Supplementary Materials for
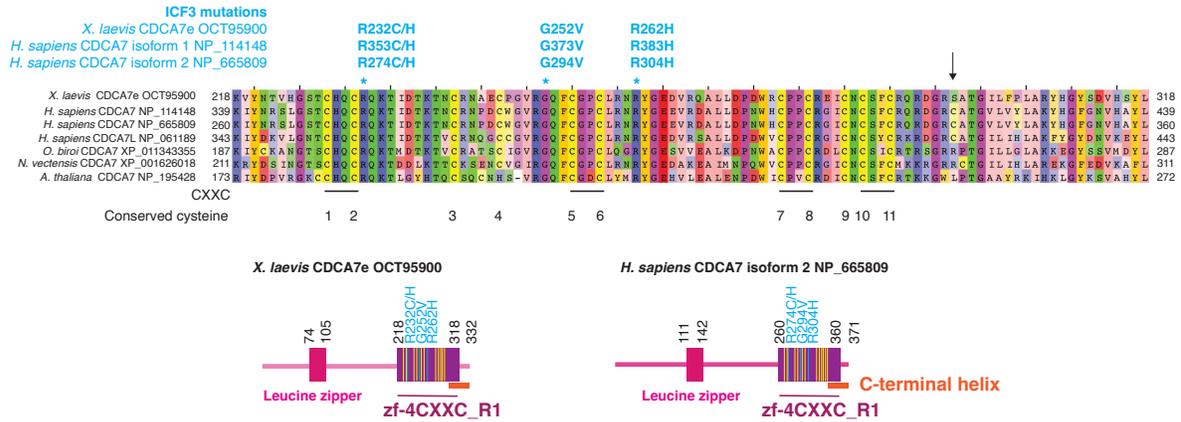
CDCA7 is a hemimethylated DNA adaptor for the nucleosome remodeler HELLS

Wassing, Nishiyama et al.

*Corresponding authors. Email: funabih@rockefeller.edu, uanishiyama@g.ecc.u-tokyo.ac.jp, aritak@yokohama-cu.ac.jp

**This PDF file includes:**

Figs. S1 to S10
Tables S1 to S3

Fig. S1. Evolutionary conservation of the zf-4CXXC_R1 domain of CDCA7 homologs
ClustalW multi-sequence alignment of CDCA7 zf-4CXXC_R1 domain, characterized by eleven conserved cysteine (yellow) and three ICF3 patient-associated (cyan) residues. *X. laevis*, *Xenopus laevis* (African clawed frog); *H. sapiens*, *Homo sapiens* (human); *O. biroi*, *Ooceraea biroi* (clonal raider ant); *N. vectensis*, *Nematostella vectensis* (starlet sea anemone); *A. thaliana*, *Arabidopsis thaliana* (thale cress). Amino acid positions of ICF3 associated mutations in *X. laevis* CDCA7e, *H. sapiens* CDCA7 isoform 1 (NP_114148) and the shorter isoform 2 (NP_665809) are indicated. An arrow indicates the position of cysteine 339 of *H. sapiens* CDCA7 isoform 2, the site that was mutated to serine in Fig. 3, fig. S3D, fig. S4 and fig. S5. Schematics showing the domain composition of *X. laevis* CDCA7e and *H. sapiens* CDCA7 isoform 2 are also shown.
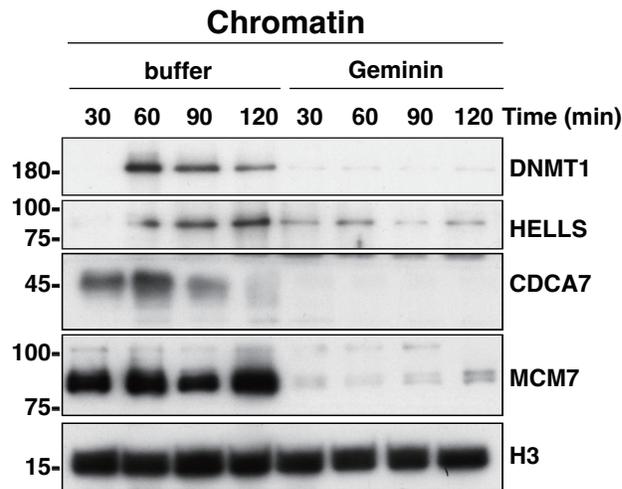
Fig. S2. DNA replication promotes chromatin association of CDCA7e and HELLS .

X. laevis sperm nuclei were incubated with interphase *Xenopus* egg extracts in the presence or absence of 0.5 μM recombinant geminin. At each indicated time point, chromatin was isolated and analyzed by western blotting.
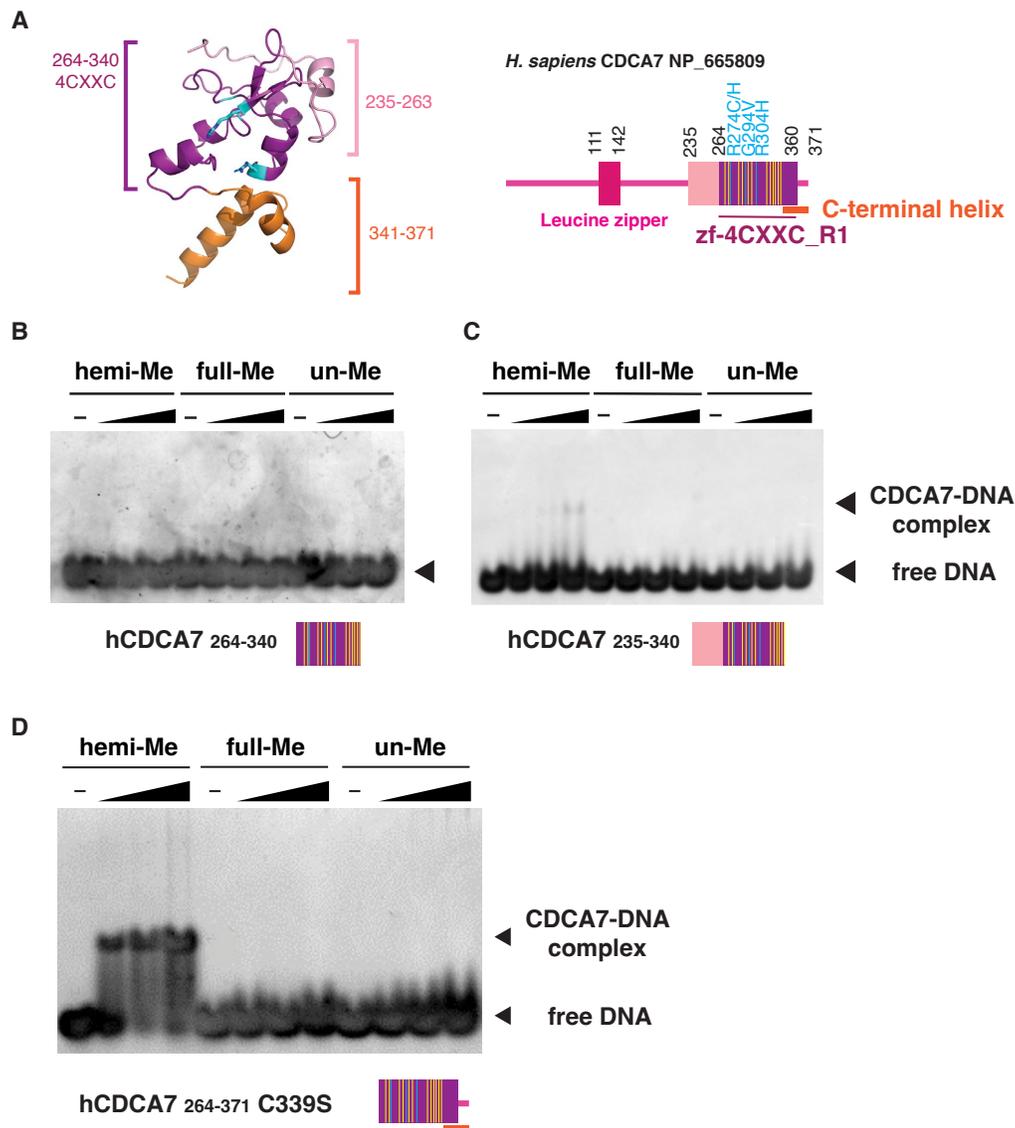
Fig. S3. Characterization of the minimum hemimethylated DNA-binding domain of human CDCA7.

(**A**) AlphaFold2-modeled structure of *H. sapiens* zf-4CXXC_R1 domain and schematic of full-length *H. sapiens* CDCA7. Yellow lines indicate the position of conserved cysteine residues. Orange bar indicates the conserved C-terminal helix. (**B-D**) Native gel electrophoresis mobility shift assay for detecting the interaction of hCDCA7 264-340 (**B**) hCDCA7 235-340 (**C**), and hCDCA7 264-371 C339S (**D**) with double stranded DNA oligonucleotides with an unmethylated, hemi-methylated or fully-methylated CpG.

3
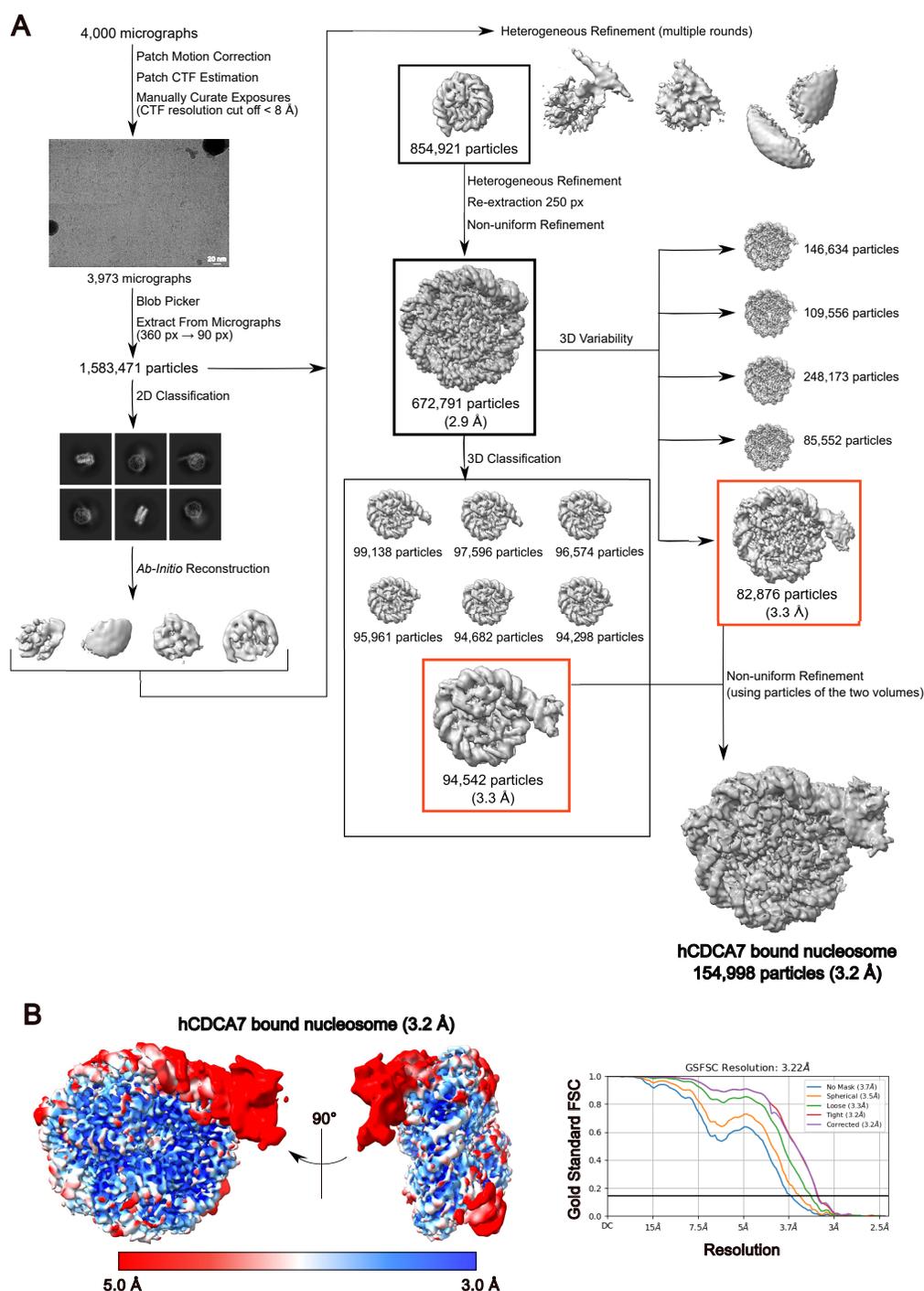
Fig. S4.
Cryo-EM single particle analysis of hCDCA7 bound to nucleosome.
(A) Cryo-EM data particle processing and refinement workflow of hCDCA7$_{264-371}$ C339S bound to nucleosome. **(B)** Local resolution of cryo-EM map of hCDCA7$_{264-371}$ C339S bound to nucleosome (left). Fourier shell correlation (FSC) curve of hCDCA7$_{264-371}$ C339S bound to nucleosome (right).

4
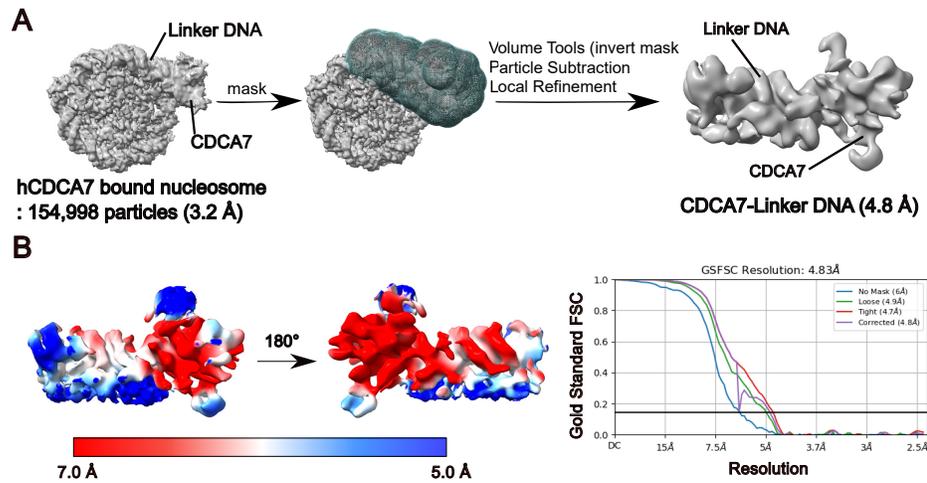
Fig. S5. Refinement workflow for cryo-EM map of linker DNA bound to hCDCA7 density. **(A)** Focused refinement of the linker DNA bound by hCDCA7$_{264-371}$ C339S moiety. Left figure shows the cryo-EM map of hCDCA7$_{264-371}$ C339S bound to nucleosome. The mask file is shown as a green mesh (center) covering the hCDCA7$_{264-371}$ C339S moiety bound to the linker DNA. The cryo-EM map corresponding to the hCDCA7$_{264-371}$ C339S moiety bound to the linker DNA was improved by local refinement at 4.83 Å resolution (right). **(B)** Local resolution of the hCDCA7$_{264-371}$ C339S moiety bound to linker DNA. (left). FSC curve of hCDCA7$_{264-371}$ C339S bound to nucleosome (right)
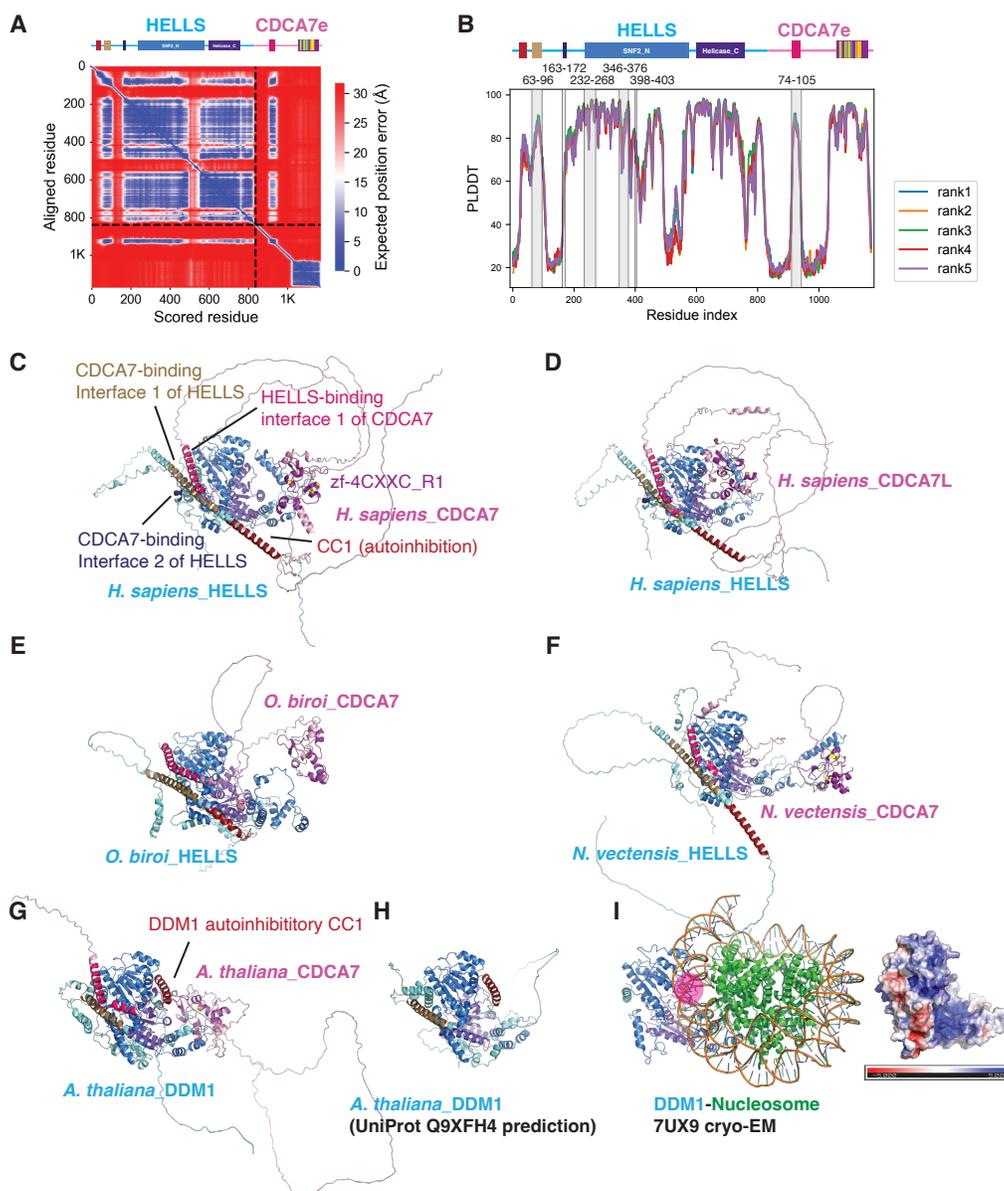
Fig. S6. Alphafold2 structure prediction of the CDCA7-HELLS/DDM1 complex
(**A**, **B**) AlphaFold2 structure prediction analysis of *X. laevis* HELLS and CDCA7e. (**A**) The predicted aligned error map of the best model, with the minimum inter-chain predicted aligned error of 1.7 Å. (**B**) PLDDT scores of the top five predicted models, with the interface highlighted on top of the figure. The top five predictions were converged (region is shaded in gray), and the interface has relatively high PLDDT scores, with the average value of 63. (**C-H**) AlphaFold2 structure prediction of HELLS/DDM1 of indicated species in complex with CDCA7 and CDCA7 paralogs . (**I**) Left; atomic model of DDM1-nucleosome complex cryo-EM structure (7UX9). Right; surface electrostatic potential of DDM1. The DNA-binding positively charged groove, which is predicted to be occupied by the autoinhibitory CC by AlphaFold2 models, is marked with a pink circle.

6
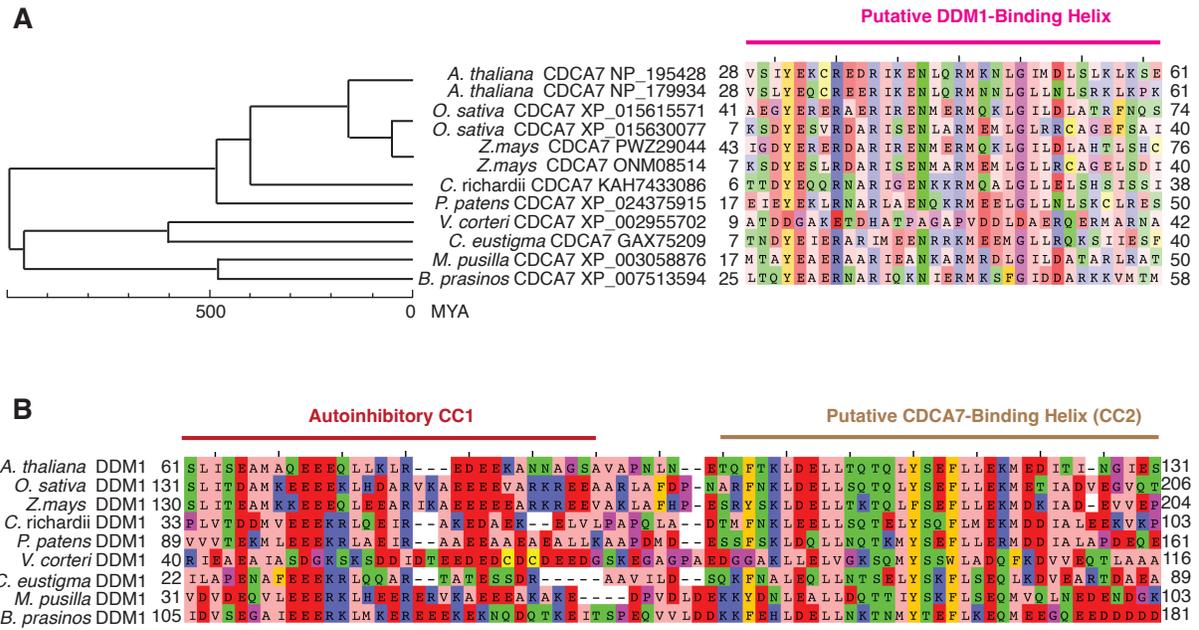
Fig. S7. Evolutionary conservation of putative DDM1-CDCA7 interaction interfaces in green plants

(A) Sequence alignment of putative DDM1-binding helix of CDCA7 homologs in green plants. *O. sativa*, *Oryza sativa* (rice); *Z. mays*, *Zea mays* (corn); *C. richardii*, *Ceratopteris richardii* (fern); *P. patens*, *Physcomitrium patens* (moss); *V. carteri*, *Volvox carteri* (colonial green alga); *C. eustigma*, *Chlamydomonas eustigma* (unicellular green alga); *M. pusilla*, *Micromonas pusilla* (unicellular green alga); *B. prasinos*, *Bathycoccus prasinos* (marine green alga). **(B)** Sequence alignment of the putative autoinhibitory CC1 and the CDCA7-binding CC2 of DDM1 homologs in green plants.

7
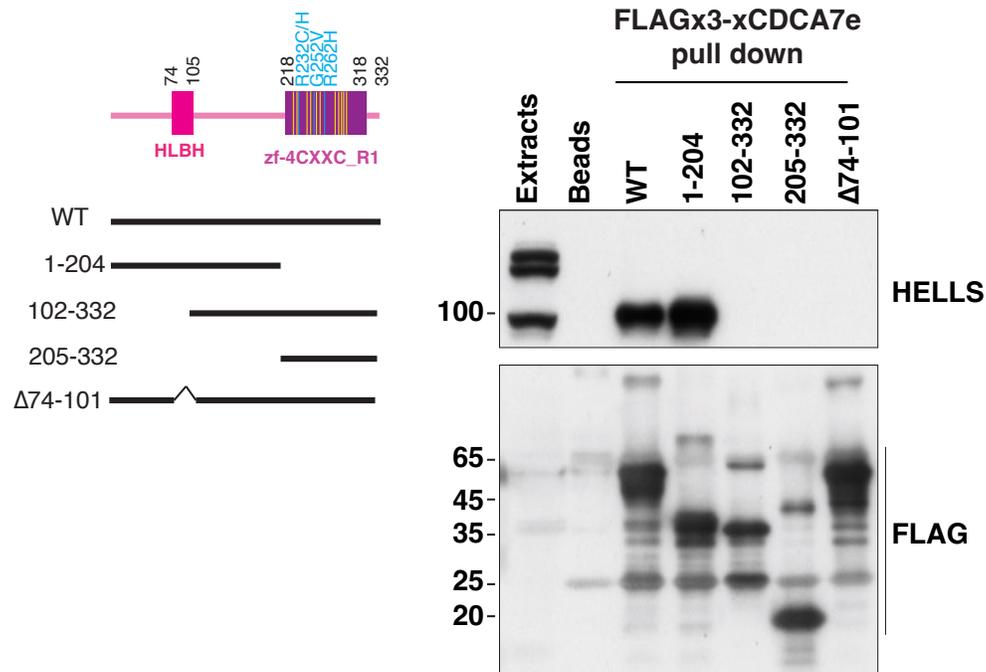
Fig. S8. N-terminal CDCA7 segment lacking the zf-4CXXC_R1 domain is sufficient for HELLS binding

Wildtype (WT) or truncated versions of recombinant FLAG3-tagged *X. laevis* CDCA7e proteins were incubated with *Xenopus* egg extracts, followed by immunoprecipitation with anti-FLAG coupled beads. Isolated proteins were analyzed by western blotting using anti-HELLS and anti-FLAG antibodies.
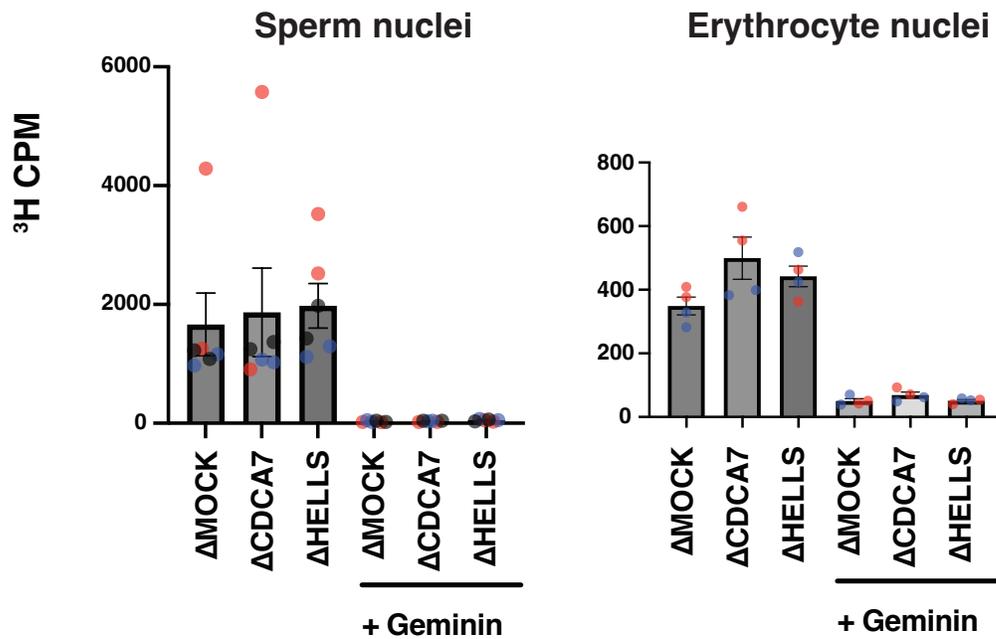
Fig. S9. CDCA7 and HELLS are not required for global maintenance DNA methylation in *Xenopus* egg extracts

X. laevis sperm nuclei **(A)** or erythrocyte nuclei **(B)** were incubated with egg extracts for 60 min with *S*-[methyl-^{3}H]-adenosyl-L-methionine with or without geminin, which inhibits DNA replication initiation. Radioactivity associated with chromosomal DNA is measured. Results include three biological replicates **(A)** or two biological replicates **(B)**, each of which includes two technical replicates (shown in the same color). Geminin effectively inhibited DNA incorporation of ^{3}H, demonstrating that DNA methylation of sperm chromatin depends on DNA replication.
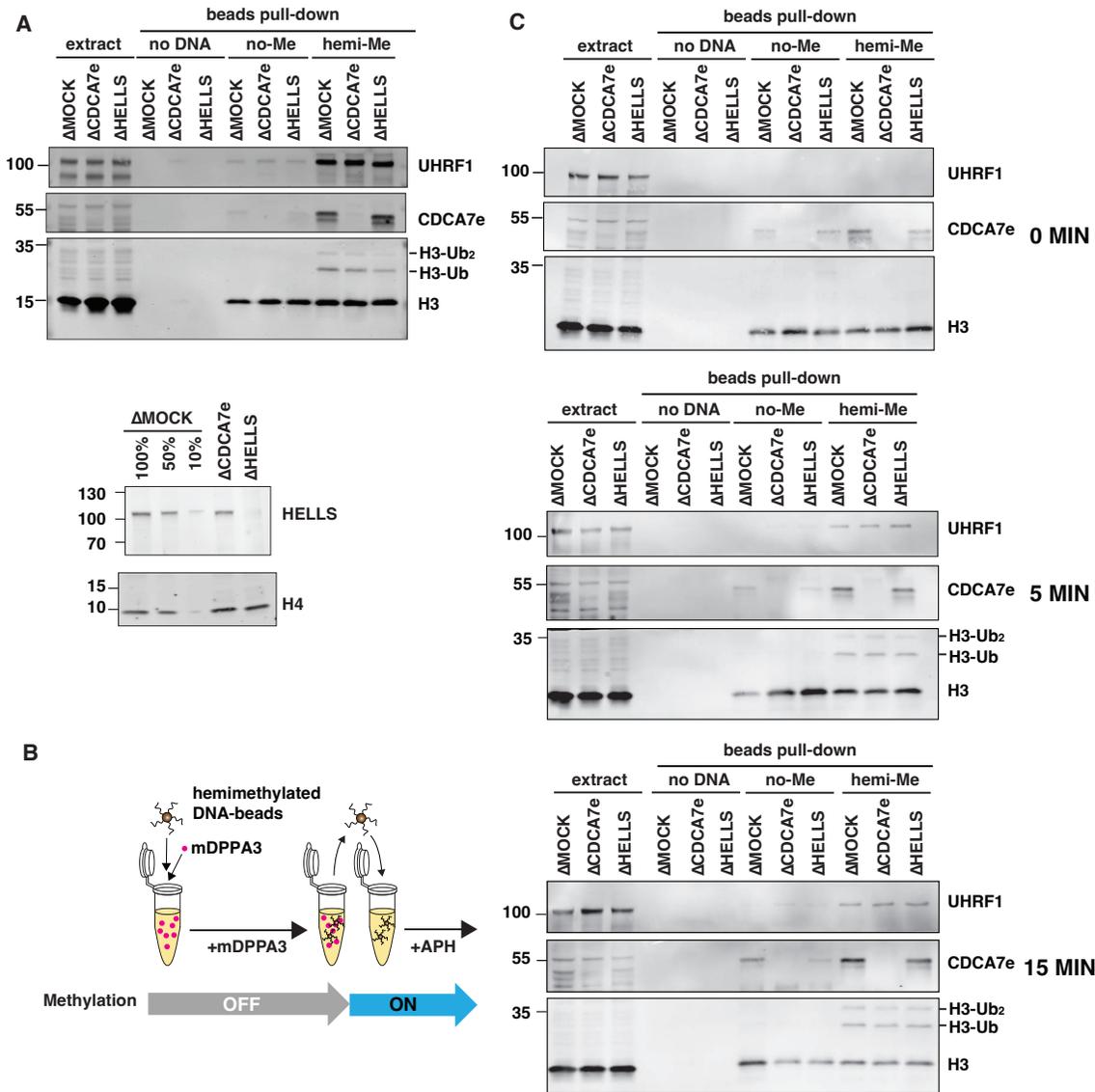
Fig. S10. CDCA7e and HELLS are not required for histone H3 ubiquitylation on hemimethylated DNA-beads in *Xenopus* egg extracts

(**A**) Beads coated with unmethylated pBlueScript DNA or hemimethylated pBlueScript DNA were incubated with interphase mock IgG-depleted (ΔMOCK), CDCA7e-depleted (ΔCDCA7e), or HELLS-depleted (ΔHELLS) *Xenopus* egg extracts for 60 min. Beads were collected and analyzed by western blotting. Bottom panel shows effective HELLS depletion. (**B, C**) Beads coated with unmethylated pBlueScript DNA or hemimethylated pBlueScript DNA were incubated with ΔMOCK, ΔCDCA7e, or ΔHELLS egg extracts for 60 min in the presence of 1.3 μM mDPPA3, which inhibits binding of UHRF1 and H3 ubiquitylation. During this preincubation, nucleosomes assemble on DNA beads without DNA methylation. Beads were then transferred to corresponding depleted interphase extracts that contained aphidicolin (APH) but not mDPPA3. After 0-, 5-, or 15-min incubation, beads were collected and analyzed by western blotting.

10

Table S1. DNA ultramer sequence used for DNA pull-downs

Name	Sequence (sense strand)
200 bp unmethylated DNA Widom601	/5Biosg/TCGGGTTATGTGATGGACCCTATACGCGGGCGCCCTGGAGAATCCTGCAGCCGAGGCCGCTCAATTGGTCGTAGCAAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGTGCATGTATTGAACAGCGAC
200 bp hemimethylated DNA Widom601	/5Biosg/T**M**GGGTTATGTGATGGACCCTATA**M**G**M**GGG**M**GCCCTGGAGAATCCTGCAGC**M**GAGGC**M**GCTCAATTGGT**M**GTAGCAAGCTCTAGCAC**M**GCTTAAA**M**GCA**M**GTA**M**G**M**GCTGTCCCC**M**GM**GTTTTAAC**M**GCCAAGGGGATTACTCCCTAGTCTCCAGGCA**M**GTGTCAGATATATACATCCTGTGCATGTATTGAACAG**M**GAC

*M:5-methylcytosine

**/5Biosg/: 5' biotin modification

Table S2. DNA sequence used for nucleosome reconstruction

Name	Sequence (sense strand) and primers
Hemimethylation site in 5'-linker DNA Widom601	ATCTGGGCC**M**GCCAT<u>ATCAGAATCCCGGTGCCGAGGC CGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAA CGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAA GGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATAT ATACATCGAT</u> (160 bp) **Primer** Forward: 5'-ATCTGGGCC**M**GCCATATCAGAATCCCGGTGCCGA GGCCG Reverse: 5'-ATCGATGTATATATCTGACACGTGC
Hemimethylation in 3'-linker DNA Widom601	<u>ATCAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGT AGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCT GTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTA GTCTCCAGGCACGTGTCAGATATATACATCGAT</u>CC**M**GC AGGCC (157bp) **Primer** Forward: 5'- ATCAGAATCCCGGTGCCGAGGCCGC Reverse: 5'- ATCCGTCTCCATCGATGTATATATC **Oligo nucleotide** Forward: 5'-CGATCC**M**GCAGGGCAG Reverse: 5'-CTGCCCTGCGGG
Hemimethylation in 3'-nucleosomal DNA Widom601	<u>ATCAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGT AGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCT GTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTA GTCTCCAGGCACGTGTCAGATATA**M**GCATCGAT</u>GCAG G (150 bp) **Primer** Forward: 5'- ATCAGAATCCCGGTGCCGAGGCCGC Reverse: 5'-TCTCAGATATCCCGTCTCGCGTATATCTGA CACGTGCCTG **Oligo nucleotide** Forward: 5'- TA**M**GCATCGATGCAGG Reverse: 5'-CCTGCATCGATG

*M:5-methylcytosine

12

Table S3. Cryo-EM data collection statics for hCDCA7:nucleosome

	hCDCA7:nucleosome	hCDCA7:linker DNA (focused map)
EMDB number	EMD-38198	EMD-38199
Microscope	Krios G4 (RIKEN BDR)	
Voltage (keV)	300	
Camera	K3/BioQuantum	
Magnification	105,000	
Pixel size at detector (Å)	0.83	
Total electron exposure (e$^-$/Å^2)	60.725	
Exposure rate (e$^-$/pixel/sec)	18.987	
Exposure time (sec)	2.2	
Defocus range (µm)	0.6-1.6 (interval: 0.2)	
Number of frames	48	
Energy filter slit width	15	
Micrographs collected (no.)	4,000	
Initial particle images (no.)	1,652,465	672,791
Final particle images (no.)	154,998	154,998
Map resolution (Å) FSC threshold	3.18	4.83
Automation software	EPU	