Electronic Supplementary Information

An expedient, CAET-assisted synthesis of dual-monoubiquitinated histone H3 enables evaluation of its interaction with DNMT1

Zichen Li,‡[a,b] Zebin Tong,‡[a] Qingyue Gong,‡[b,c] Huasong Ai,*[a] Shuai Peng,[b] Cong Chen,[b] Guo-Chao Chu,[a] Jia-Bin Li,*[b]

[a] Department of Chemistry, Tsinghua-Peking Center for Life Sciences, Tsinghua University, Beijing 100084, China.

[b] Jiangsu Key Laboratory of Neuropsychiatric Diseases and College of Pharmaceutical Sciences, Jiangsu Key Laboratory of Preventive and Translational Medicine for Geriatric Diseases, Suzhou Key Laboratory of Drug Research for Prevention and Treatment of Hyperlipidemic Diseases, Soochow University, Suzhou, 215123, China. E-mail: lijiabin@suda.edu.cn

[c] School of Life Sciences, University of Science and Technology of China, Hefei 230026, China

‡ These authors contributed equally to this work

Experimental Procedures

1. Reagents and Instruments

Reagents used in this work were all commercially available and purchased from following companies: e.g., Tris (2-carboxyethyl) phosphine hydrochloride (TCEP·HCl), sodium 2-mercaptoethanesulfonate (MesNa), dithiothreitol (DTT) from Aladdin Biochemical Technology (Shanghai, China); CAET-Acm from Peptide Biotechnology Co., Ltd (Nantong, China); 4-mercaptopheylacetic acid (MPAA) from Alfa Aesar; Acetonitrile, trifluoroacetic acid (TFA) from J. T. Baker (Phillipsburg, NJ, USA); Perchloric acid, $MgCl_2$, NaCl, NaH_2PO_4, $PdCl_2$ from Sinopharm Chemical Reagent Co., Ltd. Reverse-phase high performance liquid chromatography (RP-HPLC) was performed on a Shimadzu Prominence HPLC system (LC-20AT). The analysis and purification of peptides were conducted on an analytical Welch Ultimate XB-C4 column (250 × 4.6 mm, 5 μm, 300 Å) and a semi-preparative Welch Ultimate XB-C4 column (250 × 10 mm, 5 μm, 300 Å), respectively. Buffers for RP-HPLC: buffer A was 0.1% TFA in CH_3CN and buffer B was 0.1% TFA in water. ESI-MS was measured on the Shimadzu LC/MS-2020, which was processed by the UniDec software to generate the deconvoluted mass spectra. Size-exclusion chromatography and ion-exchange chromatography were performed on an AKTA Purifier from GE Healthcare Life Science. LC-MS/MS analysis was performed on an LTQ Orbitrap Velos instrument (Thermo Scientific) coupled to an Ultimate nanoflow HPLC system (Dionex).

2. Protein expression and purification

2.1. Preparation of ubiquitin thioester Ub-MesNa **2**

Ubiquitin thioester **2** was prepared according to the reported method with minor modifications.[1] The gene of full-length ubiquitin (Ub) was cloned into a pET-22b vector by choosing NcoI and XhoI as restriction enzymes and transformed into BL21(DE3) *E. coli* Cells. A single clone was picked into 20 mL LB medium containing 100 μg/mL ampicillin. After culturing at 37 ℃ for 12 h, the grown cells were transferred to 1 L LB medium containing 100 μg/mL ampicillin and cultured at 37 ℃ for several hours. When

the OD_{600} reached 1.0-1.2, isopropyl-β-D-thiogalactopyranoside (IPTG) with a final concentration of 0.4 mM was added to induce expression at 37 ℃ for 16 h. The cells were harvested by centrifugation (4000 rpm for 30 min) at 4 ℃ and ultra-sonicated to lyse in a lysis buffer (25 mM HEPES, pH 7.5, 150 mM NaCl). Perchloric acid ($HClO_4$) with a final concentration of 1%(v/v) was added into the lysate and the mixture was stirred for 5 min. The supernatant was collected by centrifugation (14000 rpm) at 4 ℃ for 30 min and subsequently dialyzed into an activation buffer (25 mM HEPES, pH 7.5, 100 mM NaCl). The solution containing ubiquitin was then concentrated to 5 mg/mL and added by TCEP (5 mg/ml), sodium 2-mercaptoethanesulfonate (MesNa) (100 mM), $MgCl_2$ (10 mM) and ATP-Na_2 (10 mM). The pH was adjusted to 7.5, followed by adding the E1 enzyme UBA1 (0.5 μM). The mixture was stirred at 37 ℃ for *in situ* thiolysis and the reaction was monitored by analytical RP-HPLC. After 12-18 h, the Ub-MesNa (**2**) was purified by the preparative RP-HPLC equipped with a C4 column. After characterization by ESI-MS, the desired **2** was lyophilized and stored at -20 ℃.

2.2. Expression and purification of histones

Wild type human histones (H2A, H2B, H3, and H4) and H3 mutants (H3K18C, H3K23C and H3K18CK23C) were expressed in *E. coli* according to the reported method with minor modifications.[2,3] Briefly, the histone genes were cloned into the pET-22b vectors and transformed into BL21(DE3) *E. coli* Cells. The cells were cultured in LB medium containing 100 μg/mL ampicillin at 37 ℃ to an OD_{600} of 1.0-1.2. A final concentration of 1.0 mM IPTG was added to induce histone expression at 37 ℃ for 16 h. After centrifugation (4000 rpm) at 4 ℃ for 30 min, the cells were harvested and subsequently ultra-sonicated to lyse in a lysis buffer (25 mM HEPES, pH 7.5, 150 mM NaCl). Then, the precipitation was collected by ultra-centrifugation (12000 rpm for 30 min), washed by ddH_2O for three times and dissolved in a buffer of 6.0 M Gn·HCl, 0.2 M NaH_2PO_4, pH 3.0. Ultra-centrifugation (12000 rpm) was conducted to discard the precipitation, and the supernatant containing histone was collected and dialyzed into ddH_2O containing 0.1 % TFA. After 12 h, the supernatant was purified by the preparative RP-HPLC equipped with a C4 column. The desired histones characterized

by ESI-MS were lyophilized and stored at -20 ℃.

3. Installation of CAET-Acm (1a) on histones

In a typical reaction,[4,5] CAET-Acm (**1a**) solved in DMSO (1.0 M) was added into 2.0 mL of a denatured buffer (6.0 M Gn·HCl, 0.2 M NaH_2PO_4, 5 mg/mL TCEP, pH 7.0) with a final concentration of 100 mM (for **3a** or **3b**) or 200 mM (for **3c**). After adjusting pH to 8.5, 2 μmol H3 mutant (**3a**, **3b** or **3c**) was added to the above mixture. The reaction was monitored by analytical RP-HPLC using a gradient of 20-70% buffer A over 30 min. For **3a** or **3b**, the reaction was almost completed within 2 h. For **3c**, the reaction was terminated after 12 h. The desired proteins were obtained by RP-HPLC purification and then characterized by ESI-MS.

4. Acm removal[6]

$PdCl_2$ was dissolved in the denatured buffer (6.0 M Gn·HCl, 0.2 M NaH_2PO_4, 5 mg/mL TCEP, pH 7.0) with a final concentration of 10 mM and the pH was adjusted to 7.0-7.5 till $PdCl_2$ was completely dissolved. In a typical reaction, CAET-Acm modified H3 (0.75 μmol) was dissolved in 750 μL of the $PdCl_2$-containing denatured buffer. After stirring at 37 ℃ overnight, the reaction was quenched by 7.5 mL of denatured buffer containing 1.0 M DTT. Then, centrifugation was conducted to remove the precipitation and the supernatant was purified by RP-HPLC. Finally, the desired peptides (**5a**, **5b** and **5c**) characterized by ESI-MS were obtained by lyophilization.

5. CAET-assisted ligation of ubiquitin thioester with H3 mutant

In a typical reaction,[4,5] 1.0 equiv. of CAET-modified H3 (**5a**, **5b** or **5c**, 1.2 μmol) and ubiquitin thioester **2** (1.2 equiv. for **5a** and **5b**, and 2.5 equiv. for **5c**) were dissolved in 1.2 mL of the ligation buffer (6.0 M Gn·HCl, 0.2 M NaH_2PO_4, 200 mM MPAA, pH 6.4). The reaction was stirred at 30 ℃ and monitored by analytical RP-HPLC using a gradient of 25-50% buffer A over 40 min. After the reaction was completed, the reaction was treated by a TCEP-containing buffer (6.0 M Gn·HCl, 0.2 M NaH_2PO_4, 0.25 M TCEP, pH 7.0) and subsequently purified by RP-HPLC. Finally, the desired H3K18$_C$ub, H3K23$_C$ub and H3K18$_C$ubK23$_C$ub were characterized by ESI-MS and

lyophilized.

6. Preparation of hemi-methylated Widom 601 DNA

Hemi-methylated primers were synthesized by Genewiz® company (mC in the sequence corresponds to the methylated cytosine):

Forward Primer: 5'-TACmCpGAAmCpGTTmCpGAACCATGATGCmCpGGAT-3'

Reverse Primer: 5'-TAmCpGmCpGAATTCCAAGmCpGACACmCpGGCACT-3'

A 207-bp of Widom 601 DNA flanked on each side by 30 bp of CG-containing sequence was synthesized and cloned into pUC-19 (GenScript® company). The resulting plasmid was transfected into *E. coli* DH5α cells. Then, the cells were cultured in the LB medium containing ampicillin (100 μg/mL) at 37 ℃ for 12 h, followed by plasmid extraction. The extracted plasmids were diluted to 500-1000 ng/μL and subsequently used for PCR. A total volume of 5.0 mL PCR solution containing 2.5 mL 2× PrimeStar Premix solution, forward primer, reverse primer and 100 μL of plasmids was split into 96× joint PCR tube, and subjected to amplify in a PCR system (95 ℃ for 5 min., 45× (95 ℃ for 15 s, 60 ℃ for 30 s, 72 ℃ for 20 s), and 72℃ for 10 min). The reaction solution in each tube was then collected, which was purified by ion-exchange chromatography equipped with a Mono Q 10-150 column (GE Healthcare Life Science).

7. Reconstitution of nucleosome core particle (NCP)

Histone octamers and NCPs were reconstituted using standard protocols.[7] Briefly, Histones H2A, H2B, H4 and H3 or ubiquitinated H3 at a stoichiometric ratio of 1.1: 1.1: 1: 1 were dissolved in the unfolding buffer (6.0 M Gn·HCl, 25 mM Tris-HCl, 1 mM DTT, pH 7.5) with a final concentration of 1 mg/mL. The mixture was dialyzed against the refolding buffer (10 mM Tris-HCl, 2.0 M NaCl, 1 mM DTT, 1 mM EDTA, pH 7.5) at 4 ℃ for 3 rounds (3×12 h). Then, the refolded octamer was concentrated to less than 1.0 mL and subsequently purified by size-exclusion chromatography on a

Superdex^{TM} 200 Increase 10/300 GL column. After characterization by SDS-PAGE, the desired fractions were collected and frozen for storage at -80 ℃.

The refolded octamer and DNA were mixed into a dialysis tube with a mole ratio of 1:1, which was initially suspended in the refolding buffer. The salt concentration was slowly reduced to less than 150 mM by pumping TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.5) into the refolding buffer. Then, the reconstituted NCP was purified with ion-exchange chromatography on a TSK DEAE-5PW column. The desired fractions confirmed by native gel were combined and further dialyzed against the HEPES buffer (25 mM HEPES, 25 mM NaCl, 1 mM EDTA, pH 7.5) for 6-8 h. The resulting NCP was kept at 4 ℃ for further use.

8. Electrophoretic mobility shift assay (EMSA)

All the purified NCPs were concentrated to 100 nM (judged by measuring the absorption of DNA at 260 nm) and the initial concentration of mDNMT1 was 3.33 μM. Subsequently, a series of different concentrations of mDNMT1 were prepared by using the double dilution method. To initiate the test, 1 μL NCP and 3 μL mDNMT1 were incubated at 4 ℃ for 15 min. Then, each sample mixed with 1 μL non-denaturation buffer was analyzed on the 4.5% native-page and stained with SYBR-Gold dye (Thermo Fisher Scientific) and visualized using a ChemiDoc MP Imaging System (Bio-Rad). The quantities of mDNMT1-bound NCPs at different mDNMT1 concentrations were calculated from the free NCPs quantified by ImagLAB. The resulting data were finally fitted by Sigmoidal (X is Log) curves in GraphPad Prism to obtain the Kd values.

9. *In vitro* DNA methylation assay

DNA methylation assay was conducted by detecting the production of S-adenosyl-L-homocysteine (SAH) with MTase-Glo methyltransferase kit (Promega, cat. no. V7601). 0.4 μM mDNMT1 (291-1620) in the reaction buffer (20 mM HEPES, 50 mM NaCl, 1 mM EDTA, 3 mM $MgCl_2$, 0.1 mg/ml BSA, 20% Glycerol, pH 7.5, and 20 μM SAM) was mixed with 0.4 μM nucleosomes (unmodified or H3K18$_C$UBK23$_C$Ub) in equal volumes. The reaction was started at 37°C. Samples were quenched by adding 0.5%

TFA at 0 and 15 min, respectively. The one-quarter volume of 5× MTase-GloTM reagent was then added and incubated at room temperature for 30 min to convert SAH to ADP and transferred to solid white flat bottom 96 well plates. An equal volume of MTase-GloTM Detection Solution was finally added. Monitoring the luminescence of the reaction product SAH using a microplate reader (BioTek) The data was processed using GraphPad Prism.

10. Chemical cross-linking of proteins coupled with mass spectrometry analysis

The H3K18$_C$ubK23$_C$ub-NCP and mDNMT1 were concentrated to 4 μM and 16 μM, respectively. To conduct chemical cross-linking, 10 μL nucleosome and 4 μL mDNMT1 were incubated in the HEPES buffer (25 mM HEPES, 25 mM NaCl, 1 mM EDTA, pH 7.5) with a total volume of 20 μL. After incubation on ice for 1 h, equal volume of BS3 solution (1.0 mM) were added to the mixture and incubated at 4 ℃ for another 15 min. The crosslinking reaction was quenched with 10 μL Tris-HCl buffer (500 mM Tris-HCl, pH 7.5) and separated by SDS-PAGE. For the unmodified NCP, the NCP and mDNMT1 were concentrated to 2.6 μM and 26.6 μM, respectively. Then 10 μL nucleosome and 4 μL mDNMT1 were incubated in the HEPES buffer (25 mM HEPES, 25 mM NaCl, 1 mM EDTA, pH 7.5). This was followed by the addition of an equal volume of 1 mM BS3 for crosslinking and finally quenched by the addition of 2 μl of 2M Tris. The band corresponding to the mDNMT1-H3K18$_C$ubK23$_C$ub-NCP complex or mDNMT1-NCP complex was excised, alkylated and digested by trypsin. The digested peptides were extracted and subjected to LC-MS/MS analysis. The data was processed by the pXtract software and searched using the pLink2 software with the following parameters: enzyme: trypsin; up to 3 missed cleavages; peptide mass: 400-10000 Da; peptide length: 4-100 aa; precursor tolerance: ± 20 ppm; fragment tolerance: ± 20 ppm; MS/MS tolerance: 0.8 Da. The crosslinking map was formed by XiNET and the model was made using the ChimeraX software.

Supplementary figures

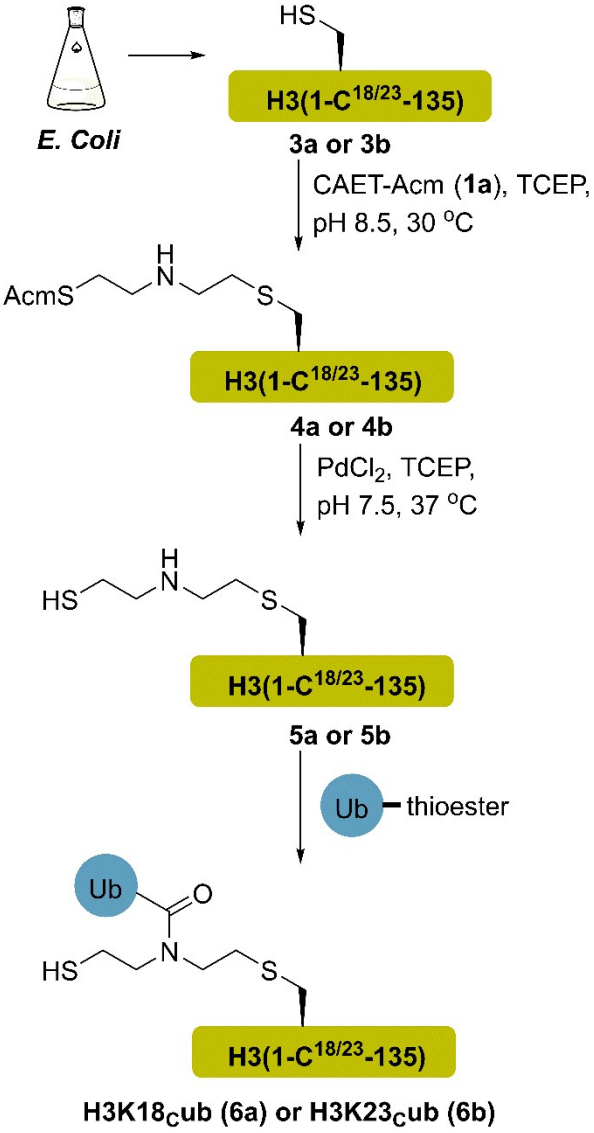


Fig. S1 Schematic depiction of CAET-assisted synthesis of single-monoubiquitinated histone H3 ($H3K18_Cub$, **6a** and $H3K23_Cub$, **6b**).

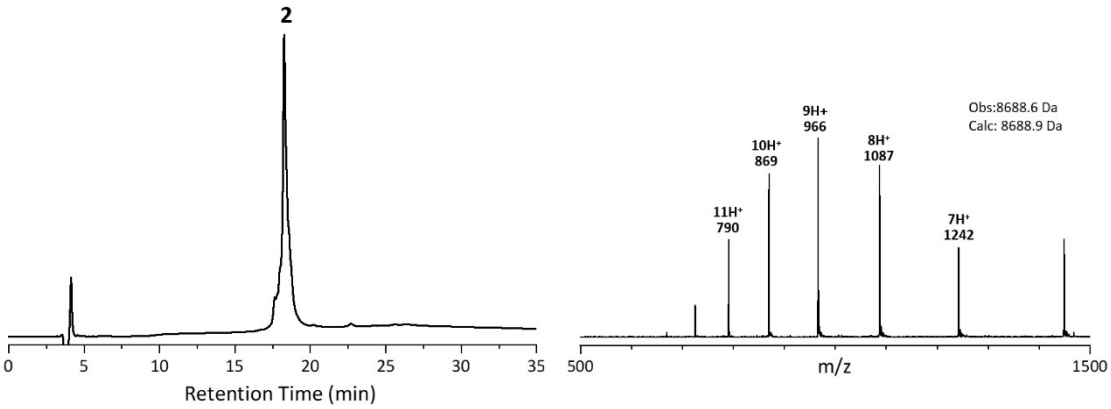


Fig. S2 RP-HPLC (214 nm) and ESI-MS characterization of ubiquitin thioester **2**.

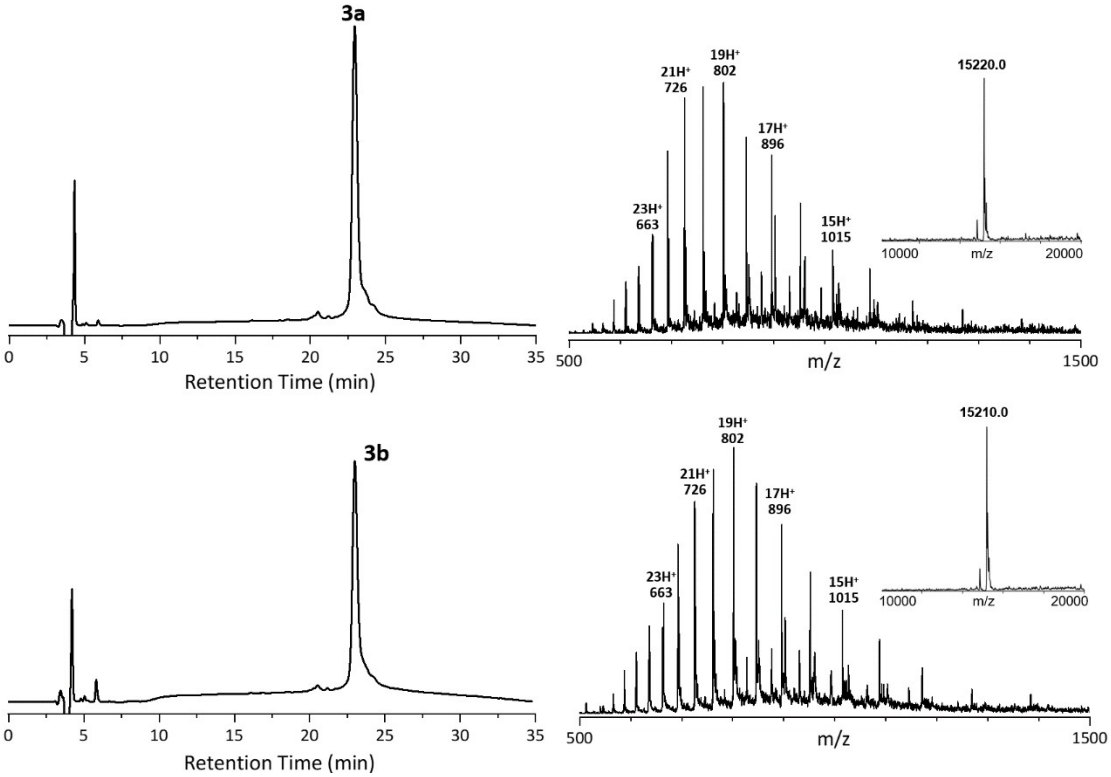


Fig. S3 RP-HPLC (214 nm) and ESI-MS characterization of H3 mutants H3K18C (**3a**, Obs. 15220.0 Da, Calc. 15215.7 Da) and H3K23C (**3b**, Obs. 15210.0 Da, Calc. 15215.7 Da). The UniDec software[8] was used to generate the deconvoluted mass spectra and provided the observed (Obs.) molecular weight.

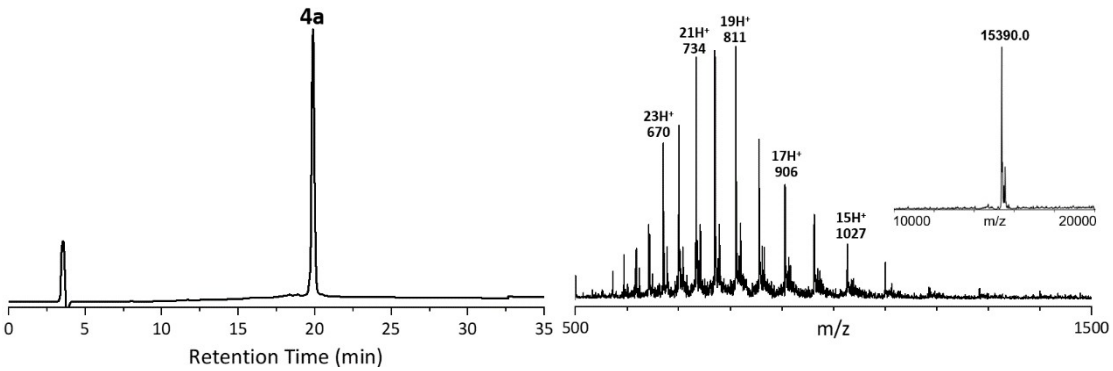


Fig. S4 RP-HPLC (214 nm) and ESI-MS characterization of peptide **4a**. Obs. 15390.0 Da, Calc. 15388.9 Da. The UniDec software[8] was used to generate the deconvoluted mass spectra and provided the observed molecular weight.

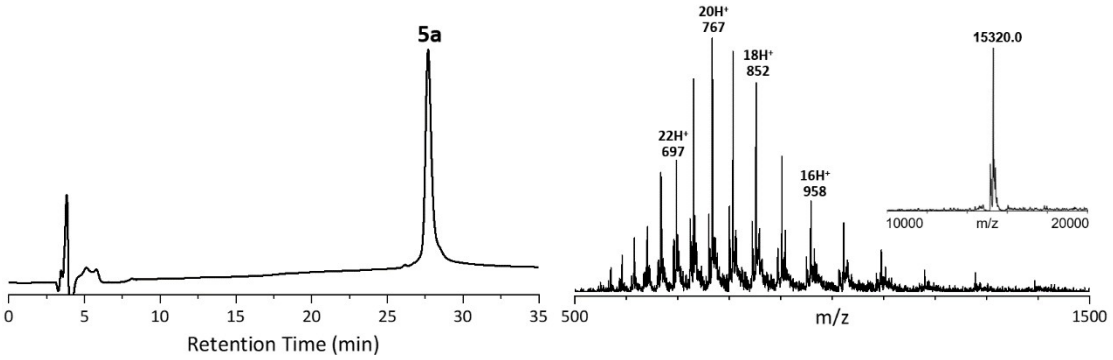


Fig. S5 RP-HPLC (214 nm) and ESI-MS characterization of peptide **5a**. Obs. 15320.0 Da, Calc. 15317.9 Da. The UniDec software[8] was used to generate the deconvoluted mass spectra and provided the observed molecular weight.

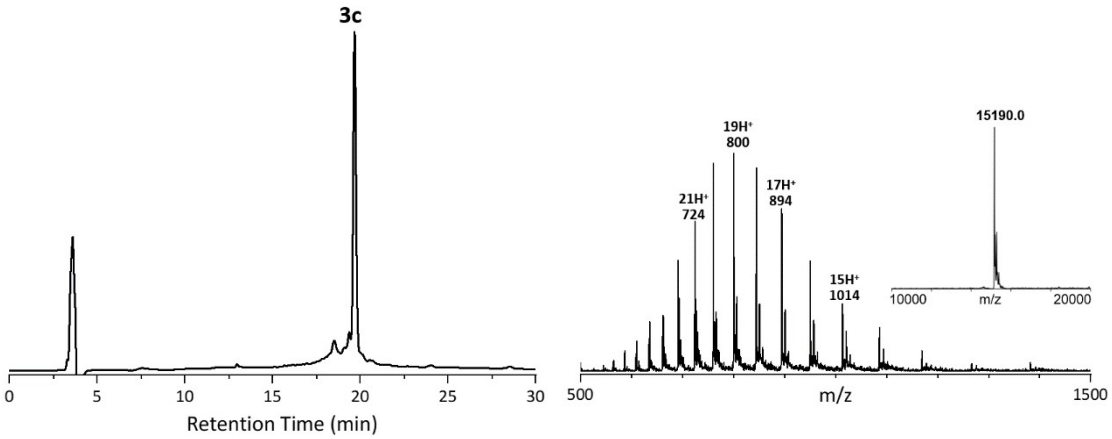


Fig. S6 RP-HPLC (214 nm) and ESI-MS characterization of histone H3 mutant H3K18CK23C (**3c**). Obs. 15190.0 Da, Calc. 15190.7 Da. The UniDec software[8] was used to generate the deconvoluted mass spectra and provided the observed molecular weight.

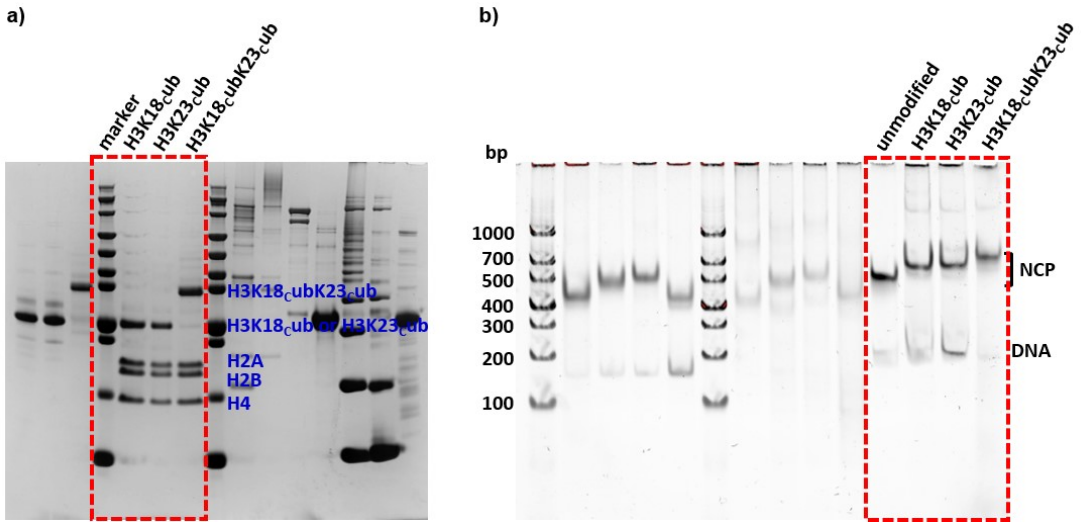


Fig. S7 The raw images of octamer and NCP reconstitution. a) SDS-PAGE analysis of the purified octamers (in the red box), corresponding to Fig. 2i. b) Native gel of the purified NCPs (in the red box), corresponding to Fig. 2k.

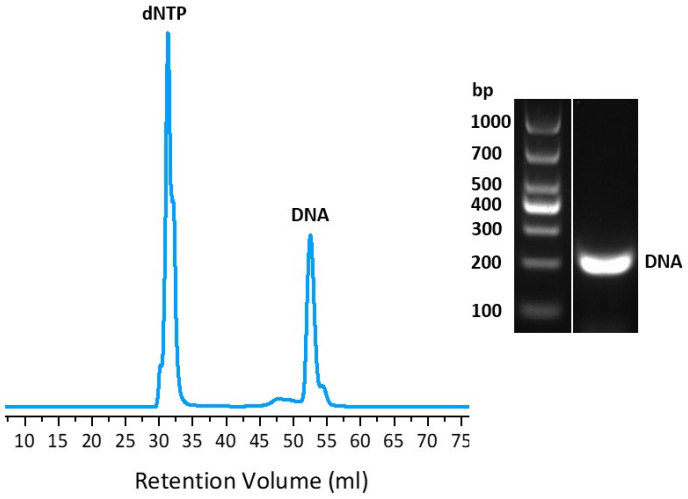


Fig. S8 Purification of the hemi-methylated DNA (He601) by ion-exchange chromatography (left) and the purified DNA was characterization by agarose gel electrophoresis (right). dNTP: deoxy-ribonucleoside triphosphate.

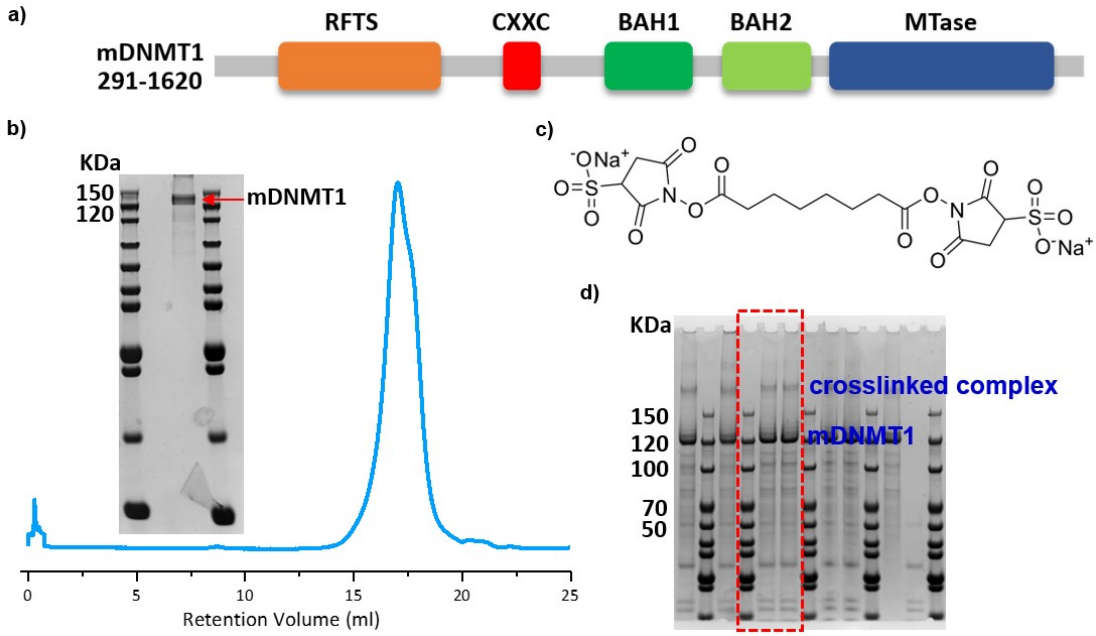


Fig. S9 Expression and purification of mDNMT1 (residues 291-1620). a) Domain architecture of mDNMT1 used in this study. b) mDNMT1 purification by size-exclusion chromatography and SDS-PAGE analysis of the purified mDNMT1, stained with Coomassie brilliant blue. c) The chemical structure of Bis (sulfosuccinimidyl) suberate sodium salt (BS3). d) Separation of BS3-crosslinked mDNMT1-H3K18$_C$ubK23$_C$ub-NCP complex by SDS-PAGE (in the red box).

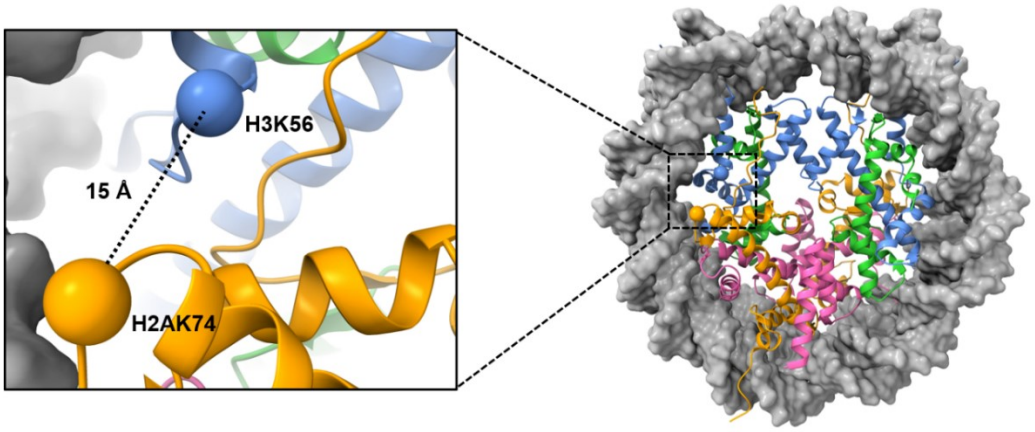
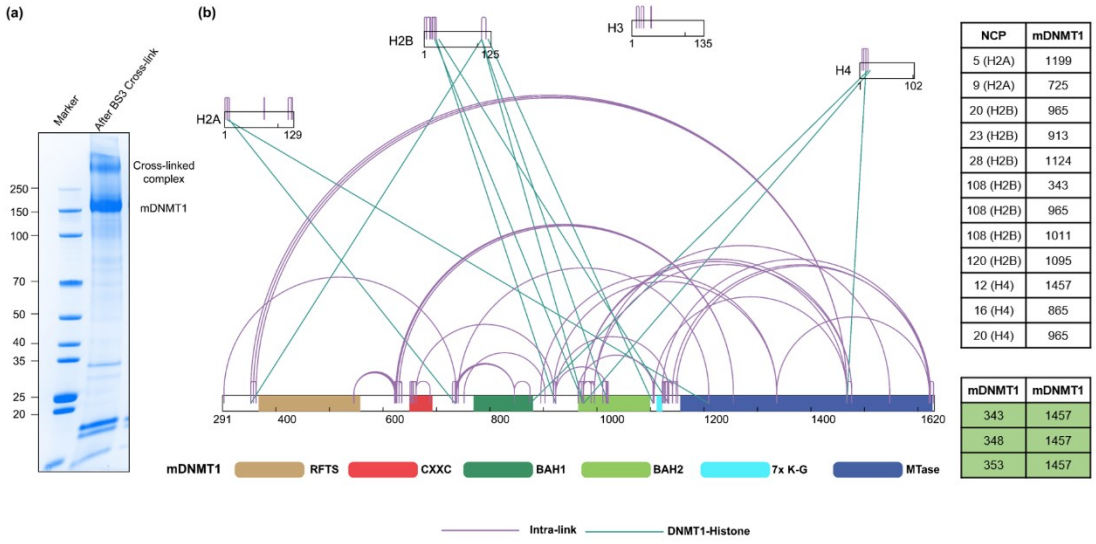


Fig. S10 The representative lysine-lysine cross-link between histones (H3K56-H2AK74) was mapped into the structure of NCP (PDB: 7XD1)[3]. The distance measured by the Cα atoms is 15 Å, which was consistent with the distance constraint of BS3.



NCP	mDNMT1
5 (H2A)	1199
9 (H2A)	725
20 (H2B)	965
23 (H2B)	913
28 (H2B)	1124
108 (H2B)	343
108 (H2B)	965
108 (H2B)	1011
120 (H2B)	1095
12 (H4)	1457
16 (H4)	865
20 (H4)	965

mDNMT1	mDNMT1
343	1457
348	1457
353	1457

Fig. S11 Cross-links of mDNMT1 in complex with the unmodified-NCP. a) Separation of BS3-crosslinked mDNMT1-NCP complex by 4%-12% SDS-PAGE. b) Interprotein cross-links and intraprotein cross-links were indicated by lines (Left) and the representative cross-linked Lys pairs were listed on the right.

References

1. Q. Zheng, T. Wang, G. Chu, C. Zuo, R. Zhao, X. Sui, L. Ye, Y. Yu, J. Chen, X. Wu, W. Zhang, H. Deng, J. Shi, M. Pan, Y. Li and L. Liu, *Angew. Chem. Int. Ed.*, 2020, **59**, 13496–13501.

2. H. Ai, G.-C. Chu, Q. Gong, Z.-B. Tong, Z. Deng, X. Liu, F. Yang, Z. Xu, J.-B. Li, C. Tian and L. Liu, *J. Am. Chem. Soc.*, 2022, **144**, 18329–18337.
3. H. Ai, M. Sun, A. Liu, Z. Sun, T. Liu, L. Cao, L. Liang, Q. Qu, Z. Li, Z. Deng, Z. Tong, G. Chu, X. Tian, H. Deng, S. Zhao, J.-B. Li, Z. Lou and L. Liu, *Nat. Chem. Biol.*, 2022, **18**, 972–980.
4. M. Pan, Q. Zheng, S. Ding, L. Zhang, Q. Qu, T. Wang, D. Hong, Y. Ren, L. Liang, C. Chen, Z. Mei and L. Liu, *Angew. Chem. Int. Ed.*, 2019, **58**, 2627–2631.
5. Q. Zheng, T. Wang, J. Mao, G. Chu, L. Liang, Y. Jing, C. Zuo, Y. Yu, H. Hu and M. Pan, *Nat. Protoc.*, 2022, doi: 10.1038/s41596-022-00761-z.
6. S. K. Maity, M. Jbara, S. Laps and A. Brik, *Angew. Chem. Int. Ed.*, 2016, **55**, 8108–8112.
7. K. Luger, T. J. Rechsteiner, T. J. Richmond, *Methods Mol. Biol.* 1999, **119**, 1-16.
8. M. T. Marty, A. J. Baldwin, E. G. Marklund, G. K. A. Hochberg, J. L. P. Benesch and C. V. Robinson, *Anal. Chem.* 2015, **87**, 4370-4376.