Engineering Bromodomains with a Photoactive Amino Acid by Engaging 'Privileged' tRNA Synthetases

SCHOLARONE™
Manuscripts

COMMUNICATION

Engineering Bromodomains with a Photoactive Amino Acid by Engaging 'Privileged' tRNA Synthetases†

Shana Wagner,[a] Babu Sudhamalla,[a,b] Philip Mannes,[a] Sushma Sappa,[a] Sam Kavoosi,[a] Debasis Dey,[a] Sinan Wang,[a] Kabirul Islam*[a]

Site-specific placement of unnatural amino acids, particularly those responsive to light, offers an elegant approach to control protein function and capture their fleeting 'interactome'. Herein, we have resurrected 4-(trifluoromethyldiazirinyl)-phenylalanine, an underutilized photo-crosslinker, by introducing several key features including easy synthetic access, site-specific incorporation by 'privileged' synthetases and superior crosslinking efficiency, to develop photo-crosslinkable bromodomains suitable for 'interactome' profiling.

Dynamic biological processes rely on transient protein-protein and protein-nucleic acid interactions occurring in precise space and time. Gaining molecular insights into such processes requires characterization of the specific interacting partners in a physiologically relevant environment. Proteins and oligonucleotides carrying photoactive building blocks have been developed to capture transient interactions.[1] A particularly important approach includes site-specific introduction of a photo-crosslinkable amino acid (PCAA) or a photo-controlled chemical crosslinker by the read-through of amber stop codon using an evolved synthetase-tRNA pair (Fig. 1A).[2, 3,4, 5]

While PCAAs 4-Benzoylphenylalanine (BzF) **1**, 4-azidophenylalanine (AzF) **2** and 4-(trifluoromethyldiazirinyl)-phenylalanine (tmdF) **3** have been used to probe protein interactions, each possesses distinct advantages and limitations (Fig. 1B-D).[3] **1** provides the maximum degree of crosslinking owing to its regeneration potential in aqueous medium; however, having a larger size, it often interferes with the binding of interacting partners. In contrast, **2** is closest in size to phenylalanine and tyrosine, but it suffers in crosslinking efficiency due to intramolecular insertion of the nitrene intermediate resulting in ring expansion. Compared to **1** and **2**, PCAA **3**, which generates a carbene upon irradiation, is the least utilized photo-crosslinker mainly because of the challenges in synthetic access and site-specific incorporation into proteins.

Furthermore, a comparative analysis of crosslinking efficiency of these PCAAs in the context of protein-protein interaction is largely unexplored.
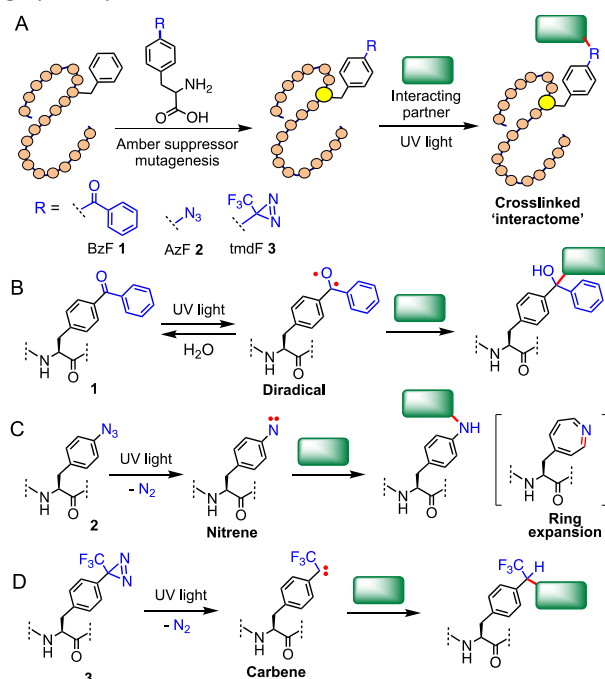

Fig. 1. Structure and reactivity of photo-crosslinkable amino acids (PCAAs). (A) Incorporation of PCAAs such as **1**, **2** and **3** at a specific site into proteins is achieved using amber suppressor mutagenesis to form a covalent bond with interacting partners upon photo-irradiation. (B) In presence of UV light, BzF generates a diradical which, in addition to undergoing crosslinking with interactome, could regenerate BzF through hydration. (C) AzF generates highly unstable nitrene which primarily decomposes to a ring-expanded inactive intermediate. (D) Upon photo-irradiation, tmdF furnishes a carbene stabilized by the phenyl and trifluoromethyl groups.

We recently developed a chemoproteomic approach called 'interaction-based protein profiling' (IBPP) by introducing **2** into bromodomain-containing protein BRD4 that recognizes acetylated histones as well as non-histone proteins such as ILF3, GATA1, RelA to regulate gene expression.[6,7] Interestingly, BRD4 carrying **1** failed to bind its cognate interacting partners, such as histone H4, likely due to increased steric bulk of **1**. To expand

a. Department of Chemistry, University of Pittsburgh, Pittsburgh, PA USA 15260
   Email: kai27@pitt.edu
b. Present address: Department of Biological Sciences, Indian Institute of Science Education and Research-Kolkata, Mohanpur, West Bengal, India 741246

the repertoire of PCAAs and compare their crosslinking efficiency, we sought to introduce **3** in a site-specific manner into a set of bromodomain-containing proteins.

Schultz and colleagues have introduced **3** into a Z domain protein using a pBK-based expression vector carrying the evolved synthetase under the constitutively active glnS promoter.[8] The suppressor tRNA and gene of interest were inserted into pLei vector under the *lpp* promoter. Schultz group later developed a pEVOL-based single vector carrying both the evolved synthetase and the suppressor tRNA under constitutive as well as inducible promoters.[9, 10] It has been shown that pEVOL-based vector leads to 250% increase in expression of the desired protein compared to the pBK-based construct. We reasoned that employing a pEVOL-based expression system for **3** would simplify its incorporation into proteins and accelerate crosslinking-based 'interactome' characterization.

We also recognized that previous attempts to access **3** involved up to twelve chemical steps, occasionally with an enzymatic hydrolysis of an acyl group, to furnish the amino acid.[11-13] We realized that use of compound **3** as a routine photo-crosslinker would benefit from an improved synthesis. Our scheme commenced from toluene **4** that was converted to trifluorodiaziridine **8** via the intermediacy of **5**, **6** and **7** in four steps following reported methods (Scheme 1).[11, 12] Although oxidation of **8** to diazirine **9** can be accomplished in multiple ways, in our hands, the previous approaches resulted in inconsistent yields often with no product formation. We found that pyridinium dichromate (PDC) oxidation of **8** was smooth and furnished **9** in excellent yield. We modified the subsequent steps, particularly the C-C bond formation between **10** and protected glycine **11** under kinetically controlled condition to afford protected, fully-assembled diazirine **12**.[14] The final deprotection proceeded smoothly to furnish (±)-**3** in eight linear steps with 10% overall yield with high purity as judged by analytical data (Scheme 1, Figure S10-22). We confirmed the structure of (±)-**3** by X-ray crystallography (CCDC# 1978317) (Fig. S1).
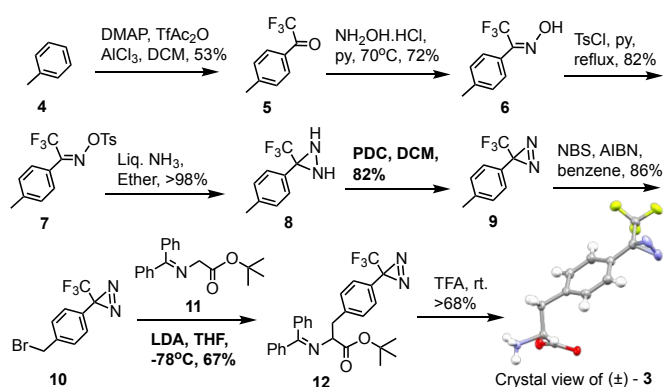
**Scheme 1**. Synthesis of (±)-tmdF **3**

After accessing **3** in a large quantity, we set out to uncover a pEVOL-based system for site-specific incorporation of **3** into proteins. It has been demonstrated that certain evolved synthetases are capable of charging their cognate tRNAs with multiple unnatural amino acids (UAAs), one at a time, while maintaining high fidelity by not recognizing the canonical amino

acids.[15] Polysubstrate specificity (often called 'permissivity') of such a 'privileged' synthetase is the result of lack of evolutionary pressure against other UAAs during in-vitro selections to evolve the synthetase.[16, 17] We envisioned exploiting the 'permissivity' of the *M. jannaschii* synthetases developed by the Schultz group to introduce **3** into bromodomain-containing proteins.

We selected a panel of four engineered synthetases and their cognate tRNAs cloned in the pEVOL-based expression system. These evolved pairs have been shown to incorporate bulky *para*-substituted tyrosine analogues including 4-acetylphenylalanine (AcF), 4-cyanophenylalanine (CNF), BzF **1** and AzF **2** into proteins.[9, 18, 19] Given the structural similarity of **3** to these analogues, we hypothesized that **3** could be accepted by some of these 'privileged' synthetases to acylate the orthogonal tRNAs with **3**. For an initial screening, we used BRD4-W81TAG mutant. Under the optimized expression condition, we observed efficient read-through of the amber codon by the CNF and AzF synthetases only in the presence of **3**
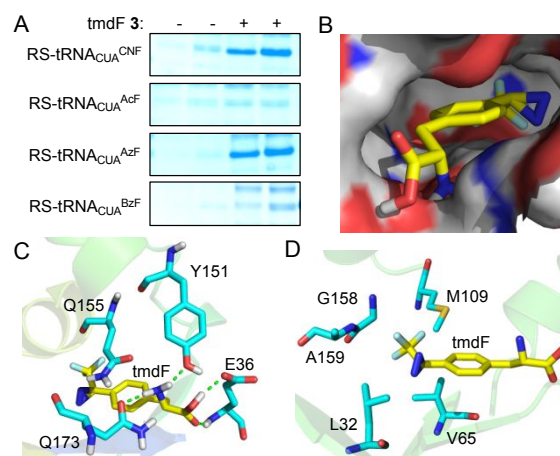
(Fig. 2A, S2).

Fig. 2. Efficient incorporation of **3** into BRD4 bromodomain by 'privileged' synthetases. (A) Expression of BRD4-W81tmdF mutant by a set of evolved *M. jannaschii* TyrRS-tRNA$_{CUA}^{Tyr}$ pairs either in absence or presence of tmdF **3** as judged by Coomassie blue staining (N=2). (B) Energy-minimized structure of S enantiomer of **3** in the catalytic pocket of evolved CNF synthetase. (C) Polar contacts between tmdF **3** and the active site residues. (D) Binding of tmdF in the conserved, extended hydrophobic pocket of evolved synthetase.

It has been shown that CNF-specific tRNA synthetase exhibits broader substrate scope due to a larger binding pocket, while maintaining its ability to discriminate against the 20 canonical amino acids.[15] To further understand the binding of **3** into this 'privileged' synthetase, we performed a docking study using Autodock energy minimization software. For this purpose, we used the coordinates from CNF-bound synthetase crystal structure (PDB 3qe4). In the most stable conformations, both the enantiomers of **3** occupy the CNF binding site and forms an array of polar contacts, similar to CNF, with residues in the active site (Fig. 2B, C S3). The CF$_3$ group is situated in the hydrophobic pocket created by L32, V65, M107, G158 and A159 (Fig. 2D). The mode of binding is consistent with the observation that these residues are important for accommodating UAAs carrying *para*-substituted phenylalanine analogues.[15]

COMMUNICATION

We next examined generality of the approach and established **3** as an efficient photo-crosslinker. The bromodomain and PHD finger containing transcription factor (BPTF) has been shown to remodel chromatin structure.[20] The bromodomain and the PHD finger bind acetylated histone H4 and methylated histone H3, respectively (Fig. 3A). We recently developed a BPTF bromodomain analogue carrying AzF **2** at 2950th position and demonstrated its ability to crosslink histone variant H2A.Z.[21] We sought to incorporate **3** into this bromodomain using the 'privileged' synthetase, examine the binding potential of engineered BPTF to histone protein and evaluate the effectiveness of **3** as a photo-crosslinker.
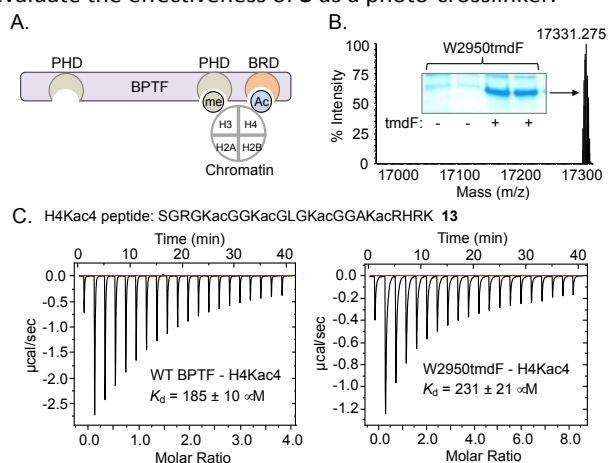

Fig. 3. Incorporation of **3** into BPTF bromodomain by evolved synthetases. (A) Domain structure of BPTF. (B) Bacterial expression of BPTF-W2950tmdF mutant as observed with Coomassie blue staining and confirmed by ESI-HRMS. (C) Dissociation constants (K_d) of BPTF and its W2950tmdF mutant from tetra-acetylated H4Kac4 peptide **13** as measured by ITC.

An amber stop codon (TAG) was introduced to replace W2950 in the BPTF bromodomain. This residue is situated in the aromatic cage known to house acetylated lysines. We observed efficient incorporation of **3** into W2950 site by both CNF- and AzF-specific synthetases (Fig. 3B, S4). Integrity of the mutant protein was confirmed by LC-HRMS (Fig. 3B, S5). This result demonstrates that incorporation of **3** into multiple proteins (BRD4, BPTF and likely others) is achievable using the 'privileged' synthetases evolved to incorporate related UAAs.

We then tested the BPTF analogue for its ability to bind interacting partners. To compare the binding and crosslinking efficiency, we also generated W2950AzF mutant carrying **2** using AzF-specific synthetase and characterized by LC-HRMS (Fig. S5). Using a synthesized tetracetylated histone H4 peptide **13** (Fig. 3C, S6), we measured dissociation constants (K_d) of the BPTF proteins from the peptide using isothermal titration calorimetry (ITC). Wild-type BPTF bound the peptide with a K_d value of 185±10 µM, closely agreeing with reported values,[6, 20] while the mutants displayed varied degrees of binding (Fig. 3C, S7). W2950AzF and W2950tmdF mutants dissociated from the peptide with K_d of 309±23 µM and 231±21 µM, respectively (Fig. 3C, S7).

Next, to examine photo-crosslinking efficiency of the BPTF analogues, we synthesized tetracetylated histone H4 peptide **14** carrying a tetramethylrhodamine (TAMRA) moiety at the N terminal (Fig. 4A, S6). The tracer peptide was incubated

separately with wild type and mutant BPTF proteins and then exposed to UV light at 365nm wavelength. Subsequent in-gel fluorescence experiment revealed a protein band corresponding to the crosslinked species of BPTF mutant and the peptide only in the presence of UV light (Fig. S8). This crosslinking ability of the mutants must have been acquired through incorporation of tmdF in the binding pocket because fluorescent labeling of wild type BPTF was undetectable in spite of its strong affinity towards the peptide (Fig. S8).

A careful time-dependent experiment revealed that tmdF-containing BPTF underwent enhanced crosslinking compared to the AzF counterpart over a period of photo-irradiation (Fig. 4B, C, S8). This is likely due to an extremely short-lived nitrene intermediate generated from photolysis of the azido moiety, making pAzF a less effective crosslinker (Fig. 1C).[3] In contrast, tmdF generates a benzyl radical which could gain further stability by the trifluoromethyl group (Fig. 1D). Furthermore, TAMRA-attached peptide **15** with a sequence identical to **14,** but lacking the acetyl moieties, did not undergo crosslinking with the BPTF mutants (Fig. S8). To the best of our knowledge, this is the first demonstration of comparative crosslinking efficiency of **2** and **3** in context of protein-protein interaction.
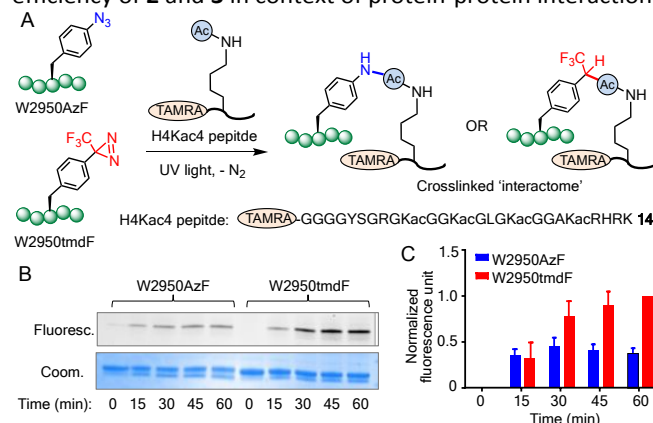

Fig. 4. Crosslinking of BPTF analogues with histone peptide. (A) Schematic depicting binding of engineered BPTF proteins to TAMRA-H4Kac4 peptide **14** followed by crosslinking. Fluorescently labeled crosslinked BPTF was visualized by in-gel fluorescence. (B) In-gel fluorescence showing crosslinking of BPTF analogues to the TAMRA-H4Kac4 peptide. Coomassie staining of the same gel showed presence of proteins in all the samples (For biological replicates, see Fig. S8C). (C) Relative quantification of crosslinking, based on the results shown in 4B and Fig. S8C, represented as normalized fluorescence intensity (Error bar = standard deviation).

We next sought to probe if the engineered BPTF is capable of interacting with a full-length histone protein. To access site-specifically acetylated histone H4, we employed a thermal thiol-ene reaction to produce acetylated thialysine and synthesized full-length H4Kc8ac and H4Kc12ac proteins (Fig. 5A, B, S9).[22] The synthetic full-length H4 proteins were allowed to bind with W2950tmdF mutant followed by UV-irradiation. The crosslinked species were separated in SDS-PAGE and visualized using anti-H4 and anti-6xhis antibodies. We observed chemiluminescent signals at a molecular weight higher than that of the individual proteins (H4 and BPTF) when irradiated with UV light, confirming successful crosslinking (Fig. 5C). No crosslinking was observed in absence of photo-irradiation or when the individual proteins were exposed to UV light. Finally, we extracted

endogenous hyperacetylated histones from HEK39T cells and subjected them to crosslinking with the BPTF analogues. Western blotting with anti-H4 antibody indeed showed crosslinking, albeit to a lesser extent, for both the mutants with W2950tmdF having marginally increased crosslinked signal compared to the W2950AzF variant (Fig. 5D). Collectively, these results demonstrate that the BPTF analogue with site-specifically introduced PCAA **3** is capable of interacting and establishing covalent bond with full-length histone H4 upon irradiation by UV light.
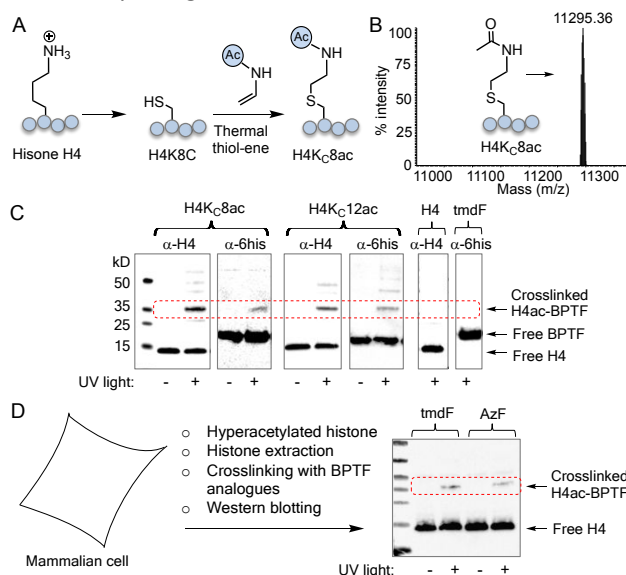

Fig. 5. Crosslinking of BPTF-W2950tmdF with full-length histone. (A) Scheme showing the synthesis of H4 carrying site-specific acetylated thialysine (K_Cac) using thermal thiol-ene reaction (Ref. 22). (B) ESI LC-HRMS spectra of full-length H4K_C8ac. (C) Western blot showing crosslinking of BPTF-W2950tmdF to H4K_C8ac and H4K_C12ac (dotted box) using anti-H4 and anti-6xhis antibodies. (D) Crosslinking of hyperacetylated H4 isolated from HEK293T cells with BPTF mutants as confirmed by H4 antibody.

In this work, we established tmdF as an excellent photo-sensitive amino acid which has remained largely underutilized in 'interactome' profiling. We repurposed a set of 'privileged' synthetases to incorporate tmdF into bromodomain containing proteins. This approach to expand substrate polyspecificity of an already evolved synthetase has the benefit of accessing engineered translational machinery without having to perform additional steps involved in directed evolution. We also improved the existing synthetic strategy towards tmdF by modifying a few critical steps and reducing the number of protection-deprotection maneuvers. Finally, we demonstrated the utility of tmdF group in crosslinking acetylated histone protein and its superior crosslinking ability over arylazide functionality. These improvements coupled with the recent finding that tmdF could be introduced into mammalian cells[23] would accelerate the application of this particular amino acid for profiling transient interacting partners of the bromodomains and other chromatin modifiers.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1. N. D. Pham, R. B. Parker and J. J. Kohler, *Curr. Opin. Chem. Biol.* 2013, **17**, 90-101.
2. H. Neumann, P. Neumann-Staubitz, A. Witte and D. Summerer, *Curr. Opin. Chem. Biol.* 2018, **45**, 1-9.
3. T. A. Nguyen, M. Cigler and K. Lang, *Angew. Chem. Int. Ed.* 2018, **57**, 14350-14361.
4. Y. Tian, M. P. Jacinto, Y. Zeng, Z. Yu, J. Qu, W. R. Liu and Q. Lin, *J. Am. Chem. Soc.* 2017, **139**, 6078-6081.
5. J. Liu, S. Li, N. A. Aslam, F. Zheng, B. Yang, R. Cheng, N. Wang, S. Rozovsky, P. G. Wang, Q. Wang and L. Wang, *J. Am. Chem. Soc.* 2019, **141**, 9458-9462.
6. B. Sudhamalla, D. Dey, M. Breski, T. Nguyen and K. Islam, *Chem. Sci.* 2017, **8**, 4250-4256.
7. J. P. Lambert, et al. *Mol. Cell*, 2019, **73**, 621-638.e617.
8. E. M. Tippmann, W. Liu, D. Summerer, A. V. Mack and P. G. Schultz, *ChemBioChem*. 2007, **8**, 2210-2214.
9. T. S. Young, I. Ahmad, J. A. Yin and P. G. Schultz, *J. Mol. Biol.* 2010, **395**, 361-374.
10. A. Chatterjee, S. B. Sun, J. L. Furman, H. Xiao and P. G. Schultz, *Biochem.* 2013, **52**, 1828-1837.
11. M. Nassal, *J. Am. Chem. Soc.* 1984, **106**, 7540-7545.
12. G. Baldini, B. Martoglio, A. Schachenmann, C. Zugliani and J. Brunner, *Biochem.* 1988, **27**, 7951-7959.
13. K. Masuda, A. Koizumi, T. Misaka, Y. Hatanaka, K. Abe, T. Tanaka, M. Ishiguro and M. Hashimoto, *Bioorg. Med. Chem. Lett.* 2010, **20**, 1081-1083.
14. H. Nakashima, M. Hashimoto, Y. Sadakane, T. Tomohiro and Y. Hatanaka, *J. Am. Chem. Soc.* 2006, **128**, 15092-15093.
15. D. D. Young, T. S. Young, M. Jahnz, I. Ahmad, G. Spraggon and P. G. Schultz, *Biochem.* 2011, **50**, 1894-1900.
16. S. J. Miyake-Stoner, C. A. Refakis, J. T. Hammill, H. Lusic, J. L. Hazen, A. Deiters and R. A. Mehl, *Biochem.* 2010, **49**, 1667-1677.
17. A. L. Stokes, S. J. Miyake-Stoner, J. C. Peeler, D. P. Nguyen, R. P. Hammer and R. A. Mehl, *Mol. BioSys.* 2009, **5**, 1032-1038.
18. 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.
19. J. W. Chin, A. B. Martin, D. S. King, L. Wang and P. G. Schultz, *Proceed. Natl. Acad. Sci. USA* 2002, **99**, 11020-11024.
20. A. J. Ruthenburg, H. Li, T. A. Milne, S. Dewell, R. K. McGinty, M. Yuen, B. Ueberheide, Y. Dou, T. W. Muir, D. J. Patel and C. D. Allis, *Cell*, 2011, **145**, 692-706.
21. G. T. Perell, N. K. Mishra, B. Sudhamalla, P. D. Ycas, K. Islam and W. C. K. Pomerantz, *Biochem.* 2017, **56**, 4607-4615.
22. A. Dhall, S. Wei, B. Fierz, C. L. Woodcock, T. H. Lee and C. Chatterjee, *J. Biol. Chem.* 2014, **289**, 33827-33837.
23. N. Hino, M. Oyama, A. Sato, T. Mukai, F. Iraha, A. Hayashi, H. Kozuka-Hata, T. Yamamoto, S. Yokoyama and K. Sakamoto, *J. Mol. Biol.* 2011, **406**, 343-353.