Supplemental Information

Top-down and middle-down protein analysis reveals that intact and clipped human histone proteins differ in post-translational modification patterns

Andrey Tvardovskiy, Krzysztof Wrzesinski, Simone Sidoli, Stephen J. Fey, Adelina Rogowska-Wrzesinska*, Ole N. Jensen*#

Telephone: +45 6550 2368; E-mail: jenseno@bmb.sdu.dk

Supplemental Figures

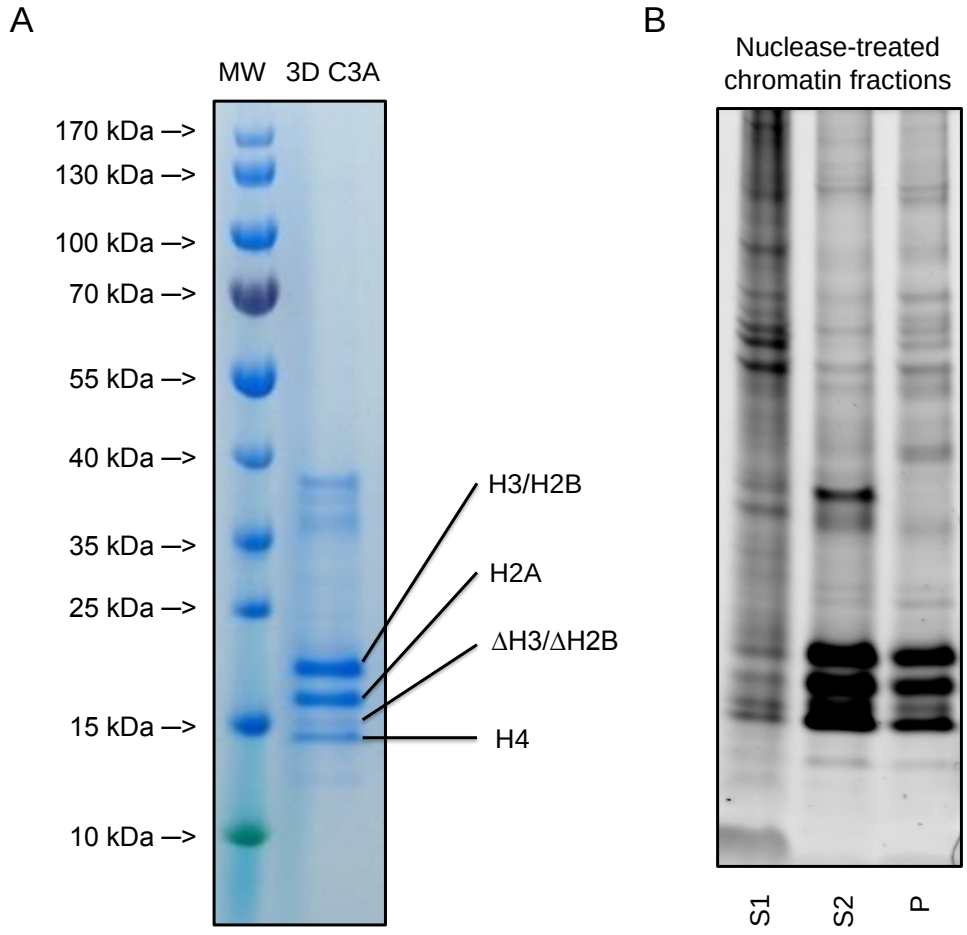


Figure S1: Evaluation of the faster migrating histone band by LC-MS/MS; chromatin fractionation by sequential salt extraction

(A) Histones were acid-precipitated from spheroid C3A (3D) culture, grown in 3D system for 21 days, as described (1). The resulting protein precipitate was resuspended in water and the protein concentration was determined with the BCA protein assay (Pierce). Five micrograms of histone precipitate were resolved by SDS–PAGE and resulting gel was stained with Coomassie Blue. The faster migrating H3 band was localized on the gel based on the anti-H3 immunoblot result. The gel piece corresponding to the faster migrating H3 band was isolated and partially digested with trypsin. The resulting peptide mixture was then analyzed by reverse-phase chromatography coupled to a Thermo LTQ-Orbitrap Velos mass-spectrometer. The MS/MS spectra were searched with MASCOT against a human protein database. By using this setup we identified peptides from both histones H2B and H3 in faster migrating H3 sub-band. Other visible protein bands were analyzed in the same manner.

(B) Chromatin was isolated from 21 days old spheroid C3A culture, partially digested with MNase and subsequently fractionated into euchromatin (S1), heterochromatin (S2) and matrix-associated chromatin (P) subfractions as described (2). Equal amount of proteins from each fraction were separated by SDS–PAGE and visualized on gel by SYPRO Ruby staining.

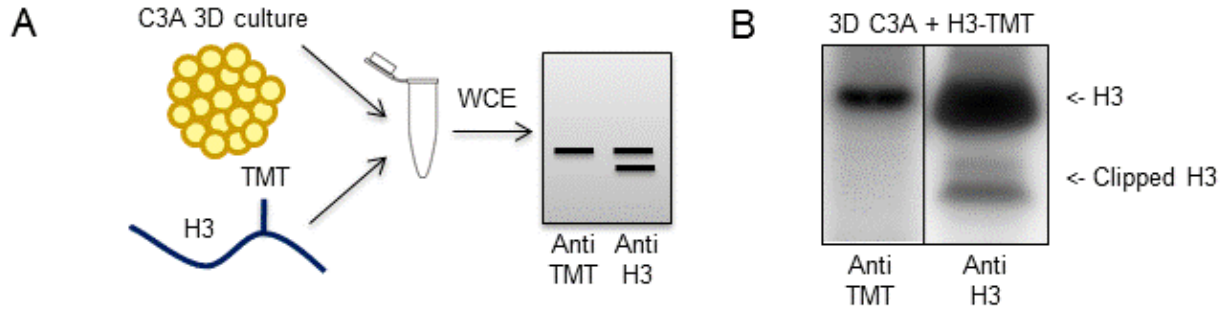


Figure S2: H3 clipping does not occur during sample preparation
(A)Schematic representation of the experiment used for eliminating the possibility of histone H3 degradation being an artifact during the WCEs preparation. Intact H3 purified from 2D culture was labeled with cysteine specific TMT reagent and added to sample buffer prior to 3D culture WCEs preparation. WCEs were analyzed by immunoblotting with anti-TMT and anti-H3 antibodies.
(B) WCEs from C3A cell line grown in 3D culture were prepared with endogenous TMT-labeled H3 and analyzed by immunoblotting.

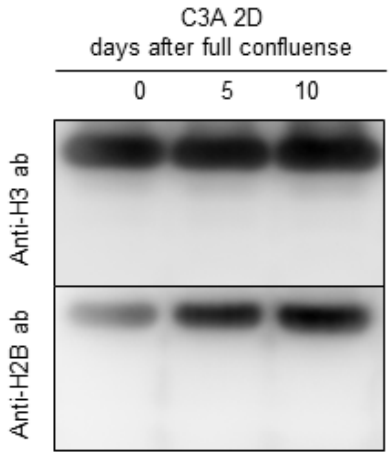


Figure S3: **Histone clipping is accumulated in 2D culture grown in arrested duplication state**
C3A cell line was grown in 2D culture to confluence and then cultivated for a further 10 days. WCEs were produced from C3A cells, collected at 1, 5 and 10 days post confluence, and probed with anti-H2B and anti-H3 Abs.

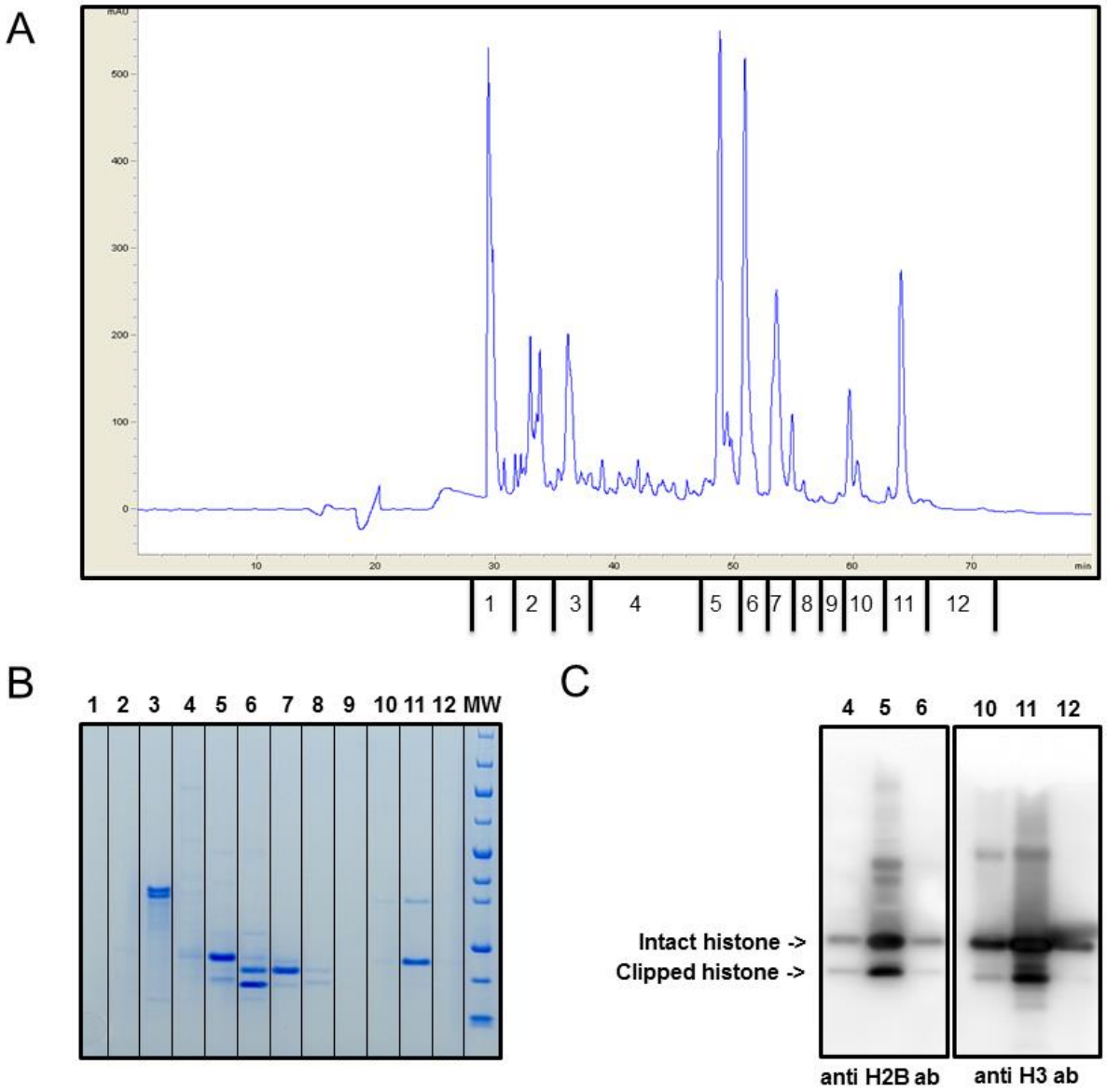


Figure S4: Histone fractionation for top-down LC-MS/MS analysis

(A) A One hundred µg portion of histones isolated from 42 days old spheroid C3A (3D) culture was separated by reversed phase (RP) chromatography using a C_{18} column (250 mm × 4.6 mm, Jupiter, 300 A, Phenomenex, Torrance, CA, USA) on an Agilent 1200 series HPLC system as described (3). Twelve fractions were collected between 28-72 minutes, dried in SpeedVac and stored at -80 °C until further analysis.

(B) Dried HPLC purified histone fractions were resuspended in water (40 µl) and 5 µl from each fraction were subjected to SDS-PAGE (NuPAGE Novex 4–12% Bis-Tris gel) followed by Coomassie Blue staining.

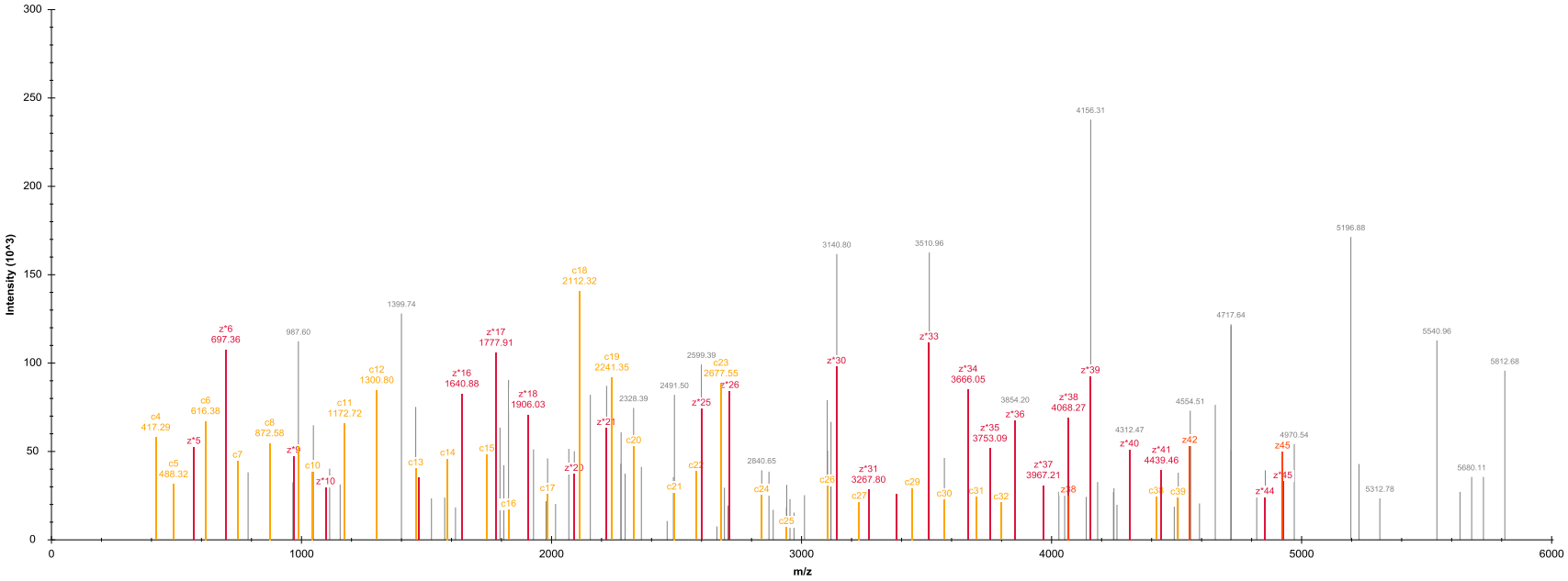
(C) Histone fractions (5 µl from each fraction) were subjected to immunoblot analysis of with anti-H2B and anti-H3 Abs.

A

H2BΔK17

A V T K A Q K K D G K K R K R S R K E S Y S V Y V Y K V L K Q V H P D T G I S S K A M G I M N S F V N D I F E R I A G E A S R L A H Y N K R S T I T

S R E I Q T A V R L L L P G E L A K H A V S E G T K A V T K Y T S S K

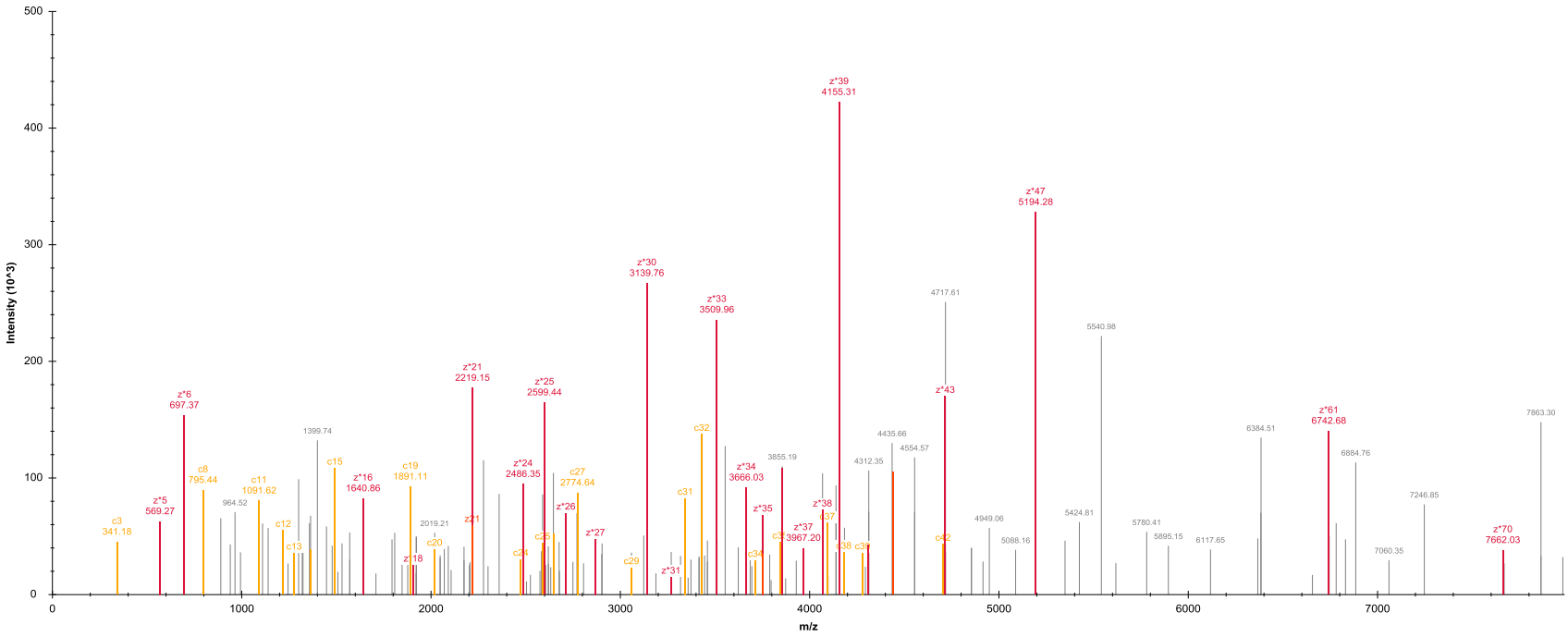


B

H2B

P E P A K S A P A P K K G S K K A V T K A Q K K D G K K R K R S R K E S Y S V Y V Y K V L K Q V H P D T G I S S K A M G I M N S F V N D I F E R I A

G E A S R L A H Y N K R S T I T S R E I Q T A V R L L L P G E L A K H A V S E G T K A V T K Y T S S K



C

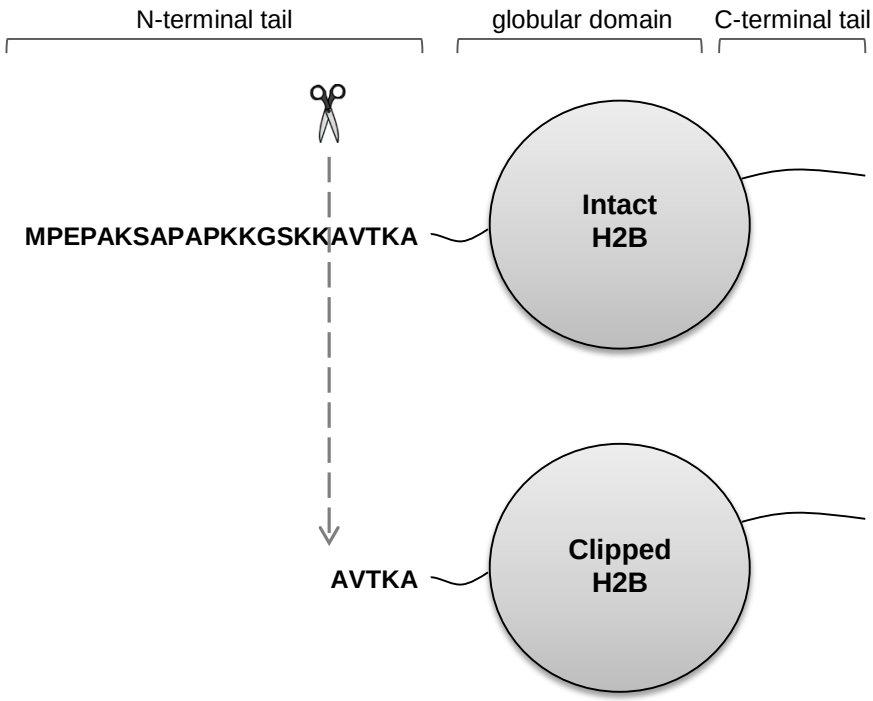


Figure S5: Top-down protein analysis identifies H2B clipping site
(A) MS/MS spectrum of clipped H2BΔK17 proteoform. The raw MS data was converted to mzXML using MSConvert tool (ProteoWizard, version 3.0.5245). The MS/MS spectra were then deconvoluted using MS-Deconv softvare and annotated using SeeMS softvare (ProteoWizard, version 3.0.5245) (4). c-type fragment peaks indicated by yellow color, z, z+1 and z+2-type fragment peaks indicated by red color.
(B) MS/MS spectrum of intact H2B
(C) Schematic representation of H2B clipping

A

Histone proteoform	N-terminal peptide sequence	Number of PSMs
H3.3ΔK9	STGGKAPRKQLATKAARKSAP**S**TGGVKKPHRYRPGTVALRE	11
H3.3ΔS10	TGGKAPRKQLATKAARKSAP**S**TGGVKKPHRYRPGTVALRE	5
H3.3ΔG12	GKAPRKQLATKAARKSAP**S**TGGVKKPHRYRPGTVALRE	16
H3.3ΔK23	AARKSAP**S**TGGVKKPHRYRPGTVALRE	6
H3.1/H3.2ΔK9	STGGKAPRKQLATKAARKSAP**A**TGGVKKPHRYRPGTVALRE	2
H3.1/H3.2ΔS10	TGGKAPRKQLATKAARKSAP**A**TGGVKKPHRYRPGTVALRE	1

B

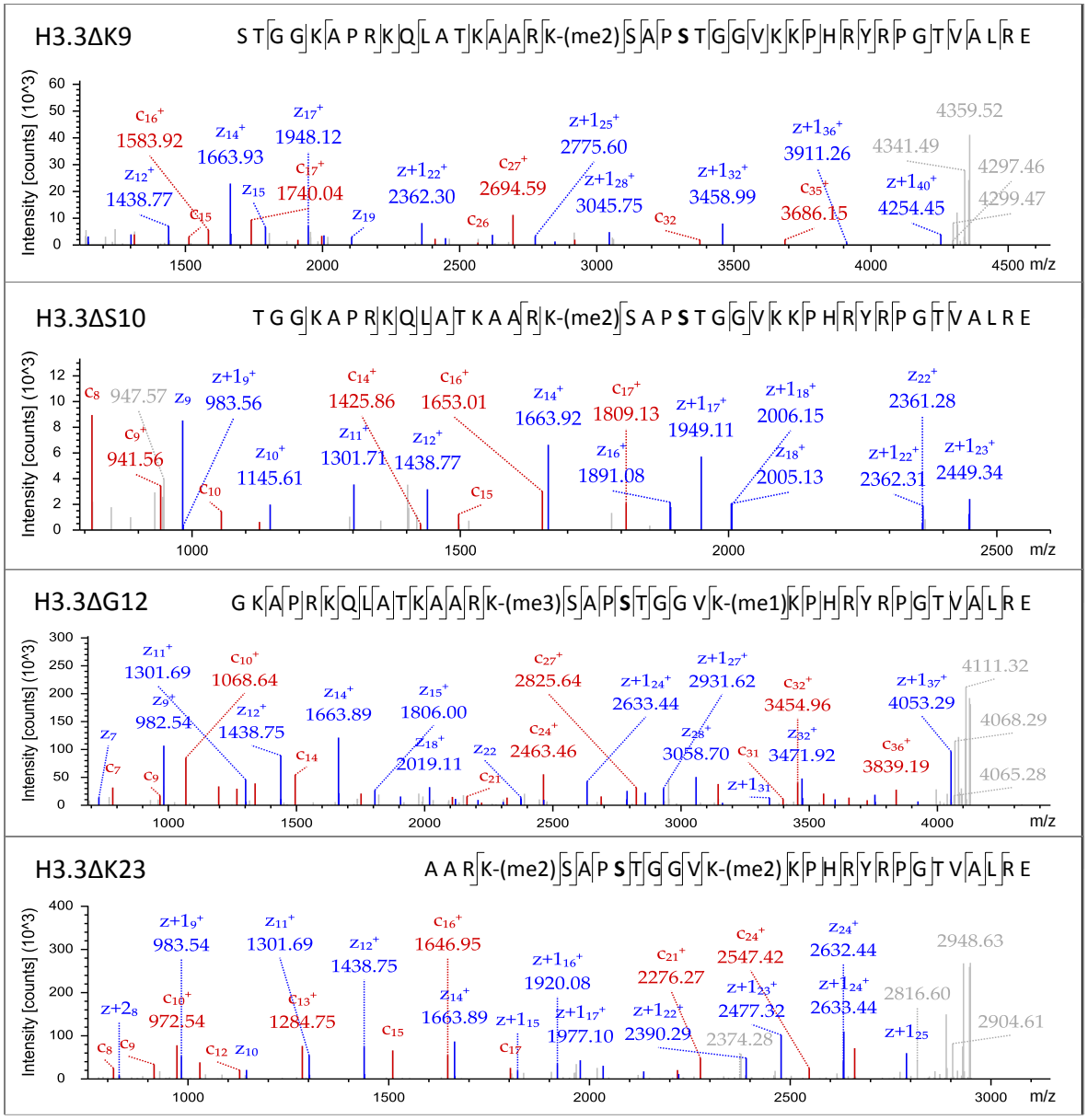


Figure S6: Determination of H3 clipping sites in primary hepatocytes by mass spectrometry

(A) Summary of clipped H3 proteoforms identified by middle-down MS analysis of histone isolated from primary hepatocytes.

(B) Charge deconvoluted ETD-MS/MS spectra of intact H3.3 N-terminal tail peptide and four distinct clipped H3.3 N-terminal tail peptides (H3.3ΔK9, H3.3ΔS10, H3.3ΔG12 and H3.3ΔK23) obtained by middle-down MS/MS analysis of histones isolated from primary hepatocytes. Matched z, z+1, z+2 - type and c - type fragment ions are indicated by blue and red colors, respectively.

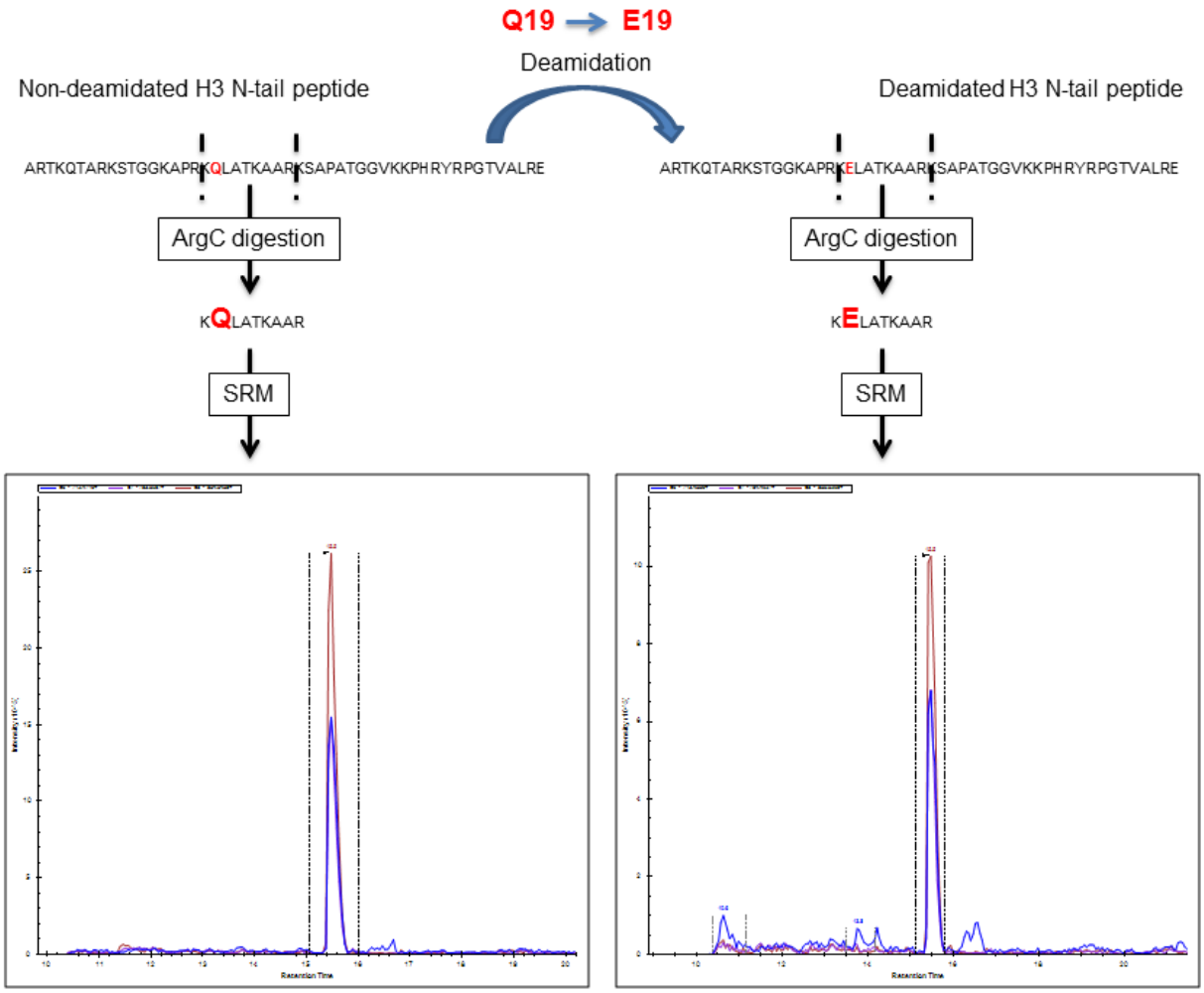


Figure S7: Evaluation of H3 Q19 to E19 deamination
Histones were isolated from 3D spheroid C3A culture and digested with ArgC. Single reaction monitoring (SRM) was used for monitoring intact (KQLATKAAR) and deamidated (KELATKAAR) H3 peptides.

Supplemental Tables

Protein name	Protein mass, Da	Number of PSMs
Histone H2B type 1-C/E/F/G/I	13898	277
H2BΔK17 type 1-C/E/F/G/I	12164	17
Histone H2B type 2-E	13912	43

Table S1: H2B proteoforms identified by top-down LC-MS/MS
H2B-containing RP HPLC fraction was analyzed by WCX/HILIC-MS/MS. Two distinct H2B proteoforms, H2B type 1-C/E/F/G/I and H2B type 2-E, were identified. The clipped H2B proteoform was only observed for H2B type 1-C/E/F/G/I.

	Intact H3					
Repl. no.	№1_1	№1_2	№1_3	№2_1	№2_2	№2_3
K4me1	0.82	1.67	3.52	3.00	1.48	1.94
K9me1	32.24	28.63	29.81	31.42	30.26	26.63
K9me2	46.08	45.04	42.70	40.09	47.36	45.38
K9me3	1.25	3.33	0.60	1.88	1.22	2.17
K9ac	1.39	1.75	1.38	1.93	1.48	2.64
K14me1	19.76	15.27	15.32	18.42	17.94	18.11
K14me2	1.66	2.83	1.76	4.30	2.53	3.64
K14me3	0	0.24	0	0.07	0	0
K14ac	9.61	6.67	8.95	9.79	7.41	9.16
K18me1	2.13	1.21	1.88	2.56	1.98	0.94
K18ac	0.07	0.45	0.85	2.59	2.62	1.26
K23me1	1.63	0.10	0	1.36	1.28	1.96
K23me2	1.02	1.05	0.85	0.86	2.57	0.89
K23me3	4.03	2.15	1.64	2.06	1.09	2.67
K23ac	19.22	29.48	18.28	22.91	18.76	13.80
K27me1	2.84	3.88	4.05	5.48	4.77	7.07
K27me2	48.24	40.04	40.84	44.45	38.11	39.07
K27me3	39.81	52.11	51.61	47.49	50.25	46.93
K27ac	3.82	2.38	2.78	0.97	3.65	2.54
K36me1	0.59	1.28	1.58	3.58	1.02	3.00
K36me2	3.66	2.46	1.02	1.31	1.85	2.11
K36me3	4.82	2.93	2.96	2.67	6.20	7.00
	H3ΔS10					
Repl. no.	№1_1	№1_2	№1_3	№2_1	№2_2	№2_3
K14me1	0	0	0	0	0	0
K14me2	0	0	0	0	0	0
K14me3	0	0	0	0	0	0
K14ac	8.53	21.03	23.40	18.19	29.86	2.32
K18me1	0	0	0	0	0	0
K18ac	9.50	21.85	6.19	3.98	4.19	1.72
K23me1	0	0	0	0	0	0
K23me2	0	0	0	0	0	0
K23me3	0	0	0	0	0	5.69
K23ac	0.68	12.17	8.68	13.45	17.72	18.35
K27me1	1.71	0	1.55	0	0	3.77
K27me2	64.00	56.70	58.15	60.91	62.78	47.47
K27me3	32.58	43.30	36.41	39.09	37.22	39.20
K27ac	0	0	0	0	0	5.69
K36me1	6.03	0.91	5.25	4.30	8.85	10.33
K36me2	11.74	7.07	6.49	11.81	4.75	7.20
K36me3	0	0	1.55	0	0	3.77

	H3ΔG12					
Repl. no.	№1_1	№1_2	№1_3	№2_1	№2_2	№2_3
K14me1	0	0	0	0	0.34	0
K14me2	0	0	0	0	0	0
K14me3	0	0	0	0	0	0
K14ac	1.93	10.45	6.30	5.11	5.20	7.85
K18me1	0	0	0	0	0	0
K18ac	0	0	0	0.71	0.47	1.62
K23me1	2.11	0	0	0	0	0
K23me2	0	0	0	0	0	0.52
K23me3	0	0	0	0	0	0
K23ac	11.03	19.03	16.96	19.64	15.68	16.87
K27me1	10.32	7.94	2.76	2.23	10.95	6.21
K27me2	59.40	55.79	60.26	63.00	59.43	54.52
K27me3	29.81	36.27	36.55	34.78	28.97	36.52
K27ac	0	0	0	0	0	0
K36me1	7.79	12.48	12.00	10.15	6.58	10.50
K36me2	13.83	6.35	6.38	7.37	11.05	9.83
K36me3	0.98	5.22	0.82	2.55	7.41	4.07

Table S2: PTM relative abundances in intact and clipped H3 proteoforms isolated from 3D C3A culture

The relative abundance of PTMs (indicated in percentage) reveled in six WCX/HILIC-MS/MS runs on N-terminal peptides of intact and clipped H3.

Supplemental methods

Preparation of the TMT-labeled H3

Histone H3 was purified from flat C3A culture by acid precipitation and RP-HPLC fractionation. Dried HPLC purified H3 fraction was resuspended in water and protein concentration was measured using Pierce BCA Protein Quantitation kit. Then twenty five µg portion of H3 was diluted with sodium dodecyl sulfate (SDS) to the final protein concentration of 1 µg/µl and SDS concentration of 0.5%. The cysteine-reactive Tandem Mass Tag (TMT) reagent (iodoTMTzero Label Reagent, Thermo Scientific) was added to the final concentration of 4mM and resulted mixture was incubated in a dark at 37^0C for one hour. The reaction was quenched by adding SDS to the final concentration of 20mM. The efficiency of labeling was determined by using western blot analysis with anti-TMT and anti-H3 Abs. TMT-labeled H3 was added to modified RIPA buffer (to the final concentration of 0.1 µg/µl) and used for elimination of H3 degradation during spheroid C3A culture WCEs preparation.

Single reaction monitoring (SRM)

Peptide mixture generated by histone digestion with endopeptidase ArgC was analyzed by using TSQ Vantage Triple Stage Quadrupole Mass Spectrometer coupled online to an Easy-nLC-system (both from Thermo-Fisher Scientific). Non-deamidated (KQLATKAAR) and deamidated (KELATKAAR) H3 peptides as well as their acetylated (+42 Da) forms were chosen for the analysis. Five SRM transitions for each peptide were monitored. The doubly charged precursor ions were monitored in Q1 and five specific singly charged fragment ions (b8, b7, b6, b5 and y8) were monitored in Q3. Q1 and Q3 resolution were set to 0.6 and 0.7 amu, respectively. For each SRM trial, the dwell time was set to 100 ms. Transitions were created and evaluated by using Skyline software.

Data processing and analysis for the top-down workflow

The raw data from the Orbitrap Fusion mass spectrometer were converted to mzXML format using MSConvert tool (ProteoWizard, version 3.0.5245). The MS/MS spectra were then deconvoluted and converted to mgf format using MS-Deconv softvare (4). The resulting files were searched against human histone database exported from Uniprot. The search was performed using Mascot v2.3 (Matrix Science, London, UK) as search engine with the following parameters: MS mass tolerance: 2.1 Da, to include possible errors in isotopic recognition; MS/MS mass tolerance 0.1 Da; acetylation (42.011 Da) on N-terminal of proteins and lysine; monomethylation on lysine (14.016) and phosphorylation on serine (79.966 Da) were selected as a variable modifications.

Interplay calculation

The “interplay score” was used to evaluate the likelihood of two histone marks to coexist on the same polypeptide (5). Interplay score was calculated by using the following equation: $Interplay_{xy} = \log2 (F_{xy} / (F_x * F_y))$, where xy is the coexistence frequency of the binary mark, Fx is the frequency or relative abundance of the mark x and Fy is the frequency of the mark y.

Histone Coder overview

Histone Coder verifies whether the PTM localization assigned by Mascot is supported by fragment ions present in the MS/MS spectrum. In order to do so, the software counts the number of `site determining ions´ between the assigned PTM localization and the two closest alternative sites. If ions are found, the Mascot assignment is considered unambiguous. Histone Coder considers all 10 alternatives that Mascot calculates for each query and assigns site determining ions for all the PTM combinations provided. The MS/MS spectrum might contain multiple identified peptides. Spectra containing multiple hits validated by Histone Coder are candidates for isoScale quantification (see "isoScale overview" chapter).
The output includes information about number and type of site determining ions detected in the MS/MS spectrum.

isoScale overview

IsoScale retrieves data for peptide quantification directly from the Mascot .csv output. The software uses the output of Histone Coder, to manage only unambiguously assigned combinatorial PTMs, and the .csv file of Mascot. In case of co-fragmented peptides the intensity of the fragment ions is adopted to estimate the relative abundance of different isobaric co-fragmented species, as previously demonstrated (6). To do so, isoScale extracts the unique ions for each of the two identified peptides, where for `unique ions´ is intended all the fragment ions that univocally discriminate one peptide from the other. After extracting such information, isoScale generates the fragment ion relative ratio (FIRR) and divides the total ion intensity of the MS/MS spectrum for such ratio between the two considered peptides. Once the total ion intensity is extracted for each identified and validated peptide the total abundance of a given peptide is calculated by summing all total ion intensities obtained from all spectra where this peptide was identified.

Statistical analysis

A paired-sample t-test was used to compare differences in PTM relative abundances Six datasets corresponding to six WCX/HILIC-MS/MS runs were used for the comparison. A p value < 0.05 was considered statistically significant.

SUPPLEMENTAL REFERENCES

1. Shechter, D., Dormann, H.L., Allis, C.D. and Hake, S.B. (2007) Extraction, purification and analysis of histones. *Nature protocols*, **2**, 1445-1457.
2. Torrente, M.P., Zee, B.M., Young, N.L., Baliban, R.C., LeRoy, G., Floudas, C.A., Hake, S.B. and Garcia, B.A. (2011) Proteomic interrogation of human chromatin. *PloS one*, **6**, e24747.
3. Jung, H.R., Sidoli, S., Haldbo, S., Sprenger, R.R., Schwammle, V., Pasini, D., Helin, K. and Jensen, O.N. (2013) Precision mapping of coexisting modifications in histone H3 tails from embryonic stem cells by ETD-MS/MS. *Analytical chemistry*, **85**, 8232-8239.
4. Liu, X., Inbar, Y., Dorrestein, P.C., Wynne, C., Edwards, N., Souda, P., Whitelegge, J.P., Bafna, V. and Pevzner, P.A. (2010) Deconvolution and database search of complex tandem mass spectra of intact proteins: a combinatorial approach. *Molecular & cellular proteomics : MCP*, **9**, 2772-2782.

5. Schwammle, V., Aspalter, C.M., Sidoli, S. and Jensen, O.N. (2014) Large-scale analysis of co-existing post-translational modifications on histone tails reveals global fine-structure of crosstalk. *Molecular & cellular proteomics : MCP*.
6. Pesavento, J.J., Mizzen, C.A. and Kelleher, N.L. (2006) Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: human histone H4. *Analytical chemistry*, **78**, 4271-4280.