Supporting Information

Bailey et al. 10.1073/pnas.1300325110

SI Materials and Methods

Affinity Purification of Centromere Protein A-Containing Nucleosomes. Approximately 6×10^9 HeLaS3 cells stably expressing localization and purification (LAP)–tagged centromere protein A (CENP-A) (Fig. S1) were harvested and resuspended in hypotonic lysis buffer (3.75 mM Tris at pH 7.5, 20 mM KCl, 0.5 mM EDTA, 0.5 mM DTT, 0.05 mM spermidine, 0.125 mM spermine, 1 mM PMSF, and 0.1% digitotin). All buffers contained 200 mM sodium vanadate, 50 mM b-glycerophosphate, and 10 mM sodium butyrate. Cells were homogenized to release the nuclei and then pelleted at 300 × *g*. The pellets containing the nuclei were resuspended in Wash Buffer A (20 mM Hepes at pH 7.5, 20 mM KCl, 0.5 mM EDTA, 0.5 mM DTT, and 0.5 mM PMSF), pelleted, and resuspended in Wash Buffer B (Wash Buffer A supplemented with 300 mM NaCl). Nuclei were resuspended in Wash Buffer B supplemented with 3 mM $CaCl_2$. Chromatin was digested for 2 h at 37 °C using 1,500 units of micrococcal nuclease (Roche). The micrococcal nuclease digestion was stopped by adding 5 mM EGTA and 0.05% Triton X-100 and was pelleted at 10,000 × *g* for 15 min at 4 °C. The supernatant contained predominantly mononucleosomes as confirmed by agarose gel electrophoresis.

Affiprep Protein-A (Bio-Rad) was cross-linked to αGFP antibodies (affinity purified) using protocol from Cheeseman and Desai (1). Nucleosomes released after the micrococcal nuclease digestion were bound to αGFP beads for 2 h at 4 °C. The beads were then washed twice with Bead Buffer A (20 mM Hepes, pH 7.4, 300 mM KCl, 0.5 M EDTA, 1 mM DTT) followed by a wash with Bead Buffer B (Bead Buffer A supplemented with 0.1% Tween 20). Beads were then resuspended in two volumes of Bead Buffer B supplemented with 2 mM $MgCl_2$ and 135 μg of PreScission Protease and incubated at 4 °C for 2 h. The supernatant was harvested, and beads were washed with one volume of the Bead Buffer. All supernatants were pooled and used for MS analysis.

Mass Spectrometry and Data Analysis of CENP-A Purifications. Purified CENP-A nucleosomes were largely insoluble when buffer-exchanged to detergent-free conditions, which is generally preferred for comprehensive MS analysis. To overcome this limitation, purified CENP-A and associated proteins were precipitated using ice-cold trichloroacetic acid (33% final volume) followed by successive washing using ice-cold 100% acetone. Speed vac-dried, precipitated protein pellets were reconstituted with 100 μL of buffer containing 100 mM NH_4HCO_3, pH 8.0, and 0.1% of the acid-labile detergent RapiGest SF Surfactant (Waters), which is compatible with liquid chromatography (LC)-MS electron-transfer dissociation (ETD) analysis. Resuspended proteins were reduced using DTT, alkylated with iodoacetamide, and digested using endoproteinases trypsin, LysC, GluC (Roche) (1:20, enzyme:protein) for 16 h at ambient room temperature. Trifluoroacetic acid was added to 0.1% total volume to quench enzymatic proteolysis and hydrolyze RapiGest. Samples were centrifuged at 20,000 × *g* for 10 min to precipitate the hydrophobic product of RapiGest hydrolysis. The entire volume of liquid was then transferred to a new tube and stored at −40 °C until time of analysis.

Digested protein mixtures were separated using in-house-fabricated microcapillary high performance liquid chromatography (HPLC) columns. Precolumns were assembled using 360 μm outer diameter (OD) × 75 μm internal diameter (ID) fused silica (Polymicro Technologies), packed with 4–5 cm of 3 μm C4 beads (Vydac) or 6–8 cm of 15 μm C18 beads (YMC). The 360 μm OD × 50 μm ID fused silica analytical columns were packed with 6–8 cm of 3 μm C4 (Vydac) beads and featured a laser-pulled, 2- to 5-μm-wide electrospray ionization (ESI) emitter (Sutter Instruments). Approximately 5–20 pmol of digested protein (peptides) and 100 fmol of angiotensin II peptide (DRVYIHPFHL) and vasoactive peptide (HSDAVFTDNYTR) was pressure bomb-loaded directly onto a precolumn. After loading, pre- and analytical-columns were assembled and connected to an Agilent 1100-series HPLC in-line with a high-resolution mass spectrometer: an LTQ-fourier transform (FT), an LTQ-Orbitrap, or a Velos-Orbitrap (Thermo Fisher Scientific). HPLC buffers for reversed phase separation consisted of 0.1% acetic acid in water for Buffer A, and 70% acetonitrile and 0.1% acetic acid in water for Buffer B. Peptides were eluted using a gradient of 0–60% B in 120 min, 60–100% B in 10 min, constant 100% B for 10 min, and 100–0% B in 10 min. Electrospray ionization was operated at 2,000 V and was heated at 200 °C. Mass spectrometers were operated in data-dependent mode such that the full MS scan from *m/z* 300–2,000 was acquired in the ion cyclotron resonance cell (R = 100,000 at *m/z* 400) or orbitrap (R = 60,000 at *m/z* 400), and was followed by multiple successive MS/MS scans; 6 scans for ETD (3 *m/z* isolation window, 30 ms reaction time with azulene radical anion), and 10 scans for collisionally activated dissociation (CAD) (3 *m/z* isolation window, 35% normalized collision energy) were taken using a dynamic exclusion of 30 s and a repeat count of 3. FT MS^2 targeted analysis of trimethylated, doubly phosphorylated CENP-A G1–K48 was performed using a target of *m/z* = 563.30 and an isolation width of 3.0 *m/z*.

MS spectra were searched using the Open Mass Spectrometry Search Algorithm (OMSSA, version 2.1.1) against the RefSeq human database (downloaded June 2009). Cysteine carbamidomethylation (Δ 57.021464 Da) was used as a static modification. Variable modifications considered were acetylation (Δ 42.010565 Da) on K; methylation (Δ 14.015650 Da) on K and R; oxidation (Δ 15.994915 Da) on M; and phosphorylation (Δ 79.966331 Da) on S, T, and Y. Spectra were further inspected manually using Qual Browser (Thermo Electron Corp.) software for combinations of posttranslational modifications on CENP-A resulting from the specific enzymatic digests above. Peptide abundances were quantified by generating extracted ion chromatograms (XIC) for all charge states observed, and summing peak areas. For peptides longer than 20 residues in length, Isotope Pattern Calculator software (Pacific Northwest National Laboratories; http://omics.pnl.gov/) was used to calculate the expected dominant isotopic peak (based on natural abundance of carbon as 1.070% ^{13}C and nitrogen as 0.368% ^{15}N) for plotting XICs. Binding partners were quantified by selecting peptides that were not extreme in length, pI, and hydrophobicity such that ionization or chromatography would dramatically influence detection. Binding partner peptides were confirmed by CAD MS^2 spectra.

Nucleosomal Array Reconstitution and Analytical Ultracentrifugation. Nucleosomal arrays were reconstituted from purified components using salt dialysis (2). Human histones H3, H4, H2A, H2B were purified as monomers and mixed to form $(H3/H4)_2$ tetramer and (H2A/H2B) dimer complexes whereas human $(CENP\text{-}A/H4)_2$ was purified from a bicistronic vector as a tetramer (3–5). The 601 12 × 200-bp DNA template was purified by anion exchange chromatography (6, 7). Histone tetramers and dimers were combined with the DNA in 2 M NaCl and taken through stepwise dialysis against 1× TE (10 mM Tris, pH 7.8, 0.25 mM

EDTA) with 1 M NaCl, followed by 1× TE with 0.75 M NaCl, and completing dialysis against 1× TEN (1× TE with 2.5 mM NaCl). Reconstituted nucleosomal arrays were characterized by sedimentation velocity in a Beckman Coulter XLA analytical ultracentrifuge. Boundaries were analyzed according to Demeler and van Holde to obtain the integral distribution of sedimentation coefficients (8). The extent of DNA template saturation with nucleosomes was determined from the sedimentation coefficient at 50% boundary fraction and by AvaI digestion. Array folding was measured in the presence of 1.8 mM $MgCl_2$.

Affinity Purification of CENP-A from the Prenucleosomal Pool. To arrest cells in mitosis, HeLaS3 cells stably expressing hemagglutinin (HA)-LAP–tagged CENP-A were treated with 400 ng/mL nocodazole for 18 h before harvesting. After homogenizing and pelleting the cells (as described above), the supernatant was used for αGFP immunoprecipitation. All steps were done as in the case of nucleosome immunoprecipitation except, after bead binding, the beads were washed with Bead Buffer that contained 1 M NaCl.

Immunofluorescence. HeLaS3 cells stably expressing HA-LAP–tagged CENP-A were grown on glass coverslips. The cells were fixed in a 4% formaldehyde solution for 10 min, permeabilized with 0.1% Triton X-100 and processed for immunofluorescence. Human anticentromere antibodies (ACA) (Antibodies, Inc.) were used at 1:500, and anti-HA.11 (Covance) mAb was used at 1:1,000. All Cy3 and C5 conjugated secondary antibodies were obtained from Jackson ImmunoResearch Laboratories and used at 1:200.

In Vitro N-Terminal RCC1 Methyltransferase Reactions. Ten-amino acid N-terminal fusion of CENPA-EGFP wild-type and mutant substrate proteins was expressed from a modified pET15b (Novagen) expression vector in BL21 *Escherichia coli* and purified on Ni-NTA beads (Qiagen), then cleaved using Factor X (Sigma-Aldrich). N-terminal RCC1 methyltransferase (NRMT) reactions were conducted as described previously (9).

Trypsin Digestion of CENP-A E10–G24 Peptide. Synthetic peptides resembling CENP-A E10–G24 and terminating in cysteine were synthesized in doubly phosphorylated [EAPRRRS(ph)PS(ph)PTPTPGC] and unmodified [EAPRRRSPSPTPTPGC] forms. Lyophilized peptides were reconstituted in 100 mM NH4HCO3, were reduced using DTT, and alkylated with iodoacetamide. Peptides were treated with trypsin (1:20, trypsin:peptide) and incubated at 37 °C for 0, 0.25, 1, 4, 8, and 19 h. Proteolytic reactions were quenched by adding glacial acetic acid to 0.1% final volume. The 10-pmol peptide digest solution was bomb-loaded directly onto an analytical column (prepared as above) and gradient eluted into an LTQ-FT (Thermo Fisher Scientific).

Ion Mobility Mass Spectrometry. Electrospray (ESI) Ion Mobility Mass Spectrometry (IM-MS) experiments were performed using a SYNAPT G2 HDMS mass spectrometer (Waters Corp.) equipped with a nano-ESI source and a traveling-wave ion mobility cell (TW-IMS) maintained at 3 mbar of nitrogen. The TW-IMS separation used a wave velocity of 450 m/s and wave amplitude of 20 V. Collision cross section (CCS) measurements were performed by using the method described previously by Ruotolo et al. (10). Literature values for all CCS calibrant peptides were obtained from the CCS database (11). The phosphorylated and unmodified E10–G24 synthesized peptides were diluted to 1 μM with 50% acetonitrile containing 0.1% formic acid. The sample was directly infused at a flow rate of 0.4 μL/min.

Transient Transfection and Imaging. Wild-type and CENP-A containing serine to alanine mutations at positions 16 and 18 were C-terminally GFP tagged. HCT116 cells were cotransfected with 1 μg of GFP-tagged CENP-A and 50 ng of histone H2B-mRFP in a six-well plate using lipofectamine (Invitrogen). Forty-eight hours posttransfection, cells were blocked with 2 mM thymidine, released for 10.5 h, and fixed in 4% formaldehyde. Images were collected using a 100× objective on a DeltaVision deconvolution microscope using a Photometrics CoolSNAP HQ^2 camera. Images were deconvolved and presented as stacked images.

1. Cheeseman IM, Desai A (2005) A combined approach for the localization and tandem affinity purification of protein complexes from metazoans. *Sci STKE* 2005(266):pl1.
2. Carruthers LM, Tse C, Walker KP, 3rd, Hansen JC (1999) Assembly of defined nucleosomal and chromatin arrays from pure components. *Methods Enzymol* 304:19–35.
3. Sekulic N, Bassett EA, Rogers DJ, Black BE (2010) The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. *Nature* 467(7313):347–351.
4. Luger K, Rechsteiner TJ, Richmond TJ (1999) Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol* 304:3–19.
5. Black BE, et al. (2004) Structural determinants for generating centromeric chromatin. *Nature* 430(6999):578–582.
6. Huynh VA, Robinson PJ, Rhodes D (2005) A method for the in vitro reconstitution of a defined "30 nm" chromatin fibre containing stoichiometric amounts of the linker histone. *J Mol Biol* 345(5):957–968.
7. Lowary PT, Widom J (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol* 276(1):19–42.
8. Demeler B, van Holde KE (2004) Sedimentation velocity analysis of highly heterogeneous systems. *Anal Biochem* 335(2):279–288.
9. Tooley CE, et al. (2010) NRMT is an alpha-N-methyltransferase that methylates RCC1 and retinoblastoma protein. *Nature* 466(7310):1125–1128.
10. Ruotolo BT, Benesch JL, Sandercock AM, Hyung SJ, Robinson CV (2008) Ion mobility-mass spectrometry analysis of large protein complexes. *Nat Protoc* 3(7):1139–1152.
11. Valentine SJ, Counterman AE, Clemmer DE (1999) A database of 660 peptide ion cross sections: Use of intrinsic size parameters for bona fide predictions of cross sections. *J Am Soc Mass Spectrom* 10(11):1188–1211.

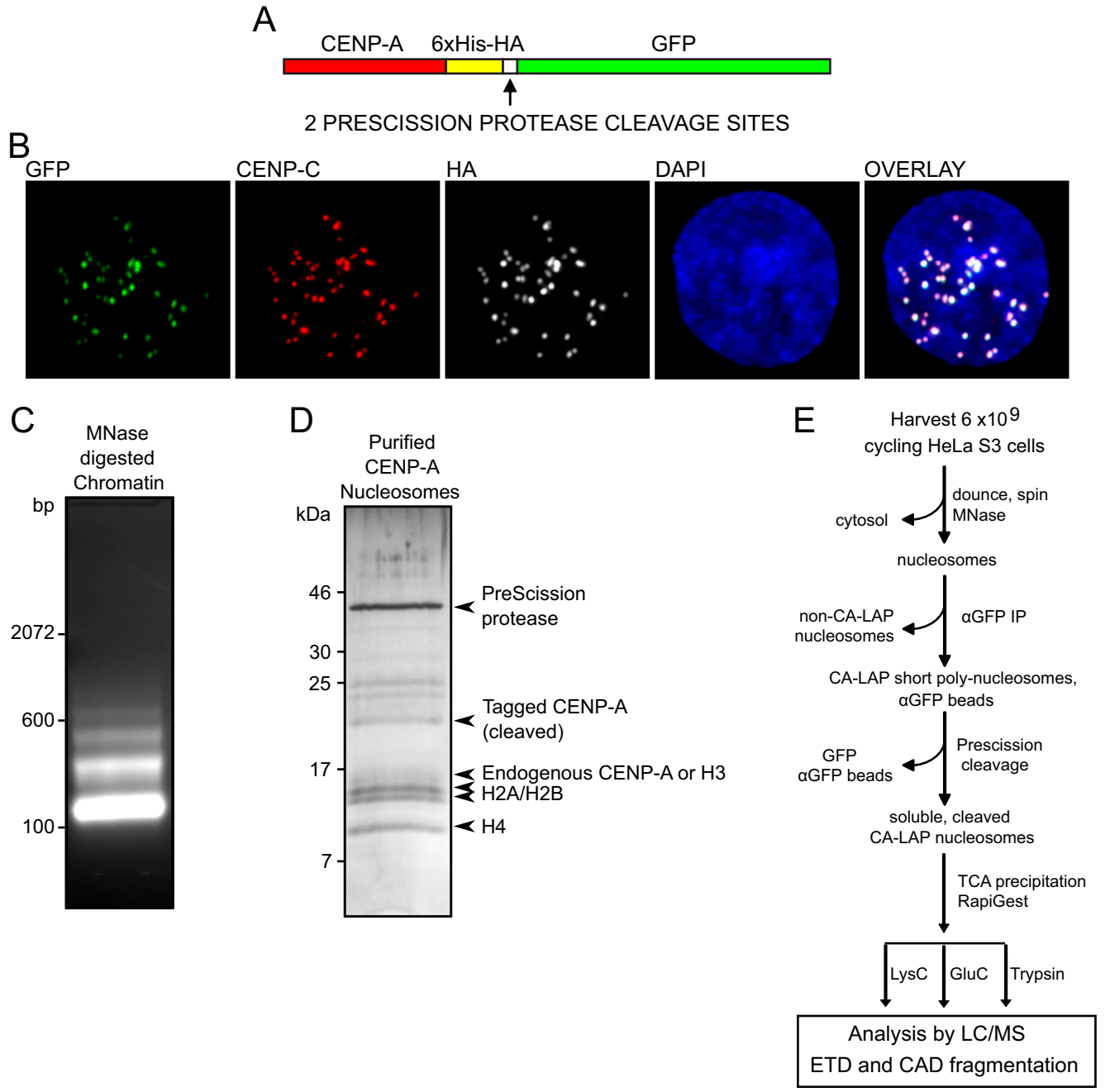

Fig. S1. CENP-A nucleosomes are purified from HeLa centromeres. (*A*) HeLa S3 cells stably express a tandem epitope-tagged CENP-A (CENP-A-LAP) comprised of human CENP-A with C-terminal additions of 6xHis, hemagglutinin (HA), two PreScission protease cleavage sequences, and green fluorescent protein (GFP). (*B*) Fluorescence microscopy shows colocalization of αHA (Cy5, yellow) and GFP (green) signals with CENP-C (Cy3, red), within DAPI-stained (blue) nuclei. (*C*) Chromatin preparations were treated with micrococcal nuclease (MNase) to generate mononucleosome-enriched material for immunoprecipitation (IP). (*D*) Post-PreScission cleaved CENP-A-LAP complexes observed by SDS/PAGE and silver staining. (*E*) IP buffer detergents were removed by trichloroacetic acid (TCA) precipitation followed by rinsing in acetone, and proteins were resuspended in a solution containing 0.1% RapiGest. Peptides were generated using three proteases: trypsin, LysC, or GluC. Each digest was analyzed using high-resolution MS and a combination of electron transfer (ETD) and collision-activated (CAD) fragmentation methods.

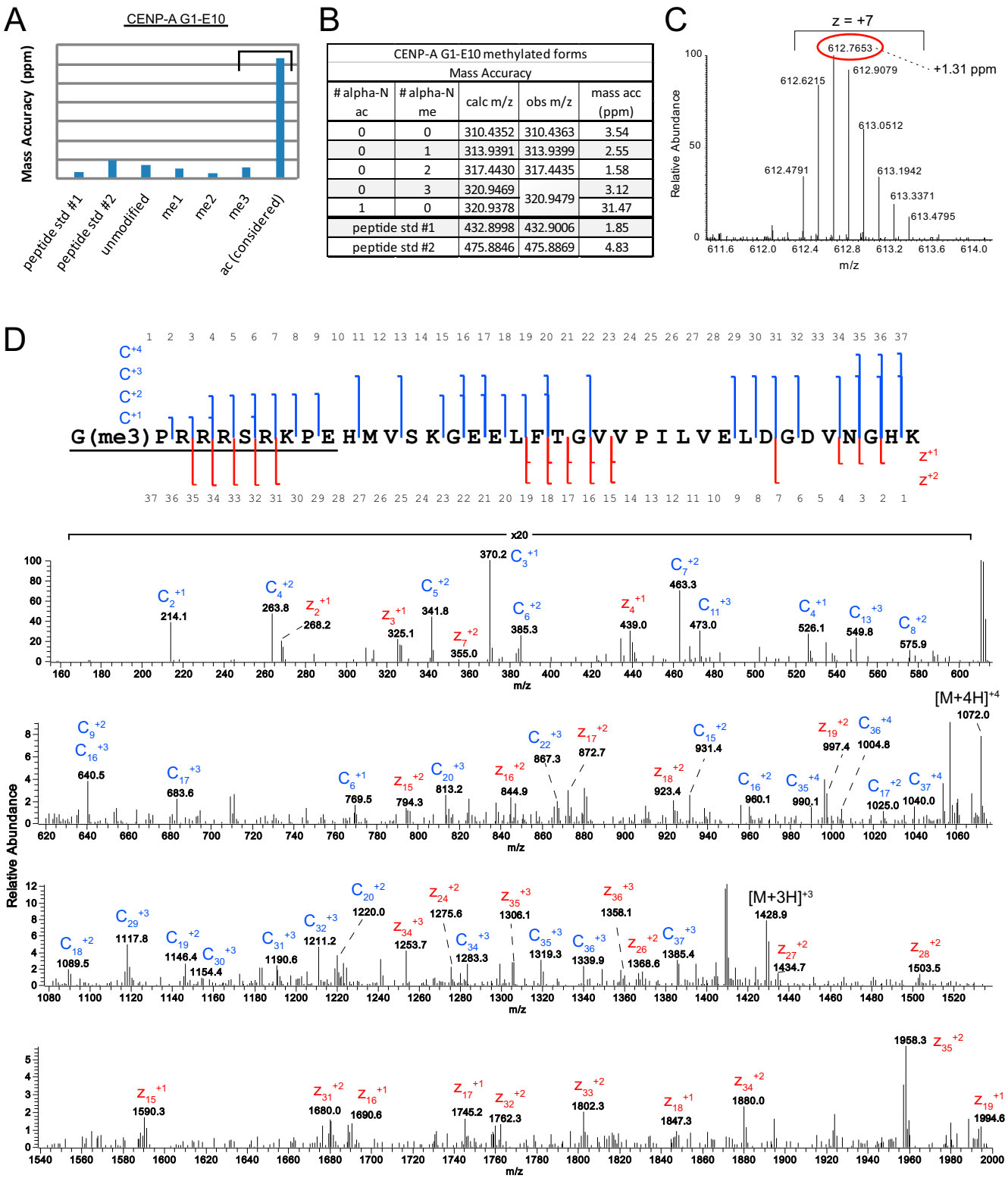

Fig. S2. CENP-A is N-terminally trimethylated, not acetylated; catalyzed by NRMT in vitro. (*A*) Mass accuracy in parts per million (ppm) is plotted for two internal standard peptides as well as for each of the posttranslationally modified (PTM) forms of CENP-A G1–E10 peptides. Using high-resolution MS, peptides could be measured within 5 ppm accuracy of the calculated masses, allowing unambiguous identification of trimethylated and acetylated forms. CENP-A G1–E10 observed at *m/z* 320.9479 was trimethylated (3.12 ppm) and not acetylated (31.47 ppm) (bracket). (*B*) Mass accuracy values plotted in *A*. (*C*) Engineered protein featuring the N-terminal 10 aa of CENP-A fused to EGFP was reacted in vitro with 6His-NRMT. Digestion with LysC revealed a trimethylated N-terminal peptide. The ^{13}C x 2 isotope of the $[M + 7H]^{+7}$ ion was observed at 1.31 ppm mass accuracy. (*D*) ETD MS^2 of LysC-digested peptide in *C*. Value at bracket indicates magnification level for region of spectrum. me, methylation; ac, acetylation.

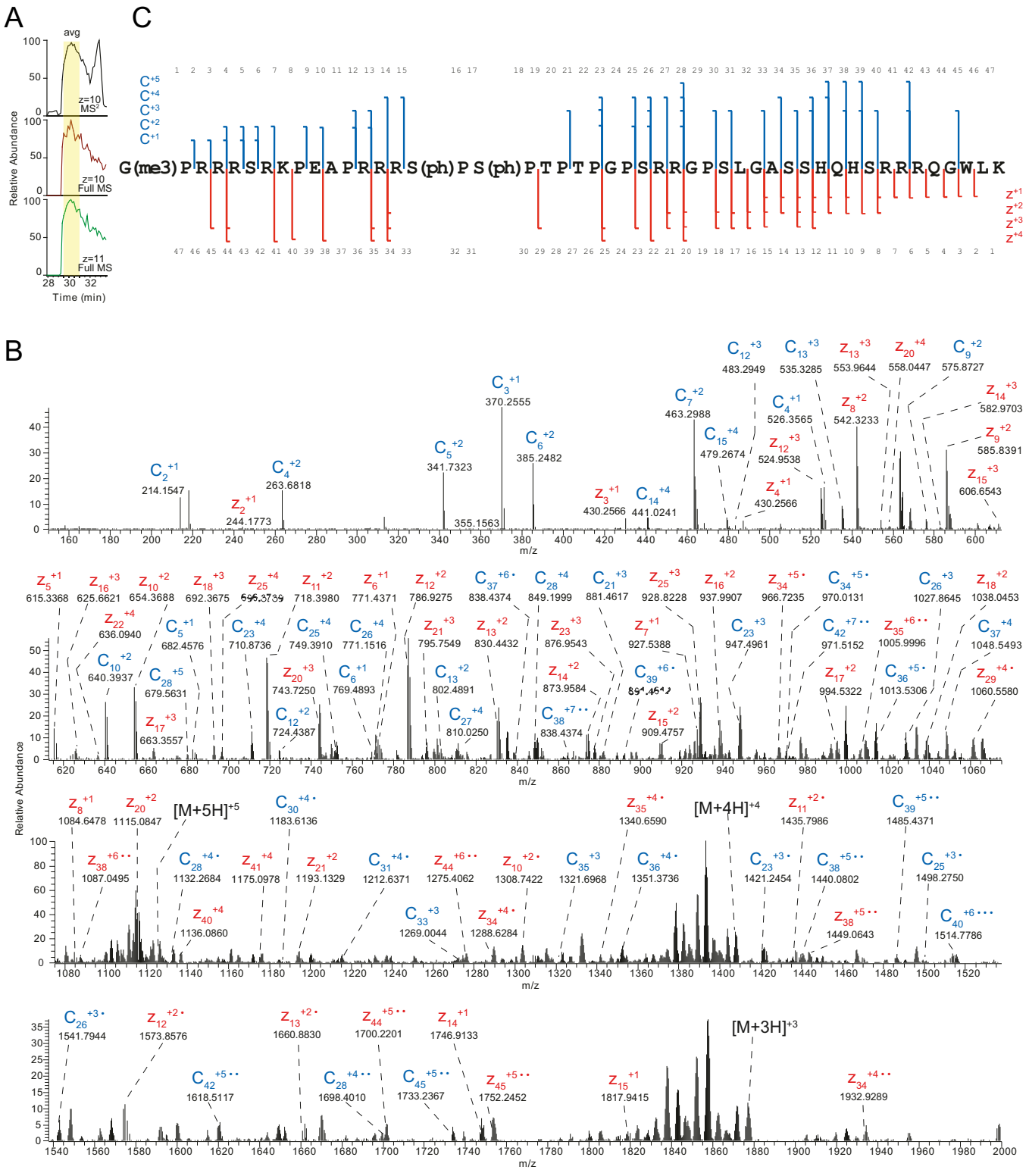

Fig. S3. High-resolution ETD MS^2 spectrum of trimethylated, doubly phosphorylated CENP-A G1–K48. (*A*) Chromatographic plots for trimethylated, doubly phosphorylated CENP-A G1–K48. Targeted $[M + 10H]^{+10}$ MS^2 scans (*Top*), XIC of $[M + 10H]^{+10}$ (*Middle*), and $[M + 11H]^{+11}$ ions (*Bottom*) are plotted. (*B*) Sixteen high-resolution ETD-MS^2 scans of the targeted $[M + 10H]^{+10}$ were averaged. High-resolution MS^2 allows unambiguous detection of nondissociated charge-reduced fragments, which are observed as having an apparent increase in mass equal to one proton and are denoted by • and the original charge state. (*C*) Sequence coverage of the ETD MS^2 spectrum of trimethylated, doubly phosphorylated CENP-A G1–K48. Nondissociated charge-reduced species are reported as the apparent charge state (e.g., +4• is reported as +3 and +6•• is reported as +4).

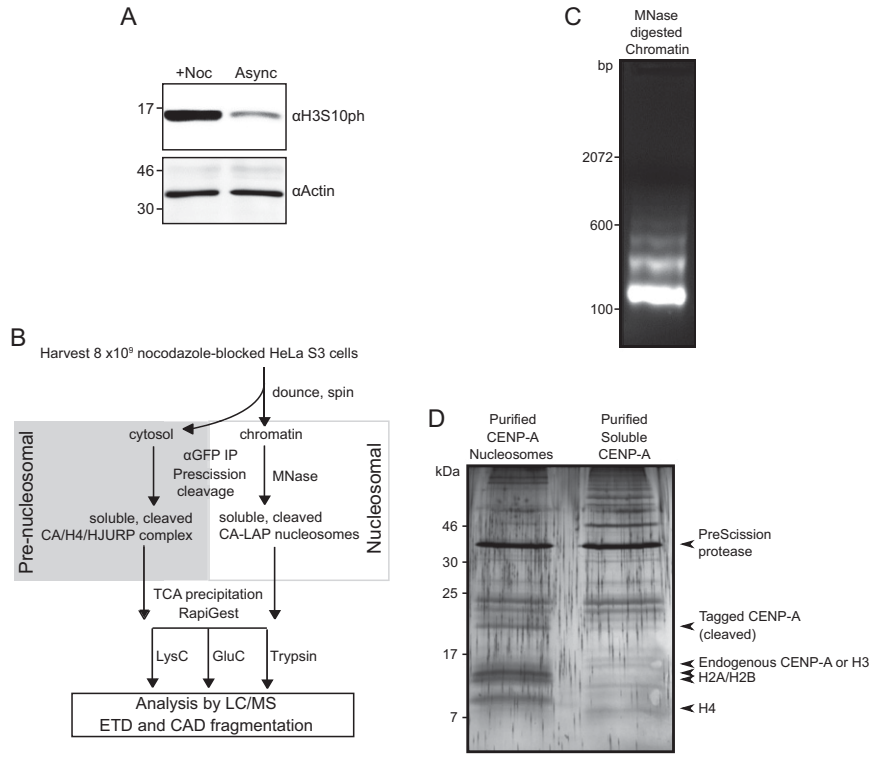

Fig. S4. Mitotic cells are harvested to generate nucleosomal and prenucleosomal fractions. (*A*) Western blot analysis of lysates from nocodazole (Noc)-blocked and asynchronous cells. The mitotic marker, phosphorylated Ser10 on H3, is enriched upon nocodazole treatment. (*B*) Schematic for fractionation of nocodazole-blocked cells and LC-MS analysis. (*C*) MNase digestion of mitotic chromatin results in a mononucleosome-rich material for immunoprecipitation. (*D*) Nucleosomal and prenucleosomal immunoprecipitation samples are observed in silver-stained SDS/PAGE gels. αH3S10ph, antiserine 10 phophorylated histone H3; HJURP, holliday junction recognition protein; CA-LAP, CENP-A-LAP.

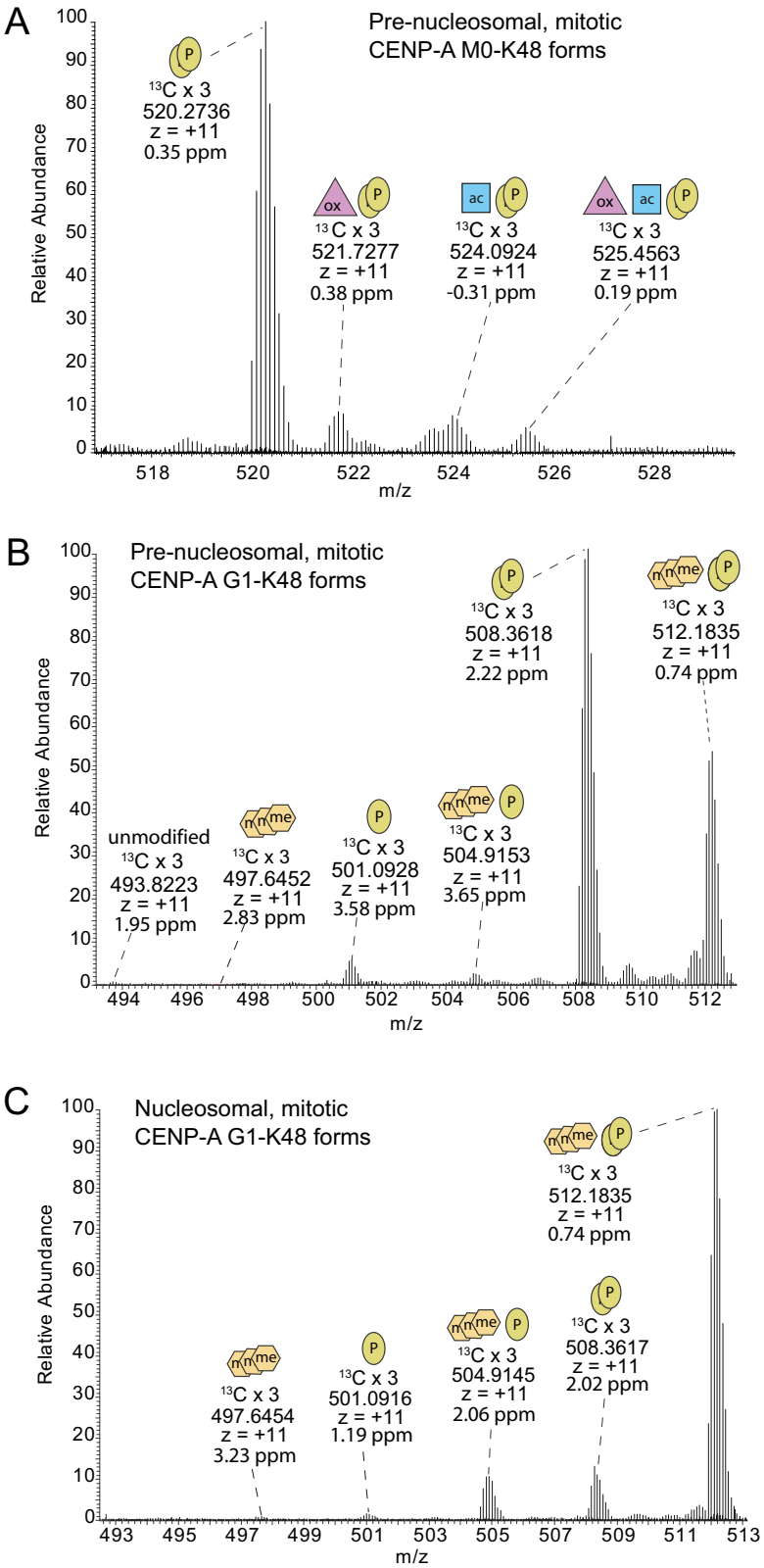

Fig. S5. Full MS spectra of mitotic CENP-A G1–K48 forms. (*A*) Initiator methionine-containing forms of unmodified and αN-acetylated prenucleosomal CENP-A (M0–K48 + 2 ph and M0–K48 + 2 ph + ac), were shown in 115 averaged full MS scans. A small fraction of both unmodified and αN-acetylated forms of M0–K48 were oxidized. Methionine oxidation is a common artifact of sample preparation. (*B*) Mitotic prenucleosomal CENP-A N-terminal tail PTM forms were observed in 17 full MS scans averaged. (*C*) Mitotic nucleosomal CENP-A N-terminal tail PTM forms were observed in 15 full MS scans averaged. Yellow "me" hexagon, methylation; green "P" oval, phosphorylation; blue "ac" square, acetylation; purple "ox" triangles, M0 oxidation.

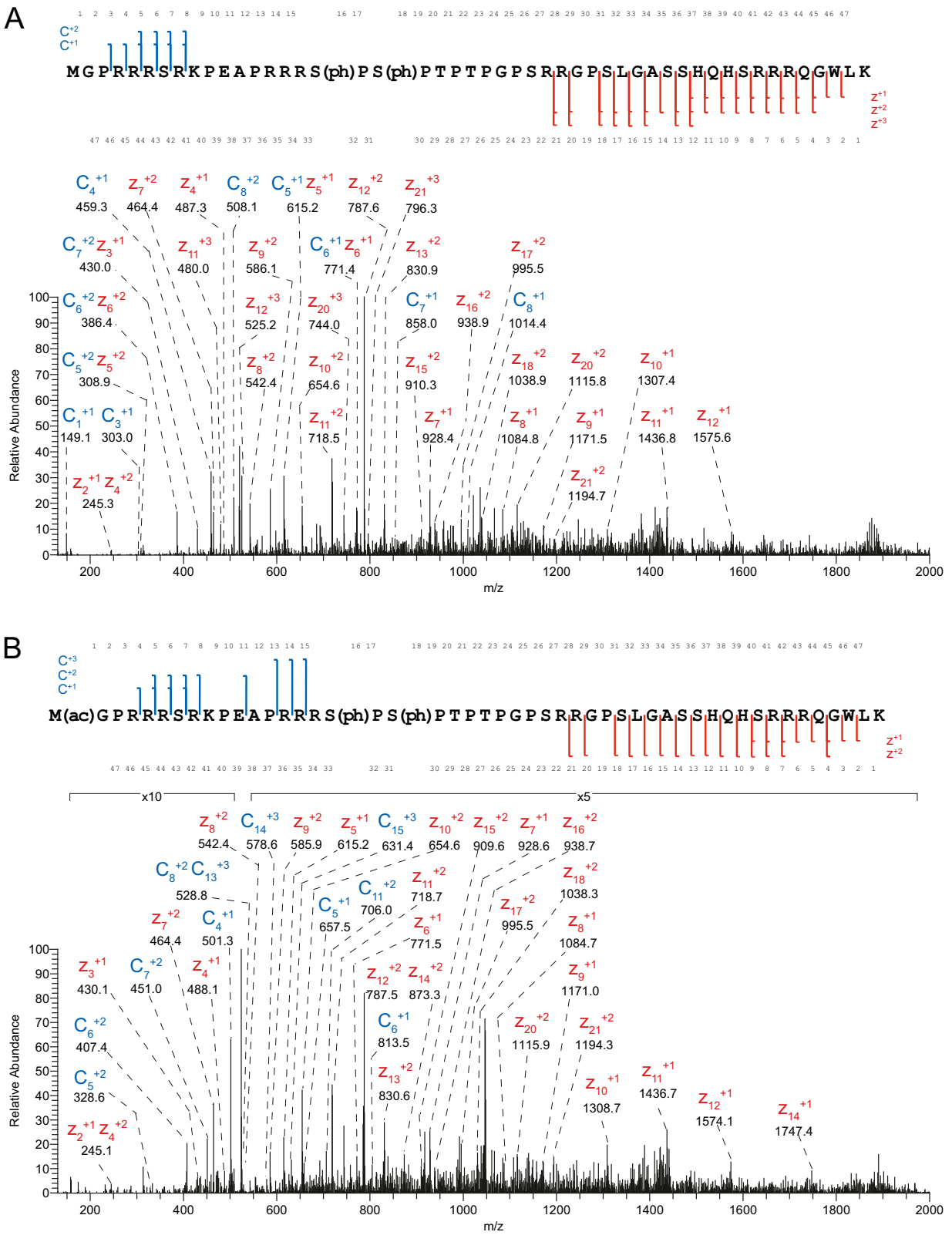

Fig. S6. ETD MS^2 spectra of CENP-A M0–K48 forms. (*A*) ETD MS^2 of CENP-A M0–K48 + 2 ph. (*B*) ETD MS^2 of CENP-A M0–K48 + 2 ph + ac. Value at bracket indicates magnification level for region of spectrum. ph, phosphorylation.

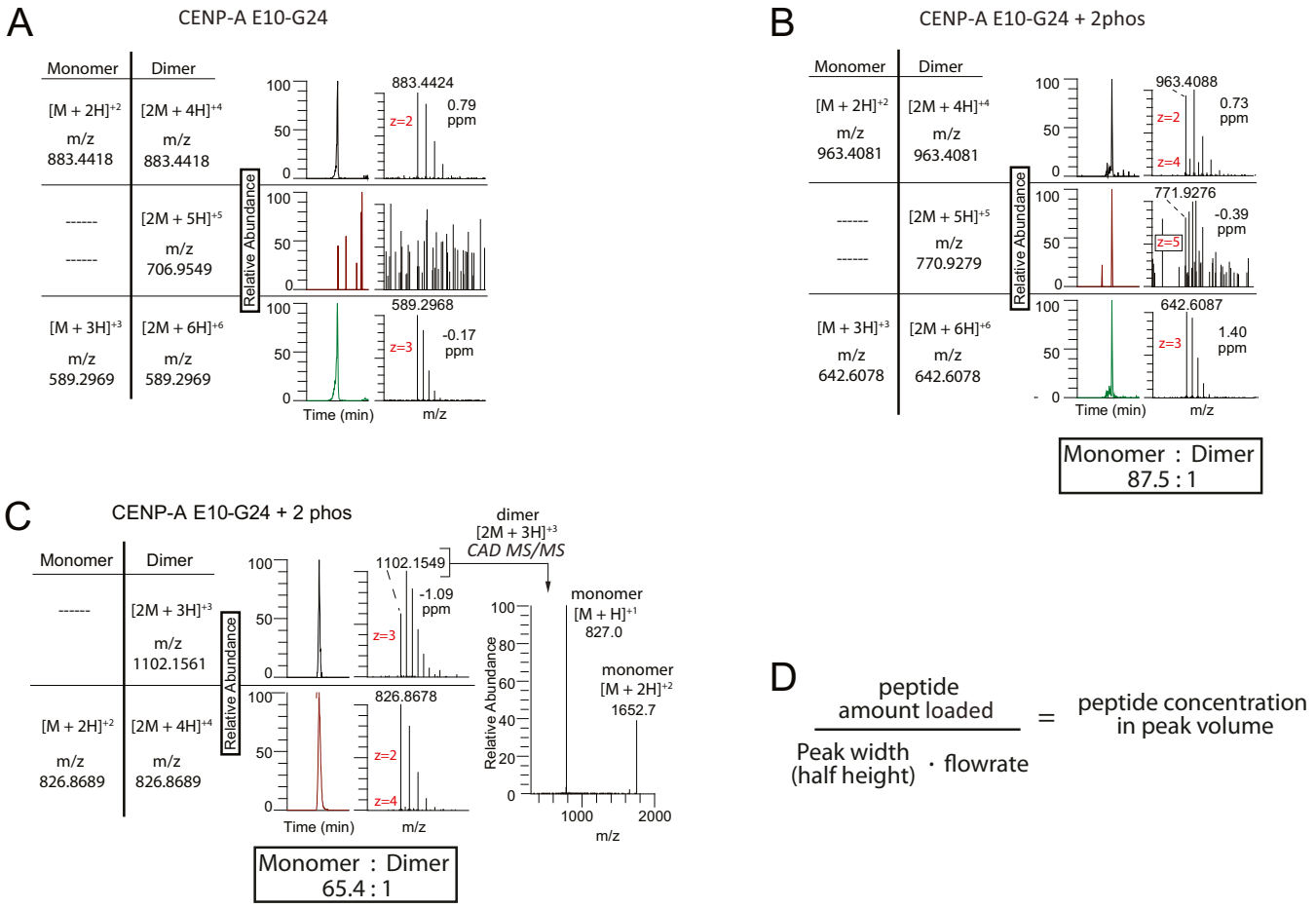

Fig. S7. CENP-A phosphorylation causes low-affinity multimerization. (*A*) Noncovalent dimers were resolved by HPLC and detected by electrospray MS. Unmodified CENP-A E10–G24 peptide was detected in +2 and +3 charge states as the monomeric form only, and not as dimeric. (*B*) Doubly phosphorylated E10–G24 was analyzed similar to *A* and detected in +2, and +3 charge states as a monomer and in +4 and +5 charge states as a dimer. (*C*) Doubly phosphorylated CENP-A R14–R27 was eluted as an HPLC peak and observed by electrospray MS. Noncovalent dimers exists as *m/z* values, which are either redundant with monomers or unique to higher-order multimers. R14–R27 was detected in monomeric and dimeric forms. Dimeric phosphorylated R14–R27 $[2M + 3H]^{+3}$ was fragmented by CAD to produce monomeric $[M + 2H]^{+2}$ and $[M + 2H]^{+2}$ species, conclusively demonstrating peptide dimerization. (*D*) Calculation for

Legend continued on following page

estimating the concentration of peptide in an HPLC-eluted peak. (*E*) Peptide concentration in the eluted peak of doubly phosphorylated E10–G24 peptide described in *B*. (*F*) Peptide concentration in the eluted peak of doubly phosphorylated R14–R27 described in Fig. 5*D*. (*G*) Calculation to estimate the local concentration of nucleosomes in condensed chromatin. The 11.0-nm center-to-center distances of nucleosomes were reported previously by Scheffer et al. (1). N_A, Avagadro's number.

1. Scheffer MP, Eltsov M, Frangakis AS (2011) Evidence for short-range helical order in the 30-nm chromatin fibers of erythrocyte nuclei. *Proc Natl Acad Sci USA* 108(41):16992–16997.

Dataset S1. Accounting and semiquantitation of identified peptides

Dataset S1

(*A*) Full MS peak areas for representative peptides of CENP-A–binding proteins observed in a trypsin digest of nucleosomal CENP-A affinity purified from the chromatin fraction of asynchronously dividing cells (Fig. S1 *A*–*E*). (*B*) Three independent biological replicates of nucleosomal CENP-A isolated from asynchronously cycling cells were digested with LysC, trypsin, or GluC. Protein sequences were digested *in silico*, and calculated products were arranged in sequential order (columns B and C) according to protease (columns D–F). Posttranslational modification status is indicated by an integer in any of four possible categories. ac, acetylation, me, methylation, ph, phosphorylation (columns G–I). For each peptide considered (column J), the expected most abundant isotopic peak (column K) and charge states (column L) were used to calculate corresponding *m/z* values (column M). XICs were plotted using calculated *m/z* values within a 10-ppm window. Peptides less than 5 aa in length were not interrogated. Observed *m/z* values of peptides were compared with calculated values to determine mass accuracy in parts per million (ppm). For *m/z* values that did not result in peptide identification "ND" denotes not detected. Successful data-dependent acquisition of MS^2 spectra is denoted with "Y", yes, or "N", no. For ETD analyses, integrated peak area for XIC plots are summed for all observed charge states for each peptide identified. (*C*) Levels of representative peptides of CENP-A–binding proteins observed in a trypsin digest of prenucleosomal CENP-A affinity purified from the soluble fraction of nocodazole-blocked cells. (*D*) Mitotic prenucleosomal CENP-A fraction was digested with LysC or trypsin, and peptides were quantified as in *B*. (*E*) Levels of representative peptides of CENP-A–binding proteins observed in a trypsin digest of nucleosomal CENP-A affinity purified from the chromatin fraction of nocodazole-blocked cells. (*F*) Mitotic nucleosomal CENP-A was digested with LysC or trypsin and peptides were quantified as in *B* and *D*.