



ChemComm

**Engineering a Methyllysine Reader with Photoactive Amino Acid in Mammalian Cells**

Journal:	<i>ChemComm</i>
Manuscript ID	CC-COM-05-2020-003814.R1
Article Type:	Communication

SCHOLARONE™  
Manuscripts

## COMMUNICATION

Engineering a Methyllysine Reader with Photoactive Amino Acid in Mammalian Cells<sup>†</sup>

Received 00th January 20xx,  
Accepted 00th January 20xx

Simran Arora,<sup>a</sup> Sushma Sappa,<sup>a,b</sup> Kathryn Hinkelman,<sup>a,b</sup> Kabirul Islam<sup>\*a</sup>

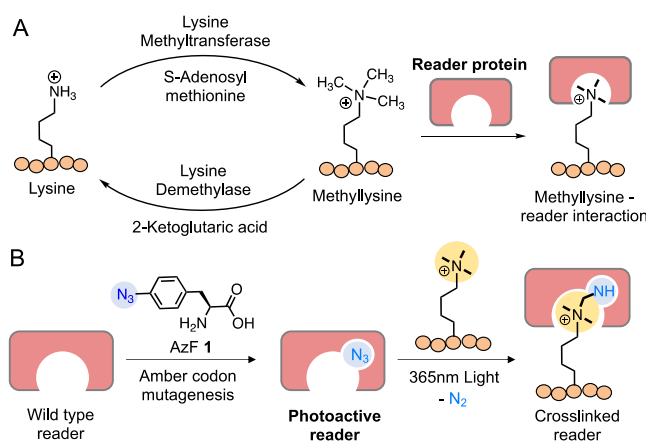
DOI: 10.1039/x0xx00000x

**Methyllysine sites in proteins are recognized by an array of reader domains that mediate protein-protein interactions for controlling cellular processes. Herein, we engineer a chromodomain, an essential methyllysine reader, to carry 4-azido-L-phenylalanine (AzF) via amber suppressor mutagenesis and demonstrate its potential to bind and crosslink methylated proteins in human cells. We further develop a first-of-its kind chromodomain variant bearing two AzF units with enhanced crosslinking potential suitable for profiling the transient methyllysine interactome.**

Site- and degree-specific methylations of  $\epsilon$ -amino group of lysine residues in histones have emerged as an important regulatory mark for controlling chromatin structure and function.<sup>1, 2</sup> While the steady-state pattern of lysine methylation is carefully maintained by reverse-acting lysine methyltransferases (KMTs) and demethylases (KDMs), its biological function is primarily manifested by the reader and effector modules that are often subunits of a multimeric protein complex (Fig. 1A).<sup>3, 4</sup> These binding domains recognize methylated lysine through cation- $\pi$  interactions in a conserved hydrophobic pocket known as the 'aromatic cage'.<sup>5, 6</sup>

The human genome encodes for more than 15 unique methyllysine readers in about 200 proteins.<sup>7</sup> These include chromodomains, plant homeodomains (PHD), Tudor domains and ankyrin repeat units, to name a few.<sup>6</sup> The reader-mediated recognition of methyllysine is highly dynamic and responsive to intrinsic and extrinsic stimuli leading to context-dependent biological outcomes.<sup>8, 9</sup> Furthermore, recent proteomic studies have identified more than 5000 methyllysine sites beyond the canonical histones.<sup>1, 10</sup> Many of these non-histone methylated sites could crosstalk synergistically or antagonistically with other modifications and are recognized by the same set of methyllysine effectors for controlling protein stability and localization.<sup>11</sup> For example, dimethylation at the evolutionarily conserved lysine residue in DNA ligase 1 (K126 in human LIG1) by G9a and GLP (G9a-like protein) is recognized by the Tandem

Tudor domain of UHRF1 for recruitment of UHRF1 to replicating genome to maintain DNA methylation in daughter cells.<sup>12</sup>



**Fig. 1.** Engineering of methyllysine reader. (A) A lysine methyltransferase (KMT) establishes site-specific methylation pattern that is recognized by an array of reader proteins. A lysine demethylase (KDM) removes the mark for steady-state composition of methyllysine. (B) Engineering of methyllysine readers to introduce a photoactive amino acid for crosslinking with interacting partners. The exact site and nature of the crosslinking are undefined.

Potential readout of 5000 methyllysine sites by 200 readers could establish an extensive array of protein-protein interactions. Structural studies have detailed the interaction between methyllysine and its readers.<sup>12, 13</sup> Due to the transient and weak nature of such associations, it has remained a significant challenge to characterize context-dependent methyllysine interactions mediated by a specific reader. The available methods based on immunoprecipitation suffer in efficiency to enrich weak interactions because the dissociation constants of the majority of such interactions are in the range of tens of micromolar. These techniques oftentimes lack temporal control as the interacting partners are netted in homogenous cellular extracts. To address such limitations, we envisioned developing photo-responsive methyllysine readers to capture weak and transient interacting partners (Fig. 1B). We recently developed photo-crosslinkable bromodomains, a reader module for acetylated lysine, to identify reader-specific

<sup>a</sup> Department of Chemistry, University of Pittsburgh, Pittsburgh, PA USA 15260  
Email: kai27@pitt.edu

<sup>b</sup> These authors contributed equally

<sup>†</sup> Electronic Supplementary Information (ESI) available. See DOI: 10.1039/x0xx00000x

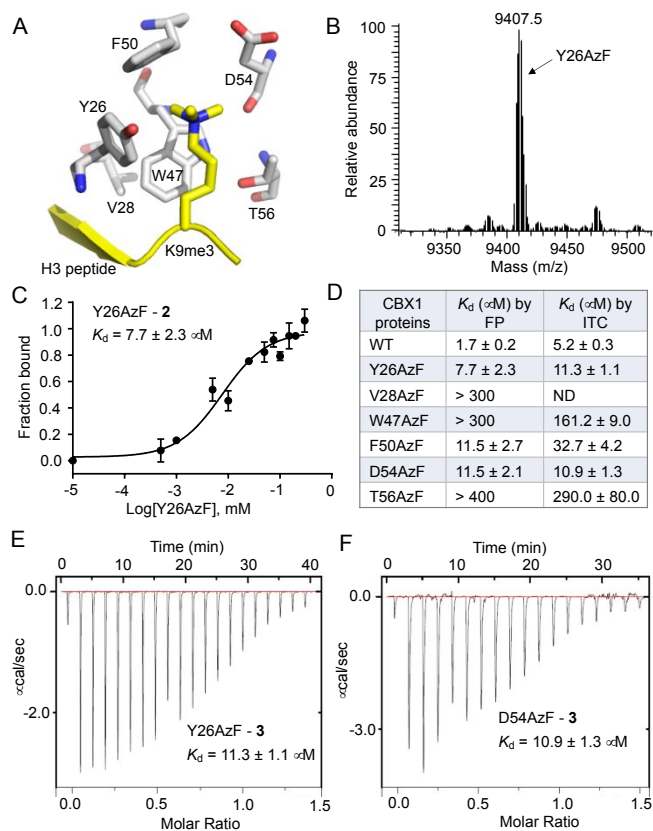
acetylome.<sup>14,15</sup> Such an approach for methyllysine readers is yet to be developed.

To access photo-crosslinkable methyllysine reader analogues, we focused on chromodomains that have shown to regulate cellular processes by recognizing methylated histone and non-histone proteins.<sup>16,17</sup> We have determined a high-resolution crystal structure (PDB: 6D07) of H3K9me3-bound chromodomain of Chromobox protein 1 (CBX1).<sup>18</sup> The structure revealed presence of a canonical hydrophobic pocket binding the Kme3 moiety via cation- $\pi$  interaction mediated by conserved aromatic residues (Fig. 2A). Earlier structural studies on chromodomains have revealed a unique mode of sequence recognition with the histone tail inserted between two  $\beta$ -strands of the chromodomain forming an antiparallel three-stranded  $\beta$ -sheet that imparts sequence selectivity such as H3K9me3 over H3K27me3 lacks connection.<sup>5, 19-21</sup> Given that engineering the amino acids involved in the  $\beta$ -sheet interactions may lead to a change in sequence preference, we specifically focused on residues in the aromatic cage. We surmised that replacing these amino acids, one at a time, with light-sensitive amino acids would lead to photo-crosslinkable CBX1 variants that could be employed to capture the CBX1 interactome.

The selected residues include Y26, V28, W47, F50, D54 and T56. We focused on 4-azido-L-phenylalanine (AzF, **1**), an analogue of phenylalanine, as a potential crosslinker because of its high crosslinking efficiency, minimal structural perturbation and availability of suitable methods for site-specific incorporation into proteins (Fig. 1B).<sup>22, 23</sup> For this purpose, we employed amber nonsense codon (TAG) mutagenesis using the evolved orthogonal *M. jannaschii* TyrRS-tRNA<sub>CUA</sub><sup>Tyr</sup> pair developed by Schultz and colleagues.<sup>24</sup> Using a bacterial expression system, we successfully obtained all the desired photo-responsive variants (Y26AzF, V28AzF, W47AzF, F50AzF, D54AzF and T56AzF) in high yields (Fig. S1). ESI LC-MS analysis confirmed the integrity of the intact proteins, each carrying one AzF unit (Fig. 2B, S2, Table S1).

To examine the ability of these mutants to bind H3K9me3 peptide, we employed a fluorescence polarization (FP) assay using a tracer peptide **2** carrying tetramethyl rhodamine (TAMRA) at the N-terminus (Fig. S3). In the assay, wild type chromodomain of CBX1 bound to the peptide with a dissociation constant ( $K_d$ ) of  $1.7 \pm 0.2 \mu\text{M}$ , similar to previously reported values.<sup>18, 21</sup> While the mutants displayed varied degrees of binding, the affinity was consistently lower than that of their wild type counterpart (Fig. 2D, S4). Among the tested mutants, Y26AzF, F50AzF and D54AzF retained significant binding towards the peptide with  $K_d$  values of  $7.7 \pm 2.3 \mu\text{M}$ ,  $11.5 \pm 2.7 \mu\text{M}$  and  $11.5 \pm 2.1 \mu\text{M}$ , respectively (Fig. 2C, D, S4). Given the structural similarity between tyrosine/phenylalanine and AzF, Y26AzF and F50AzF mutations encountered minimum steric and electronic perturbations as reflected by their binding efficiency being comparable to the wild type reader. Although the affinity of D54AzF towards the tracer peptide appears to be counterintuitive as this mutation resulted in a loss of electrostatic interactions between the carboxylic group of D54, water molecule in the aromatic cage and the trimethylammonium ion of K9me3, we reasoned that the recovery of binding is likely due to new cation- $\pi$  interaction between methylated lysine and AzF54.<sup>13, 25</sup> Consistently, we have noted that certain methyllysine readers such as PHD fingers carry an aromatic residue in that particular site, which can engage in cation- $\pi$  interaction with methyllysine.<sup>26</sup>

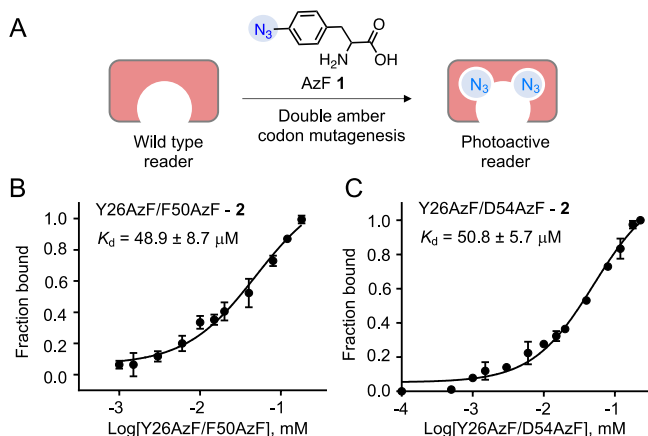
In order to obtain accurate values of the binding constants, we measured the thermodynamic parameters of binding using isothermal titration calorimetry (ITC). We synthesized a non-TAMRA peptide **3** with sequence identical to that of **2** (Fig. S3, Table S2). The  $K_d$  of wild type chromodomain with H3K9me3 peptide was measured to be  $5.3 \pm 0.4 \mu\text{M}$  (Fig. 2D, S5). We next determined dissociation constants for the selected mutants which showed moderate to strong binding in the initial FP assay. Y26AzF and D54AzF indeed exhibited strong affinity towards the peptide with  $K_d$  of  $11.3 \pm 1.1 \mu\text{M}$  and  $10.9 \pm 1.3 \mu\text{M}$ , respectively, closely matching with the data obtained through our FP assay (Fig. 2D, E, F). V28AzF, W47AzF and T56AzF mutants all showed significantly weaker binding affinity than the wild type CBX1, consistent with the fluorescence anisotropy-based results (Fig. 2D, S5). Collectively, these results obtained from two independent binding experiments using different techniques demonstrate robust and reliable binding of



**Fig. 2.** Engineering of CBX1 chromodomain with photo-crosslinkable amino acid AzF **1**. (A) Crystal structure showing chromodomain of CBX1 bound with H3K9me3 peptide (PDB: 6D07). The indicated residues are replaced, one at a time, with AzF. (B) ESI LC-MS spectrum of the Y26AzF mutant. (C) Dissociation constant of Y26AzF mutant from peptide **2** as measured by fluorescence anisotropy. (E, F) Dissociation constants of Y26AzF and D54AzF mutants, respectively, as measured by isothermal titration calorimetry.

the engineered CBX1 chromodomains carrying AzF towards methylated histone peptide.

Encouraged by the above results, we decided to explore other possibilities to engineer the CBX1 chromodomain. We noted that trimethyllysine readers carry multiple aromatic residues in the hydrophobic pocket to maximize cation- $\pi$  interaction. We reasoned that site-specific incorporation of two AzF units in the chromodomain is expected to enhance crosslinking efficiency without compromising the integrity of the aromatic cage (Fig. 3A).<sup>27</sup> We introduced amber codon (TAG) to generate Y26TAG/F50TAG and Y26TAG/D54TAG mutants by combining the variants that showed strong binding towards H3K9me3 peptide as noted above. Quite remarkably, we observed efficient readthrough of the nonsense codons by the orthogonal *M. jannaschii* TyrRS-tRNA<sub>CUA</sub><sup>Tyr</sup> pair to access the combined mutants, albeit in a lower yield than the respective single mutant. We confirmed the double incorporation of AzF in these mutants by ESI LC-MS (Fig. S6). In the FP assay, both Y26AzF/F50AzF and Y26AzF/D54AzF showed moderate binding to the peptide with  $K_d$  values of  $48.9 \pm 8.7 \mu\text{M}$  and  $50.8 \pm 5.7 \mu\text{M}$ , respectively (Fig. 3B, C). These results show that CBX1, and likely other chromodomains, can be engineered to bear multiple photoactive amino acids.



**Fig. 3.** Engineering of CBX1 chromodomain with AzF 1. (A) Schematic showing site-specific incorporation of two AzF units in CBX1 chromodomain. (B, C) Dissociation constant of the double mutants from peptide 2 as measured by fluorescence anisotropy.

We proceeded to examine if the selected mutants were capable of crosslinking with the methylated interacting partners upon exposure to ultraviolet (UV) light. A mixture of TAMRA-attached peptide 2 and individual mutants was subjected to photo irradiation and resolved in SDS-PAGE. Successful crosslinking is expected to generate a fluorescently labeled protein band in polyacrylamide gel. Indeed, Y26AzF, F50AzF and Y26AzF/F50AzF mutants underwent crosslinking with 2 only when the engineered reader and peptide complexes were exposed to 365 nm light (Fig. S7). No fluorescent band was observed for the wild type CBX1 in spite of its strong association ( $K_d \sim 2 \mu\text{M}$ ) with the peptide, thus demonstrating that crosslinking ability of the mutants must have been acquired through incorporation of AzF in the binding pocket (Fig. S7). We further noted that Y26AzF/F50AzF, although a weak binder, has

higher crosslinking efficiency than corresponding single mutants, due to the presence of two AzF groups that enhances the likelihood of crosslinking via either or both of the photoactive moieties (Fig. S7). We next synthesized TAMRA-H3K9me1 4 and TAMRA-H3K9me2 5 peptides and photo-irradiated in presence of the mutants (Fig. S3, S7). We noted a varied degree of crosslinking, consistent with the observation that CBX1 chromodomain is also capable of binding with the mono- and di-methyllysine marks.

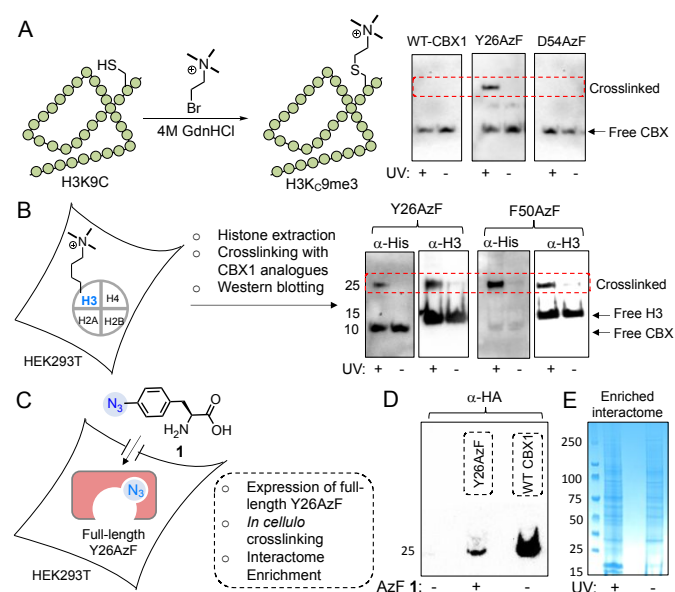
To assess the ability of the mutants to crosslink full-length proteins, we prepared semisynthetic histone H3 carrying trimethylated thialysine derivative at position 9 (H3K<sub>9</sub>me3) (Fig. 4A, S8).<sup>28</sup> This site-specifically methylated H3 protein was incubated with selected mutants, one at a time, followed by UV exposure. Western blot analysis using anti 6xHis antibody (CBX1 chromodomain carries 6xHis tag) confirmed robust crosslinking of Y26AzF with H3 only in the presence UV light as evident from the appearance of higher molecular weight protein bands (Fig. 4A). The D54AzF mutant, although a strong binder, consistently failed to crosslink with H3 as observed with peptide 3 as well, likely because of an orientation of the AzF moiety towards solvent-exposed protein surface. As expected, wild type CBX1 did not undergo crosslinking with the modified H3 (Fig. 4A).

We next tested the ability of the mutants to crosslink nucleosomal H3 which is a biologically more relevant interacting partner of reader proteins.<sup>29-31</sup> Upon UV irradiation followed by western blot analysis with anti 6xHis antibody, the selected mutants underwent varied degrees of crosslinking with histone (Fig. 4B, S8). To further confirm that the crosslinked species contained H3, we repeated the western blot using anti H3 antibody that recognizes C-terminal of human H3. We indeed observed the protein band at higher molecular weight for the successful mutants (Fig. 4B). It is worth noting that F50AzF which failed to crosslink methylated peptide 2 or the semisynthetic histone, did show noticeable crosslinking with the functionally relevant endogenous H3, likely due to enhanced association between the engineered reader and H3 (Fig. 4B). Crosslinking efficiency of the double mutant (Y26AzF/F50AzF) was comparable or marginally higher than that of the individual mutants in spite of its weaker binding capacity ( $K_d \sim 50 \mu\text{M}$ ), further demonstrating the potential benefit of two AzF groups in the hydrophobic pocket (Fig. S8). No crosslinking was noticed in samples not exposed to UV light.

Finally, we examined the expression of CBX1 analogues in cultured human cells and their ability to crosslink cellular proteins upon photoactivation (Fig. 4C). Employing the orthogonal TyrRS-tRNA<sub>CUA</sub><sup>Tyr</sup> pair developed by Sakmar and colleagues,<sup>32</sup> full-length CBX1 Y26AzF mutant was successfully expressed in HEK293T cells, only when treated with AzF, as determined by western blot with anti-hemagglutinin (HA) antibody (Fig. 4D). Subsequently, AzF-treated cells were irradiated with 365nm light, lysed and the crosslinked proteins were enriched via affinity tag. Coomassie staining of the pull-down samples revealed distinct protein bands present only in the irradiated cells (Fig. 4E). Collectively, the results demonstrate that CBX1 mutant with photoactive amino acid can be expressed in human cells and that these mutants are

capable of covalently binding with putative interacting partners upon UV irradiation.

In this work, we introduced an amino acid carrying arylazide moiety into the aromatic cage of CBX1 chromodomain. We demonstrated that setting up an azide-methyllysine photoreaction deep inside the hydrophobic pocket allows covalent bond formation between the reader and a methylated protein to help capture the interacting dimer. We also developed a first-of-its kind chromodomain variant with two AzF units introduced into the aromatic cage that could undergo robust crosslinking with nucleosomal H3 in spite of its relatively weaker affinity towards methylated peptide. Furthermore, a CBX1 mutant was expressed in human cells to capture in-cellulo binding partners for profiling context-dependent interactome. By reason of the ubiquitous presence of an aromatic cage in methyllysine readers, the current approach will find widespread application to more than 200 such reader modules in humans.



**Fig. 4.** Photo-crosslinking of engineered of CBX1 with histone and non-histone proteins. (A) Semi-synthetic H3 carrying trimethylated thialysine. Western blot with anti 6xHis antibody demonstrating successful crosslinking of semisynthetic H3 with Y26AzF mutant. (B) Crosslinking of CBX1 mutants with nucleosomal H3 isolated from HEK293T cells. Western blot with anti 6xHis and H3 antibodies showing successful crosslinking of H3 with selected mutants. (C) Schematic showing expression of full-length CBX1-Y26AzF in HEK293T cells using evolved tRNA-synthetase pair. (D) Western blot with anti HA antibody confirmed the expression of full-length Y26AzF mutant in HEK293T cells. Expression of wild type CBX1 (Mol. wt. 21.4 KDa including affinity tags) does not require AzF. (E) Coomassie gel showing putative interacting partners of CBX1 upon affinity enrichment following photoirradiation of HEK293T cells expressing Y26AzF variant.

We thank the University of Pittsburgh, the National Institutes of Health (R01GM123234, R01GM130752) and the National Science Foundation (MCB-1817692) for financial support; Prof. T. Sakmar for the evolved tRNA and synthetase pair; Dr. D. Dey, Dr. M. Waldman, S. Albright for reagents; Dr. D. Chakraborty and members of our laboratory for editing of the manuscript.

## Conflicts of interest

There are no conflicts to declare.

## Notes and references

- J. Murn and Y. Shi, *Nat. Rev. Mol. Cell Biol.* 2017, **18**, 517-527.
- T. Kouzarides, *Cell*, 2007, **128**, 693-705.
- A. J. Ruthenburg, C. D. Allis and J. Wysocka, *Mol. Cell*, 2007, **25**, 15-30.
- J. C. Black, C. Van Rechem and J. R. Whetstone, *Mol. Cell*, 2012, **48**, 491-507.
- P. R. Nielsen, D. Nietlispach, H. R. Mott, J. Callaghan, A. Bannister, T. Kouzarides, A. G. Murzin, N. V. Murzina and E. D. Laue, *Nature*, 2002, **416**, 103-107.
- S. D. Taverna, H. Li, A. J. Ruthenburg, C. D. Allis and D. J. Patel, *Nat. Struct. Mol. Biol.* 2007, **14**, 1025-1040.
- M. Yun, J. Wu, J. L. Workman and B. Li, *Cell Res.* 2011, **21**, 564-578.
- M. Vermeulen, et al. *Cell*, 2007, **131**, 58-69.
- X. Shi, et al. *Nature*, 2006, **442**, 96-99.
- S. V. Frye, *Future Med. Chem.* 2015, **7**, 1831-1833.
- K. K. Biggar and S. S. Li, *Nat. Rev. Mol. Cell Biol.* 2015, **16**, 5-17.
- L. Ferry, et al. *Mol. Cell*, 2017, **67**, 550-565.e555.
- S. A. Baril, A. L. Koenig, M. W. Krone, K. I. Albanese, C. Q. He, G. Y. Lee, K. N. Houk, M. L. Waters and E. M. Brustad, *J. Am. Chem. Soc.* 2017, **139**, 17253-17256.
- B. Sudhamalla, D. Dey, M. Breski, T. Nguyen and K. Islam, *Chem. Sci.* 2017, **8**, 4250-4256.
- S. Wagner, B. Sudhamalla, P. Mannes, S. Sappa, S. Kavooosi, D. Dey, S. Wang and K. Islam, *Chem. Commun.* 2020, **56**, 3641-3644.
- T. Cheutin, A. J. McNairn, T. Jenuwein, D. M. Gilbert, P. B. Singh and T. Misteli, *Science* 2003, **299**, 721-725.
- H. Liu, M. Galka, E. Mori, X. Liu, Y. F. Lin, R. Wei, P. Pittock, C. Voss, G. Dhami, X. Li, M. Miyaji, G. Lajoie, B. Chen and S. S. Li, *Mol. Cell*, 2013, **50**, 723-735.
- S. Arora, W. S. Horne and K. Islam, *J. Am. Chem. Soc.* 2019, **141**, 15466-15470.
- R. J. Eisert, S. A. Kennedy and M. L. Waters, *Biochemistry*, 2015, **54**, 2314-2322.
- R. Hard, N. Li, W. He, B. Ross, G. C. H. Mo, Q. Peng, R. S. L. Stein, E. Komives, Y. Wang, J. Zhang and W. Wang, *Sci. Adv.* 2018, **4**, eaau1447.
- L. Kaustov, H. Ouyang, M. Amaya, A. Lemak, N. Nady, S. Duan, G. A. Wasney, Z. Li, M. Vedadi, M. Schapira, J. Min and C. H. Arrowsmith, *J. Biol. Chem.* 2011, **286**, 521-529.
- N. D. Pham, R. B. Parker and J. J. Kohler, *Curr. Opin. Chem. Biol.* 2013, **17**, 90-101.
- T. A. Nguyen, M. Cigler and K. Lang, *Angew. Chem. Int. Ed.* 2018, **57**, 14350-14361.
- J. W. Chin, S. W. Santoro, A. B. Martin, D. S. King, L. Wang and P. G. Schultz, *J. Am. Chem. Soc.* 2002, **124**, 9026-9027.
- R. J. Eisert and M. L. Waters, *ChemBioChem* 2011, **12**, 2786-2790.
- D. J. Patel, *Cold Spring Harbor Perspec. Biol.* 2016, **8**, a018754.
- Y. Zheng, M. J. Gilgenast, S. Hauc and A. Chatterjee, *ACS Chem. Biol.* 2018, **13**, 1137-1141.
- M. D. Simon, F. Chu, L. R. Racki, C. C. de la Cruz, A. L. Burlingame, B. Panning, G. J. Narlikar and K. M. Shokat, *Cell*, 2007, **128**, 1003-1012.
- F. H. Andrews, J. Gatchalian, K. Krajewski, B. D. Strahl and T. G. Kutateladze, *ACS Chem. Biol.* 2016, **11**, 547-553.
- W. Fischle, B. S. Tseng, H. L. Dormann, B. M. Ueberheide, B. A. Garcia, J. Shabanowitz, D. F. Hunt, H. Funabiki and C. D. Allis, *Nature*, 2005, **438**, 1116-1122.
- Z. Wang and D. J. Patel, *J. Biol. Chem.* 2011, **286**, 18363-18368.
- A. Grunbeck, T. Huber, P. Sachdev and T. P. Sakmar, *Biochemistry*, 2011, **50**, 3411-3413.