SHORT THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

(PHD)

Histones and histone modifying proteins in a biophysical aspect

by

Szabolcs Hetey

Supervisor: Dr. Lóránt Székvölgyi

UNIVERSITY OF DEBRECEN

DOCTORAL SCHOOL OF MOLECULAR CELL AND IMMUNE

BIOLOGY

DEBRECEN, 2019

Histones and histone modifying proteins in a biophysical aspect

by **Szabolcs Hetey, MSc**

Supervisor: Lóránt Székvölgyi, PhD

Doctoral School of Molecular Cell and Immune Biology, University of Debrecen

Examination committee:

Head:	Prof. László Fésüs PhD, DSc, MHSC
Members:	Prof. Ferenc Erdődi, PhD, DSc
	Tamás Arányi, PhD

The examination took place at the Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen

Debrecen; at 11 AM, 24th of November, 2016

Reviewers of the thesis:

Reviewers:	Prof. Sándor Bíró, PhD, DSc
	Tamás Orbán, PhD

Defense Committee:

Head:	Prof. László Fésüs PhD, DSc, MHSC
Members:	Attila Bácsi, PhD, Dsc
	Tamás Arányi, PhD

The PhD Defense takes place at the Lecture Hall of Building A, Department of Internal Medicine, Faculty of Medicine, University of Debrecen

Debrecen; at 1 PM, 19th of March, 2019.

1. INTRODUCTION

1.1 Eukaryotic chromatin at the level of nucleosomes and histones

In 1974, prominent studies of Kornberg and Olins & D. Olins proposed that the repeating units of chromatin are histone oligomers forming flexible chain with chromatin fiber. It has been revealed that these units are nucleosomes consisting of two copies of histone H2A, H2B, H3 and H4 assembled into octamer and wrapped by 1.65 superhelical turns of 145-147 bp DNA. In the octamer, the core histones form an $(H3\text{-}H4)_2$ tetramer flanked by the H2A-H2B dimer. The repeating nucleosomes are connected with a 10-90 bp long linker DNA, and further assembled into higher-order structure of chromatin that is stabilized by linker histone H1.

The nucleosome core particle and the DNA organization around it was determined by X-ray crystallography in 1997. This study explained the interactions between the minor groove of the DNA double helix and the histone core in details. There are electrostatic interactions, hydrogen bonds and non-polar interactions between the histones and the DNA backbone. A typical characteristics of the nucleosome core particle is the exposed histone tails passing over and between the gyres of superhelix. These flexible N-terminals of the nucleosome have regulatory roles through histone post-translational modifications and play mechanistic role in the regulation of DNA dynamics and nucleosome conformation.

Based on their proteomic and genomic characteristics, histone proteins are classified into two groups: canonical histones and histone variants. Canonical histone proteins include the major core histones: H2A, H2B, H3 (H3.1 and H3.2 in human) and H4. Genes encoding these histones of animals are found clustered in repeat arrays and their transcription is tightly coupled to DNA replication. This intensive expression allows their rapid deposition behind the replication fork. In contrast, non-canonical histone variants are encoded by single or low copy genes, which are expressed in a replication-independent manner throughout the cell

cycle. Histone variants have roles in a range of processes, including DNA repair, meiotic recombination, chromosome segregation, transcription initiation and termination or sex chromosome condensation.

1.2 Histone post-translational modifications and the lysine methylations

Post-translational modifications of histones introduce meaningful variations into chromatin and provide a regulatory platform for controlling and/or fine-tuning many important DNA-templated processes, including gene transcription, the repair of DNA damage, DNA replication and meiotic processes. There are at least eight distinct types of modifications found on histones (acetylation, lysine methylation, arginine methylation, phosphorylation, ubiquitylation, sumoylation, ADP ribosylation, deimination and proline isomerization). The biological outcome of these histone modifications is manifested by direct physical modulation of nucleosomal structure or by providing a signalling platform to recruit downstream 'reader' or 'effector' proteins. The modifications have been found to be dynamic and specific enzymes are able to remove them. Preferences of these histones modifying enzymes may depend on the complex, where they are present. The proteins that operate with the enzyme complex may affect the residue to be modified, the degree of the methylation status (mono-, di-, or tri-) or influence the substrate preference (nucleosomal or free histones) of the enzyme complex.

Modification of lysine and arginine side chains by methyl groups plays an important role in many biological processes. In case of mammalian histones, evolutionarily conserved lysine methylation marks have been observed at multiple sites including K4, K9, K27, K36 and K79 of histone H3 and K20 of histone H4. The most studied modifications are the H3K4 and H3K27 trimethylations. H3K4me3 is localized at the 5′ end of the genes and considered as a universal hallmark of active transcription from yeasts to humans. Their occurence around transcriptional start sites highly correlates with transcriptional activation. H3K4me3 is

supposed to facilitate transcription by the recruitment of nucleosome remodelling proteins and histone modifying enzymes. Furthermore, H3K4me3 is associated with the initiation of meiotic recombination in yeasts and V(D)J recombination by binding the PHD domain of the RAG2 protein. In contrast to H3K4me3, H3K27me3 is frequently connected to gene silencing, particularly in the repression of unwanted differentiation pathways during lineage specification. H3K27me3 plays a role in maintaining gene transcriptional repression, and it is deposited preferentially at CpG-dense promoters by the lysine methyltransferase activity of the Polycomb repressive complex 2 (PRC2). H3K27me3 is also linked to silencing processes including homeotic gene silencing, genomic imprinting and X inactivation.

1.3 Trithorax and polycomb group proteins

Trithorax group (TrxG) and Polycomb group (PcG) genes were discovered in *Drosophila melanogaster* as repressors and activators of Hox genes, a set of transcription factors that specify cell identity along the anteroposterior axis of segmented animals. TrxG and PcG genes have opposite roles in homeotic gene regulation and many of them encode proteins bearing the 130- to 140-amino acid motif called SET domain. Most of the TrxG and PcG proteins are evolutionarily conserved in mammals and function within similar pathways to those of their *Drosophila* counterparts. Both of these complexes have long been linked to the occurrence of different forms of cancer.

The PcG proteins assemble into functionally distinct complexes that belong to two protein families: PRC1 and PRC2. Both complexes have catalytic activity inducing histone modifications. PRC1 with its E3 ligase activity monoubiquitinate the histone H2A at lysine 119 (H2AK119u1) whereas PRC2 is involved in the mono/di/trimethylation of H3K27 (H3K27me2/3) through its Ezh2 catalytic unit. The composition of PRC2 complex is dynamic, containing subunits responsible for the H3K27me3 mark and several accessory

regulatory subunits controlling the enzymatic activity and holoenzyme function. The core components of mammalian PRC2 complex include Suz12, Eed and Ezh2 (mutually exclusive with Ezh1 isotype).

In budding yeasts (*Saccharomyces cerevisiae*), all H3K4 methylation is established by a single Set1 complex called COMPASS (Complex Proteins Associated with Set1), which is an identified MLL homolog. The COMPASS protein complex consists of seven polypeptides ranging from 25 to 130 kDa. One of the most remarkable COMPASS subunit is the Spp1, which not only regulates the catalytic activity of the enzyme, but also interacts with the deposited mark, and mediates its biological effect (meiotic double-strand breaks formation). Spp1 interacting with H3K4me3 and Mer2 promotes the recruitment of potential meiotic double-strand breaks (DSBs) to the chromosomal axis allowing their subsequent cleavage by transesterase Spo11.

1.4 Implications of histone proteins and PRC2 in brain tumor progression

Approximately 70 %-80 % of pediatric gliomas are characterized by the same histone mutations manifested in the variant H3.3 or H3.1. The mutations cause amino acid substitutions at two critical residues within the H3 histone tail (K27M and G34R/G34V) that are specific to glioblastoma multiforme (GBM) and highly prevalent in children and young adults. They are both spatially and temporally restricted. K27M tumors are primarily found in the midline locations (spinal cord, thalamus, pons, brainstem) and G34R/V tumors are mostly located in the cerebral hemispheres (frontal, parietal, occipital, and temporal lobes). K27M plays a dominant role in blocking the accumulation of repressive H3K27 methyl marks. The explanation of this inhibition is that methionine binds and stabilizes PRC2 thus it prevents the deposition of methyl groups. The binding and interaction between H3K27M and Ezh2

catalytic subunit of PRC2 is proved by immunoprecipitation and binding partner analysis of photoreactive K27M containing peptides.

1.5 Implications of MLL complex in brain tumor progression

Inactivating mutations in MLL2 and MLL3 have been identified in 16% of pediatric medulloblastoma patients. The majority of these mutations (nonsense mutations, out-of-frame indels, or splice site mutations) results in protein products lacking the key methyltransferase domain. *Parsons et al.* have demonstrated that the MLL pathways are important to medulloblastomas and that Mll2 and Mll3 play a tumor suppressor role. Additionally, there is an identified Mll-Homeobox axis that significantly contributes to tumorigenic behavior of glioblastoma cancer stem cells. Mll directly activates the Homeobox gene HOXA10 that activates a downstream Homeobox network and other genes previously characterized by playing a role in tumorigenesis. Expression level of MLL1 has also been found higher in glioma stem cells than matched non-stem tumor cells.

2. AIMS

My major goals were the followings:

• *In vitro* reconstitution of H3K27M containing nucleosomes and study the methionine mutation and Ezh2 effect on nucleosome conformation using Förster resonance energy transfer-based approaches.

• Analysis of H3.3K27M nuclear distribution in living cells and describe the salt-dependent elution profile of chromatin incorporated H3.3K27M using laser-scanning cytometry (LSC).

• Test the viability of *Saccharomyces cerevisiae* strains expressing H3K27M histones.

- Study the diffusional properties of H3.3 wild type and H3.3K27M histones using FCS and FRAP techniques, and describe the transcriptional relationships of mutant histones and Ezh2 protein.
- Diffusional characterization of Spp1 and Set1 COMPASS subunits with FCS and FRAP techniques.
- Quantitative analysis of Spp1 residence dynamics and turnover rate in meiosis using competition ChIP-Seq.

3. MATERIALS AND METHODS

3.1. *In vitro* nucleosome reconstitution

Fluorescently labeled Widom-601 positioning PCR amplicon (with length of 170 bp) and recombinant *Xenopus laevis* histones were reconstituted into nucleosomes containing histone H3 wild type or H3K27M. Histone octamers were mixed with positioning DNA fragment in 2 M NaCl-TE buffer and reconstituted into nucleosomes with slow dialysis down to 5 mM NaCl-TE. In our energy transfer experiments the positioning DNA fragment was labelled at thymine nucleobases via C6-linker at -53 (Alexa 594) and +41 positions (Alexa 488) with respect to the dyad axis. In case of Ezh2/nucleosomes, the nucleosomes reconstituted by end-labeled DNA fragments were also analysed. Where nEeded, positioning DNA to histone octamer ratio was optimized between 1:1.5 and 1:1.8 molar ratios in order to avoid aggregation. Nucleosome concentration following reconstitution was determined using a Cary 4E spectrophotometer. Nucleosome quality was validated on native PAGE.

3.2. Bulk FRET (μpsFRET) analysis using microplate scanning

In order to determine the average proximity ratios at the level of nucleosome populations Typhoon 9400 (GE Healthcare) variable mode fluorescence scanner was used. Samples in experimental buffer with different NaCl concentrations were incubated in 384-well microplates prior to the measurement (μpsFRET). The final concentration of labeled nucleosomes was approximately 1 nM. Fluorescent emission was detected in three spectral channels: donor channel (excitation at 488 nm, detection at 500-540 nm); acceptor channel (excitation at 532 nm, detection at 595-625 nm); energy transfer channel (excitation at 488 nm, detection at 595-625 nm). Detection voltages of the two photomultiplier tubes (PMT) were set between 600 V and 700 V. Proximity ratios were calculated based on intensity values

of each acquired image. As with spFRET measurements, correction factors were determined prior to each measurement.

3.3. Single-pair FRET experiments

For spFRET measurements intact nucleosomes were freshly diluted in a 0.02 µm-filtered experimental buffer with different concentrations of NaCl in 10 mM Tris-HCl, 0.1 mM EDTA, pH 7.5, supplemented with 0.01 % Nonidet P40 and 1 mM ascorbic acid in order to avoid photobleaching. Nucleosomes were placed into 384-well microplates (SensoPlate Plus, Greiner Bio-One) previously passivated with Sigmacote® (Sigma-Aldrich). The final concentration of labeled nucleosomes was approximately 50 pM supplemented with 250 pM of unlabeled nucleosomes to prevent dissociations caused by low nucleosome concentration. Experiments were carried out using a specific confocal system illuminated continuously with a 491 nm laser (Cobolt) for excitation. Prior to burst analysis the nucleosomes were incubated with or without the Ezh2 complex (Active Motif, containing Ezh2, Eed1, Suz12 proteins) for 60 minutes at room temperature in experimental buffer. After donor excitation fluorescent emission was separated into two detection windows for donor (520-560 nm) and acceptor (>600 nm). Emitted photons were collected by two avalanche photodiodes (APD, Perkin Elmer Optoelectronics). Single molecule bursts were collected by TimeHarp2000 (PicoQuant), and analyzed by the software Frettchen where one burst was defined as a group of at least 50 photons with a mutual separation of less than 120 µs. Proximity ratio (P) histograms were plotted based on the selected single events and analyzed by IGOR Pro software (WaveMetrics). P is related to the energy transfer efficiency depending on the distance between the fluorophores. In both spFRET and bulk systems the correction factors (background and cross talk) needed for P determination were defined in independent measurements.

3.4. *In situ* salt elution assay using laser scanning cytometry

Embedding of live HeLa cells constitutively expressing pEGFP-N1-H3.3 or pEGFP-N1-H3.3K27M into 8-well chambers (Ibidi, Martinsried, Germany) and their quantitative microscopy after salt treatment were based on the protocol of *Imre et al.* The salt elution range was 0 mM-1400 mM NaCl in histone eviction assay and 0-400 mM NaCl in EZH2 assay. Intensity of the remaining EGFP-tagged histones or immunolabeled Ezh2 after treatment with increasing concentration of NaCl was quantitatively analyzed by LSC.

3.5. Growth and viability of control and H3K27M mutant *Saccharomyces cerevisiae*

Control (H3 wild type), H3K27M, and K27R/Q mutant budding yeast cells were established in the SK1 genomic background by plasmid shuffling technique. Cells were grown to an optical density (OD600) of 0.7-1.0 and ten-fold serial dilutions were spotted onto fresh plates of various metabolite and drug composition. Colony sizes were checked after two days of growth at 30 ºC. Kinetic growth measurements were performed in various liquid cultures, and optical densities were tracked for 24 hours. OD600 values higher than 1.5 were extrapolated from diluted cultures, corrected with the dilution factor. Sporulation was induced in diploid plasmid shuffle strains using 1 % (w/v) potassium acetate as non-fermentable carbon source. After 24 hours of sporulation at 30 ºC, ascus sacs were digested by Zymolyase 20T (MP Biomedicals) at 30 ºC for 10 minutes, and tetrads were dissected using a micromanipulator. Fertility (spore viability) was determined by counting the number of viable spores after two days of growth on YPD plates.

3.6. Cell culture and transfection

HeLa cells were grown in RPMI-1640 (Sigma, R5886) supplemented with 10% (v/v) fetal calf serum, 2 mM glutamine, penicillin and streptomycin, in 5% CO_2 humidified chamber. H3.3K27M point-mutation was introduced into a pEGFP-N1-H3.3 and pmCherry-N1-H3.3 plasmid, using the quick change mutagenesis technique. EZH2 was PCR amplified from the NM_004456 (EZH2) Human cDNA ORF Clone (OriGene), and the amplicon was cloned into a pEGFP-N1 and pmCherry-N1 plasmid, respectively. Transient and stable transfections of pEGFP-N1-H3.3 / H3K27M, pmCherry-N1-H3.3 / H3K27M, alone or in combination with pEGFP-N1-EZH2 or pmCherry-N1-EZH2, were carried out by Lipofectamine 2000 (Invitrogen) or polyethylenimine PEI-B according to the manufacturer's recommendations. Where indicated, cells were pre-treated for 60 minutes before each measurement with actinomycin D (act D, 5 µg/ml), cycloheximide (CHX, 20 µg/ml) and flavopiridol (flav, 100 nM).

3.7. Confocal laser scanning microscopic (CLSM) analysis

HeLa cells were fixed with 1% formaldehyde prior to each measurement. CLSM images were acquired using Olympus FluoView 1000 confocal microscope supplied with a 60x oil immersion objective (NA 1.35). Excitation and emission filters were as follows: EGFP, 488 nm excitation, 500-540 nm detection; mCherry, 543 nm excitation, 600-680 nm detection. Ten optical slices having a thickness of 0.7-1.1 µm were collected from each nucleus, applying the Kalman filter mode to reduce noise and alternative excitation to exclude crosstalk. Colocalization between EZH2-mCherry and H3.3-EGFP or H3.3K27M-EGFP was computed by the JACoP plugin in ImageJ.

3.8. Fluorescence recovery after photobleaching

FRAP measurements were performed in HeLa cells using Olympus FluoView 1000 confocal microscope, based on an inverted IX-81 stand with an UPlanAPO 60x (NA 1.2) water immersion objective. EGFP was excited by the 488 nm Argon-ion laser line and fluorescence was detected through a 500-550 nm band-pass filter. In the histone FRAP measurements, H3.3-EGFP or H3.3K27M-EGFP HeLa cells were randomly selected and five pre-bleach images were taken (256×256-pixel area, 10x zoom, ~9 μW laser power at the objective) which was followed by a 500 ms bleach period of 100% laser power (900 μW). Rectangular areas were selected as bleach ROIs. In the first 90 minutes, images were acquired every ten minutes and then every 30 minutes, up to 420 minutes (7 hours). Transcription was inhibited by flavopiridol (100 nM) or actinomycin D (5 μg/ml), while translation was repressed by cycloheximide (20 μg/ml). Drugs were added 60 minutes before the onset of measurements. FRAP experiments on EZH2-EGFP was performed under similar conditions to those on histones (photobleaching parameters and confocal setup were the same), except that the fast-recovering EZH2-EGFP fluorescent signal was tracked for only ten seconds.

FRAP measurements were performed in sporulating yeast cells (between 0-6 hrs in SPM) using an Olympus FluoView 1000 confocal microscope, based on an inverted IX-81 stand with an UPlanAPO 60x (NA 1.2) water immersion objective. Samples were taken every hour from standard liquid sporulation cultures and measurements were carried out on microscope slides covered with 1 % potassium acetate pad. GFP was excited by the 488 nm Argon-ion laser line and fluorescence was detected through a 500-550 nm band-pass filter. Cells expressing the Set1-GFP or Spp1-GFP proteins were randomly selected after $CuSO_4$ induction (100 μM) and five pre-bleach images were taken (256 × 256-pixel area, 15x zoom, ~9 μW laser power at the objective) followed by a 500 ms bleach period of 100 % laser power (900 μW). Images were taken every second up to 1 minute.

3.9. Fluorescence correlation spectroscopy

HeLa cells transfected by H3.3-EGFP or H3.3K27M-EGFP were analysed by a special fluorescence fluctuation microscope (FFM) that combines fluorescence correlation spectroscopy (FCS) and confocal laser scanning microscopy (CLSM). The FFM consists of an FCS module combined with inverted IX-70 microscope (Olympus, Hamburg, Germany), supplied with an UplanApo / IR 606 water immersion objective lens (NA 1.2) and a 5% CO_2 humidified, constant temperature chamber. Fluorescence excitation of EGFP was elicited by a Cobolt laser (at 491 nm, 5-15 μW outgoing power, 1 μW excitation power at the objective). EGFP emission was detected through a 515-545 nm band-pass filter using an avalanche photodiode (APD). Measurements were conducted at 37 °C.

HeLa cells transfected with EZH2-EGFP were analysed by Olympus FluoView 1000 confocal microscope (described in the FRAP experiments). Autocorrelation curves were calculated by an ALV-5000E correlation card at three randomly selected points of the nucleus of each cell, with 10 × 8 s runs. All measurements were performed at room temperature (22 °C).

Set1 and Spp1 FCS measurements were all performed by Olympus FluoView 1000 confocal microscope. Sporulating yeast cells were taken every hour from liquid sporulation cultures and FCS was performed on microscope slides covered with 1 % potassium acetate pad. Autocorrelation curves were calculated by an ALV-5000E correlation card at three randomly selected points of each nuclei, with 10 × 8 s runs. In case of rich medium, cells were grown in YPD medium up to A260 = 1.0 and were measured on pre-coated microscope slide covered with YPD agar. Set1 / Spp1 expression was induced using 100 μM $CuSO_4$.

3.10. Spp1 c-ChIP experiments in *Saccharomyces cerevisiae*

50 mL of meiotic yeast cells (4×10^7 cells/ml) were collected at the indicated timepoints and crosslinked with 1 % formaldehyde for 20 min at room temperature. Formaldehyde was quenched with 125 mM glycine for 5 min at room temperature, and cells were washed three times with ice-cold 1x TBS, pH 7.5 (20 mM Tris HCl at pH 7.5, 150 mM NaCl). Cells were resuspended in 500 µL of lysis buffer (50 mM Hepes KOH at pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate, 1 tablet of complete inhibitor cocktail (Roche) in 50 mL solution) and lysed with acid-washed glass beads for 10 min in a FastPrep bead beater machine. Chromatin samples were fragmented to an average size of 300 bp by sonication (Bioruptor, Diagenode). In order to obtain whole-cell extract (WCE), a 50 µL pre-IP sample was removed and centrifuged at full spEed for 10 sec to separate cell debris (supernatant = WCE). The rest of the samples was also centrifuged at 12,000 rpm (4 °C) for 20 sec to separate cell debris. IP was performed by adding the 450-µl extract to a pellet of magnetic protein G dynabeads (Dynal), corresponding to 50 µl or 2×10^7 beads, which were preincubated with 9E11 (monoclonal mouse anti-myc, ab56, Abcam) or anti-GFP (polyclonal rabbit, ab290, Abcam) antibodies overnight at 4 °C. IP samples were washed twice with lysis buffer, twice with lysis buffer supplemented with 360 mM NaCl, twice with washing buffer (10 mM Tris HCl at pH 8.0, 250 mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate, 1 mM EDTA), and finally once with 1x TE at pH 7.5, using the magnetic device supplied by Dynal. Cross-linking was reversed by heating in TE-1% SDS overnight at 65 °C. Afterwards, proteins were digested with proteinase K (12 µl of 20 mg/ml stock) for 3 h at 65 °C. Nucleic acids were purified using a PCR clean up kit and RNA digestion (10 µg RNase) was carried out for 1 h at 37 °C. DNA was finally resuspended in 50 µl nuclease-free dH_2O.

Sequencing libraries were prepared according to the Illumina's TruSeq ChIP Sample Preparation protocol. The libraries were sequenced using 50 bp single end reads with Illumina

HiScan SQ (Genomic Medicine and Bioinformatics Core Facility of the University of Debrecen); or with Illumina HiSeq 2500 (EMBL Genomics Core Facility, Heidelberg, Germany).

3.11. Turnover rate estimation from c-ChIP data

Average coverage (i.e. the occupancy) of the Spp1 binding sites were calculated using both the GFP- and MYC competition ChIP-seq data for each timepoint separately. Next, GFP/MYC occupancy ratios were calculated and the same exponential model was fitted according to *Deal et al.* After fitting the model, the standard error of the estimates was calculated and a t-test was performed to evaluate the goodness of fit of the model. In total, 977 binding sites could be described with the model (p-value < 0.05).

4. RESULTS

4.1. Characterization of the structure of K27M nucleosomes by Förster resonance energy transfer (FRET)

We measured the equilibrium stability of wild type nucleosomes and nucleosomes with histone H3K27M by microplate-scanning FRET (µpsFRET). Energy transfer proximity ratios (P) were computed at gradually increasing salt concentrations that elicited nucleosome disassembly in well-controlled way. The comparison of the dissociation kinetics of wild type and mutant samples showed a small reduction in salt-dependent stability of K27M nucleosomes; however, the difference was not statistically significant. Next, supplementation of the reconstituted nucleosomes with a recombinant Ezh2-complex (contains Ezh2, Eed and Suz12 proteins) resulted in no detectable changes regarding nucleosome stability.

We repeated the salt dissociation measurements using spFRET that allows us to track potential subpopulations of nucleosomes. At the level of individual nucleosomes, the

disassembly process did not reveal a significant difference or structural heterogeneity between wild type and H3K27M nucleosomes.

4.2. *In situ* salt elution assay of H3.3K27M and Ezh2

With this current approach, we were able to compare the stability of chromatin-incorporated nucleosomes containing the H3.3K27M histones to the incorporated wild type nucleosomes. Moreover, in a modified salt range we observed the elution profile of EZH2 in HeLa cells transfected with H3.3 wild type or H3.3K27M transgene.

Based on three independent experiments we concluded that K27M substitution did not alter the incorporation property of the H3.3 histone. Similarly, the EZH2 elution profile was not affected by the K27M mutation either. Our assay implies that this phenomenon does not influence the chromatin incorporation of H3.3K27M histones.

4.3. The effect of K27M mutation histones on the proliferative capacity and stress tolerance of H3K27M-expressing budding yeast cells

We introduced the K27M point mutation into the HHT2 gene (coded by a pCEN-ARS-hht2K27M-HHF2-TRP1 plasmid), and transformed it into an hht1Δ recipient strain expressing histone H3 (Hht2) from a URA3 plasmid (pCEN-ARS-HHT2-HHF2-URA3). Wild type ura+ and mutant trp+ plasmids were exchanged by plasmid shuffling, and trp+/ura+ colonies were selected. Growth and survival of the test strains were monitored in kinetic and end-point growth assays that were performed under various metabolic and stress conditions. The process of sporulation and meiosis were tested in diploid K27M strains to assess differentiation capacity and fertility. We found similar growth characteristics and stress tolerance in all tested conditions, suggesting that the basic metabolic processes, repair

mechanisms and meiotic differentiation of yeast cells were not perturbed by the homozygous expression of histone H3K27M.

4.4. Microscopic analysis of the nuclear distribution of wild type and H3K27M nucleosomes in relation to Ezh2

We tagged EZH2 with mCherry and co-expressed the fusion protein with H3.3-EGFP or H3.3K27M-EGFP in live HeLa cells. We used confocal laser scanning microscopy (CLSM) to analyze the subcellular distribution of EZH2-mCherry in relation to H3.3-EGFP and H3.3K27M-EGFP. Based on the Manders' colocalization coefficients about half of the Ezh2 pool overlapped with histone H3.3 or H3.3K27M, while the other half occupied distinct nuclear compartments. K27M mutation or transcriptional inhibition elicited by actinomycin D and flavopiridol treatments did not change significantly the pattern of colocalization and genomic distribution of Ezh2.

4.5. Kinetics study of Ezh2 and H3.3K27M at various spatial and temporal resolutions

4.5.1. Measuring the kinetics of H3.3K27M nucleosomes and Ezh2 in live cells using FRAP

We performed kinetics measurements to assess the diffusional properties of Ezh2, H3.3 and H3.3K27M histones at various spatial and temporal resolutions. We studied HeLa cells stably expressing histone H3.3-EGFP or H3.3K27M-EGFP and EZH2-EGFP in the presence of wild type or K27M mutant histones by FRAP analysis. Transcription dependence of H3.3/H3.3K27M-EGFP and EZH2-EGFP mobility was assessed by comparing control and transcriptionally stressed HeLa cells.

The recovery of H3.3/H3.3K27M-EGFP fluorescence did not reach the initial pre-bleach value since most histone molecules remained in the slowly exchanging fraction. Actinomycin

D increased the immobile (unrecovered) fraction of H3.3 and H3.3K27M (from 55 % to 85 %), suggesting strong and direct chromatin binding depending on the process of transcription. Interestingly, the pTEFb (transcriptional elongation factor) inhibitor flavopiridol did not change significantly the repopulation rate of H3.3/H3.3K27M histones implicating that nascent RNA elongation was nonessential for the chromatin binding of H3.3. The same general pattern was observed after cycloheximide treatment.

Contrary to more static histone molecules, flavopiridol and actinomycin D treatments induced the formation of a stable immobile fraction of Ezh2 (15 % and 35 %, respectively).

4.5.2. Measuring the kinetics of H3.3K27M nucleosomes and Ezh2 in live cells using FCS

The average ratio of fast components, corresponding to the fraction of molecules freely diffusing across the nucleoplasm, was 61 % (SD=19) for H3.3-EGFP and 56 % (SD=20) for H3.3K27M-EGFP, respectively. The proportion of mobile pool increased by 10 % after actinomycin D treatment ($p<0.05$) in case of H3.3K27M. The average diffusion coefficient (D) was the same for wild type and K27M mutant histones (31±19 $\mu m^2/s$ and 29±16 $\mu m^2/s$, measured at 37 ºC) and it was not affected by transcription inhibition (act D) or impaired protein synthesis (CHX). The average fraction of the fast component was approximately 60 % for EZH2-EGFP, which slightly increased in the presence of H3.3K27M-mCherry ($p<0.05$) and after transcription inhibition by flavopiridol ($p<0.05$). The latter pool of EZH2-EGFP also had a higher mobility, reflected by the increased diffusion coefficients of flavopiridol inhibited cells ($p<0.05$). There was no significant difference between the real and apparent masses of H3.3/H3.3K27M-EGFP molecules (55.9 kDa / 49.26 kDa vs. 42.26 kDa). However, EZH2-EGFP gave ~10-fold larger average molecular mass than the real molecular mass of the fusion protein (1.221 kDa / 1.900 kDa vs. 102.36 kDa).

4.6. Kinetics assay of Spp1 and Set1 COMPASS subunits using FRAP and FCS methods

We carried out kinetics analysis on two COMPASS subunits, Spp1 and Set1, involving FRAP and FCS techniques. Using FRAP, whole live-cell nuclei were bleached and fluorescence recovery was examined in the first five hours of meiosis in SPM and in the exponential phase in YPD rich medium. GFP-Spp1 and GFP-Set1 recovered within less than 50 seconds. About half of the FRAP signal did not return after the initial bleach pulse, indicating that ~50 % of Spp1 and Set1 remained tightly bound to chromatin representing the immobile fraction.

Next, the distribution of fast and slow components did not show any difference between Spp1 and Set1, however, the average diffusion coefficient (D) of Spp1 was significantly slower compared to Set1. The average apparent molecular mass of GFP-Set1 was equal to the expected molecular mass of COMPASS (379 kDa) while GFP-Spp1 gave an approximately 43-fold higher molecular mass (1764 kDa) compared to the real molecular mass of the fusion protein.

4.7. Quantitative analysis of Spp1 chromatin binding by competition ChIP

We differentially tagged a constitutive and an inducible isoform of Spp1 with 9 x myc and GFP epitopes, respectively, and turned on the expression of the inducible allele (driven by a pCUP1 promoter similarly to FCS and FRAP experiments) with copper addition during the meiotic time course. By sampling dense meiotic timepoints (4.5, 5.0, 5.5, 6.0, 6.5 hours in SPM), both Spp1 isoforms were immunoprecipitated using anti-myc and anti-GFP antibodies and binding sites were determined similarly to conventional ChIP-Seq pipelines. Then, Spp1 turnover rates were calculated by determining the ratio of GFP (new Spp1) and myc (old Spp1) ChIP signal and fitting the data with an exponential model. Our analysis revealed that Spp1-binding sites exhibited different replacement dynamics compared to common (Spp1 & Bre2) binding sites. Spp1-only sites were significantly slower than canonical sites over gene

bodies, TTS, RPG/Ribi genes, and Mer2/Red1 sites. If Spp1 binding sites were grouped by their kinetic behaviour, disappearing and appearing Spp1 sites would sharply stood apart based on the distribution of turnover rates and occupancies like disappearing Spp1 sites tended to have higher turnover rates and higher occupancies compared to appearing sites. Appearing Spp1 sites were strongly associated with high Mer2 occupancy, low H3K4me3 and Bre2 occupancies and low turnover rates whereas disappearing Spp1 sites showed low Mer2 occupancy, high H3K4me3 and Bre2 occupancies and high turnover rates.

5. DISCUSSION

Performed spFRET and μpsFRET experiments, we found that dissociation mechanism of nucleosomes containing H3K27M histones is the same as dissociation of wild type nucleosomes. These results suggest that H3K27M nucleosomes maintain their wild type molecular architecture and stability, indicating that this single amino acid substitution of the N-terminal H3 tail domain does not cause detectable rearrangements in the structure of nucleosome core particle. Furthermore, the salt elution profile of incorporated H3.3K27M histones serves a good evidence that these mutant nucleosomes show similar structural features in the native chromatin as their wild type counterparts. We assumed that the collaboration between Ezh2 and H3K27M nucleosomes results in structural alterations. Therefore, we repeated the μpsFRET experiments supplemented the reconstituted nucleosomes with Ezh2 complex. Although we probed various conditions including the optimalization of Ezh2/nucleosome molecular ratio, incubation time and fluorophore position, we could not find any nucleosome alteration caused by the Ezh2 complex. We also quantified the remaining Ezh2 in salt treated HeLa nuclei in vicinity of H3.3K27M-EGFP and concluded that the mutant histone does not affect the elution profile of Ezh2 in the range of 0-400 mM NaCl.

To examine the impact of H3K27M on cell viability and chromatin-templated processes we engineered H3K27 mutant yeast strains and carried out a genetic screen with a growth assay under various stress and metabolic conditions. The aim of this assay was to gain a mechanistic insight how H3K27M mutation may promote genomic instability in glioma. The experiments demonstrated that *Saccharomyces cerevisiae* cells with the mutant form of histone H3 did not show any growth defect and stress sensitivity neither in vegetative nor in meiotic conditions. As the PRC2 complex does not exist in budding yeasts, these results might suggest that the involvement of PRC2 pathway is a crucial point of the K27M phenotype and the mutation alone cannot account for remarkable alteration in nuclear processes.

Our further aim regarding H3K27M was to accomplish an *in vivo* functional analysis in human cells, in which we image the nuclear distribution of mutant histones in relation to wild type histones and Ezh2, and probe the diffusional properties of these proteins at various spatial and temporal resolutions using FRAP and FCS. The nuclear distribution of Ezh2 and H3.3 / H3.3K27M based on CLSM analysis did not show convincing differences in the colocalization patterns neither in non-treated nor in transcription inhibited condition. Manders' coefficient was a good indicator that Ezh2 was not changed much by the K27M mutation or due to transcription inhibition. However, Ezh2 distribution (e.g. Ezh2 accumulation near K27M nucleosomes) seemed to be different likely due to Ezh2-K27M methionine binding.

In live human cells, H3.3K27M histones followed similar diffusion kinetics to their wild type counterparts and fully recapitulated the slow kinetics and tight chromosome binding of canonical (replication-dependent) H3 molecules. Furthermore, we found that act D treatment (but not flavopiridol) significantly increased the immobile fraction of H3.3 histones suggesting transcription-dependent H3.3 mobile component. By FRAP and FCS we found a remarkable differential recovery of Ezh2 in response to transcriptional stress that was

accompanied by a significantly increased immobile fraction and faster diffusion rate of the mobile fraction. The differential recovery of Ezh2 was rather dependent on transcription than K27M mutation status. Except the significant difference between Ezh2 fast component ratio of H3.3 and H3.3K27M co-transfected cells, the transcription inhibition perturbed the Ezh2 kinetics in each case. An explanation for the faster Ezh2 diffusional rate observed by FCS might be that the inhibited transcription allows a more mobile, possibly a PRC2 independent function for Ezh2. The estimation of the apparent molecular weight of Ezh2 showed a slight, but not significant, increased molecular weight in H3.3K27M-EGFP transfected cells. This increase might be a result of the "gluelike" property of methionine resulting in non-specific bindings between the PRC2 subunit and nucleosomes bearing H3.3K27M.

Moreover, we were particularly interested how the Spp1 protein diffusional behaviour differs from the Set1 catalytic subunit in meiotic conditions. The results raised the possibility, that Spp1 has a COMPASS-independent function on chromosome axial sites and contributes to chromatin changes preparing potential recombination initiation sites for meiotic DNA break formation. On the other hand, the diffusional properties of the COMPASS protein complex, similarly to mobility of TrxG and PcG proteins was unexplored so far. We observed a similar recovery kinetics in both Spp1 and Set1 using FRAP. The intensity curves reached the plateau phase within 50 seconds during the measurements without significant differences. The estimated, approximately 50 % immobile fraction indicates that the half of the whole fraction is tightly bound either to the chromatin or to another high molecular weight complex. These experiments were repeated in YPD medium and the profile showed high similarity. A modest difference between Spp1 and Set1 recovery in this condition was detected.

Moreover, in order to characterize these two subunits, we applied FCS in the first six hours of sporulation. Similar amount of the fast component could be measured, but distinct diffusion coefficient between Spp1 and Set1 was determined. Furthermore, to better understand the

mechanism behind slower Spp1 mobility, we estimated the apparent molecular weight of GFP-Spp1 and GFP-Set1 using a similar method as in our Ezh2 and H3.3 experiments. Based on our results, the apparent molecular weight of GFP-Spp1 was increased and resulted in a remarkable difference between the expected and observed molecular weight of Spp1. These data supported the previously proposed model in which the interaction between Spp1 and Mer2 brings potential meiotic DSB sites to the axis, thereby allowing their cleavage by Spo11 at axis-proximal regions that are depleted in nucleosomes.

Genome-wide competition ChIP-Seq was an excellent approach to investigate Spp1-DNA dynamics and to complete our biophysical results with binding site turnover across the entire genome. Differential turnover of COMPASS-associated and COMPASS-independent Spp1 peaks prove the presence of two kinetic Spp1 pools that distribute differentially between the Set1 complex and meiotic DSB proteins. We showed that association of Spp1 with Mer2 axial sites reduces *bona fide* turnover rates of Spp1 upon chromatin binding. This increased residence time is also supported by the reduced nuclear dynamics of Spp1 in live meiotic cells. In conclusion, the binding kinetics measured by c-ChIP was in accordance with the FCS estimates and with the hypothesized loop-axis model. These data delineate the remarkable multi-functionality of Spp1 and help us to better understand how Spp1 switch function during meiosis and selects potential meiotic DSB sites independently from COMPASS.

6. SUMMARY

The molecular biological consequences of H3K27M driver mutation and how the substitution affects pathways contributing tumor initiation and progression are not well understood yet. In the present thesis, we investigated the H3K27M mutation using *in vivo* and *in vitro* experimental systems and determined its biophysical characteristics at chromatin, nucleosome and nuclear diffusion levels. Furthermore, we were interested in the examination of Ezh2

protein, which is the catalytic subunit of the PRC2 polycomb group protein and the interaction of PRC2 with the K27M residue alters the epigenetic state of the tumor cells. Finally, we examined the diffusional properties of COMPASS protein complex which belongs to the trithorax group proteins and its mammalian homolog is involved in hematological and brain malignancies as well. We observed that Spp1 switches to COMPASS-independent function in meiosis and it helps to prepare potential recombination initiation sites.

Our experimental results supported the biophysical characterization of H3K27M oncohistone and gave insights into the diffusion properties of Ezh2 methyltransferase therefore it helps us to better understand the pathogenesis of pediatric gliomas. We found that H3.3K27M histones follow similar diffusion kinetics as histone H3.3 and the mutant nucleosomes maintain the wild type molecular architecture and chromatin incorporation profile. Additionally, the K27M mutation does not manifest in growth defect in budding yeast strains bearing H3K27M histone pool. We demonstrated for the first time the *in vivo* nuclear mobility of H3.3 histone variant in various time and spatial resolution and estimated the diffusional parameters of Ezh2 observed transcription dependent mobility. The experiments on COMPASS subunits revealed differences between Set1 and Spp1 diffusion. These results support our hypothesis that Spp1 might have a COMPASS-independent regulatory role and collaborates with high molecular weight macromolecular complexes such as chromosome axial sites. We revealed that dynamic turnover of Spp1 is important in the establishment of transient chromatin changes during meiosis and prepare potential recombination initiation sites for DNA break formation.

7. KEYWORDS

histone, histone modification, H3K27M, glioblastoma, Ezh2, recombination, COMPASS, Spp1, double-strand break.

8. PUBLICATIONS

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/10/2019.PL
Subject: PhD Publikációs Lista

Candidate: Szabolcs Hetey
Neptun ID: CNDDBL
Doctoral School: Doctoral School of Molecular Cellular and Immune Biology

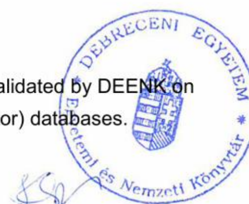
List of publications related to the dissertation

1. Karányi, Z., Halász, L., Acquaviva, L., Jonás, D., **Hetey, S.**, Boros-Oláh, B., Peng, F., Chen, D., Klein, F., Géli, V., Székvölgyi, L.: Nuclear dynamics of the Set1C subunit Spp1 prepares meiotic recombination sites for break formation.
   J. Cell Biol. 217 (10), 3398-3415, 2018.
   DOI: http://dx.doi.org/10.1083/jcb.201712122
   IF: 8.784 (2017)

2. **Hetey, S.**, Boros-Oláh, B., Kuik-Rózsa, T., Li, Q., Karányi, Z., Szabó, Z., Roszik, J., Szalóki, N., Vámosi, G., Tóth, K. Á., Székvölgyi, L.: Biophysical characterization of histone H3.3 K27M point mutation.
   Biochem. Biophys. Res. Commun. 490 (3), 868-875, 2017.
   DOI: http://dx.doi.org/10.1016/j.bbrc.2017.06.133
   IF: 2.559

Total IF of journals (all publications): 11,343
Total IF of journals (publications related to the dissertation): 11,343

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of Web of Science, Scopus and Journal Citation Report (Impact Factor) databases.



10 January, 2019

9. ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Dr. Lóránt Székvölgyi, for the excellent training, continuous support, motivation, immense knowledge and for scientific discussions during my Ph.D. I am also grateful to Dr. Katalin Tóth (Division Biophysics of Macromolecules, DKFZ, Heidelberg, Germany) who was my tutor in Heidelberg and introduced me into FRET and FCS experiments. Their guidance helped me through all aspects of research, writing this thesis and the related papers.

I am grateful to Professor Dr. József Tőzsér, head of the Department of Biochemistry and Molecular Biology, and my advisory committee: Dr. Laszló Fesüs and Dr. Attila Bácsi.

Special thanks to Professor Dr. Gábor Szabó, Dr. György Vámosi, Dr. László Imre and to our colleagues from Cell biology research group of Department of Biophysics and Cell Biology for the valuable and constructive suggestions.

My gratitude also goes to all the past and current members of Genome architecture and recombination research group for their help and scientific discussions, specially to Zsolt Karányi and to my lab mates, László Halász, Dr. Éva Sipos and Dr. Ágnes Mosolygó-Lukács. I should also acknowledge the endless help I received from technicians: Adél Nagy Vezendiné and Ibolya Fürtös.

Last but not least, I would like to express my deepest gratitude to my beloved family, who have supported me throughout the entire process.