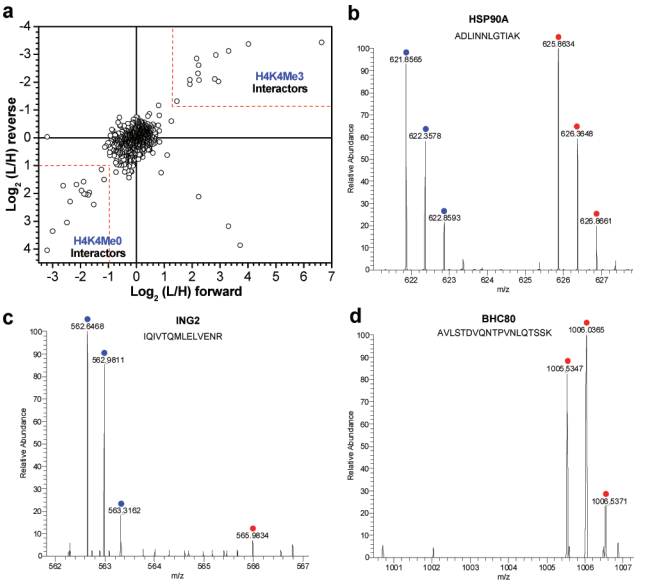


Figure 2. Profiling of $H3K4Me_3$ "readers" by CLASPI. (a) A 2D plot showing the Log2 values of the SILAC ratios (L/H) of each identified protein for the "forward" (*x* axis) and "reverse" (*y* axis) experiments. The $H3K4Me_3$-binding and unmodified H3 (H3K4Me0)-interacting partners are indicated in the top right and bottom left quadrants, respectively. Representative MS spectra for peptides from (b) nonspecific binder (HSP90A), (c) H3K4Me3 binder (ING2), and (d) H3K4Me0 binder (BHC80). The "light" and "heavy" peptide isotopes are indicated by blue and red dots, respectively.

Table 1. CLASPI-Identified Proteins That Were Enriched or Excluded by the $H3K4Me_3$ Probe (1), Relative to the Unmodified Probe (1C)[a]

$H3K4Me_3$ interactors	ratio (L/H)	$H3K4Me_0$ interactors	ratio (L/H)
SGF29[b]	99.60	BHC80[c]	0.11
ING1[c]	16.21	RPL4	0.13
PHF23[c]	9.89	HAT1	0.16
ING2[c]	7.69	RBBP7	0.18
JMJD2A[b]	7.19	TRIM33[c]	0.19
CXXC1[c]	7.01	JADE2[c]	0.23
MORC3	4.75	JADE1[c]	0.26
PYGO2[c]	4.70	RBBP4	0.27
JADE3[c]	4.63	DPF2[c]	0.27
UBR7[c]	4.47	MTA2	0.28
SPIN1[b]	3.79	MTA1	0.30
DIDO1[c]	3.78	UHRF2[c]	0.31
SIN3A	2.47	RPL17	0.34
		UHRF1[c]	0.45

[a]The criteria used to select these proteins was >1.5-fold change in both the "forward" and "reverse" experiments and also statistically significant, such that the product of outlier probabilities of both the experiments ($P = P^{(for)}P^{(rev)}$) < 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). [b]Tudor domain-containing protein. [c]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

dx.doi.org/10.1021/ja210528v | J. Am. Chem.Soc. 2012, 134, 1982−1985

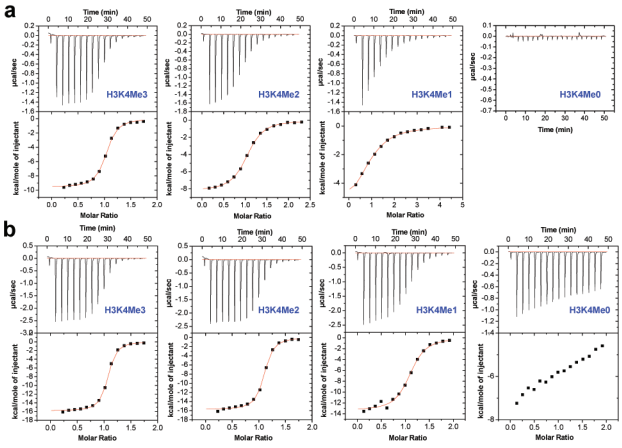
likely excluded by H3K4Me$_3$. This data set includes proteins that are known to bind preferentially unmodified histone H3 "tail" such as BHC80[18] (Figure 2d), TRIM33,[19] and JADE1[20] (Table 1). Furthermore, nonlabeled exogenous contaminants such as keratin and proteins derived from the medium did 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 π-cation interactions.[9] 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 Figure S2 (top), primary sequence alignment revealed that all putative H3K4Me$_3$-binding proteins except JADE3, 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 (Figure S2, bottom). These analyses suggest that these CLASPI-identified proteins that have conserved aromatic cage residues are likely to recognize the H3K4Me$_3$ "mark".

We next examined whether H3K4Me$_3$-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 recently been reported as a potential H3K4Me$_3$ binder in a SILAC-based proteomics study.[21] However the interaction between SPIN1 and H3K4Me$_3$ 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$_3$ peptide with K_d = 1.1 μM. Additionally, we found that SPIN1 bound more tightly to a trimethylated H3K4 peptide, compared to mono- or dimethylated peptides (Figure 3a,c).

We next characterized MORC3, which has not previously been linked to H3K4Me$_3$ 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.[22,23] 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$_3$ peptide with K_d = 0.6 μM, (Figure 3b,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 (Figure 3c). Together, these results validate the specific interactions between histone H3K4Me$_3$ 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$_3$ in whole-cell proteomes. We believe this approach complements conventional pull-down approaches and offers the following advantages. First, relatively weak (μM) interactions can be identified. Second, when compared with nonspecific chemical



c

Protein	K_d (μM)			
	H3K4Me3	H3K4Me2	H3K4Me1	H3K4Me0
SPIN1	1.08 ± 0.04	3.43 ± 0.08	29.7 ± 3.5	No binding
MORC3	0.61 ± 0.01	0.68 ± 0.02	2.45 ± 0.21	> 100

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 ± standard deviation.

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 multiprotein complexes. For example, eight subunits of the SAGA complex were detected in association with H3K4Me$_3$ peptide using a pull-down approach,[16] whereas our CLASPI approach captured the one SAGA subunit responsible for H3K4Me$_3$ binding, SGF29. On the other hand, while we identified several known H3K4Me$_3$ 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.[24] 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.

■ ASSOCIATED CONTENT

Supporting Information

Experimental method, complete ref 13, and supplementary figures and tables. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

kapoor@rockefeller.edu

Present Address
‡Department of Chemistry, The University of Hong Kong, Pokfulam Road, Hong Kong, China

■ ACKNOWLEDGMENTS

T.M.K. acknowledges support by the NIH (GM98579). B.T.C. acknowledges support by the NIH (RR00862).

■ 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.; et al. *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.; Tsuda, K.; Matsuda, T.; Aoki, M.; Seki, E.; Kobayashi, N.; Guntert, P.; Yokoyama, S.; Muto, Y. *Structure* **2010**, *18*, 1127.
(23) Hoppmann, V.; Thorstensen, T.; Kristiansen, P. E.; Veiseth, S. V.; Rahman, M. A.; Finne, K.; Aalen, R. B.; Aasland, R. *EMBO J.* **2011**, *30*, 1939.
(24) van Ingen, H.; van Schaik, F. M.; Wienk, H.; Ballering, J.; Rehmann, H.; Dechesne, A. C.; Kruijzer, J. A.; Liskamp, R. M.; Timmers, H. T.; Boelens, R. *Structure* **2008**, *16*, 1245.