Research

A Quantitative Proteomic Analysis of *In Vitro* Assembled Chromatin*[S]

Moritz Carl Völker-Albert‡‖, Miriam Caroline Pusch‡‖, Andreas Fedisch‡, Pierre Schilcher§, Andreas Schmidt§, and Axel Imhof‡§¶

The structure of chromatin is critical for many aspects of cellular physiology and is considered to be the primary medium to store epigenetic information. It is defined by the histone molecules that constitute the nucleosome, the positioning of the nucleosomes along the DNA and the non-histone proteins that associate with it. These factors help to establish and maintain a largely DNA sequence-independent but surprisingly stable structure. Chromatin is extensively disassembled and reassembled during DNA replication, repair, recombination or transcription in order to allow the necessary factors to gain access to their substrate. Despite such constant interference with chromatin structure, the epigenetic information is generally well maintained. Surprisingly, the mechanisms that coordinate chromatin assembly and ensure proper assembly are not particularly well understood. Here, we use label free quantitative mass spectrometry to describe the kinetics of *in vitro* assembled chromatin supported by an embryo extract prepared from preblastoderm *Drosophila melanogaster* embryos. The use of a data independent acquisition method for proteome wide quantitation allows a time resolved comparison of *in vitro* chromatin assembly. A comparison of our *in vitro* data with proteomic studies of replicative chromatin assembly *in vivo* reveals an extensive overlap showing that the *in vitro* system can be used for investigating the kinetics of chromatin assembly in a proteome-wide manner. *Molecular & Cellular Proteomics 15: 10.1074/mcp.M115.053553, 945–959, 2016.*

DNA replication, transcription and repair continuously disturb the conformation of chromatin, which results in a relatively high rate of histone turnover (1) and poses a constant threat to the maintenance of epigenetic information (2, 3). Therefore, chromatin assembly has to be controlled thoroughly to ensure a proper chromatin structure. It is well appreciated that chromatin assembly is a highly regulated multistep process involving synthesis, storage and nuclear transport of histones followed by their deposition onto DNA. Immediately after translation and before the assembly onto DNA, histones are bound by a number of chaperones that assist their folding, posttranslational modification, nuclear transport and prevent nonspecific association with negatively charged cellular molecules (4–6). Once histones are deposited, chromatin adopts a particular conformation containing specific histone modification patterns (7–9) and a defined composition of associated proteins (10–13). Crosslinking experiments show that histones H3 and H4 are first deposited as a tetramer, whereas two dimers of H2A and H2B are added at a subsequent stage (14, 15). A similar assembly pathway is also observed in an *in vitro* assembly system where the process of histone deposition and chromatin contraction occurs within 30 s (16, 17). Regardless of this apparent rapid compaction, it takes much longer for new chromatin to become indistinguishable from the bulk chromatin *in vivo* (9, 13).

Recent systematic studies revealed that mature chromatin adopts a complex molecular structure containing a large variety of binding factors that go way beyond a simple aggregate of DNA and histones (11, 12, 18, 19). This observation raises the question of how this structure is assembled, in which order individual factors bind to the DNA, whether distinct intermediates during chromatin assembly exist and which key players mediate chromatin maturation. Many of those questions are extremely difficult to address experimentally because of the high complexity of chromatin assembly and maturation *in vivo* and its high level of cooperativity. Particularly, the analysis of functionally important components of chromatin synthesis will be difficult to decipher *in vivo*, as they are expected to have a severe impact on cell division and viability. Therefore, key aspects of chromatin assembly are better accessible by an *in vitro* reconstitution system. Embryonic extracts are extremely rich sources for factors required in chromatin assembly such as storage chaperones (20–22) and can therefore support chromatin assembly *in vitro* (20, 23, 24). Although it has been shown that such extracts recapitulate several aspects of chromatin assembly *in vivo* and can therefore be used to investigate this process

From the ‡BioMedical Center and Center for Integrated Protein Sciences Munich, Ludwig-Maximilians-University of Munich, Großhaderner Straße 9, 82152 Planegg-Martinsried, Germany; §Zentrallabor für Proteinanalytik (Protein Analyis Unit), Ludwig-Maximilians-University of Munich, Großhaderner Straße 9, 82152 Planegg-Martinsried, Germany

Received July 14, 2015, and in revised form, January 17, 2016
Published, MCP Papers in Press, January 25, 2016, DOI 10.1074/mcp.M115.053553
Author contributions: M.C.V., M.C.P., A.S., and A.I. designed research; M.C.V., M.C.P., A.F., P.S., and A.S. performed research; M.C.V., A.S., and A.I. analyzed data; M.C.V., M.C.P., A.S., and A.I. wrote the paper.

(23–25), a systematic comparative study has not been done so far. With the recent development of methods like iPOND (10, 26) and NCC (13) to investigate replicative chromatin assembly *in vivo* and improved techniques of label free MS based quantitation of proteins in complex samples (27) such comparative studies became feasible.

In this study, we used immobilized linear DNA to rapidly isolate *in vitro* assembled chromatin at different time points and determined its protein composition in a time resolved manner using sequential window acquisition of all theoretical fragment ions (SWATH)[1]-MS-based label-free protein quantitation. A comparison with the proteomic investigation of chromatin assembled *in vivo* (13) reveals an almost 80% overlap with the orthologue proteins assembled *in vitro*. Interestingly, we observe very similar binding kinetics, as proteins enriched in nascent chromatin *in vivo* also bind preferentially during early time points of *in vitro* chromatin assembly. The similarities of protein identity, binding kinetics and the largely sequence independent protein binding to *in vitro* assembled chromatin further support the usability of such *in vitro* assembly systems to dissect the general mechanisms of chromatin assembly.

EXPERIMENTAL SECTION

Preparation of Drosophila Embryonic Extract (DREX)—D. melanogaster embryos were collected on agar trays with yeast paste 0–90 min after egg-laying. Using a brush and sieves with descending mesh size (0.71 mm, 0.355 mm, 0.125 mm), embryos were rinsed with cold tap water and allowed to settle into ice-cold embryo wash buffer (0.7% NaCl, 0.05% Triton X-100) to arrest further development. After five successive collections, the wash buffer was decanted and replaced with wash buffer at room temperature. For dechorionation of the embryos, the volume was adjusted to 200 ml and 60 ml of 13% hypochlorite solution was added. The embryos were stirred vigorously for 4 min on a magnetic stirrer, poured back into the collection sieve (0.125 mm), and rinsed with tap water for 5 min. Embryos were allowed to settle in 200 ml of wash buffer for about 3 min. Afterward the supernatant containing the chorions was removed. Following two more settlings in 0.7% NaCl and in extract buffer (10 mM HEPES (pH 7.6), 10 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM EGTA, 10% glycerol, 10 mM 3-glycero-phosphate, 1 mM dithiothreitol (DTT), and 0.2 mM phenylmethylsulfonyl fluoride (PMSF), added freshly) at 4 °C, the embryos were settled in extract buffer in a 60 ml glass homogenizer on ice. The volume of the packed embryos was estimated before the supernatant was aspirated, leaving packed embryos and additional 2 ml buffer on top. Homogenization was performed with one stroke at 3000 rpm and 10 strokes at 1500 rpm with a pestle connected to a drill press. The homogenate was supplemented with $MgCl_2$ to a final $MgCl_2$ concentration of 5 mM. Nuclei were pelleted by centrifugation for 10 min at 10,000 rpm in a SS34 rotor. (Sorvall, Thermo-Fisher Scientific, San Jose, CA). The supernatant was centrifuged again for 2 h at 45,000 rpm in a chilled SW 56 rotor (Beckman-Coulter, Munich, Germany). The clear extract was isolated with a syringe, avoiding the top layer of lipids. Extract aliquots were frozen in liquid nitrogen. Protein concentration was determined by Nanodrop measurement and titration with chromatin assembly experiments.

Biotinylation of DNA—To obtain linearized and biotinylated DNA, we used a plasmid DNA that contains oligomers of the sea urchin 5S rDNA positioning sequence. Five hundred micrograms plasmid DNA were linearized using the restriction enzyme SacI. Completion of the digest was analyzed by agarose gel electrophoresis. Upon completion of the plasmid digestion, we added the restriction enzyme XbaI to the reaction and incubated for at least 3 h at 37 °C. Subsequently, the DNA was precipitated and purified, followed by incubation with 80 mM dCTP and dGTP, 3 mM biotinylated dUTP and dATP and the Klenow Polymerase. To purify DNA from excessive nucleotides and enzyme, we used G50 Sepharose columns (Roche, Penzberg, Germany) according to the manufacturers protocol. Finally, DNA concentration was measured and adjusted to 200 ng/μl. The same procedure was applied for the heterochromatic 359bp repeat sequence using a pBluescript plasmid containing 4 repetitive elements of the 359bp repeat.

Chromatin Assembly on Immobilized DNA—Four micrograms DNA was immobilized onto 120 μl M280 paramagnetic streptavidin beads (Invitrogen, Carlsbad, CA) in EX100 buffer (10 mM HEPES (pH 7.6), 100 mM NaCl, 1.5 mM $MgCl_2$, 0.5 mM EGTA, 10% (v/v) glycerol, 0.2 mM PMSF, 1 mM DTT) for 1 h. Beads were extensively washed and blocked with BSA for 30 min (1.75 g/l) in EX100. After another washing step in EX-Nonidet P-40 (10 mM Hepes pH 7.6, 1.5 mM $MgCl_2$, 0.5 mM EGTA, 10% (v/v) Glycerol, 0.05% Nonidet P-40) beads were resuspended in a total volume of 240 μl containing 45 μl DREX and an ATP regenerating system (3 mM ATP, 30 mM creatine phosphate, 10 μg creatine kinase/ml, 3 mM $MgCl_2$, and 1 mM DTT). To study the influence of protein acetylation,

[1] The abbreviations used are: SWATH, sequential window acquisition of all theoretical fragment ions; Aa, amino acid; Ac, acetylation; CAF, chromatin assembly factor; CE, collision energy; CHRAC, Chromatin Accessibility Complex; DDA, data-dependent acquisition; DIA, data-independent acquisition; Dre4, Dras3-Roughened-ecdysoneless region 4; DREX, Drosophila Embryonic Extract; FDR, false discovery rate; Gnf1, Germ line transcription factor 1; GO, Gene Ontology; H1/H2A/H2B/H3/H4, histone 1/2A/2B/3/4; HDAC, histone deacetylase; HMG, High mobility group protein; Hsp, Heat shock protein; iBAQ, intensity based absolute quantitation; Irbp, Inverted repeat-binding protein; ISWI, Imitation SWI; LFQ, label-free quantitation; MBT complex, Myb-MuvB/DREAM transcriptional regulatory complex; MNase, Micrococcal nuclease; MCM, Minichromosome Maintenance; MRM-HR, multiple reaction monitoring; Nurf, Nucleosome remodeling factor; PCNA, Proliferating cell nuclear antigen; PTM, post-translational modification; Rfc38, Replication factor C 38kD subunit; RpA, Replication protein A; SSRP, structure specific recognition protein; Tcp1, T-complex protein 1; TSA, Trichostatin A; XIC, extracted ion chromatogram.

TSA was added to a final concentration of 50 μM during chromatin assembly experiments as indicated. For time-resolved studies, the assembly reaction was incubated at 26 °C for 15 min, 1 h and 4 h, respectively. After two stringent wash steps with EX200, beads were resuspended in elution buffer (EX100 with 0.5 U/μl MNase and 2 mM $CaCl_2$). The supernatant after 10 min of MNase-mediated elution was subjected to mass spectrometry-based protein quantitation.

Micrococcal Nuclease Digestion—Chromatin from 1 μg circular DNA assembled for 15 min, 1 h or 4 h was resuspended in EX50 containing 5 mM $CaCl_2$ and 100 Boehringer units/μl of MNase (Sigma). After incubation at room temperature for 30 s and 90 s, respectively, a 110 μl fraction of the digestion was stopped by adding 40 μl MNase stop solution (2.5% N-Lauroylsarcosine, 100 mM EDTA pH 8.0). The suspension was subjected to RNase A and proteinase K treatment and precipitated DNA was separated with a 1.3% agarose gel. A 100bp ladder (Invitrogen) was used as a size marker.

Primary Antibodies and Plasmids—PCNA (Abcam, Cambridge, UK; ab29, clone PC10, 1:1000), Histone 3 (Abcam ab1791, 1:5000), Asf1 ((28), 1:5). 359bp repeat was isolated by PCR from genomic DNA using cDNA from S2-cells from Drosophila cells with following primers: FW: 5′-CGGTCATCAAATCATTTATTTTGC-3′, RV: 5′-CGAAATTTGGAAAAACAGACTCTG-3′. For cloning, pBluescript SK (-) plasmid was used as the basic backbone vector. Multimerization of repeats was archived by Gibson cloning.

Preparation of MS Samples for Proteomics Analysis—Assembled chromatin was subjected to mass spectrometry analysis. 10% of the chromatin-bound proteins were separated on a 4–20% gradient SDS-PAGE and analyzed by silver staining (29). 90% of the chromatin bound proteins were subjected to MS-sample preparation. Proteins were denatured in 3 M Urea, 1 M Thiourea and 25 mM DTT for 2 h at 20 °C followed by an incubation for 30 min in a dark place with a final concentration of 25 mM iodoacetamide at 20 °C to carbamidomethylate sulfhydryl groups of free cysteine. Subsequently, DTT was added to a final concentration of 50 mM and incubated for 30 min at 20 °C. The samples were diluted with 100 mM ammonium bicarbonate to lower the urea concentration below 1 M for tryptic cleavage with 200 ng of trypsin (Promega, Madison, WI) in 50 mM ammonium bicarbonate. Digestion was completed after 14 h at 25 °C. 10% of the tryptic peptide mixture were acidified using trifluoroacetic acid (TFA) and desalted using C18 stage tips prior to mass spectrometry analyses and redissolved in 0.2% TFA (30). The resulting liquid, containing the digested peptides, was dried and redissolved in 17 μl of 0.2% TFA and stored at −20 °C until further processing.

Sample Preparation for Histone Modification Analysis by MS—Nuclear-enriched fractions were separated by SDS-PAGE, stained with Coomassie (Brilliant blue G-250) and protein bands in the molecular weight range of histones (15–23 kDa) were excised as single band/fraction. Gel slices were destained in 50% acetonitrile/50 mM ammonium bicarbonate. Lysine residues were chemically modified by propionylation for 30 min at RT with 2.5% propionic anhydride (Sigma) in ammonium bicarbonate, pH 7.5 to prevent tryptic cleavage. This step added a propionyl group only to unmodified and monomethylated lysines, whereas lysines with other side chain modification will not obtain an additional propionyl-group. A set of 30 precursors of heavy SILAC-R10 labeled standard peptides, (spiketides) coding for common histone modifications was added prior to tryptic digestion. Spiketide abundance was used to normalize for different sample amounts. Subsequently, proteins were digested with 200 ng of trypsin (Promega) in 50 mM ammonium bicarbonate overnight and the supernatant was desalted by carbon Top-Tips (Glygen) according to the manufacturer's instructions.

Proteomic Analysis via LC-MS/MS on Orbitrap Mass Spectrometer—The peptide mixture resulting from tryptic cleavage was injected onto an Ultimate 3000 HPLC system equipped with a C18 trapping column (C18 PepMap, 5 mm × 0.3 mm × 5 μm, 100Å) and an analytical column (C18RP Reposil-Pur AQ, 120 mm x 0.075 mm x 2.4 μm, 120 Å, Dr. Maisch, Ammerbuch-Entringen, Germany) packed into an ESI-emitter tip (New Objective, Woburn, MA). First, the peptide mixture was desalted on the trapping column for 7 min at a flow rate of 25 μl/min (0.1% FA). For peptide separation a linear gradient from 5–40% B (HPLC solvents A: 0.1% FA, B: 80% ACN, 0.1% FA) was applied over a time of 120 min. The HPLC was online coupled to an LTQ Orbitrap XL mass spectrometer (Thermo-Fisher Scientific).

The mass spectrometer was operated in DDA-mode employing a duty cycle of one survey scan in the orbitrap at 60,000 resolution followed by up to 6 tandem MS scans in the ion trap. Precursors were selected when they had a minimal intensity of 10,000 counts and a charge state of 2+ or higher. Previously analyzed precursors were excluded for 20 s within a mass window of −1.5 to + 3.5 Da.

Proteomic Analysis via LC-MS/MS on Q-TOF Mass Spectrometer—Samples were injected into an Ultimate 3000 HPLC system (Thermo Fisher Scientific) for nano-reversed phase separation of tryptic peptide mixtures before MS analysis. Peptides were desalted on a trapping column (5 × 0.3 mm inner diameter; packed with C18 PepMap100, 5 μm particle size, 100 Å pore diameter, Thermo-Fisher Scientific). The loading pump flow of 0.1% formic acid (FA) was set to 25 μl/minute with a washing time of 10 min under isocratic conditions. Samples were separated on an analytical column (150 × 0.075 mm inner diameter; packed with C18RP Reposil-Pur AQ, 2.4 μm particle size, 100 Å pore diameter, Dr. Maisch) using a linear gradient from 4% to 40% B in 170 min with a gradient flow of 270 nl/minute. Solvents for sample separation were A: 0.1% FA in water and B: 80% ACN, 0.1% FA in water. The HPLC was directly coupled to the 6600 TripleTOF mass spectrometer using a nano-ESI source (both Sciex, Framingham, MA). A data-dependent method was se-

lected for MS detection and fragmentation of eluting peptides comprising one survey scan for 225 ms from 300 to 1800 *m/z* and up to 40 tandem MS scans for putative precursors (100–1800 *m/z*). Precursors were selected according to their intensity. Previously fragmented precursors were excluded from reanalysis for 30 s.

Data Analysis of Data-dependent LC-MS Experiments—DDA-MS data recorded on the LTQ Orbitrap mass spectrometer were processed with MaxQuant (version 1.2.2.5) using standard settings with the additional options LFQ and iBAQ (log fit) selected. Data were searched against a combined forward/reversed database (special amino acids: KR) including common contaminants for false-discovery rate filtering of peptide and protein identifications (Dmel_all translation r5.57, 30305 entries). The mass deviation for the precursor mass was set 20 ppm; fragment ions were matched within 0.5 Da mass accuracy. Fixed modifications of cysteine (Carbamidomethyl (C)) were included as well as variable modifications by oxidation of methionine and acetylation (Acetyl (Protein N-term); Oxidation (M)). Matches were filtered setting false peptide and protein (PSM FDR and protein FDR) hits to 1%. The minimum peptide length was allowed to be 6 amino acids, the minimum score for modified peptides was set to 40. For protein identification, one non-unique razor peptide was required, whereas protein quantitation was only performed if at least two razor peptides were associated with the protein hit. Prior to statistical analysis in Perseus, protein hits associated with the reversed database or common contaminants were filtered in the protein.groups.txt file (supplemental Table S1).

Data-dependent experiments performed on the Q-TOF mass spectrometer were analyzed in MaxQuant (version 1.5.1.2) using the Andromeda search engine and the same flybase database as for Orbitrap data. The settings for database search were as follows: fixed modification carbamidomethyl (C), variable modification oxidation (M) and acetyl (protein N-term); Δmass = 20 ppm for precursors, Δmass = 50 ppm for TOF fragment ions. Peptide hits required a minimum length of seven amino acids and a minimum score of 20 for unmodified and 40 for modified peptides. Resulting protein hits were FDR filtered for 1% false discoveries on the PSM level and up to 5% false protein hits. Settings for protein identification and quantitation were identical as for orbitrap data (see above).

SWATH Data Acquisition—Peptides from tryptic digestion were resuspended in 10 μl 0.1% TFA and injected into an Ultimate 3000 nano-chromatography system equipped with trapping column (C18 AcclaimPepMap, 5 $\times$ 0.2 mm, 5 μm 100 Å) and a separation column (C18RP Reposil-Pur AQ, 150 $\times$ 0.075 mm x 2.4 μm, 100 Å, Dr. Maisch) poured into a nano-ESI emitter tip (New Objective). After washing for 10 min on the precolumn with 0.05% TFA, peptides were separated by a linear gradient from 4% to 40% B (solvent A 0.1% FA in water, solvent B 80% ACN, 0.1% FA in water) for 150 min at a flow rate of 270 nl/min. Eluting peptides were detected on a 6600 Triple TOF quadrupol-TOF hybrid mass spectrometer (Sciex, Framingham, MA). First, a mixture of all conditions was run in data-dependent mode to generate an ion library for the data-independent SWATH measurements and optimize the isolation window distribution over the mass range for SWATH-data acquisition. Data-dependent acquisition consisted of a survey scan and up to 40 tandem MS scans for precursors with charge 2–5 and more than 200 cps abundance. Rolling collision energy was set to generate peptide fragments. The overall cycle time for the DDA experiment was 2.676 s. Previously analyzed precursors were excluded from repeated fragmentation for 30 s employing a mass window of 20 ppm around the precursor mass.

MS data with data-independent SWATH acquisition were generated using the same HPLC conditions as used for the generation of the ion library. Based on the distribution of the m/z values of identified peptides in the ion library, the mass range from 300–1200 *m/z* was split into 40 SWATH mass windows to optimize the number of precursor ions per window. First, precursors were monitored from 300–1500 *m/z* in a survey scan of 50 ms, followed by the SWATH data acquisition for 65 ms/mass window, resulting in an overall cycle time of 2.7 s. The fragmentation energy was adjusted to fragment 2+ charged ions in the center of the mass window and a collision energy spread over seven units was allowed. For data analysis, SWATH data were mapped to a protein database containing RT, peptide precursor and fragment ion information, that was generated in ProteinPilot 4.5 (Sciex) against the previously described drosophila database using database search settings described for MaxQuant 1.5.1.2. Settings for SWATH peak extraction in the Peak View 2.1 software (Sciex) were five peptides/protein with at least six mapping fragment ion signals, a confidence interval of 99% and FDR rate of 2%.

Histone Modification—Following carbon stage tip, the dried peptides were resuspended in 5 μl of 0.1% TFA and the complete sample was directly injected onto the reversed-phase separation column (C18RP Reposil-Pur AQ, 120 $\times$ 0.075 mm x 2.4 μm, 100 Å, Dr. Maisch) of an Ultimate 3000 nano-chromatography system (Thermo-Fisher Scientific), coupled by a nanoESI source. A separation gradient from 5% B to 30% B (solvent A 0.1% FA in water, solvent B 80% ACN, 0.1% FA in water) over 32 min was applied to separate the histone peptides at a flow rate of 325 nl/min. Because the column was poured into the nano-ESI emitter tip, peptides were directly analyzed by mass spectrometry using a TripleTOF 6600 q-TOF mass spectrometer (Sciex). A targeted MS/MS method was selected for detection and quantitation of Nterminal peptides of histone 3.1 and histone 4 with specific modifications. The first scan monitored the abundance of the precursor ion for 225 ms, the MRM scans for individual modifications were acquired for 35 ms per precursor. The overall cycle time was 2.05 s. For the list of precursor ions,

peptide sequences, modifications and fragmentation conditions see supplemental methods (supplemental Table S2).

Data Analysis—Peptide fragment masses of heavy and light peptide variants were calculated in silico using GPMAW 5.0 software (GPMAW3) and applied to filter the MRM data for abundance of specific modifications using MultiQuant software (Sciex, version 3.0). Peptides with similar precursor masses containing either trimethylation or acetylation of K were distinguished based on the mass accuracy of the instrument and the difference in retention time.

PTM-data analysis was performed with PeakView software (version 2.1, Sciex) by using doubly and triply charged peptide masses for extracted ion chromatograms (XICs). XICs were checked manually and values were exported to Excel for further calculations. Standard deviation of the mean was used for error bar calculation. The mass spectrometry raw data are deposited to the ProteomeXchangeConsortium with the data set identifier submission number PXD002537 and PXD003445. Annotated spectra can be viewed at MS-Viewer (http://prospector2.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msviewer). Spectra from DREX samples can be accessed with search key: hlyzvamoxh. Spectra from chromatin samples can be accessed with search key: 3keubvytc7.

Statistical Methods—Data were handled with Perseus software. Three biological replicates acquired with DDA-MS were analyzed for chromatin and DREX. All three biological replicates from 1 h and 4 h assembled chromatin or wtDREX were Log2(x) transformed. Missing values were replaced by random numbers from a standard deviation (width 0.3, shift −1.8).

DIA-MS SWATH intensities were normalized to the total area sum within the MarkerView software (version 1.2.1.1). In order to compare different assembly times, the median of three biological replicates was calculated for each protein and each time point. To determine early and late binding proteins, medians of SWATH intensities after 15 min assembly were divided by medians at 4 h for each protein. Resulting ratios were Log2(x) transformed and plotted by Scatter Graph tool sorted from largest to smallest. Ratios higher than 1 were regarded as early binding proteins. Ratios below −1 were regarded as late binding proteins. Proteins with ratios in between both thresholds were regarded as not changing between both time points and therefore as constant binding proteins. A similar technique with different threshold was applied to chromatin assembly samples treated and untreated with TSA. Medians of SWATH intensities for each protein were calculated and the values for TSA-untreated samples were divided by medians of TSA-treated samples for each time point separately to determine the enrichment upon TSA treatment during chromatin assembly. After Log2(x) transformation of the resulting ratios, proteins with ratios higher than 1 were regarded as enriched in unperturbed chromatin assembly whereas proteins with ratios lower than 0.8 were regarded as enriched upon TSA-treatment. Both protein groups were subjected to GO-term analysis. Functional Annotation Clustering for GO-term analysis was performed by means of DAVID Bioinformatics Resources 6.7 (31).

For comparison of sequence specific protein binding, *t* test-based statistics was applied on SWATH intensities. First, the logarithm (log 2) of the SWATH intensities was taken, resulting in a Gaussian distribution of the data. Statistical outliers for the 5S-rDNA compared with the 359 repeat DNA were then determined using two-tailed *t* test. Multiple testing correction was applied by using a permutation-based false discovery rate (FDR) method in Perseus. A Similar technique was used for the statistical evaluation of chromatin enriched proteins in comparison to beads-only bound control proteins.

For comparison between *in vitro* and *in vivo* data, Nascent Chromatin Capture (NCC) repository data were used as *in vivo* data (13). NCC data are illustrated by the quotient of Log2(x) transformed nascent SILAC ratios divided by Log2(x) transformed mature SILAC ratios. *In vitro* data are based on median-averaged SWATH-intensities of three biological replicates of 15 min divided by median-averaged SWATH-intensities of three biological replicates of 4 h. Error bars represent standard error of the mean (S.e.m) with $n = 3$.

Experimental Design and Statistical Rationale—Chromatin assembly experiments have been performed in three biological replicates with three independently collected DREX. As negative controls, beads-only were incubated in three biological replicates with DREX. Silver gels and agarose gels show a representative example out of three replicates. A pilot study in our lab revealed that three biological replicates enable us for a precise and statistical valid conclusion between chromatin assembly experiments and the composition of proteins during different time points of assembly. Based on biological function of the identified proteins, we altered the initial settings for statistical analysis to be (s(0) = 3 and FDR = 0.5%). Using these parameters, we were able to recover almost all chromatin-assembly factors and reduce the unspecific background.

RESULTS

To investigate factors involved during *in vitro* chromatin assembly and to study proteome dynamics of chromatin maturation, we analyzed the proteins binding to an immobilized array of a nucleosome positioning sequence derived from the sea urchin 5S rRNA gene (32) using a well-characterized S-150 chromatin-assembly extract prepared from early *Drosophila* embryos (DREX) (24) (Fig. 1*A*). It has been shown that this protein extract is able to assemble large DNA fragments into an ordered array that closely resembles the chromatin structure seen in early embryos with regards to nucleosome spacing (33) and histone modifications (34, 35) (Fig. 1*B*).

Chromatin matures *in vitro* over a time of ~4 h, which can be observed by the increased regularity of the nucleosomal array generated by MNase digestion (Fig. 1*C*). After assembly and extensive washing, newly synthesized chromatin was

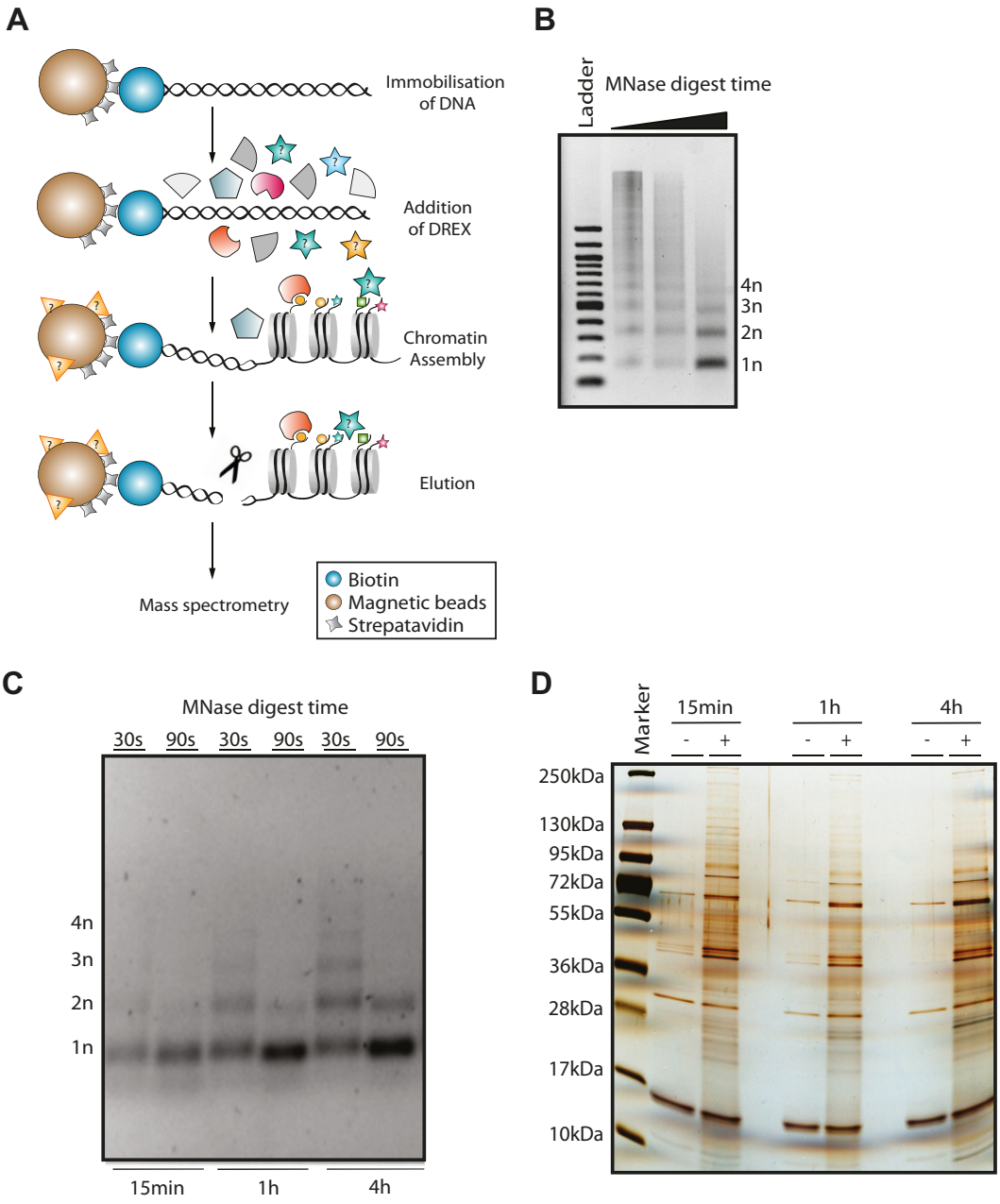

FIG. 1. **Workflow of DREX-mediated chromatin assembly.** *A*, Linearized and biotinylated DNA is immobilized on streptavidin-coated, paramagnetic beads. It is incubated with *Drosophila* embryo extract. The assembly is stopped in regularly time intervals to follow protein binding kinetics. *B*, One microgram of DNA was reconstituted into chromatin by an incubation for 6 h at 26 °C with 40 μl of *Drosophila* embryo extract. It was digested with MNase for 30 s, 90 s and 240 s. M: 100 bp DNA *C*, Two microgram of DNA were incubated for 15 min, 1 h and 4 h at 26 °C with 40 μl *Drosophila* embryo extract and digested with MNase for 30 s, 90 s. *D*, SDS-PAGE of proteins of newly-synthesized chromatin. Beads with (+) and without (-) immobilized DNA were incubated for 15 min, 1 h and 4 h at 26 °C with 45 μl of *DREX*. DNA was cleaved from beads by MNase digestion for 10 min at 26 °C. 10% of each sample was loaded onto a 4–20% gradient gel and proteins were visualized by silver staining.

eluted by a nuclease digestion step, which selectively released all chromatin proteins bound during chromatin assembly. This nuclease-mediated elution had the advantage of a lower background compared with an elution at low pH or high-detergent concentrations (Fig. 1*D*), because proteins that non-specifically interact with the beads are not released (36).

Proteins present in the DREX as well as proteins bound to chromatin were digested with trypsin and the resulting peptides were analyzed on a 2 h LC-MS gradient on an Orbitrap XL mass spectrometer operated in a data-dependent acquisition mode. We identified a total of 977 proteins in the assembly extract with 530 of them being present in all three

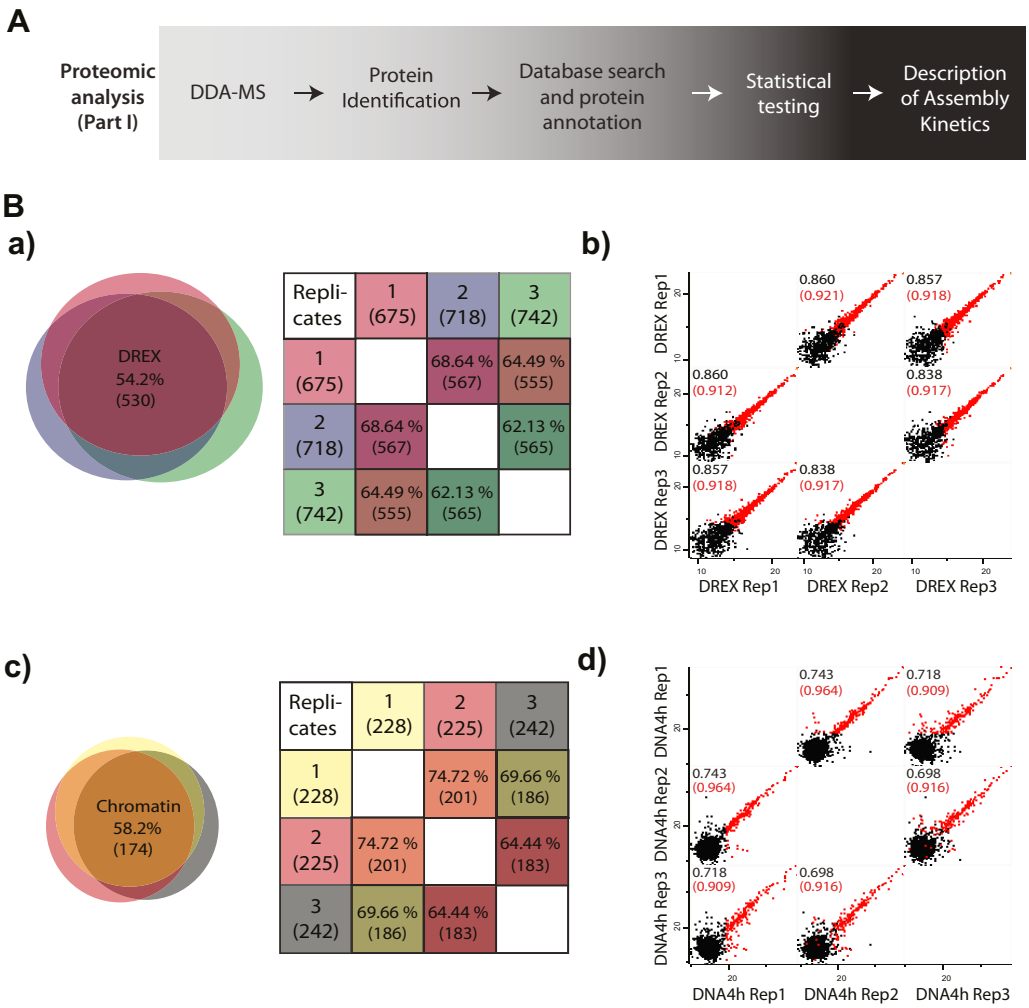

FIG. 2. **Data-dependent acquisition of chromatin assembly.** *A,* Workflow of proteomic analysis for DDA-MS mode. *B*, Venn diagram with respective table showing the overlap of proteins identified by DDA-MS. (a) Overlap of proteins in three biological replicates of DREX with replicate 1 (675 proteins), replicate 2 (718 proteins) and replicate 3 (742 proteins). Over 54% (530 proteins) have been identified in all DREX samples. (b) Multiscatter plots showing the correlation of replicates after imputation of missing values for DREX samples. Red dots show correlation of proteins identified in all three replicates. Black dots display the correlation of proteins identified at least once per data set and after imputation of missing values. Black and embraced red numbers indicate the coefficient of determination in R^2. (c) Overlap of proteins of three 4 h assembly reactions with replicate 1 (228 proteins), replicate 2 (225 proteins) and replicate 3 (242 proteins). Over 58% (313 proteins) are identical in the chromatin samples. (d) Multiscatter plots showing the correlation of replicates after imputation of missing values for chromatin samples as described in Fig. 2*B*, panel (b).

replicates (54.2%) (Fig. 2*B*, panel (a)). When analyzing the proteins bound to chromatin after 4 h, we identified a total of 299 proteins using the OrbitrapXL with 174 (58.2%) being present in all three replicates (Fig. 2*B*, panel (c)). Proteins were quantified using LFQ values provided by the MaxQuant analysis suite. Consistent with a high reproducibility of the assembly system, the intensities of the reproducible proteins correlated well with regards to the individual replicates (red data points Fig. 2*B*, panels (b) and (d)). Proteins that were not identified in all replicates had the tendency to be of low intensity even in the replicates from which they were identified. In order to get more accurate quantitative values and to reduce the number of missing values we generated a high quality ion library of the *Drosophila* chromatin assembly system that could be used for SWATH based label free quantitation (37). To generate this library, we measured the three replicates of the DREX assembly extract and the proteins bound to DNA after 4 h on a 6600 TTOF mass spectrometer in a data dependent acquisition run using a Top25 fragmentation scheme. Proteins were identified using the ProteinPilot software, which was also used to generate the ion library. On the 6600 TTOF instrument, we identified 1050 proteins in the DREX and 943 proteins were associated with chromatin after 4 h with a general overlap of 56.9% or 45.87% respectively (Fig. 3*B*, panels (a) and (c)). Upon retention time calibration using ten conserved peptides that were distributed over the entire LC gradient we could quantify 1035 proteins in the SWATH runs performed on triplicate samples

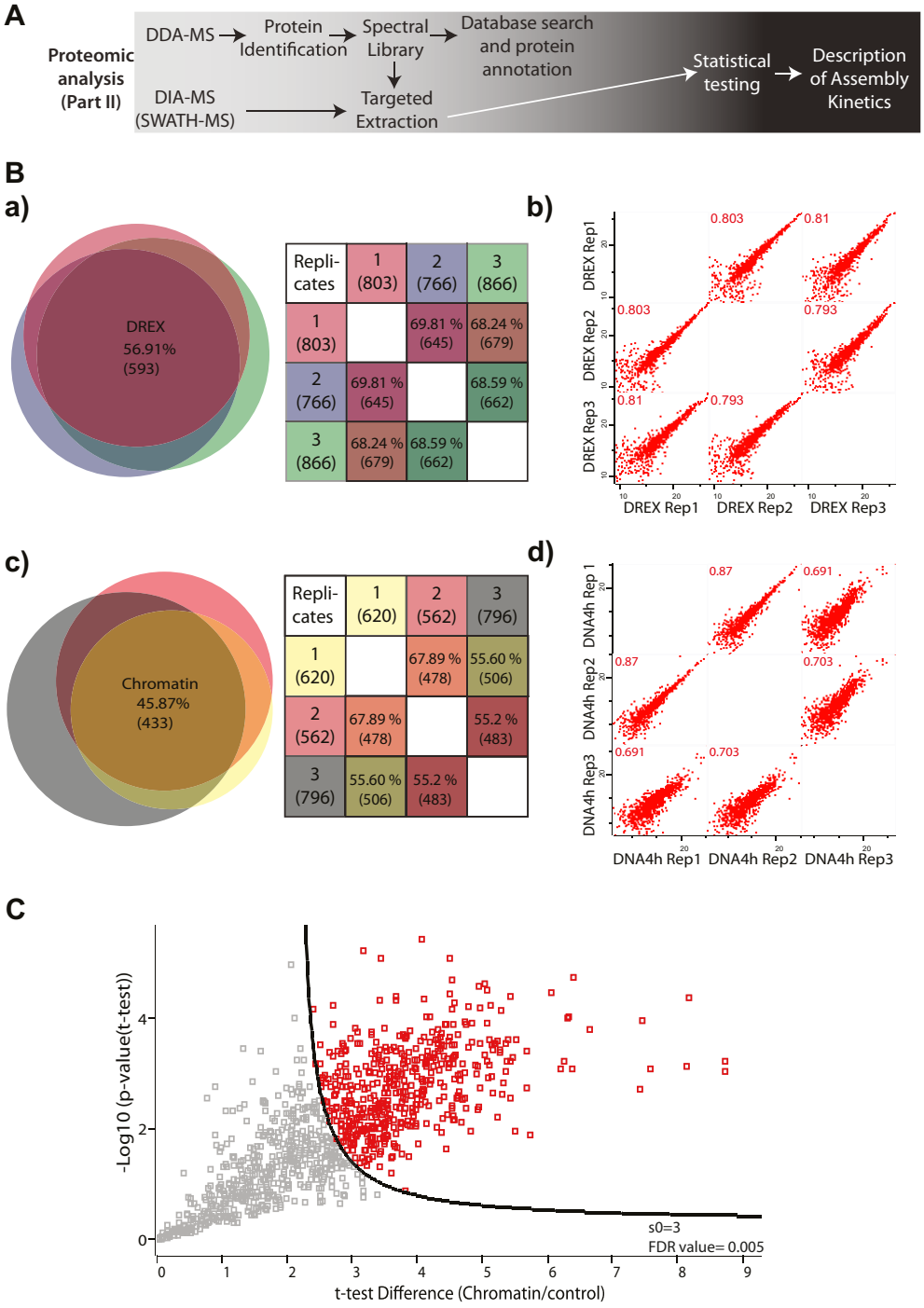

FIG. 3. **Data-dependent acquisition of chromatin assembly for library generation and quantitation with data-independent acquisition.** *A*, Workflow of proteomic analysis for DIA-MS mode (SWATH-MS). *B*, Venn diagram with respective table showing the overlap of proteins identified by DDA-MS. (a) Overlap of proteins in three biological replicates of DREX with replicate 1 (803 proteins), replicate 2 (766 proteins) and replicate 3 (866 proteins). Over 56% (593 proteins) have been identified in all DREX samples. (b) Multiscatter plots showing the correlation of three replicates for DREX samples acquired in DIA-MS mode. All proteins have been identified in all three biological replicates indicated by red color of the dots as described in Fig. 2*B*, panel (b). (c) Overlap of proteins of three 4 h assembly reactions with replicate 1 (620 proteins), replicate 2 (562 proteins) and replicate 3 (796 proteins). Over 45% (433 proteins) are identical in the chromatin samples. (d) Multiscatter plots showing the correlation of three replicates for DREX samples acquired in DIA-MS mode. All proteins have been identified in all three biological replicates indicated by red color of the dots as described in Fig. 2*B*, panel (b). *C*, Identification of chromatin-specific binding proteins by a two-sample, FDR-corrected *t* test of the mean averages of SWATH-values of beads-only control experiments *versus* chromatin assembly experiments (threshold: $p = 0.005$ (FDR from permutation) and $s0 = 3$). Red illustrated dots illustrate chromatin specific proteins (480).

of chromatin assemblies after 15, 60, and 240 min (Fig. 3*B*, panel (d)).

Despite the fact that we selectively eluted the assembled chromatin using a digestion with micrococcal nuclease, which shows a relatively low background (Fig. 1*D*) we still identified and quantified proteins that were eluted from the beads in the absence of chromatin. As we only intended to study proteins specifically assembled on chromatin, we focused all further analysis on the 480 proteins that showed a significant enrichment on chromatin over beads-only control experiments (red dots Fig. 3*C*). Most of these proteins assemble on chromatin independent of the underlying DNA sequence as we only observe a small number of proteins that show differential binding between the 5S rDNA repeat and a plasmid containing a repeated 359bp sequence from *Drosophila* pericentromeres (38) (supplemental Fig. S1*A*).

A hierarchical clustering of averaged log2-transformed SWATH intensities for each sample condition revealed the existence of three main clusters (Fig. 4*A* and 4*B*): One cluster contains 104 proteins that have a relatively high value at 15 min, which then drops with longer assembly times. This cluster containing early chromatin binding proteins contains most factors known to play a role in DNA replication (PCNA, Gnf1 Rfc3/38, etc.) or histone deposition (Ssrp, Dre4, Acf1, Caf subunits), which was also verified using an orthogonal Western blot assay (Fig. 4*C*). The largest number of proteins (774) binds early and remains bound to chromatin. This cluster, which we have termed constant binders, contains most canonical histone proteins but also BigH1/tefu (ATM kinase), DNA pol alpha, Smc5, Chrac 16 and MCM 2,3,7. Finally, 146 proteins only bind after 4 h of assembly. Examples of proteins belonging to this class are factors known to be involved in DNA repair like Irbp and Ku80 or RPA3/2 and 70.

We next wondered whether the *in vitro* chromatin assembly system resembles the general processes that occur during replication coupled chromatin assembly *in vivo*. To do this, we compared our results with the most comprehensive data on the composition of nascent chromatin, which was measured by Alabert *et al.* (13). As the nascent chromatin capture assay was done in a human cell line (HeLa) we first determined the number of chromatin associated *Drosophila* proteins that contained a unique human orthologue using the BioMart algorithm (12). 216 proteins (out of 480) that specifically assembled onto chromatin had a clear orthologue in the human proteome and 171 of them (79%) were also detected as chromatin associated in either nascent or mature chromatin (13). We next wondered if proteins that are enriched in nascent chromatin were also preferentially found at the early time points during *in vitro* chromatin assembly. Indeed many of the early *in vitro* binding proteins were also enriched in nascent (early assembled) chromatin in HeLa cells, which reveals a strong evolutionary conservation of chromatin assembly mechanisms and a high degree of similarity between *in vitro* and *in vivo* chromatin assembly. Similarly, the late binding proteins also have a preference for being enriched in mature chromatin (Fig. 4*D*).

The presence of a broad HDAC inhibitor such as Trichostatin A during replication coupled chromatin assembly *in vivo* results in a failure to establish repressive chromatin (39) and a broadening of replication initiation areas (40) suggesting that it results in a more open chromatin structure. To get a better insight into the effect of TSA on chromatin assembly, we performed a time resolved analysis of protein assembled to chromatin *in vitro* in the absence or presence of TSA. We verified the efficiency of the TSA treatment on the acetylation levels by measuring the acetylation of H4 in the absence or presence of TSA (supplemental Fig. S1*D* and S1*E*). It is worth mentioning that the presence of TSA does not result in a histone hyperacetylation, which is usually observed in tissue culture cells but induces a moderate increase in the acetylation because of a failure to efficiently remove the acetylation pattern present on histones before assembly (41). The reason for this is unclear but it is probably because of a general lack of site specific histone acetyltransferases in early embryos (34). Nevertheless, we observe a substantial change in the proteomic composition of chromatin assembled in the presence of TSA. To detect all proteins affected by TSA treatment we quantified the changes of all proteins rather than only focusing on the ones that were significantly enriched on chromatin. An unsupervised clustering revealed a clear difference between chromatin assembled in the absence or the presence of TSA (Fig. 5*A*). We found 131, 224, and 270 proteins being enriched on chromatin upon TSA treatment and 195, 71, and 121 being reduced after 15, 60, and 240 min respectively (Fig. 5*B*). A GO term analysis of the assembled chromatin in the presence of TSA suggests that the TSA treatment results in an increased association of factors that interact non-specifically with chromatin in comparison to the assembly reaction that is unperturbed (Table I), which suggests a more open and hence more error prone chromatin assembly. Alternatively, the higher degree of nonspecific binding could also be caused by a hyperacteylation and concomitant regulation of the multiple chaperones we find to be associated with *in vitro* assembled chromatin and whose job it may be to remove unwanted protein associations. We also find an increased binding of all subunits of the MCM helicase complex to chromatin that is treated with TSA (Fig. 5*C*), which is interesting in light of earlier findings that show a general broadening of DNA replication and a more widespread binding of MCMs to replication origins *in vivo* when cells are treated with TSA (40).

DISCUSSION

The combination of label-free quantitative proteomics with a well-described *in vitro* chromatin assembly system allowed us to temporally dissect the process of chromatin assembly. By choosing three different time points, we were able to describe the kinetics of the chromatin assembly *in vitro* at an

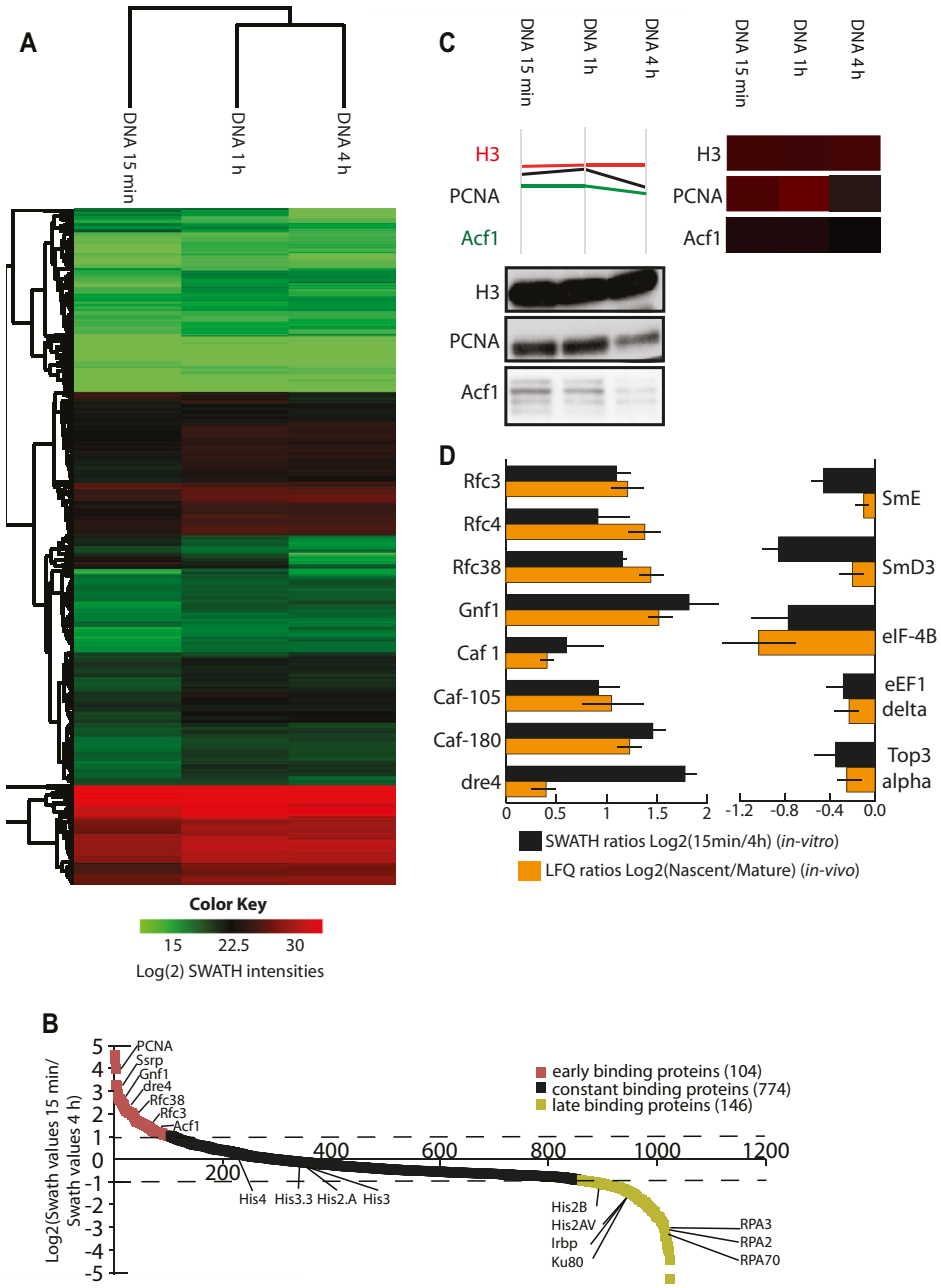

FIG. 4. **The Kinetics of Chromatin Assembly.** *A*, Heatmap illustrating Log2(x) transformed median averaged SWATH intensities from three replicates after euclidean clustering. Rows indicate proteins, columns represent different conditions of chromatin assembly. Red and green colors indicate SWATH intensities. *B*, Waterfall plot illustrates Log(2)-transformed quotient of median averaged replicates after 15 min binding divided by replicates of 4 h. Proteins were sorted from largest to smallest quotient and plotted with a scatter graph tool. 104 proteins (red) show higher values than 1 and can be identified as early binding proteins. 774 proteins (black) have quotient value between 1 and −1 and do not change between 15 min and 4 h. 146 late binding proteins (yellow) have negative values below −1. They are enriched after 4 h. Examples of proteins are spotted in the waterfall plot. *C*, Time resolved kinetics for Histone 3 (H3), Proliferating cell nuclear antigen (PCNA) and ATP-dependent chromatin assembly factor 1 (Acf1). Plots illustrate the Log2(x) transformed SWATH intensities in each assembly condition for 15 min, 1 h and 4 h. Each line illustrates the binding behavior of a single protein. Enlarged heatmap rows illustrate the kinetics of specified factors as described in legend Fig. 3*A*. Western blot of all three proteins in chromatin assembly experiments for 15 min, 1 h and 4 h. *D*, Comparison of nascent and mature chromatin enrichment of selected factors between *in vitro* and *in vivo* data. Black bars illustrate the Log2(x) transformed difference between median averaged 15 min replicates divided by median averaged 4 h replicates. Orange bars illustrate the Log2(x) transformed difference between median averaged nascent chromatin divided by median averaged mature chromatin from NCC experiments performed by Alabert *et al.* 2014. Error bars describe the standard error of the mean.

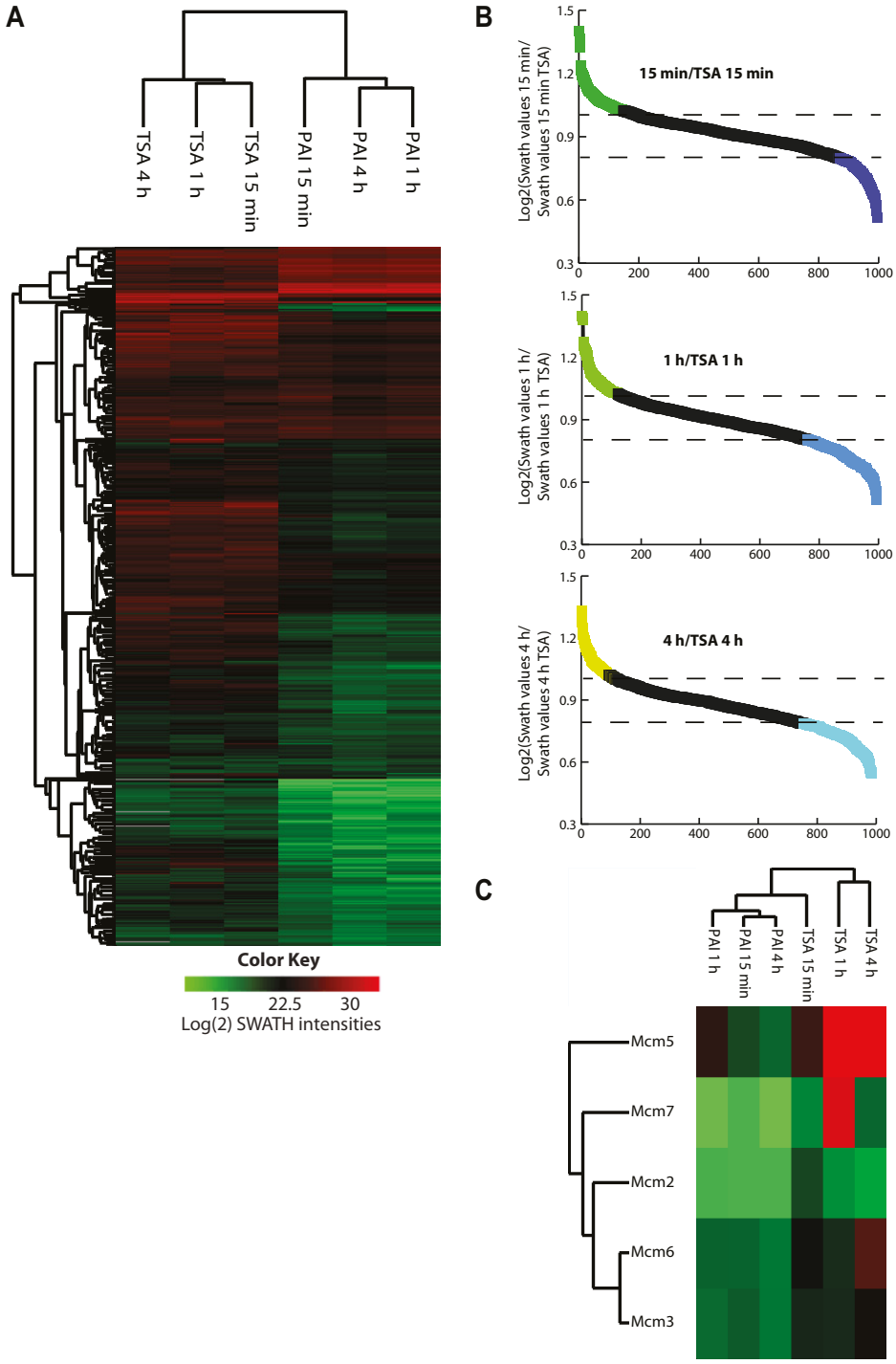

FIG. 5. **Challenge of chromatin assembly by means of chromatin decondensation.** *A*, Heatmap illustrating Log2(x) transformed median averaged SWATH intensities from three replicates after euclidean clustering. Rows indicate proteins, columns represent different conditions of chromatin assembly for unperturbed assembly and TSA-treated chromatin assembly for different time points. Red and green colors indicate SWATH intensities. *B*, Waterfall plots illustrate change between unperturbed assembly and TSA-treated assembly. For each of the three time points, Log2(x) transformed median-averaged unperturbed assembly replicates were divided by respective TSA-treated assembly replicates. Proteins were sorted from largest to smallest quotient and plotted with a scatter graph tool. Proteins with values higher than 1 are regarded as enriched in unperturbed chromatin assembly. Proteins with values lower than 0.8 are regarded as enriched upon chromatin decondensation. *C*, Heatmap illustrating Log2(x) transformed median averaged SWATH intensities from three replicates after euclidean clustering. Rows indicate Mcm proteins, columns represent different conditions of chromatin assembly for unperturbed assembly and TSA-treated chromatin assembly for different time points. Red and green colors indicate SWATH intensities.

TABLE I

Functional Annotation Clustering for factors enriched or repelled upon TSA. Colored Proteins with Log2(x) values in Fig. 5B were processed with the Functional Annotation Tool from DAVID Bioinformatics Resources 6.7, NIAID/NIH. FBgn numbers were uploaded as list and used for the search for GO-terms of biological processes according to the Drosophila melanogaster background with a threshold of 2 counts and an EASE of 0.1. Table shows GO-terms with highest enrichment scores for proteins enriched or repelled upon TSA treatment for samples after all three time points of assembly

GO-term analysis of 15 min unperturbed assembly

Enrichment Score: 9.22

Term	Count	PValue	FDR
mitotic spindle organization	25	3.80E-13	6.01E-10
microtubule cytoskeleton organization	29	3.47E-12	5.49E-09
mitotic cell cycle	28	1.36E-09	2.15E-06
cell cycle process	32	6.80E-08	1.08E-04

Enrichment Score: 6.2

Term	Count	PValue	FDR
atp-binding	31	7.92E-10	9.74E-07
nucleotide binding	53	2.67E-08	3.60E-05
ATPase activity	24	5.61E-07	7.57E-04
ribonucleotide binding	41	1.68E-06	0.002265

Enrichment Score: 5.04

Term	Count	PValue	FDR
glycolysis	8	2.90E-08	3.56E-05
gluconeogenesis	5	1.32E-06	0.001627
binding site:Substrate	6	0.016564	19.71693
sequence variant	12	0.087857	70.15314

Enrichment Score: 4.94

Term	Count	PValue	FDR
DNA packaging	13	3.84E-09	6.07E-06
nucleosome organization	11	9.58E-09	1.52E-05
chromatin assembly	9	3.71E-07	5.86E-04
nucleosome assembly	8	1.12E-06	0.001774

GO-term analysis of 15 min TSA treated assembly

Enrichment Score: 2.12

Term	Count	PValue	FDR
cytoplasm	14	1.63E-04	0.186748
alternative splicing	13	0.010174	11.04422
splice variant	13	0.256158	97.44797

Enrichment Score: 1.85

Term	Count	PValue	FDR
cell division	10	2.03E-04	0.309723
septin ring	3	9.74E-04	1.215431
cell cortex part	4	0.021168	23.53893
cell cycle	12	0.043918	49.58396

Enrichment Score: 1.76

Term	Count	PValue	FDR
cell cortex	6	0.002667	3.294388
organelle localization	4	0.039664	46.05263
organelle localization	4	0.049507	53.89628

Enrichment Score: 1.75

Term	Count	PValue	FDR
nuclear envelope	8	1.04E-04	0.131005
endomembrane system	10	0.001763	2.189531
nuclear pore	5	0.002557	3.161166
pore complex	5	0.003245	3.994777

GO-term analysis of 1 h unperturbed assembly

Enrichment Score: 9.28

Term	Count	PValue	FDR
microtubule cytoskeleton organization	29	2.47E-14	3.87E-11
mitotic spindle organization	23	4.60E-13	7.20E-10
mitotic cell cycle	26	4.97E-10	7.78E-07
M phase	27	4.20E-08	6.58E-05

Enrichment Score: 4.80

Term	Count	PValue	FDR
Chaperonin Cpn60/TCP-1	7	9.96E-09	1.37E-05
protein folding	14	3.95E-08	6.18E-05
Chaperone	8	3.84E-06	0.004638
GO:0016887~ATPase activity	14	0.007046	8.940187

Enrichment Score: 3.62

Term	Count	PValue	FDR
nucleotide-binding	30	3.63E-09	4.39E-06
atp-binding	23	6.76E-07	8.17E-04
ATPase activity, coupled	13	0.00604	7.709858
ATPase activity	14	0.007046	8.940187

Enrichment Score: 3.61

Term	Count	PValue	FDR
proteasome	11	4.55E-11	5.50E-08
proteasome complex	12	1.87E-08	2.29E-05
Proteasome, subunit alpha/beta	7	8.97E-07	0.001234
proteasome core complex	7	8.17E-06	0.010008

GO-term analysis of 1 h TSA treated assembly

Enrichment Score: 2.59

Term	Count	PValue	FDR
ligase	12	1.91E-04	0.233314
ligase activity	10	9.84E-04	1.314718
acid-amino acid ligase activity	9	0.001069	1.427017
small conjugating protein ligase activity	7	0.005496	7.1377

Enrichment Score: 2.08

Term	Count	PValue	FDR
cytoskeleton	10	8.61E-06	0.010528
cytoskeleton organization	23	3.30E-05	0.052609
cytoskeletal part	17	3.56E-04	0.464948
mitosis	10	0.001368	2.159029

Enrichment Score: 2.03

Term	Count	PValue	FDR
Transcription factor CBF/NF-Y	4	2.14E-04	0.302006
Histone-fold	4	0.008696	11.5976
sequence-specific DNA binding	6	0.407632	99.91201

Enrichment Score: 1.98

Term	Count	PValue	FDR
nuclear envelope	10	8.09E-05	0.10594
nuclear pore	7	3.54E-04	0.462561
pore complex	7	5.06E-04	0.660863
organelle envelope	13	0.064149	58.0273

GO-term analysis of 4 h unperturbed assembly

Enrichment Score: 7.66

Term	Count	PValue	FDR
microtubule cytoskeleton organization	22	1.76E-10	2.72E-07
cytoskeleton organization	26	8.11E-10	1.25E-06
mitotic spindle organization	17	3.52E-09	5.44E-06
mitotic cell cycle	22	5.03E-09	7.78E-06

Enrichment Score: 3.43

Term	Count	PValue	FDR
glycolysis	7	6.80E-08	8.09E-05
gluconeogenesis	4	2.69E-05	0.032023
binding site:Substrate	4	0.095377	71.81538
sequence variant	8	0.24148	96.95562

Enrichment Score: 2.73

Term	Count	PValue	FDR
protein folding	10	1.79E-05	0.027585
Chaperone	6	2.00E-04	0.23782
Chaperonin Cpn60/TCP-1	4	3.78E-04	0.49338
unfolded protein binding	6	0.003623	4.63903

Enrichment Score: 2.73

Term	Count	PValue	FDR
proteasome	8	1.48E-07	1.76E-04
proteasome complex	9	5.23E-06	0.006187
Proteasome, subunit alpha/beta	6	7.74E-06	0.010139
ubl conjugation pathway	3	0.232195	95.68984

GO-term analysis of 4 h TSA treated assembly

Enrichment Score: 4.07

Term	Count	PValue	FDR
chromosome	23	2.24E-05	0.029437
chromosomal part	20	2.99E-05	0.039307
non-membrane-bounded organelle	44	2.71E-04	0.356303
non-membrane-bounded organelle	44	2.71E-04	0.356303

Enrichment Score: 3.01

Term	Count	PValue	FDR
DNA metabolic process	17	5.77E-05	0.092643
DNA repair	10	6.16E-04	0.98601
cellular response to stress	13	8.41E-04	1.343486
response to DNA damage stimulus	10	0.001459	2.318664

Enrichment Score: 2.24

Term	Count	PValue	FDR
translational initiation	8	2.72E-04	0.43593
translation initiation factor activity	8	3.37E-04	0.457158
protein biosynthesis	8	0.001828	2.213803
Initiation factor	4	0.027596	28.99392

Enrichment Score: 2.18

Term	Count	PValue	FDR
cytoskeleton	10	4.34E-05	0.053145
cell cycle	10	7.91E-05	0.096712
microtubule cytoskeleton organization	15	0.010139	15.10588
spindle organization	12	0.01656	23.53692

unprecedented depth. Although the *in vitro* chromatin assembly system has been shown to recapitulate aspects of chromatin assembly *in vivo* such as chromatin spacing and the establishment of specific histone modification patterns (41, 42) or the coordinated binding and release of chromatin binding factors (41, 43, 44), the proteomic composition of *in vitro* and *in vivo* assembled chromatin has not been compared so far. The recent advancements toward a complete proteomic description of *in vivo* chromatin assembly (13, 26) have now made such comparisons feasible.

The use of a deep high quality ion library generated from the chromatin assembly extract and the chromatin bound factors enabled us to determine the kinetic of chromatin binding for 480 proteins highly enriched in chromatin. Because of the fact that the *in vitro* extract is prepared from Drosophila and the nascent chromatin capture has been performed in human cells, we could only compare 217 of these 480 factors that had a clear orthologue in the human proteome. Given this limitation, we find 171 (79%) of the proteins detected as bound to chromatin during assembly *in vitro* also on nascent chromatin. This supports the hypothesis that the general pathways for chromatin assembly are conserved among different eukaryotes.

Interestingly, not only the identity of the bound proteins is conserved but also their dynamics of binding. For example, we do find multiple components of the replication machinery like the RFC clamp loading complex, PCNA, the single strand binding complex RPA, members of the MCM helicase complex and the subunits of chromatin assembly factor CAF1 enriched at early time points of the assembly reaction, which is consistent with their enrichment in nascent over mature chromatin shown by NCC data. The possibility to perform measurements of all chromatin associated proteins at three time points using SWATH based quantitation enabled us to include an intermediate measurement of assembly (1 h) rather than only comparing nascent and mature chromatin. This dissection of the binding kinetics verifies earlier *in vitro* findings that the clamp loader complex first binds to the DNA and facilitates PCNA loading and then dissociates from the template upon loading of the sliding clamp leaving PCNA bound to the template to stabilize the polymerase. So far this loading event has only been characterized upon reconstitution in a highly purified system (45, 46). The fact that we observe a transient peak of PCNA at the intermediate time point during *in vitro* chromatin, which decreases substantially at late chromatin assembly resembling nascent chromatin findings (13), suggests that this loading process is indeed occurring during chromatin assembly. In contrast to what happens to the replication and histone deposition machinery and similar to what is seen *in vivo*, most proteins rapidly associate with chromatin and stay bound. Many of these factors are bona fide structural components of chromatin like the canonical histone proteins, the linker histone and HMG proteins like HMGD. However, we also detect proteins and protein complexes involved in protein folding like members of the TriC/CCT complex (Tcp1, TCPeta, TCPzeta, CG7033, CG8258, CCT5, and CCTgamma) or Hsp60 and Hsp70, being stably bound to chromatin during assembly. This binding of chaperone factors suggests that chromatin assembly is associated with extensive and continuous protein folding and refolding events. Although TriC/CCT has been suggested to be mainly involved in the folding of newly translated proteins (47), there is increasing evidence that it has additional functions such as the prevention of the aggregation of poly-Q proteins (48) or the formation of macromolecular protein complexes in the nucleus (49, 50). The TriC/CCT complex can also be observed on nascent and mature chromatin *in vivo* (13) and in interphase chromatin (12), which also supports its integral role in chromatin dynamics and metabolism. The function of these chaperones in chromatin has so far been enigmatic and no clear function was assigned to them. The fact that we find much more factors with no apparent chromatin function associated with chromatin upon TSA treatment suggests that chaperones may be required to prevent such proteins from binding to accessible (*i.e.* TSA treated) chromatin or remove mislocalized factors. This hypothesis would fit to previous findings that the formation of regularly spaced chromatin *in vitro* is entirely dependent on the presence of ATP (20, 24), which is a substrate of all nucleosome remodeling factors but also of the Hsp and TriC/CCT chaperones. Alternatively, chaperones might function in the remodeling of chromatin-bound multisubunit complexes. This hypothesis is supported by findings showing that molecular chaperones are responsible for the rapid exchange of hormone receptors during oscillating transcription (51, 52). An example of such a complex that gets remodeled while bound to chromatin is the CAF1 complex. The two large, CAF1 specific, subunits are removed from chromatin at later stages of assembly whereas the small CAF1 histone binding subunit can also be detected at later time points. As the small subunit is also part of the NURF chromatin-remodeling complex, whose other three subunits (ISWI, ACF1, and NURF38) are also present later, it may stay bound to chromatin serving as an additional binding platform for the other components.

In order to test whether we are able to investigate quantitative changes in the assembly kinetics of specific chromatin factors upon a challenge of the system, we added the broad histone deacetylase inhibitor TSA to a reaction and repeated the quantitative proteomic analysis of the bound proteins. We have shown in the past that TSA treatment results in a moderate increase of histone acetylation on assembled chromatin in contrast to what is observed in tissue culture cells where TSA treatment results in a strong hyperacetylation of histones (53–55). During *in vitro* assembly, the histones are deposited in a preacetylated form, which will not get deacetylated when TSA is present (41). TSA treatment has been shown to prevent the formation of transcriptionally repressed chromatin (39) in *Xenopus* oocytes and lead to an altered pattern of DNA replication origin activity in human tissue culture cells (40). The

exact mechanism of how TSA mediates these effects has so far remained unclear. Our finding that all MCM proteins, which are key regulators of eukaryotic replication having an increased binding to chromatin when TSA is present, provides a potential explanation for this finding. However, it remains to be determined whether this is an indirect effect of a more open chromatin structure or a direct effect that is mediated by MCM acetylation (55).

Our results show that the use of a quantitative method to investigate chromatin assembly *in vitro* resembles many aspects of replication dependent chromatin synthesis *in vivo* and can therefore be used to dissect key regulatory steps in chromatin assembly, which are often difficult to investigate in living cells.

Acknowledgments—We thank Peter Becker and Lisa Harpprecht for helpful discussions and sharing of reagents. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (*http://www.proteomexchange.org*) via the PRIDE partner repository (54) with the data set identifier PXD002537 and PXD003445. We thank all members of the Imhof group for critical discussions and feedback. Annotated spectra can be viewed at MS-Viewer (http://prospector2.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msviewer). Spectra from DREX samples can be accessed with search key: hlyzvamoxh. Spectra from chromatin samples can be accessed with search key: 3keubvytc7.

* This work was supported by grants from the German Research council (Deutsche Forschungsgemeinschaft (SFB1064 A16, Z4)) and the European Union (EpiGeneSys, 257082).

Ⓢ This article contains supplemental materials.

¶ To whom correspondence should be addressed: BioMedical Center and Center for Integrated Protein Sciences Munich, Ludwig-Maximilians-University, Großhaderner Straße 9, 82152 Planegg-Martinsried, Germany. Tel.: +49 89 2180 75420; Fax: +49 89 2180 75440; E-mail: imhof@lmu.de.

‖ Both authors contributed equally to this work.

Moritz Völker-Albert: Moritz.Voelker-Albert@med.uni-muenchen.de, Miriam Pusch: Miriam.Pusch@web.de, Andreas Fedisch: Andreas.Fedisch@med.uni-muenchen.de, Pierre Schilcher: Pierre.schilcher@med.uni-muenchen.de, Andreas Schmidt: A.Schmidt@med.uni-muenchen.de.

REFERENCES

1. Teves, S. S., and Henikoff, S. (2011) Heat shock reduces stalled RNA polymerase II and nucleosome turnover genome-wide. *Genes Dev.* **25,** 2387–2397
2. Groth, A., Corpet, A., Cook, A. J., Roche, D., Bartek, J., Lukas, J., and Almouzni, G. (2007) Regulation of replication fork progression through histone supply and demand. *Science* **318,** 1928–1931
3. Jasencakova, Z., Scharf, A. N., Ask, K., Corpet, A., Imhof, A., Almouzni, G., and Groth, A. (2010) Replication Stress Interferes with Histone Recycling and Predeposition Marking of New Histones. *Mol. Cell* **37,** 736–743
4. Campos, E. I., Fillingham, J., Li, G., Zheng, H., Voigt, P., Kuo, W. H., Seepany, H., Gao, Z., Day, L. A., Greenblatt, J. F., and Reinberg, D. (2010) The program for processing newly synthesized histones H3.1 and H4. *Nat. Struct. Mol. Biol.* **17,** 1343–1351
5. Alvarez, F., Munoz, F., Schilcher, P., Imhof, A., Almozuni, G., and Loyola, A. (2011) Sequential establishment of marks on soluble histones H3 and H4. *J. Biol. Chem.* **286,** 17714–17721
6. Hamiche, A., and Shuaib, M. (2013) Chaperoning the histone H3 family. *Biochim. Biophys. Acta.* **1819,** 230–237
7. Benson, L. J., Gu, Y., Yakovleva, T., Tong, K., Barrows, C., Strack, C. L., Cook, R. G., Mizzen, C. A., and Annunziato, A. T. (2006) Modifications of H3 and H4 during chromatin replication, nucleosome assembly, and histone exchange. *J. Biol. Chem.* **281,** 9287–9296
8. Loyola, A., Bonaldi, T., Roche, D., Imhof, A., and Almouzni, G. (2006) PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Mol. Cell* **24,** 309–316
9. Alabert, C., Barth, T. K., Reverón-Gómez, N., Sidoli, S., Schmidt, A., Jensen, O. N., Imhof, A., and Groth, A. (2015) Two distinct modes for propagation of histone PTMs across the cell cycle. *Genes Dev.* **29,** 585–590
10. Sirbu, B. M., Couch, F. B., and Cortez, D. (2012) Monitoring the spatiotemporal dynamics of proteins at replication forks and in assembled chromatin using isolation of proteins on nascent DNA. *Nat. Protocols* **7,** 594–605
11. Ohta, S., Bukowski-Wills, J. C., Sanchez-Pulido, L., Alves, F. de L., Wood, L., Chen, Z. A., Platani, M., Fischer, L., Hudson, D. F., Ponting, C. P., Fukagawa, T., Earnshaw, W. C., and Rappsilber, J. (2010) The protein composition of mitotic chromosomes determined using multiclassifier combinatorial proteomics. *Cell* **142,** 810–821
12. Kustatscher, G., Hégarat, N., Wills, K. L., Furlan, C., Bukowski-Wills, J. C., Hochegger, H., and Rappsilber, J. (2014) Proteomics of a fuzzy organelle: interphase chromatin. *EMBO J.* **33,** 648–664
13. Alabert, C., Bukowski-Wills, J. C., Lee, S. B., Kustatscher, G., Nakamura, K., de Lima Alves, F., Menard, P., Mejlvang, J., Rappsilber, J., and Groth, A. (2014) Nascent chromatin capture proteomics determines chromatin dynamics during DNA replication and identifies unknown fork components. *Nat. Cell Biol.* **16,** 281–293
14. Worcel, A., Han, S., and Wong, M. L. (1978) Assembly of newly replicated chromatin. *Cell* **15,** 969–977
15. Smith, P. A., Jackson, V., and Chalkley, R. (1984) Two-stage maturation process for newly replicated chromatin. *Biochemistry* **23,** 1576–1581
16. Ladoux, B., Quivy, J. P., Doyle, P., du, Roure, O., Almouzni, G., and Viovy, J. L. (2000) Fast kinetics of chromatin assembly revealed by single-molecule videomicroscopy and scanning force microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **97,** 14251–14256
17. Wagner, G., Bancaud, A., Quivy, J. P., Clapier, C., Almouzni, G., and Viovy, J. L. (2005) Compaction Kinetics on Single DNAs: Purified Nucleosome Reconstitution Systems versus Crude Extract. *Biophys. J.* **89,** 3647–3659
18. 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,* 1–13
19. Dutta, B., Adav, S. S., Koh, C. G., Lim, S. K., Meshorer, E., and Sze, S. K. (2012) Elucidating the temporal dynamics of chromatin-associated protein release upon DNA digestion by quantitative proteomic approach. *J. Proteomics* **75,** 5493–5506
20. Bulger, M., Ito, T., Kamakaka, R. T., and Kadonaga, J. T. (1995) Assembly of regularly spaced nucleosome arrays by Drosophila chromatin assembly factor 1 and a 56-kDa histone-binding protein. *Proc. Natl. Acad. Sci. U.S.A.* **92,** 11726–11730
21. Tyler, J. K., Adams, C. R., Chen, S. R., Kobayashi, R., Kamakaka, R. T., and Kadonaga, J. T. (1999) The RCAF complex mediates chromatin assembly during DNA replication and repair. *Nature* **402,** 555–560
22. Kleinschmidt, J. A., and Franke, W. W. (1982) Soluble acidic complexes containing histones H3 and H4 in nuclei of Xenopus laevis oocytes. *Cell* **29,** 799–809
23. Shimamura, A., and Worcel, A. (1989) The assembly of regularly spaced nucleosomes in the Xenopus oocyte S-150 extract is accompanied by deacetylation of histone H4. *J. Biol. Chem.* **264,** 14524–14530
24. Becker, P. B., and Wu, C. (1992) Cell-free system for assembly of transcriptionally repressed chromatin from Drosophila embryos. *Mol. Cell. Biol.* **12,** 2241–2249
25. Kamakaka, R. T., Bulger, M., and Kadonaga, J. T. (1993) Potentiation of RNA polymerase II transcription by Gal4-VP16 during but not after DNA replication and chromatin assembly. *Genes Dev.* **7,** 1779–1795
26. Kliszczak, A. E., Rainey, M. D., Harhen, B., Boisvert, F. M., and Santocanale, C. (2011) DNA mediated chromatin pull-down for the study of chromatin replication. *Sci. Rep.* **1,** 95
27. Gillet, L. C., Navarro, P., Tate, S., Rost, H., Selevsek, N., Reiter, L., Bonner, R., and Aebersold, R. (2012) Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. *Mol. Cell. Proteomics* **11**, O111.016717–O111.016717

28. Jain, D., Baldi, S., Zabel, A., Straub, T., and Becker, P. B. (2015) Active promoters give rise to false positive "Phantom Peaks" in ChIP-seq experiments. *Nucleic Acids Res.* **43,** 6959–6968
29. Blum, H., Beier, H., and Gross, H. J. (1987) Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* **8,** 93–99
30. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protocols* **2,** 1896–1906
31. Huang, da, W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protocols* **4,** 44–57
32. Hansen, J. C., van Holde, K. E., and Lohr, D. (1991) The mechanism of nucleosome assembly onto oligomers of the sea urchin 5 S DNA positioning sequence. *J. Biol. Chem.* **266,** 4276–4282
33. Blank, T. A., Sandaltzopoulos, R., and Becker, P. B. (1997) Biochemical analysis of chromatin structure and function using Drosophila embryo extracts. *Methods* **12,** 28–35
34. 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
35. Scharf, A. N., Barth, T. K., and Imhof, A. (2009) Establishment of histone modifications after chromatin assembly. *Nucleic Acids Res.* **37,** 5032–5040
36. Mellacheruvu, D., Wright, Z., Couzens, A. L., Lambert, J.-P., St-Denis, N. A., Li, T., Miteva, Y. V., Hauri, S., Sardiu, M. E., Low, T. Y., Halim, V. A., Bagshaw, R. D., Hubner, N. C., al-Hakim, A., Bouchard, A., Faubert, D., Fermin, D., Dunham, W. H., Goudreault, M., Lin, Z.-Y., Badillo, B. G., Pawson, T., Durocher, D., Coulombe, B., Aebersold, R., Superti-Furga, G., Colinge, J., Heck, A. J. R., Choi, H., Gstaiger, M., Mohammed, S., Cristea, I. M., Bennett, K. L., Washburn, M. P., Raught, B., Ewing, R. M., Gingras, A.-C., and Nesvizhskii, A. I. (2013) The CRAPome: a contaminant repository for affinity purification–mass spectrometry data. *Nat. Meth.* **10,** 730–736
37. Schubert, O. T., Gillet, L. C., Collins, B. C., Navarro, P., Rosenberger, G., Wolski, W. E., Lam, H., Amodei, D., Mallick, P., MacLean, B., and Aebersold, R. (2015) Building high-quality assay libraries for targeted analysis of SWATH MS data. *Nat. Protocols* **10,** 426–441
38. DiBartolomeis, S. M., Tartof, K. D., and Jackson, F. R. (1992) A superfamily of Drosophila satellite related (SR) DNA repeats restricted to the X chromosome euchromatin. *Nucleic Acids Res.* **20,** 1113–1116
39. Wong, J., Patterton, D., Imhof, A., Guschin, D., Shi, Y. B., and Wolffe, A. P. (1998) Distinct requirements for chromatin assembly in transcriptional repression by thyroid hormone receptor and histone deacetylase. *EMBO J.* **17,** 520–534
40. Kemp, M. G., Ghosh, M., Liu, G., and Leffak, M. (2005) The histone deacetylase inhibitor trichostatin A alters the pattern of DNA replication origin activity in human cells. *Nucleic Acids Res.* **33,** 325–336
41. Scharf, A. N., Meier, K., Seitz, V., Kremmer, E., Brehm, A., and Imhof, A. (2009) Monomethylation of lysine 20 on histone H4 facilitates chromatin maturation. *Mol. Cell. Biol.* **29,** 57–67
42. Shimamura, A., and Worcel, A. (1989) The assembly of regularly spaced nucleosomes in the Xenopus oocyte S-150 extract is accompanied by deacetylation of histone H4. *J. Biol. Chem.* **264,** 14524–14530
43. Nightingale, K., Dimitrov, S., Reeves, R., and Wolffe, A. P. (1996) Evidence for a shared structural role for HMG1 and linker histones B4 and H1 in organizing chromatin. *EMBO J.* **15,** 548–561
44. Nightingale, K. P., Wellinger, R. E., Sogo, J. M., and Becker, P. B. (1998) Histone acetylation facilitates RNA polymerase II transcription of the Drosophila hsp26 gene in chromatin. *EMBO J.* **17**, 2865–2876
45. Gomes, X. V., Schmidt, S. L., and Burgers, P. M. (2001) ATP utilization by yeast replication factor C. II. Multiple stepwise ATP binding events are required to load proliferating cell nuclear antigen onto primed DNA. *J. Biol. Chem.* **276,** 34776–34783
46. Hedglin, M., Perumal, S. K., Hu, Z., and Benkovic, S. (2013) Stepwise assembly of the human replicative polymerase holoenzyme. *eLife* **2,** e00278
47. Dunn, A. Y., Melville, M. W., and Frydman, J. (2001) Review: cellular substrates of the eukaryotic chaperonin TRiC/CCT. *J. Struct. Biol.* **135,** 176–184
48. Shahmoradian, S. H., Galaz-Montoya, J. G., Schmid, M. F., Cong, Y., Ma, B., Spiess, C., Frydman, J., Ludtke, S. J., and Chiu, W. (2013) TRiC's tricks inhibit huntingtin aggregation. *eLife* **2,** e00710
49. Freund, A., Zhong, F. L., Venteicher, A. S., Meng, Z., Veenstra, T. D., Frydman, J., and Artandi, S. E. (2014) Proteostatic control of telomerase function through TRiC-mediated folding of TCAB1. *Cell* **159,** 1389–1403
50. Guenther, M. G. Yu, J., Kao, G. D., Yen, T. J., Lazar, M. A. (2002) Assembly of the SMRT-histone deacetylase 3 repression complex requires the TCP-1 ring complex. *Genes Dev.* **16,** 3130–3135
51. McNally, J. G., Müller, W. G., Walker, D., Wolford, R., and Hager, G. L. (2000) The glucocorticoid receptor: rapid exchange with regulatory sites in living cells. *Science* **287,** 1262–1265
52. Stavreva, D. A., Müller, W. G., Hager, G. L., Smith, C. L., and McNally, J. G. (2004) Rapid glucocorticoid receptor exchange at a promoter is coupled to transcription and regulated by chaperones and proteasomes. *Mol. Cell. Biol.* **24,** 2682–2697
53. Turner, B. M., O'Neill, L. P., and Allan, I. M. (1989) Histone H4 acetylation in human cells. Frequency of acetylation at different sites defined by immunolabeling with site-specific antibodies. *FEBS Letters* **253,** 141–145
54. Bantscheff, M., Hopf, C., Savitski, M. M., Dittmann, A., Grandi, P., Michon, A. M., Schlegl, J., Abraham, Y., Becher, I., Bergamini, G., Boesche, M., Delling, M., Dümpelfeld, B., Eberhard, D., Huthmacher, C., Mathieson, T., Poeckel, D., Reader, V., Strunk, K., Sweetman, G., Kruse, U., Neubauer, G., Ramsden, N. G., and Drewes, G. (2011) Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nat. Biotechnol.* **29,** 255–265
55. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., and Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325,** 834–840