



Facile semisynthesis of histone H3 enables nucleosome probes for investigation of histone H3K79 modifications

Kun Zou^{a,b}, Yihang Xiao^{a,b}, Jinyu Yang^{a,b}, Mingxuan Wu^{b,c,*}

^a Department of Chemistry, Zhejiang University, Hangzhou 310027, China

^b Department of Chemistry, School of Science, Westlake University, Hangzhou 310030, China

^c Institute of Natural Sciences, Westlake Institute for Advanced Study, Hangzhou 310024, China

ARTICLE INFO

Article history:

Received 19 November 2023

Revised 29 December 2023

Accepted 3 January 2024

Available online 6 January 2024

Keywords:

Histone

H3K79 modifications

Methylation

Succinylation

Semisynthesis

Nucleosome probes

ABSTRACT

Histone H3K79 modifications are essential to regulate chromatin structure and gene transcription, but understanding of the molecular mechanisms is limited. Because H3K79 is at globular domain, short histone peptide cannot mimic H3K79 in chromatin. Instead, reconstituted nucleosome-based chemical tools are ideally used to investigate H3K79 modifications. In consequence, H3K79-modified histone H3 with additional chemical handles are required, but such synthesis is challenging and laborious. Here we report a facile semisynthesis method that enables multifunctional histone H3 readily available. H3K79-containing fragment is short for straight peptide synthesis that was later ligated to recombinant expressed H3 fragments for full-length product in large scale. As a result, nucleosomes with H3K79 modifications as well as photo-reactive group and affinity tag were obtained to investigate potential binding proteins. We believe this method that enhances synthetic accessibility of nucleosome probes will accelerate understanding of the underexplored H3K79 modifications.

© 2024 Published by Elsevier B.V. on behalf of Chinese Chemical Society and Institute of Materia Medica, Chinese Academy of Medical Sciences.

Histone modifications, such as lysine methylation and lysine acetylation, are important to regulate gene transcription and epigenetic status [1,2]. Dysregulation of histone modification is frequently found in cancer cells and there are U.S. Food and Drug Administration (U.S. FDA)-approved anti-cancer drugs that inhibit histone-modifying enzymes [3,4]. Therefore, studies of histone modification functions at molecular level are important for basic research and drug development.

H3K79 modifications are widely distributed in chromatin. H3K79 methylation is a highly conserved lysine methylation from yeast to human [5,6]. In mixed lineage leukaemia (MLL), H3K79 methylation level is upregulated due to fusion proteins of MLL, AF10, DOT1L, etc. [7]. As a result, inhibition of DOT1L as H3K79 methyltransferase is under development for cancer therapy [8]. Menin is the only known reader of H3K79me2 that was identified very recently [9]. In addition to methylation, there are H3K79 acetylation, succinylation, formylation and ubiquitination that were identified by mass spectrometry [10–13]. Lu and coworkers revealed that H3K79 succinylation causes cancer cell proliferation and tumor growth by alternation of gene expression that occurred in more than 7000 genes [14]. Although the significance of H3K79

modifications has been demonstrated, there is still much unknown about the binding proteins and molecular mechanisms from the modifications to phenotypes.

The major reason for slow progress of H3K79 modification research is a lack of chemical tools. Histone N-tail is usually unstructured, so short histone peptide could partially mimic the tail in chromatin [15]. Readily available peptide-based chemical tools have been widely used in histone research of sites including H3K4, H3K9 and H3K27 [16–20]. H3K79, however, is in globular domain with a well-folded structure. Synthetic peptides cannot mimic the local chemical environment because the residues are usually at an interface among histone and DNA. Instead, nucleosome-based tools are ideal for mechanistic study, but it requires full-length histone with site-specific modifications as well as analytical handle and reactive warhead.

Protein chemists have developed some methods to achieve functional full-length histone. Although chemoselective alkylation [21,22] and genetic code expansion [23–25] enable site-specific modified histone H3, it is challenging to incorporate additional chemical handles for functional probes (Figs. 1A and B) [26,27]. Protein total synthesis could overcome the limitation by ligation of synthetic peptides [28–30]. However, it is laborious for sequential ligation of four fragments, two of which are more than 45 amino acids (Fig. 1C). Recently, a semisynthesis method was reported for

* Corresponding author.

E-mail address: wumingxuan@westlake.edu.cn (M. Wu).

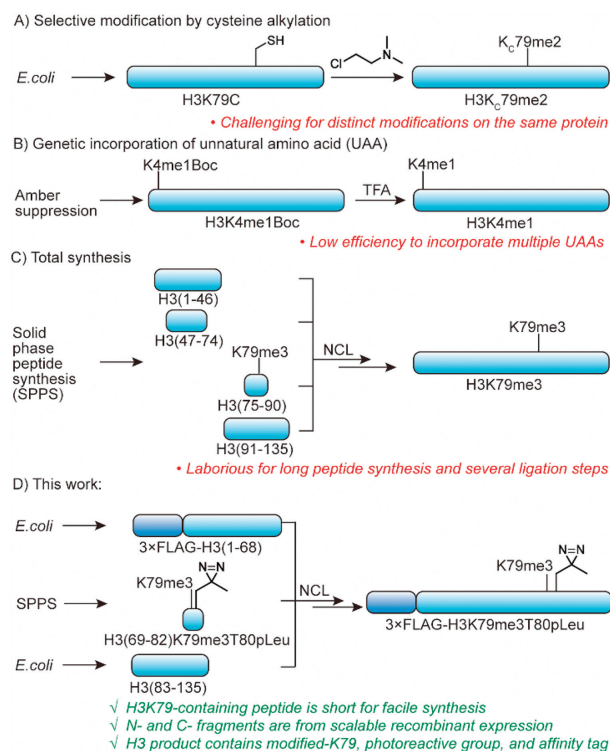


Fig. 1. Methods to prepare site-specifically modified histone H3 in previous studies and in this work. (A) Site-specific post translational modifications (PTMs) like methyllysine could be achieved via cysteine alkylation, but it is challenging to introduce different modifications on the same polypeptides. (B) Genetic code expansion enables site-selective incorporation of unnatural amino acids (UAA) but difficult for distinct UAA on the same protein. (C) Histone H3 with diverse UAA could be prepared by total synthesis, but synthesis and ligation of four fragments are laborious with low efficiency. (D) In this study, we proposed a facile semisynthesis method by ligation of a short peptide fragment and two recombinant expressed fragments to obtain K79-modified histone H3 with additional chemical handles.

histone H3 with K79me2 and proximate diazirine, but it is limited to introduce chemical handle at Asp81 position [9]. Here we aim to establish a general semisynthesis method of H3K79-modified histone H3 that offers a convenient way for investigation of H3K79 modifications using multi-functional nucleosome probes.

Since K79 is in the middle of histone H3, ideally, the K79-containing fragment is short for facile peptide synthesis and the other fragments are readily available in large scale. However, in the sequence of histone H3, there is no Cys or Ala (from Cys desulfurization) to allow short fragment (<15 aa) with K79 for native chemical ligation. As a result, a different fragmentation strategy is necessary. By evaluating histone H3 sequence, we found that there are multiple arginine residues close to H3K79 (Fig. 2A). Since Cys could be alkylated to Arg mimic [31], we proposed to select arginine as ligation site for a new synthetic route. Because the bulky isopropyl side chain may cause inefficient native chemical ligation (NCL) between Val71 and Arg72Cys [32], we designed a new semisynthetic route with Arg69 and Arg83 as ligation site for some advantages (Fig. 1D). First, the fragment of H3 (69–82) only contains 14 amino acids so the synthesis of diverse H3K79-modified peptides with chemical handle is expected to be simple. Second, the N-H3 (1–68) and C-H3 (83–135) fragments could be recombinantly expressed in large scale to avoid length limitation of peptide synthesis. The C-thioester of the N-H3 fragment is available from intein-mediated splicing, and the N-Cys of C-H3 fragment is exposed from cleavage of small ubiquitin-like modifier (SUMO). Third, after the ligations, Cys69 and Cys83 could be converted to arginine mimic so the semisynthetic histone H3 could be func-

tional for nucleosome reconstitution. Fourth, 3xFLAG as a purification and analytical tag is added to N-H3 fragment by recombinant expression.

We hence started peptide synthesis of the short H3 (69–82) fragment with N-thiopropine (Thz) and C-3,4-diaminobenzoic acid (Dbz) for sequential native chemical ligation (Fig. 2B). Because we are interested in identification of binding proteins of H3K79 modifications, additional photo-reactive group was added to the peptide. Photo-leucine (pLeu) has been applied to crosslinking histone binding proteins [33], so Fmoc-pLeu-OH was applied in solid-phase peptide synthesis (SPPS) near K79. After the synthesis of peptide and functionalization of Dbz to benzimidazole-2-one-5-carboxylic acid (Nbz) [34], the cleaved peptide intermediate was subsequently treated with sodium 2-mercaptoethanesulfonate (MesNa) to yield Thz-H3 (70–82) T80pLeu-MesNa (**a2**) with 36% yield from initial resin.

The H3 C-fragment was prepared from recombinant expression of 6xHis-SUMO-Cys-H3 (84–135) and subsequent Ulp1-mediated cleavage. Next, the peptide **a2** and the C-fragment **a6** were ligated and the N-terminal Thz was later converted to Cys for the next ligation. Meanwhile, the H3 N-fragment was prepared from recombinant expression and purification of 3xFLAG-H3 (1–68)-intein-CBD followed by thiolysis with MesNa [35]. The released fragment **a4** was applied for the next ligation with **a8** and the resulting full-length H3 was treated with acrylamidine to transform both Cys69 and Cys83 to arginine mimic (R_c) (Fig. 2C). The semisynthetic H3 product **a10** was demonstrated by electrospray ionization time-of-flight mass spectrometer (ESI-TOF) (Fig. 2D), that also indicated that diazirine group was stable through the synthesis.

Although such alkylated cysteine has been shown as a good arginine mimic in histone H4 semisynthesis [31,36], we still decided to test whether R_c69 and R_c83 interfere nucleosome reconstitution. The semisynthetic 3xFLAG-H3T80pLeu was mixed with the other core histones for octamer refolding followed by nucleosome reconstitution with 145 bp 601 DNA. Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) revealed correct histone composition of the nucleosome core particle (NCP). In addition, the mobile shift in native gel and the retention time in anion exchange chromatography are the same as the NCP with natural H3 sequence (Fig. 2E). Thus, we conclude the facile semisynthesis method could be widely applied to diverse modified H3 for research at nucleosome level.

With the initial success, we continued to prepare more histone H3 with K79 modifications. Because readers and erasers of H3K79 modifications are not much known, we selected K79 trimethylation and succinylation for exploration. Fmoc-Lys(me3)-OH is commercially available but SPPS building block of lysine succinylation is not available. Thus, we developed a robust synthetic route that started from Fmoc-Lys(Boc)-OH with 4 steps and 62% overall yield (Fig. 3A). Next, the building block of Kme3 and Ksucc were applied in SPPS and the resulting peptides were applied in the following ligations (Fig. 3B). The corresponding H3K79-modified nucleosomes were obtained from standard procedure and the quality was confirmed by SDS-PAGE and native gels (Fig. 3C).

With the reconstituted nucleosomes in hands, we set a protein profiling that bind to H3K79me3 or H3K79succ. Nuclear proteins were extracted from HeLa cells and treated with **NCP-a** or **NCP-b** or **NCP-c** as three experiment groups A, B and C. After ultraviolet (UV) irradiation to induce photo crosslinking, the resulting mixtures were applied to anti-FLAG-resin. The remaining proteins on resin after washing were digested by trypsin for mass spectrometry-based label-free quantitative proteomics (Fig. 4A).

In the volcano plots, PAF1, RPS27L, SET and ZN316 were found to be enriched by the H3K79me3 NCP probe (group B), and Lamin-B1 was enriched by the H3K79succ NCP probe (group C) (Fig. 4B). To gain a comprehensive understanding of the inter-

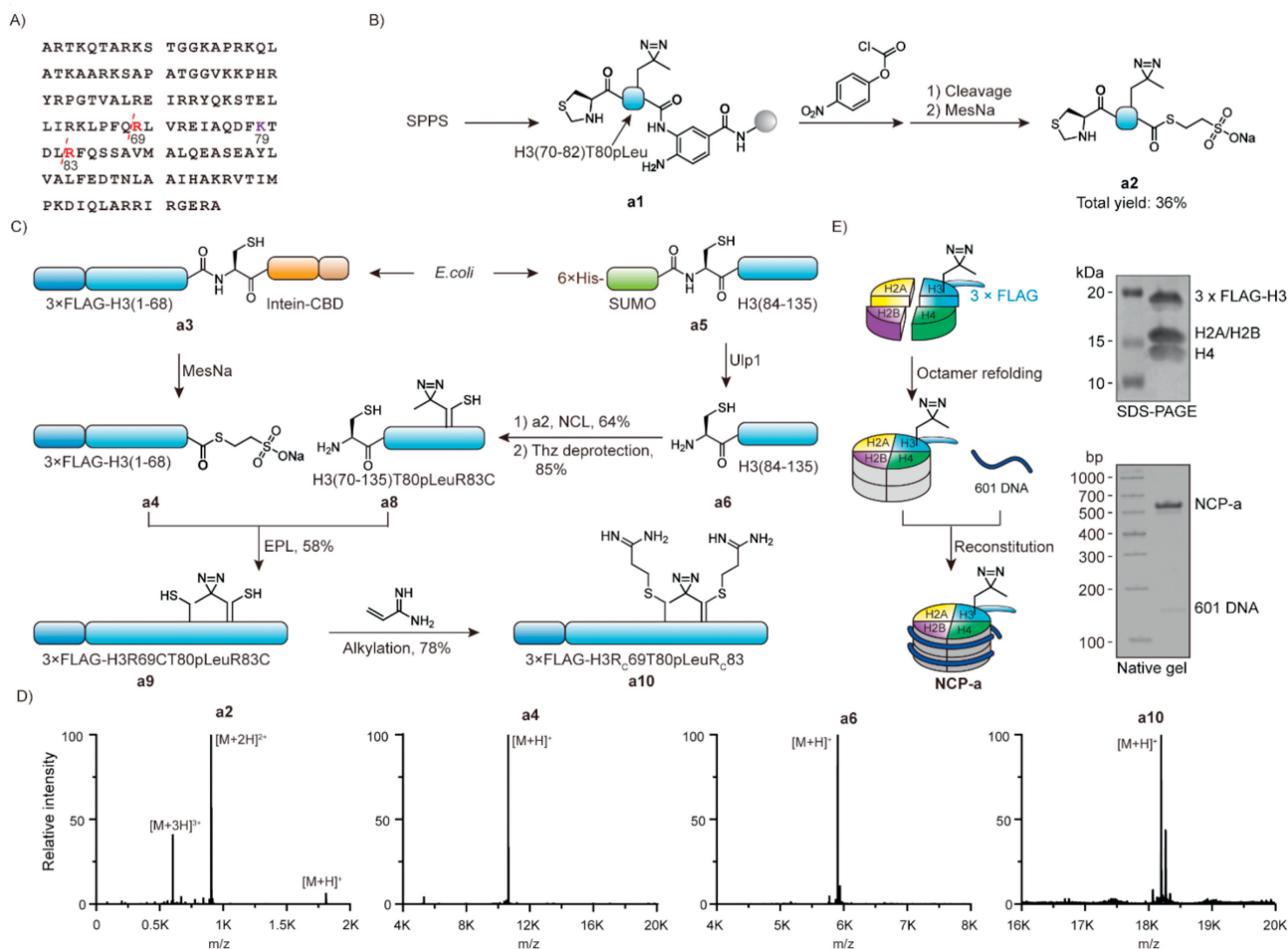


Fig. 2. Establishment of new histone H3 semisynthesis method and subsequent nucleosome reconstitution. (A) The amino acid sequence of *X. laevis* histone H3C110A that is widely used in histone research. The ligation sites at Arg69 and Arg83 are highlighted in red and Lys79 is highlighted in purple. (B) Synthesis of the H3K79-containing short peptide fragment **a2**. (C) Semisynthesis of histone H3 sequential ligation from synthetic **a2** and recombinant expressed **a4** and **a6**. (D) High resolution mass spectrometry analysis of the key intermediates in the semisynthesis. (E) Nucleosome reconstitution using **a10** *in vitro*. The quality was characterized by SDS-PAGE with Coomassie blue staining and native gel with 4S Green Plus nucleic acid staining.

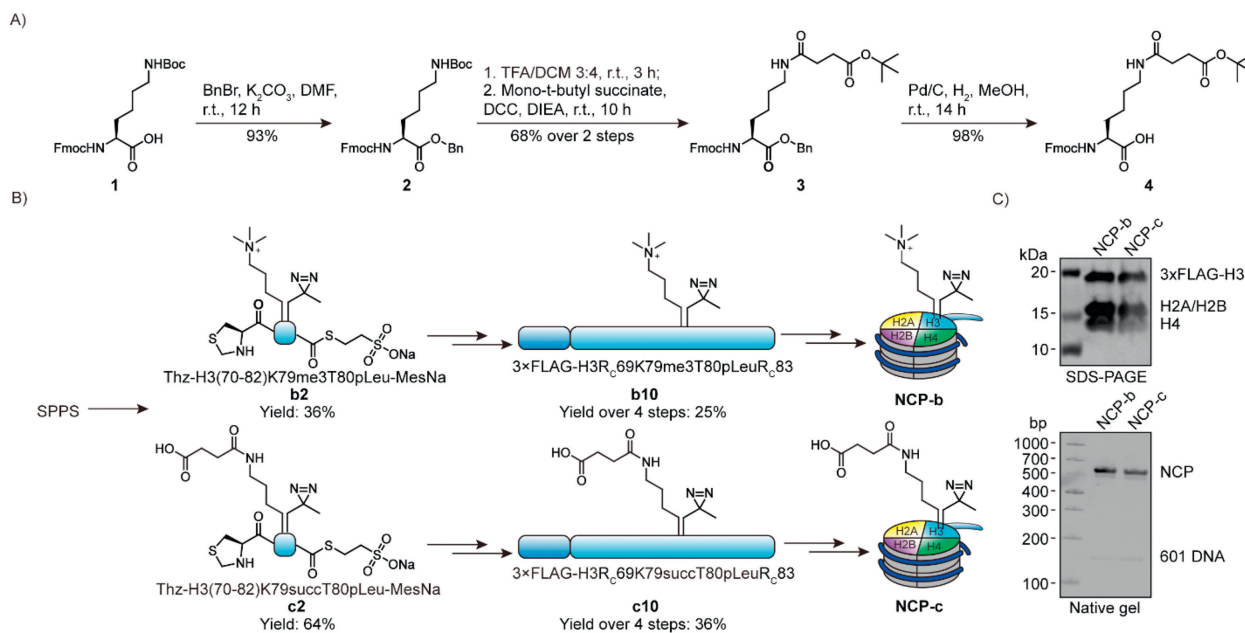


Fig. 3. Preparation of photo-reactive nucleosome probes with H3K79 trimethylation and succinylation. (A) Synthetic route of **4** as a solid phase peptide synthesis (SPPS) building block for lysine succinylation-containing peptide. (B) Scheme toward nucleosome probes with H3K79 trimethylation (**NCP-b**) and succinylation (**NCP-c**). (C) Characterization of **NCP-b** and **NCP-c** by SDS-PAGE and native gel.

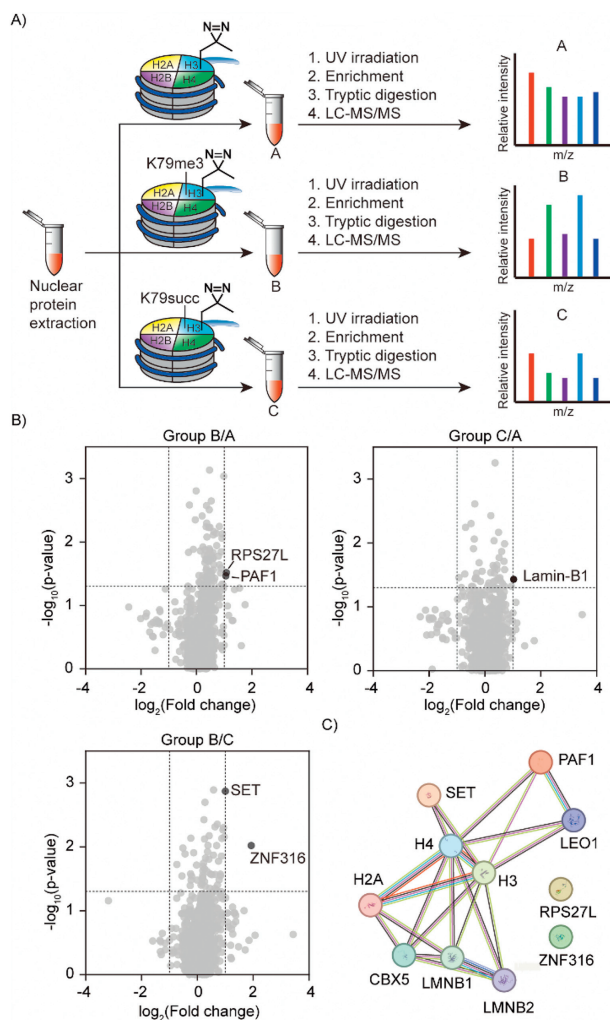


Fig. 4. Investigation of binding proteins of H3K79 modifications using nucleosome probes. (A) Nuclear protein extracts from HeLa cells are treated with each NCP probe in parallel for three groups. After UV irradiation, the crosslinked proteins were enriched via 3xFLAG tag and digested by trypsin before liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis. (B) Volcano plots to analyze H3K79me3 or H3K79succ specific binding proteins. (C) Protein-protein interaction network of the identified proteins by STRING database.

actions between the identified protein hits and H3K79-modified nucleosomes, we employed the Search Tool for the Retrieval of Interaction Gene/Proteins (STRING) database to visualize the interaction network (Fig. 4C) [37]. PAF1, SET and Lamin-B1 were found to have direct interactions with histone H3. Among these proteins, PAF1 is a core subunit of PAF1 complex, that interact with DOT1L to influence the H3K79 methyltransferase activity [38,39]. The PAF1 complex has been found to recruit DOT1L to chromatin for H3K79 methylation. SET is a multifunctional protein, involved in processes such as apoptosis, transcription, nucleosome assembly and histone chaperoning [40,41]. Lamin-B1 is close to inner nuclear membrane [42]. Therefore, they are potential H3K79me3 or H3K79succ binding proteins and further investigations will be carried out. Since there were only a few positive proteins appeared in the volcano plots, different placement of pLeu on semisynthetic histone H3 will be applied for a better comprehensive coverage of binding proteins in the future.

In summary, we have established a histone H3 semisynthesis method that enables efficient preparation of site-specific modified histone H3 for research of H3K79 modifications. The multi-

functional nucleosomes allowed us to investigate H3K79 methylation and succinylation and we believe research of other significant H3K79 modifications also could benefit from our chemical method.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the support from National Natural Science Foundation of China (Nos. 22077103 and 22161132006) and Westlake University startup. We acknowledge the Instrumentation and Service Center for Molecular Sciences for the instrument support. We also thank the Mass Spectrometry & Metabolomics Core Facility of the Biomedical Research Core Facilities at Westlake University.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ccllet.2024.109497.

References

- [1] M. Shvedunova, A. Akhtar, *Nat. Rev. Mol. Cell Biol.* 23 (2022) 329–349.
- [2] D. Husmann, O. Gozani, *Nat. Struct. Mol. Biol.* 26 (2019) 880–889.
- [3] A. Mullard, *Nat. Rev. Drug Discov.* 19 (2020) 156.
- [4] Y.X. Li, E. Seto, *Cold Spring Harb. Perspect. Med.* 6 (2016) a026831.
- [5] Q. Feng, H.B. Wang, H.H. Ng, et al., *Curr. Biol.* 12 (2002) 1052–1058.
- [6] F. van Leeuwen, P.R. Gafken, D.E. Gottschling, *Cell* 109 (2002) 745–756.
- [7] A.V. Krivtsov, S.A. Armstrong, *Nat. Rev. Cancer* 7 (2007) 823–833.
- [8] S. Ribich, D. Harvey, R.A. Copeland, *Cell Chem. Biol.* 24 (2017) 1120–1147.
- [9] J.W. Lin, Y.P. Wu, G.F. Tian, et al., *Science* 379 (2023) 717–723.
- [10] B.A. Garcia, S.B. Hake, R.L. Diaz, et al., *J. Biol. Chem.* 282 (2007) 7641–7655.
- [11] J.R. Wisniewski, A. Zougman, M. Mann, *Nucleic Acids Res.* 36 (2008) 570–577.
- [12] Z.H. Zhang, M.J. Tan, Z.Y. Xie, et al., *Nat. Chem. Biol.* 7 (2011) 58–63.
- [13] X. Zhang, B.K. Li, A.H. Rezaeian, et al., *Nat. Commun.* 8 (2017) 14799.
- [14] Y.G. Wang, Y.R. Guo, K. Liu, et al., *Nature* 552 (2017) 273–277.
- [15] R.K. McGinty, S. Tan, *Chem. Rev.* 115 (2015) 2255–2273.
- [16] X. Li, E.A. Foley, K.R. Molloy, et al., *J. Am. Chem. Soc.* 134 (2012) 1982–1985.
- [17] T.P. Yang, X.M. Li, X.C. Bao, Y.M.E. Fung, X.D. Li, *Nat. Chem. Biol.* 12 (2016) 70–72.
- [18] A.J. Bannister, P. Zegerman, J.F. Partridge, et al., *Nature* 410 (2001) 120–124.
- [19] M. Lachner, N. O'Carroll, S. Rea, K. Mechtler, T. Jenuwein, *Nature* 410 (2001) 116–120.
- [20] X.C. Guo, Y.N. Wang, Y.H. An, et al., *Chem. Eur. J.* 29 (2023) e202301624.
- [21] R. Huang, M.A. Holbert, M.K. Tarrant, et al., *J. Am. Chem. Soc.* 132 (2010) 9986–9987.
- [22] F.P. Li, A. Allahverdi, R.L. Yang, et al., *Angew. Chem. Int. Ed.* 50 (2011) 9611–9614.
- [23] M.J. Schmidt, D. Summerer, *Front. Chem.* 2 (2014) 7.
- [24] Y.P. Huang, G.J. Zhai, Y.A. Li, et al., *Anal. Chem.* 94 (2022) 10705–10714.
- [25] S.J. Elsässer, R.J. Ernst, O.S. Walker, J.W. Chin, *Nat. Methods* 13 (2016) 158–164.
- [26] D.P. Nguyen, M.M.G. Alai, P.B. Kapadnis, H. Neumann, J.W. Chin, *J. Am. Chem. Soc.* 131 (2009) 14194–14195.
- [27] M.D. Simon, F.X. Chu, L.R. Racki, et al., *Cell* 128 (2007) 1003–1012.
- [28] M. Jbara, N. Guttman-Raviv, S.K. Maity, N. Ayoub, A. Brik, *Biorg. Med. Chem.* 25 (2017) 4966–4970.
- [29] H.S. Ai, M.S. Sun, A.J. Liu, et al., *Nat. Chem. Biol.* 18 (2022) 972–980.
- [30] H.S. Ai, G.C. Chu, Q.Y. Gong, et al., *J. Am. Chem. Soc.* 144 (2022) 18329–18337.
- [31] D.D. Le, A.T. Cortesi, S.A. Myers, A.L. Burlingame, D.G. Fujimori, *J. Am. Chem. Soc.* 135 (2013) 2879–2882.
- [32] T.M. Hackeng, J.H. Griffin, P.E. Dawson, *Proc. Natl. Acad. Sci. U. S. A.* 96 (1999) 10068–10073.
- [33] T.P. Yang, Z. Liu, X.D. Li, *Chem. Sci.* 6 (2015) 1011–1017.
- [34] J.B. Blanco-Canosa, P.E. Dawson, *Angew. Chem. Int. Ed.* 47 (2008) 6851–6855.
- [35] R. Iwase, D.R. Dempsey, S.D. Whedon, et al., *J. Am. Chem. Soc.* 145 (2023) 6039–6044.
- [36] Y.H. Xiao, K. Zou, J.Y. Yang, M.X. Wu, *Cell Rep. Phy. Sci.* 4 (2023) 101638.
- [37] D. Szklarczyk, A.L. Gable, D. Lyon, et al., *Nucleic Acids Res.* 47 (2019) D607–D613.
- [38] A.G. Muntean, J.Y. Tan, K. Sitwala, et al., *Cancer Cell* 17 (2010) 609–621.
- [39] N.J. Krogan, J. Dover, A. Wood, et al., *Mol. Cell* 11 (2003) 721–729.
- [40] P.J. Beresford, D. Zhang, D.Y. Oh, et al., *J. Biol. Chem.* 276 (2001) 43285–43293.
- [41] Z. Fan, P.J. Beresford, D.Y. Oh, D. Zhang, J. Lieberman, *Cell* 115 (2003) 659–672.
- [42] J.L. Pedrosa, V. Munford, A.U. Bastos, et al., *J. Neurol. Sci.* 379 (2017) 249–252.