The proteomic investigation of chromatin functional domains reveals novel synergisms among distinct heterochromatin components

Monica Soldi and Tiziana Bonaldi

SUPPLEMENTAL FIGURE LEGENDS

Figure S1 (Related to Fig.1). **Set up and quality control of N-ChIP.** A) DNA isolated after digestion with micrococcal nuclease S7 (MNase), separated in fractions containing mono- (S1) or poly-nucleosome (S2) (left panel) and chromatin input (middle panel), resolved on a 1% agarose gel stained with ethidium bromide. SDS-PAGE of ChIP-ed S1 fraction (#1 and #2 indicate two replicates): core histones H3, H4, H2A and H2B are visible around and below the 17kDa band of the pre-stained protein marker, with H3 and H2B co-migrating (right panel). B) Competition assays for antibody specificity test: after MNase digestion chromatin is incubated with H3K9me3 antibody with or without excess (150X) of different soluble peptides bearing methylations at distinct sites (H3K9me3, H3K9me2, H3K4me3, H3K27me3 and H4K20me3): immunoprecipitated materials are SDS-PAGE separated and stained with Colloidal Comassie blue. C) Relative enrichment of unmodified, mono-, di-, and tri-methylated on K9 in flow-through (FT) as compared to input (IN). Histogram represents the averages ± SEM from three independent experiments and results in significant depletion of K9me3 and K9me2 in FT.

***Figure S2 (Related to Fig.2 and Fig. 3)*. Estimation of hPTMs frequencies in bulk chromatin (input).** Relative abundance of hPTMs on histone H3, H4 and H2A in input.

***Figure S3 (Related to Fig.3)*. Relative enrichment of modifications on peptides H3 (3-8) and (9-17) in H3K9me3 and H3K4me3 ChIPs.** A) Enrichment of K4 methylations in N-

ChIP with anti-H3K9me3. B) Relative enrichment of K9 methylations in N-ChIP with anti-H3K4me3. C) Relative enrichment (left) and MS/MS spectra (right) of R26me2 in N-ChIP with anti-H3K4me3. Relative enrichment is expressed as a log2 Ratio between the abundance of PTM in IP over the input and represents the averages ± SEM from three independent experiments.

Figure S4 (Related to Fig.3). **Western blot validation of hPTMs MS analysis.** Immunopurified nucleosomes from H3K9me3 (A) and H3K4me3 (B) are resolved on a 17.5% Bis-Tris gel, transferred onto PVDF, and probed with primary antibodies against H3K4me3, H3K36me2, H3K27me3, H4K20me3, acetylated H3 (K9/K14Acetyl), H3K79me2 and pan-acetylated H4. Unmodified histones H3 and H4 serve as loading controls (#1 and #2 indicate two replicates). 0.2 % and 0.04% of input were loaded for H3K9me3 and H3K4me3 experiments, respectively. C) WB validation of H3K18me1 enrichment in H3K9me3 nucleosomes in respect to H3K4me3. H3K9me3 and H3K4me3 are used as positive control in the corresponding ChIPs, whereas the unmodified H3 is the loading control.

Figure S5 (Related to Fig.3). **Elution profile of H3 (27-40) modified peptide**. Extracted ion chromatograms (XIC) of various 2+ charge modified forms relative to H3(27-40) peptide are reported, upon Arg-C digestion**,** for the time range corresponding to 20-34 min. Peptide ions at the specific *m/z* values: 717.4204, 724.4282, 731.4361 and 738.4439, correspond to unmodified, mono- (me1), di- (me2) and tri-(me3) methylated H3(27-40) peptides, respectively (left panel). Extracted ion chromatograms (XIC) of various 2+ charge modified forms relative to H3(27-40) peptide are reported, upon deuterated acetic anhydride alkylation, prior to trypsin digestion, for the time range corresponding to 36-50 min. Peptide ions at the

specific *m/z* values: 784.9645, 791.9723, 776.4655 and 783.4733, correspond to unmodified, mono- (me1), di- (me2) and tri-(me3) methylated H3(27-40) peptides, respectively. Peptide ion at 798.9802 *m/z* is assigned to mono-methylations at K27 and K36 (right panel). Based on the number of D_3-acetyl groups and methylations, distinct modification degrees at specific Lysines residues can be assigned unambiguously. Furthermore, with this strategy certain isobaric peptides (i.e. K27me2 and K36me2) can be efficiently resolved during chromatography by their distinct elution times.

Figure S6 (Related to Fig.4). **Quality control of X-ChroP approach.** A) DNA isolated after shearing by sonication, resolved on a 1.3% agarose gel and stained with ethidium bromide. The sample prepared for X-ChIP usually contains fragments of 300-500bp (left panel). SDS-PAGE of co-immunoprecipitated proteins: lanes were sliced in ten pieces, digested with trypsin and analyzed by LC-MS/MS (right panel). B) Representative zoomed mass spectra for the 2+ charge state of the H3K9me2 and H3K9me3 peptides (9-17), in both light and heavy forms (upper panel) and extracted ion chromatograms (XIC) constructed for each precursor *m/z* value (lower panel). C) Relative enrichment of unmodified, mono-, di- and tri-methylated on K9 after X-ChIP represents the averages ± SEM from three independent experiments. D) Validation of H3K4me3 enrichment in the corresponding Forward (For) and Reverse (Rev) ChIPs: aliquot of ChIP-ed and input samples were probed with α- H3K4me3 antibody. Unmodified H3 is used as loading control.

Figure S7 (Related to Fig.4 and Fig.5). **Features of the H3K4me3 and H3K9me3 interactomes**. A) Protein ratio distribution for the H3K9me3 ChIP Forward (For) and Reverse (Rev) experiments. Venn Diagrams of Top 40% of protein ratios for two H3K9me3 X-ChIP replicates. B) Protein ratio distribution for the H3K4me3 ChIP Forward (Dir) and

Reverse (Rev) experiments. Venn Diagrams of Top 40% of protein ratios for two H3K4me3 replicates. C) Chromatin purifications identify both distinct and common sets of proteins in the H3K4me3 and H3K9me3 interactomes: Venn diagrams show numbers of proteins identified and enriched (Ratio>1) in both data sets (upper) and present in the top 40% of proteins ratios (lower).

Figure S8 (Related to Fig.6). **MS spectra and PTMs of H1.4 and H1.5.** A) Confirmation of H3.3 enrichment in H3K4me3 domains by standard ChIP followed by WB. B) Zoomed mass spectrum of precursor ions at m/z $[637.8879]^{2+}$ and $[645.9029]^{2+}$ corresponding to the peptide SLVSKme1GTLVQTK of linker H1.2, H1.4 and H1.5 species, in light and heavy forms. C) MS/MS spectrum of the peptide H1.2/H1.4 (86-97) or H1.5 (89-100), from which sequence and methylation was detected.

Figure S9 (Related to Fig.7). **Heterochromatic enrichment of H2A.X.** A) Western blot analysis of H3K9me3 and H3K4me3 upon ChIP with α-H2A.X (#1 and #2 indicate two replicates). Unmodified H3 is the loading control. B) qRT-PCR measures the levels of α-satellite repeat regions in H2A.X, H3K9me3 and H3K4me3 domains IP-ed by X-ChIP, over the mock control. The actively transcribed genes PAICS and HSPD1 are used as negative controls.

Figure S10. **MS/MS spectra of hPTMs (Related to Fig. 2 and 3).** Fragmentation spectra were used for the site-specific assignment of modifications within the peptides; MASCOT search, with the most intense ions identified in the MS/MS spectra and relative calculated score: experimental spectra are displayed. For (3-8) and (9-17) peptide the MSMS spectra are manually annotated.

SUPPLEMENTAL TABLES

Table S1 (*related to Fig.3 and Fig.4*). Proteins identified and quantified with a least two peptides, one of which unique, in the H3K9me3 interactome. Protein Group output from MaxQuant software (I), common proteins between two experiments (II) and Top40% proteins binders (III). For the detailed explanation of each column see below.

Table S2 (*related to Fig.3 and Fig.4*). Proteins identified and quantified with a least two peptides, one of which unique, in the H3K4me3 interactome. Protein Group output from MaxQuant software (I), common proteins between two experiments (II) and Top40% proteins binders (III). For the detailed explanation of each column see below.

For each tables:

Protein IDs: Identifier(s) of protein(s) contained in the protein group. They are sorted by number of identified peptides in descending order.

Majority Protein IDs: These are the IDs of those proteins that have at least half of the peptides that the leading protein has.

Protein Names: Name(s) of protein(s).

Gene Names: Name(s) of the gene(s) associated to the protein(s).

Uniprot: UniProt (http://www.uniprot.org) ID(s) of the protein(s).

Proteins: Number of proteins contained within the group. This corresponds to the number of entries in the column 'Protein IDs'.

Peptides: The total number of peptide sequences associated with the protein group (i.e. for all the proteins in the group).

Razor Peptides: The total number of razor peptide.

Unique Peptides: The total number of unique peptides associated with the protein group (i.e. these peptides are not shared with another protein group).

Sequence Coverage [%]: Percentage of the sequence that is covered by the identified peptides of the best protein sequence contained within the group.

Mol. Weight [kDa]: Molecular weight of the best protein sequence contained within the protein group.

Sequence Length: The total length of the best protein sequence contained within the group.

PEP: Posterior Error Probability of the identification. This value essentially operates as a p-value, where smaller is more significant.

Ratio H/L A: The ratio between two heavy and light label partners in the experiment 1 (Forward).

Ratio H/L Normalized A: Normalized ratio between two medium and light label partners in the experiment 1. The median of the total ratio population was shifted to 1.

Ratio H/L Normalized A/SD: Ratio H/L Normalized divided to the standard deviation in the experiment 1.

Ratio H/L Count A: Number of redundant peptides used for quantitation.

Ratio H/L B: The ratio between two heavy and light label partners in the experiment 2 (Reverse).

Ratio H/L Normalized B: Normalized ratio between two medium and light label partners in the experiment 2. The median of the total ratio population was shifted to 1.

Ratio H/L Normalized B/SD: Ratio H/L Normalized divided to the standard deviation in the experiment 2.

Ratio H/L Count B: Number of redundant peptides used for quantitation.

Intensity: Summed up eXtracted Ion Current (XIC) of all isotopic clusters associated with the identified AA sequence. In case of a SILAC labeled experiment this is the total intensity of

all the isotopic patterns in the SILAC cluster. A and B are referred to Forward and Reverse experiment, respectively.

Intensity L: Summed up Extracted Ion Current (XIC) of the isotopic cluster linked to the light label partner. A and B are referred to Forward and Reverse experiment, respectively.

Intensity H: Summed up Extracted Ion Current (XIC) of the isotopic cluster linked to the heavy label partner. A and B are referred to Forward and Reverse experiment, respectively.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Antibodies

For Western blot analysis the following antibodies were used, according to the manufacturer's instructions: H3K9me3 (Abcam 8898, dilution 1:1000), H3K36me2 (Abcam 9049, dilution 1:1000), H3K27me3 (Upstate Millipore 07-449, dilution 1:2500), H3K4me3 (Active Motif 39159, dilution 1:1000), H4K20me3 (Abcam 9053, dilution 1:1000), acetyl-Histone H4 (Upstate Millipore 06-866, dilution 1:2000) *K5/8/12/16 acetylated*, H3K79me2 (Abcam 3594-00, dilution 1:1000), acetyl-Histone H3 (Upstate 05-599, dilution 1:5000) *K9/14 acetylated*, H3K18me1 (Active Motif 39667, dilution 1:500), Histone H2A.X (phospho Tyr142) (Upstate Millipore 07-1590, dilution 1:100; Abcam 94602, dilution 1μg/ml), histone H2A.X (Abcam 11175, dilution 1:5000), histone H3.3 (Abcam ab62642, dilution 1:1000), unmodified histone H3 (Abcam 1791, dilution 1:5000), unmodified histone H4 (Millipore Upstate 07-108, dilution 1:1000).

For immunofluorescence (IF): WSTF (Sigma W3516, dilution 1:250), HP1β (Millipore MAB34448, dilution 1:500), H3K9me3 (Abcam 8898, dilution 1:500).

Buffers composition for native chromatin immunoprecipitation (N-ChIP)

Lysis Buffer: 10% sucrose, 0.5 mM EGTA pH 8.0, 15 mM NaCl, 60 mM KCl, 15 mM HEPES, 0.5% Triton, 0.5 mM PMSF, 1mM DTT, 5 mM NAF, 5 mM Na_3VO_4, 5mM NaButyrate, 5 μg/ml Aprotinin, 5 μg/ml Pepstatin A, 5 μg/ml Leupeptin.

Digestion Buffer: 0.32 M sucrose, 50 mM Tris-HCl pH 7.6, 4 mM $MgCl_2$, 1 mM $CaCl_2$, 0.1 mM PMSF.

Dialysis Buffer: 10 mM Tris-HCl pH 7.6, 1 mM EDTA, 0.5 mM PMSF, 5 mM NAF, 5 mM Na_3VO_4, 5mM NaButyrate, protease inhibitors cocktail.

Buffers composition for crosslinking chromatin immunoprecipitation (X-ChIP)

Lysis Buffer: 50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton-100, 0.5 mM PMSF, 5 mM NAF, 5 mM Na_3VO_4, 5mM NaButyrate, 5 μg/ml Aprotinin, 5 μg/ml Pepstatin A, 5 μg/ml Leupeptin.

Washing Buffer: 10 mM Tris-HCl pH 8, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.5 mM PMSF, 5 mM NAF, 5 mM Na_3VO_4, 5mM NaButyrate, 5 μg/ml Aprotinin, 5 μg/ml Pepstatin A, 5 μg/ml Leupeptin.

ChIP Incubation Buffer: 10 mM Tris-HCl pH 8, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% sodium deoxycholate, 0.5% sodium lauroylsarcoside, 0.5 mM PMSF, 5 mM NAF, 5 mM Na_3VO_4, 5mM NaButyrate, 5 μg/ml Aprotinin, 5 μg/ml Pepstatin A, 5 μg/ml Leupeptin.

In-gel digestion of immunopurified proteins

Processing of gel-separated proteins prior MS analysis was carried out as previously described, with minor modifications (2). Briefly, slices were cut from gels and de-stained in 50% v/v acetonitrile (ACN)/50 mM NH_4HCO_3. Reduction was carried out with 10 mM DTT

in 50 mM NH_4HCO_3, followed by alkylation with 55 mM iodoacetamide in 50 mM NH_4HCO_3. In-gel digestion was performed with 12.5 ng/μL trypsin (Promega V5113) in 50mM NH_4HCO_3, overnight at 37 °C. Digested peptide were extracted with 3%TFA, 30%ACN and finally with 100% ACN, lyophilized, desalted and concentrated on C_{18} Stage Tips (3). Samples were loaded in 1% TFA and 5% ACN and eluted with high organic solvent (80% ACN). Eluted peptides were lyophilized, re-suspended in 0.1% TFA and 0.5% acetic acid in ddH2O, and subjected to LC-MS/MS analysis.

In-gel digestion of histones for MS analysis

Bands corresponding to the core histones were excised from the gel, de-stained with repeated washes in 50% acetonitrile (ACN) in H2O, alternated with dehydration steps in 100% ACN. Gel pieces were in gel chemically alkylated as previously described, by incubation with D_6-acetic anhydride (Sigma 175641) 1:9 in 1M NH_4HCO_3 and CH_3COONa solution as catalyzer (4). After 3h at 37 °C with high shaking in thermo mixer, chemically modified gel slices were washed increasing ACN % (50% and 100%). In-gel digestion was performed with 100 ng/μl trypsin (Promega V5113) in 50 mM NH_4HCO_3 at 37 °C overnight, in order to obtain an "in-gel"-like Arg-C digestion, which cleaves at the amide bond C-terminal to Arginine residues, producing peptides with an optimal length for MS analysis. Digested peptides were extracted, desalted and concentrated using a combination of reverse-phase C18/Carbon "sandwich" system and ion-exchange (SCX) chromatography, on hand-made nano-columns (StageTips) (3): digested peptides loaded on on C_{18}/C and SCX StageTips were then eluted with high organic solvent (80% ACN) and NH_4OH, respectively. Eluted peptides were lyophilized, re-suspended in 0.1% TFA and 0.5% acetic acid in ddH2O, pooled and subjected to LC-MS/MS.

Masses (in Da) of site-specific identification of PTMs:

Histone H3: Peptide 3-8 (TKQTAR, K4unmod=375.2208, K4me1=382.2286, K4me2=366.7218, K4me3=373.7296); peptide 9-17 (KSTGGKAPR, K9unmod=496.2937, K9me1=503.3016, K9me2=487.7947, K9me3=494.8025, K9unmod/K14Ac=494.7843, K9me1/K14Ac=501.7921, K9me2/K14Ac=486.2853, K9me3/K14Ac=493.2931, K9Ac/K14Ac=493.2749); peptide 18-23 (KQLATKAAR, K18/23unmod=538.8383, K18me1/K23unmod=545.8461, K18unmod/K23Ac=537.3289, K18me1/K23Ac=544.3367, K18Ac/K23Ac=535.8195); peptide 27-40 (KSAPATGGVKKPHR, K27/36/37unmod=784.9645, K27me1 or K36me1=791.9723, K27me2 or K36me2=776.4655, K27me3 or K27me2/K36me1 or K36me2/K27me1=783.4733, K27me1/K36me1=798.9802, K27me2/K36me2=767.9664, K27me3/K36me1=790.4811, peptide 73-83 (EIAQDFKTDLR, K79unmod=690.8635, K79me1=697.8713, K79me2=682.3644).

Histone H2A: peptide 4-11 (GKQGGKAR, K5/9unmod=446.2675, KAc=444.7581, Kdi-Ac=443.2487).

Histone H4: peptide 4-17 (GKGGKGLGKGGAKR K5/K8/K12/K16unmod=725.9476, KAc=724.4381, Kdi-Ac=722.9287, Ktri-Ac =721.4193, Ktetra-Ac=719.9099).

Immunoblot analysis

Input chromatin and immuno-precipitated histone octamers were separated in 17.5% SDS-PAGE and transferred to PVDF membranes (Millipore). Membranes were blocked 1 h in 5% milk in TBS supplemented with 0.1% Tween (TBS-T). After blocking, membranes were incubated at 4°C for an overnight with primary antibodies specific for histone modifications, diluted in TBS-T 5% milk. After three washes in TBS-T, binding was revealed by ECL

Plus® Immunoblotting Detection System (Amersham Biosciences); antibodies against the unmodified version of both H3 and H4 were used as loading control (described above).

Immunofluorescence analysis

Cells grown on coverslips were washed twice with PBS, fixed with 4% para-formaldehyde for 20 min, permeabilized in 0.5% Triton X-100 in PBS for 5 min and then blocked in 10% BSA for 1 h. Cells were subsequently probed with the following mix of antibodies: WSTF and HP1, H3K9me3 and HP1. After 1 h of primary antibody incubation, cells were washed three times with PBS and then incubated with either α-rabbit secondary antibody conjugated with Cy3 (diluted 1:800), or with α-rabbit secondary antibody conjugated with fluorescein isothiocyanate (FITC-conjugate), diluted 1:50. DNA was stained with DAPI, diluted 1:5000 in PBS, for 15 sec. Slides were mounted in Mowiol and images were acquired using a wide field Olympus Biosystem Microscope BX51.

Primers for quantitative PCR upon conventional ChIP

AP945 (chr4:57142864(start)-57142927(end)):

Forward primer: 5´-CGCTACTGTTGGGTGCTGG-3´

Reverse primer: 5´-GCCTGGAAAGCTGTATTTGCTG-3´

AP777 (chr2:198189648(start)-198189843(end)):

Forward: 5´-TCCATCACGTGCGACGC-3´

Reverse: 5´-GAGGCGCGGTATCCCAG-3´

α-Repeats Regions:

Forward primer: 5´-CTCAGTAACTTCCTTGTGTTGTG-3´

Reverse primer: 5´-ATTCTGTCTAGTTTCTATAAGAAG-3´.

ChIP-Sequencing: preparation of ChIP DNA libraries and sequencing

ChIP-ed DNA was treated to remove 3' overhangs and fill in 5' overhangs resulted in blunt ended DNA fragments. An A residue was added by terminal transferase to the 3' end and the resulting fragments were ligated with Illumina adapters. The resulting Adapter-modified DNA fragments were separated by agarose gel electrophoresis and the band between 120-200 bp was excised and the DNA fragments were extracted using a Qiaquick Gel Extraction Kit (Qiagen Inc). The specific DNA fragments were subjected to 18 cycles of PCR amplification; amplified fragments were then gel purified from an excess of PCR primers, using Qiagen columns. The DNA fragment library was quantified with Bioanalyzer using High Sensitivity Chip, diluted to a 10 nM working stock concentration for cluster generation. Finally, cluster generation was performed according to standard protocols of the manufacturer (Illumina) and loaded into individual lanes of a flow cell (4 picomoles/sample). ChIP-Seq data were acquired with the Illumina Genome Analyzer II, producing a fixed 36bp read length. After each base incorporation step, the flow cell surface was washed to remove reactants and then imaged by microscope objective.

ChIP-Sequencing: computational analysis

For the analysis of sequencing Illumina data, read tags passing standard Illumina quality filter (Failed-chastity < 0.6) were aligned to hg18 genome using BWA 0.5.9 with default parameters (5). H2AX data and H3K9me3 and H3K4me3 ChIP seq data from (6) (stored in the NCBI GEO SuperSeries GSE20303) were analyzed with dspchip 0.8.5 (http://code.google.com/p/dspchip). Non-duplicated tags with mapping quality higher than 15 were retained; normalized profile of IP data were subtracted from the Input; the resulting profile was processed using a Hanning Window low-pass filter (window size: 500

kbp) and negative values were removed. Correlations between ChIP profiles were calculated with wigCorrelate from UCSC Genome Browser utilities.

SUPPLEMENTAL REFERENCES

1. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 26, 1367-1372.
2. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., and Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* 1, 2856-2860.
3. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2, 1896-1906.
4. Bonaldi, T., Imhof, A., and Regula, J. T. (2004) A combination of different mass spectroscopic techniques for the analysis of dynamic changes of histone modifications. *Proteomics* 4, 1382-1396.
5. Li, H., and Durbin, R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589-595.
6. Vermeulen, M., Eberl, H. C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K. K., Olsen, J. V., Hyman, A. A., Stunnenberg, H. G., and Mann, M. (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142, 967-980.