Cell, Volume *135*

Supplemental Data

Cathepsin L Proteolytically Processes Histone H3 During Mouse Embryonic Stem Cell Differentiation

Elizabeth M. Duncan, Tara L. Muratore-Schroeder, Richard G. Cook, Benjamin A. Garcia, Jeffrey Shabanowitz, Donald F. Hunt, and C. David Allis

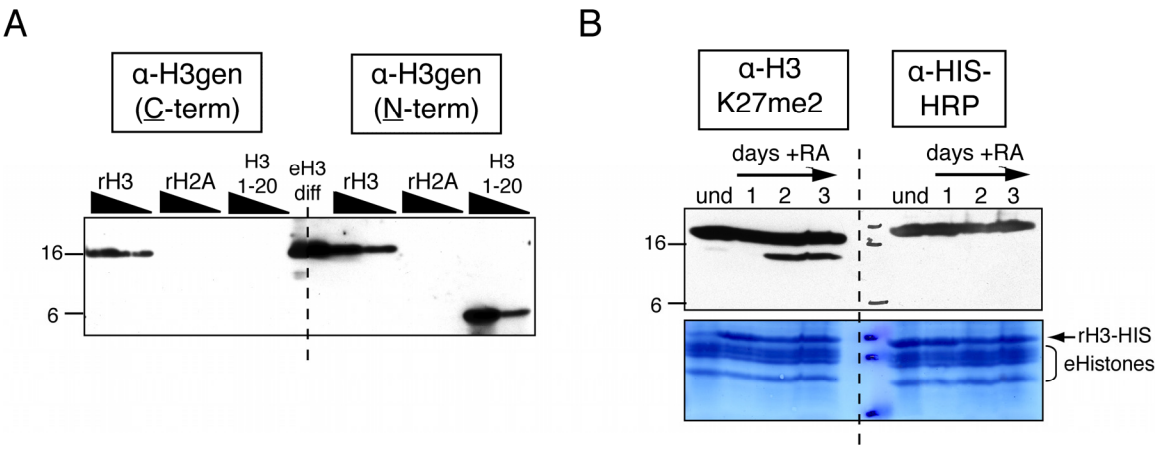

Figure S1: The faster migrating histone H3 sub species appears to be lacking its amino terminus but is not cleaved during whole-cell extract preparation.

(A) recombinant (r) histone H3, rH2A, H3 1-20 peptide and RP-HPLC purified endogenous (e) H3 from differentiated ESCs (eH3 diff) were separated by SDS-PAGE and analyzed by immunoblotting with H3-general antibodies; although the antibody generated against the H3 C-terminus recognized the faster migrating H3 sub-band, the antibody generated against the H3 N-terminus does not.

(B) To test whether uncontrolled proteolysis occurred upon cell lysis, unmodified rH3-HIS was added to SDS-Laemmli sample buffer prior to adding it to cell pellets and solubilization to generate WCEs. The unmodified rH3-HIS showed no evidence of cleavage when added to the lysing cells (right); however, endogenous H3 cleavage was detected in the lysates of 2 and 3 day differentiated ESC, as expected, which was shown by immunoblotting with an H3K27me2 antibody (left).

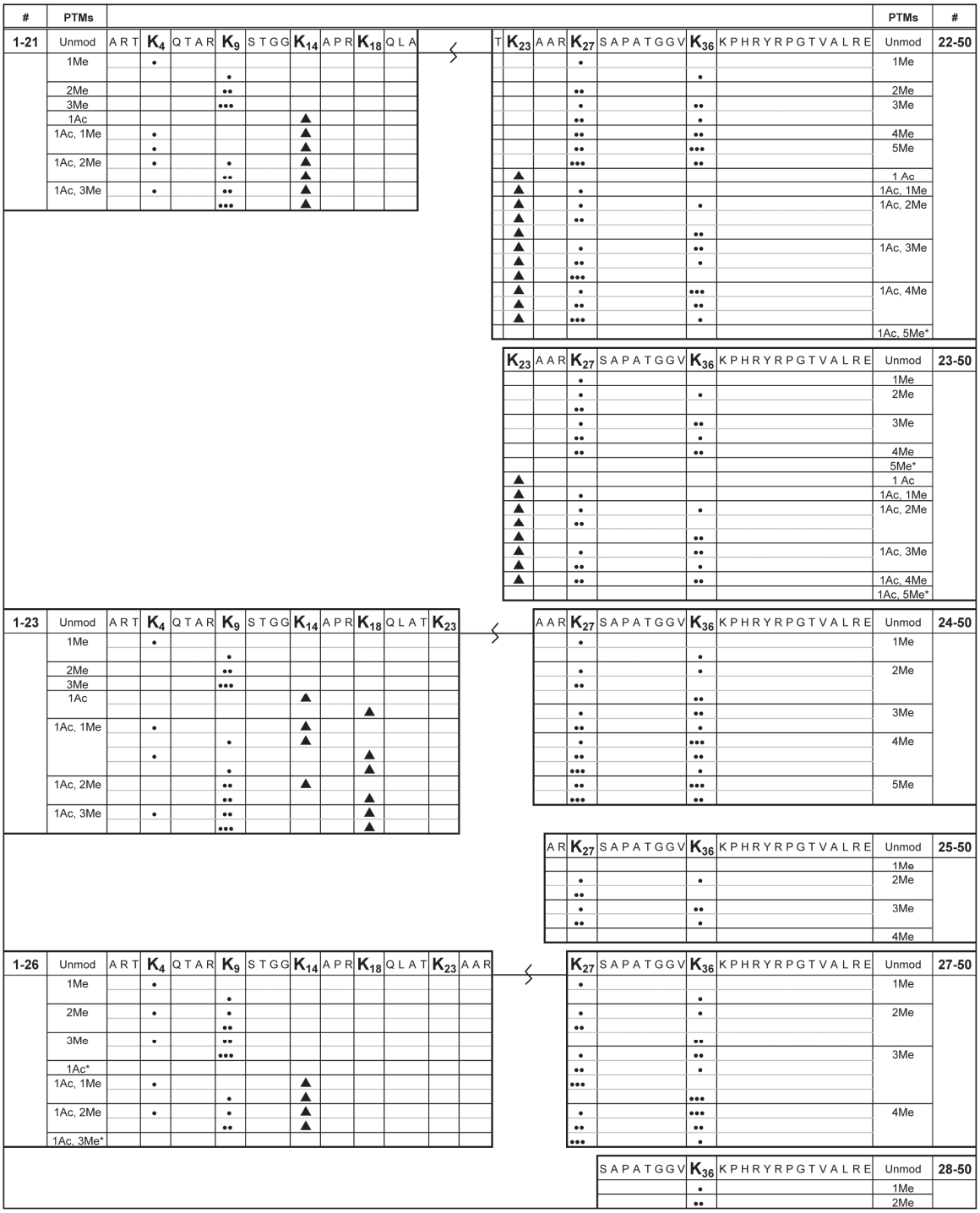

Figure S2: Summary of the post-translational modifications detected on proteolytically cleaved H3 obtained from differentiating ESCs. Material from fraction 54 (Figure 2A) was digested with GluC to produce N-terminal, H3 fragments ending in E50. The resulting mixture was then analyzed by nano-flow

HPLC interfaced with both a linear ion trap-Fourier transform mass spectrometer and a linear ion trap instrument equipped for electron transfer dissociation. Spectra recorded with the former instrument detected six, highly-modified, truncated, N-terminal, histone H3-peptides beginning at residues T22, K23, A24, A25, K27 and S28 and ending at E50 (Figure 2B, right panels). Post-translational modifications detected by recording electron transfer dissociation (ETD) mass spectra on different isoforms of these six peptides are shown above in the right panel. Sixty-seven different forms of the above six peptides were characterized. Ten contain marks associated with active transcription (H3K23Ac, H3K36me, H3K36me2 or H3K36me3), ten contain marks associated with gene silencing (H3K27me, H3K27me2 or H3K27me3), and 43 contain combinations of above active and repressive marks. Summed ion currents for all charge states and modified forms of the above sequences were employed to estimate the relative abundances of the six peptides beginning with the indicated residues as follows: T22 and A24 > than K23 and K27 > A25 and S28 (see right hand list). The peptide beginning with R26 was not detected. Three of the most abundant, complementary, N-terminal fragments generated by proteolytic cleavage of H3 (A1-A21, A1-K23, and A1-R26) were also detected in HPLC fraction 54 (Figure 2B, left panels). Post-translational modifications detected on different isoforms of these three peptides are shown above in the left panel. Monomethyl-, dimethyl-, and trimethyl-marks are depicted by one, two, and three solid circles, respectively. Acetyl marks are depicted by triangles. An asterisk indicates that

the modified isoform was detected by accurate mass measurement only. All ETD spectra were interpreted manually.

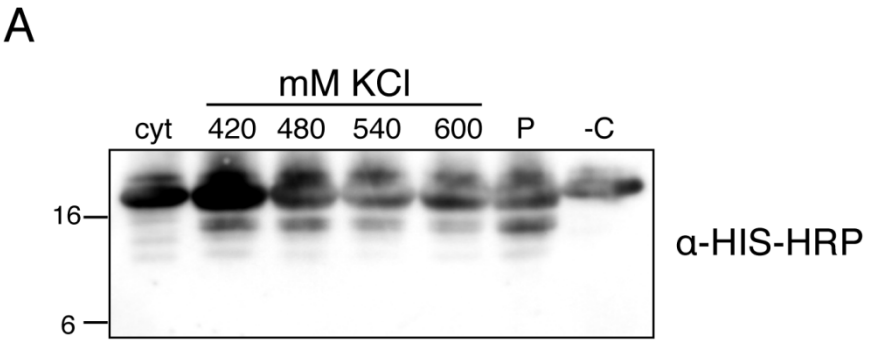

Figure S3: High salt extraction of nuclei does not efficiently extract H3 cleavage activity.

(A) Nuclear extracts were prepared as described (Dignam et al., 1983) from 3 days +RA differentiated cells; after the initial extraction of nuclear proteins with 420 mM KCl (420), the chromatin pellet (P) was further extracted by sequential 60mM increases in KCl concentration. The remaining chromatin pellet was then solubilized by sonication in buffer A. The cytosolic (cyt), high salt, and chromatin extracts were then assayed for H3 cleavage activity using the H3 cleavage assay described in Figure 3.

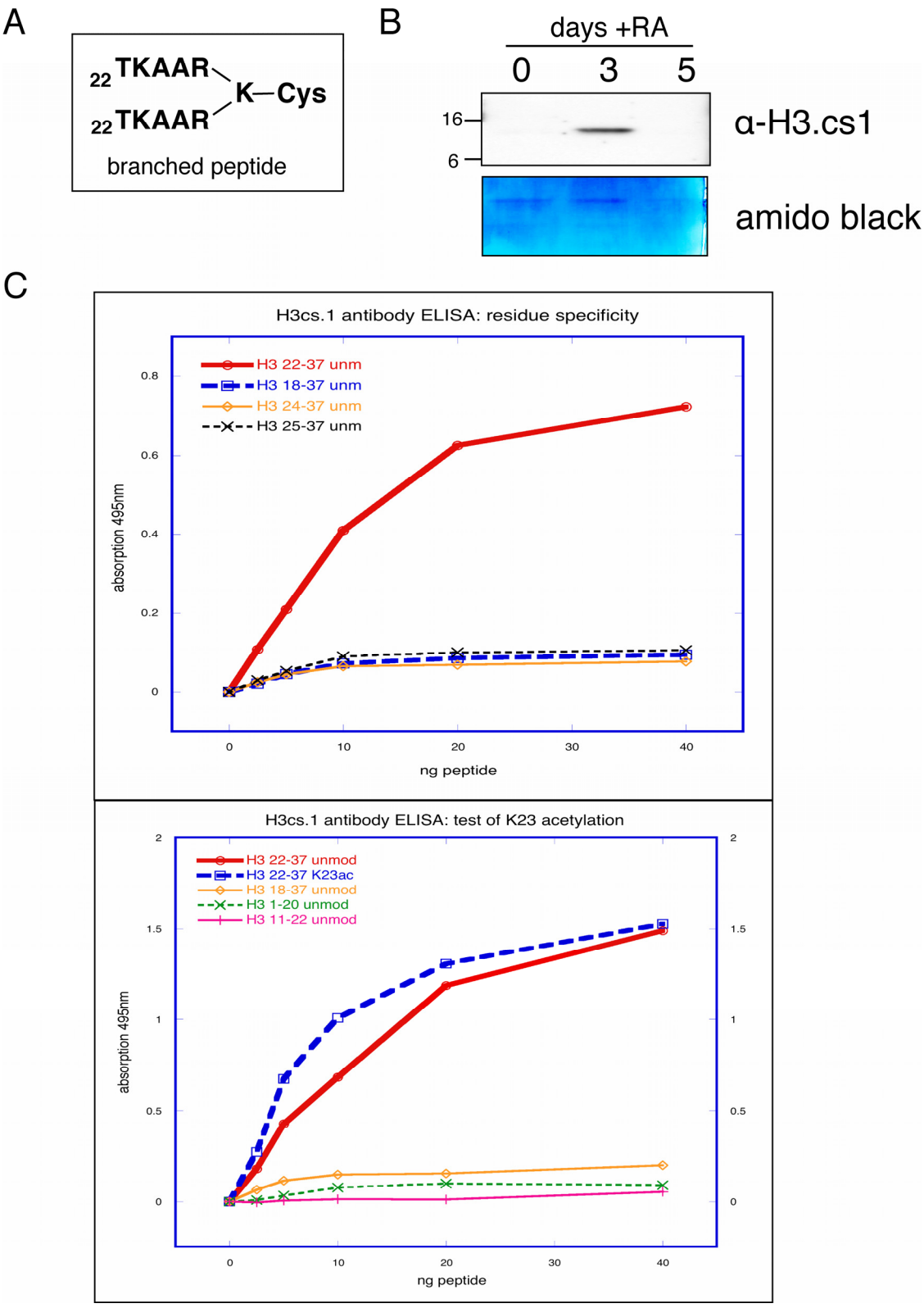

Figure S4: characterization of H3.cs1 antibody.

(A) A 2x branched peptide sequence used to generate antibody (see Experimental Procedures for details)

(B) Rabbit serum was tested for biological specificity by immunoblotting on WCEs from undifferentiated, 3 days +RA, and 5 days +RA.

(C) Rabbit serum was tested for amino acid sequence and modification specificity (H3K23ac) by ELISA

>gi|6753558|ref|NP_034114.1| cathepsin L preproprotein Mus musculus]

MNLLLLLAVLCLGTALATPKFDQTFSAEWHQWKSTHRRLYGTNEEEWRRAIWE
KNMRMIQLHNGEYSNGQHGFSMEMNAFGDMTNEEFRQVVNGYRHQKHKKG
RLFQEPLMLKIPKSVDWREKGCVTPVKNQGQCGSCWAFSASGCLEGQMFLKT
GKLISLSEQNLVDCSHAQGNQGCNGGLMDFAFQYIKENGGLDSEESYPYEAKD
GSCKYRAEFAVANDTGFVDIPQQEKALMKAVATVGPISVAMDASHPSLQFYSS
GIYYEPNCSSKNLDHGVLLVGYGYEGTDSNKNKYWLVKNSWGSEWGMEGYIK
IAKDRDNHCGLATAASYPVVN

Four Peptides Detected
ENGGLDSEESYPYEAK

NSWGSEWGMEGYIK

DRDNHCGLATAASYPVVN

DNHCGLATAASYPVVN

Figure S5: Mass spectrometry analysis of hydroxyapatite fractions identifies the presence of cysteine protease Cathepsin L specifically in those fractions exhibiting H3 cleavage activity.

To identify the putative H3 protease, proteins in two of the active fractions (#22 and #23) and one of the adjacent non-active fractions (#20) were digested with trypsin. The resulting peptides were analyzed by nano-flow HPLC interfaced with a linear ion trap-Fourier transform mass spectrometer and several thousand collision activated dissociation (CAD) mass spectra were acquired. By searching theses spectra against a database of murine proteins with the SEQUEST algorithm, we identified more than 1,000 peptides from 80-100 proteins in each fraction. Four low level tryptic peptides (0.5%) detected in each of the two active fractions but not in the inactive fraction, matched to the mature form of the

lysosomal cysteine protease, Cathepsin L. All were present at or below the 0.5% abundance level. Sequences for the detected peptides are listed below the full sequence and the corresponding residues are underlined within the full amino acid sequence of Cathepsin L. No other proteases were detected in the above analyses.

A

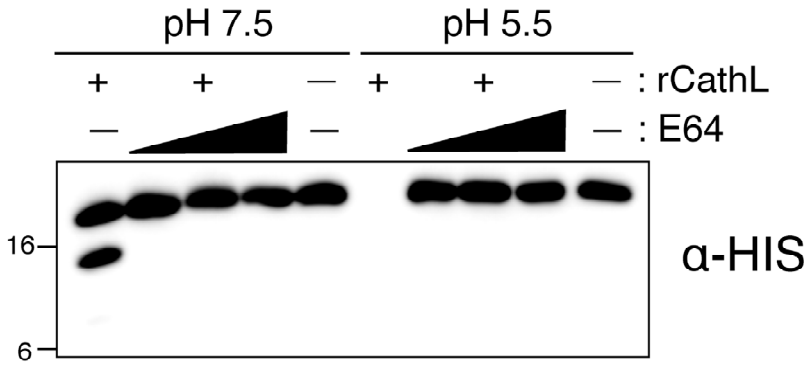

Figure S6: inhibition of rmCathepsin L by cysteine protease inhibitor E64.

(A) recombinant mouse Cathepsin L was incubated with C-terminally HIS-tagged recombinant histone H3 at both pH 7.5 (left) and pH 5.5 (right) in either buffer alone (—) or in the presence of the cysteine protease inhibitor E64 (+); the final concentration of E64 was 10μM, 50μM, and 500μM. rH3 cleavage was assayed by immunoblot, as shown in earlier figures.

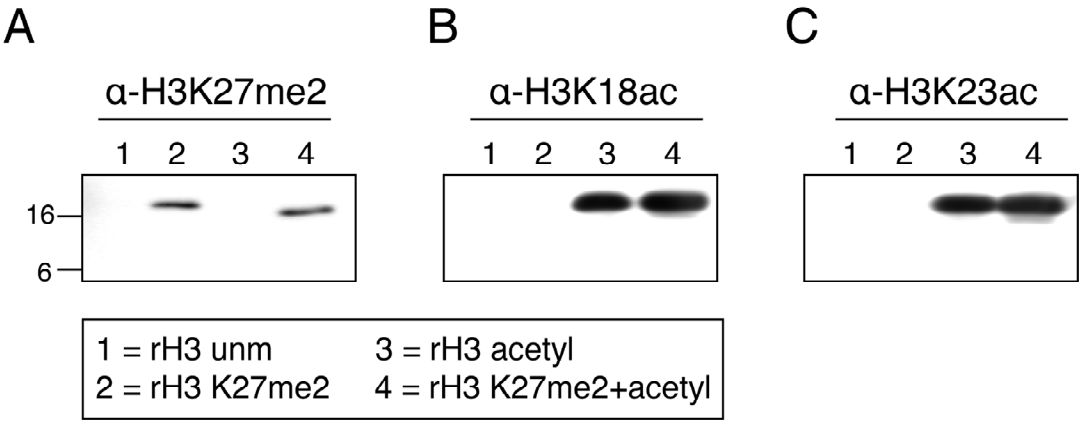

Figure S7: confirmation of methylation and acetylation status of rH3-HIS

(A) C-terminally HIS-tagged recombinant histone H3 was mutated to K27C and then alkylated to convert K27C to K27me2; an aliquot of this protein was then treated with acetic anhydride to acetylate all free lysines. These proteins were then assayed for the presence of K27me2 by immunoblot. The key below applies to panels A-C.

(B) The same recombinant histone H3 proteins were also assayed for the presence of H3K18ac and (C) H3K23ac.

Supplemental Experimental Procedures

Cell culture and differentiation

Cells were grown on gelatin-coated plates without feeder cells and maintained in an undifferentiated state through culture in KODMEM (Invitrogen 1082-9018)), 2mM L-glutamine (Sigma G7513), 15% ES grade fetal bovine serum (Gibco 10439-024), 10^{-4} mM 2-mercaptoethanol, and leukemia inhibitory factor (LIF). To differentiate, cells were plated at ~$1*10^4$ cells/cm^2 and differentiated with LIF-free media ± 100nM all-trans-RA (Sigma R2625). Embryoid bodies were formed by splitting partially trypsinized cells onto non-treated petri dishes and allowing cells to cluster. Cathepsin L was chemically inhibited by adding 10uM [final] Cathepsin L I inhibitor (Calbiochem 219402) to the cell media.

Cellular extract preparation

Briefly, cells were swelled in low salt buffer and then lysed by either mechanical disruption (Dignam et al., 1983) or detergent (Mendez and Stillman, 2000); soluble nuclear proteins were then either extracted with high-salt (Dignam) or released by incubation in no salt buffer and mechanical disruption (Stillman), leaving behind the chromatin pellet. Protease inhibitors were omitted from extract preparations used for activity assays.

Enzyme identification by MS

Both active and inactive fractions were reduced with 1mM DTT at 51^{o}C for 1 hr and alkylated with 2mM iodoacetamide in the dark at RT for 45 mins, followed by digestion with trypsin (Promega Corp., Madison, WI) at an enzyme to substrate ratio of 1:20 (wt:wt) for 6 hrs at 37 ^{o}C; digest was then acidified with glacial acetic

acid and aliquots of samples were analyzed using an LTQ-FT as described above. Data was searched against a mouse database using SEQUEST (Eng et al., 1994). All spectra of interest were manually validated.

Antibodies

The cleavage-specific H3.cs1 antibody was generated as follows: a 2x branched peptide corresponding to mammalian histone H3 sequence 22-26 was conjugated to KLH and injected into rabbits (Covance). Serum was collected and tested for specificity as described in Supplemental Figure S4. The N-terminal H3gen antibody was generated as follows: a peptide corresponding to mammalian histone H3 sequence 1-6 was conjugated to KLH and injected into rabbits (Covance). Serum was collected and specificity is shown in Supplemental Figure S1A.

Plasmid construction and recombinant protein purification

A PCR fragment of the complete mouse histone H3 ORF was cloned into the pET30a plasmid vector to add a C-terminal HIS tag to the coding sequence. Mutations were made using the Quick-Change Mutagenesis II kit (Stratagene). Plasmids were transformed into BL21 *E.coli* and rH3-HIS protein was purified from inclusion bodies using both Ni^{2+}-NTA and C8 columns (RP-HPLC). Acetic anhydride treatment was performed as described previously (Garcia et al., 2007). rH3-HIS was converted to rH3-HIS+K27me2 by first mutating K27 to C and C110 to A and then alkylating the cysteine to dimethyl by published methods (Simon et al., 2007). BPTF and CBX7 proteins were made as described in previous studies (Bernstein et al., 2006; Li et al., 2006).

MS-MS mapping of H3 cleavage sites

RP-HPLC fraction 54 was digested with endoproteinase GluC (Roche Diagnostics, Indianapolis, IN) for 4 hours at 37^{o}C (1:20 wt: wt) and loaded onto a C18 packed capillary column as previously described (Martin et al., 2000). Samples were analyzed by nanoflow HPLC-microelectrospray ionization on a linear quadrupole ion trap-Fourier transform mass spectrometer (LTQ-FT; Thermo Electron) for accurate mass and a Thermo LTQ instrument modified for electron transfer dissociation (ETD) and proton transfer charge reduction (PTR) for adequate tandem mass spectrometry (MS/MS) as previously described (Coon et al., 2005; Syka et al., 2004a; Syka et al., 2004b). Data was manually interpreted, as well as searched against an H3 database using OMSSA (Geer et al., 2004).

Fluorescence anisotropy statistical analysis

To quantify Cbx7-CD binding to K27me3 cleaved versus non-cleaved peptides, fluorescence anisotropy was performed as described (Bernstein et al., 2006) using FAM-labeled peptides. Three independent batches of protein were purified and used in 2-4 replicate binding assays per batch. Anisotropy results for each replicate were normalized and then plotted; binding curves were fit to all normalized data points (R= 0.99135 for H3 18-37 “non-cleaved” peptide and 0.98992 for H3 22-37 “cleaved” peptide) to determine K_ds and error (SEM). Normalized results for all replicates were then averaged and plotted as shown in Figure 7D, with error bars representing standard deviation. P value was calculated using the Student’s t-test.

Supplemental References

Bernstein, E., Duncan, E.M., Masui, O., Gil, J., Heard, E., and Allis, C.D. (2006). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. Mol Cell Biol *26*, 2560-2569.

Coon, J.J., Ueberheide, B., Syka, J.E., Dryhurst, D.D., Ausio, J., Shabanowitz, J., and Hunt, D.F. (2005). Protein identification using sequential ion/ion reactions and tandem mass spectrometry. Proc Natl Acad Sci U S A *102*, 9463-9468.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res *11*, 1475-1489.

Eng, J., McCormack, A.L., and Yates III, J.R. (1994). Journal of the American Society of Mass Spectrometry *5*, 976-989.

Garcia, B.A., Mollah, S., Ueberheide, B.M., Busby, S.A., Muratore, T.L., Shabanowitz, J., and Hunt, D.F. (2007). Chemical derivatization of histones for facilitated analysis by mass spectrometry. Nat Protoc *2*, 933-938.

Geer, L.Y., Markey, S.P., Kowalak, J.A., Wagner, L., Xu, M., Maynard, D.M., Yang, X., Shi, W., and Bryant, S.H. (2004). Open mass spectrometry search algorithm. J Proteome Res *3*, 958-964.

Li, H., Ilin, S., Wang, W., Duncan, E.M., Wysocka, J., Allis, C.D., and Patel, D.J. (2006). Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature *442*, 91-95.

Martin, S.E., Shabanowitz, J., Hunt, D.F., and Marto, J.A. (2000). Subfemtomole MS and MS/MS peptide sequence analysis using nano-HPLC micro-ESI fourier transform ion cyclotron resonance mass spectrometry. Anal Chem *72*, 4266-4274.

Mendez, J., and Stillman, B. (2000). Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. Mol Cell Biol *20*, 8602-8612.

Simon, M.D., Chu, F., Racki, L.R., de la Cruz, C.C., Burlingame, A.L., Panning, B., Narlikar, G.J., and Shokat, K.M. (2007). The site-specific installation of methyl-lysine analogs into recombinant histones. Cell *128*, 1003-1012.

Syka, J.E., Coon, J.J., Schroeder, M.J., Shabanowitz, J., and Hunt, D.F. (2004a). Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc Natl Acad Sci U S A *101*, 9528-9533.

Syka, J.E., Marto, J.A., Bai, D.L., Horning, S., Senko, M.W., Schwartz, J.C., Ueberheide, B., Garcia, B., Busby, S., Muratore, T.*, et al.* (2004b). Novel linear quadrupole ion trap/FT mass spectrometer: performance characterization and use in the comparative analysis of histone H3 post-translational modifications. J Proteome Res *3*, 621-626.