

Aus dem Adolf-Butenandt-Institut
Lehrstuhl: Molekularbiologie
der Ludwig-Maximilians-Universität München,
Vorstand Prof. Dr. Peter Becker

Mass Spectrometric Analysis of Global Histone Modification Profiles During *Xenopus laevis* Development



Dissertation zum Erwerb des Doktorgrades der Medizin an der Medizinischen
Fakultät der Ludwig-Maximilians-Universität zu München

vorgelegt von
Tobias Daniel Schneider
aus Völklingen

München 2012

**Mit Genehmigung der Medizinischen Fakultät
der Universität München**

Berichterstatter: Prof. Dr. Ralph A.W. Rupp

Mitberichterstatter: Prof. Dr. Dr.h.c. Hans A. Kretschmar
Prof. Dr. Stefan Kääb
Prof. Dr. Gunnar Schotta

Mitbetreuung durch den
promovierten Mitarbeiter: Prof. Dr. Axel Imhof

Dekan: Prof. Dr.med. Dr. h.c. M. Reiser, FACR, FRCR

Tag der mündlichen Prüfung: 03. Mai 2012

Ideale sind wie Sterne.

*Wir erreichen Sie niemals,
aber wie die Seefahrer auf dem Meer,
richten wir unseren Kurs nach ihnen.*

Carl Schurz

This work gave rise to the following publication: “Schneider TD, Arteaga-Salas JM, Mentele E, David R, Nicetto D, Imhof A, Rupp RAW, Stage-Specific Histone Modification Profiles Reveal Global Transitions in the *Xenopus* Embryonic Epigenome, PLoS ONE 6 (2011)”.

Table of Contents

1 Summary	1
2 Zusammenfassung	2
3 Introduction	4
3.1 Development of Vertebrates	4
3.1.1 <i>Morphological Differentiation</i>	4
3.1.2 <i>Differences concerning the Development of Amphibians and Mammals</i>	5
3.2 Regulatory Mechanisms of Early Development	6
3.2.1 <i>Differential Gene Expression in Pluripotent and Somatic Cells</i>	6
3.2.2 <i>Inductive Events and Transcription Factor Cascades</i>	7
3.2.3 <i>Establishment and Maintenance of Differential Gene Expression</i>	8
3.2.4 <i>Epigenetics</i>	9
3.3 Chromatin	9
3.3.1 <i>The Nucleosome</i>	11
3.3.2 <i>ATP-dependent Nucleosome Remodeling</i>	12
3.3.3 <i>The Higher Order Organization of the Chromatin</i>	12
3.3.4 <i>Histones and Histone-Variants</i>	14
3.3.5 <i>Post-Translational Histone Modifications</i>	15
3.3.5.1 <i>Acetylation</i>	17
3.3.5.2 <i>Methylation</i>	18
3.3.5.3 <i>Phosphorylation</i>	21
3.3.5.4 <i>Ubiquitination</i>	21
3.3.5.5 <i>SUMOylation</i>	21
3.3.5.6 <i>ADP-ribosylation</i>	22
3.3.5.7 <i>Biotinylation</i>	22
3.3.5.8 <i>Interactions among Histone Modifications</i>	22
3.4 Mass Spectrometry	23
3.5 Objectives	26
4 Material and Methods	27
4.1 Laboratory Equipment	27
4.2 Reagents	28
4.2.1 <i>Chemicals</i>	28

4.2.2 Consumables	28
4.3 Biological Material	28
4.3.1 <i>Xenopus laevis</i> (Nasco, <i>Xenopus Express</i>).....	28
4.3.2 XTC and A6 Culture Cells (ATCC, LGC-Promochem).....	29
4.3.3 ES Cells and Primary Embryonic Fibroblast Culture	29
4.4 Embryological Methods	29
4.4.1 Solutions	29
4.4.2 Superovulation of Female <i>Xenopus laevis</i>	30
4.4.3 Preparation of Testis.....	30
4.4.4 In vitro Fertilization of Eggs and Culture of the Embryos.....	30
4.4.5 Removal of the Egg Jelly Coat.....	30
4.5 Biochemical Methods	30
4.5.1 Solutions	30
4.5.2 Nuclear Preparation from Cell Cultures	31
4.5.3 Nuclear Preparation from <i>Xenopus laevis</i> Embryos	31
4.5.4 Cell Lysate from <i>Xenopus laevis</i> Embryos.....	32
4.5.5 Histone Purification by FPLC	32
4.5.6 Nuclear Preparation from <i>Xenopus laevis</i> Embryos by Density Gradient Centrifugation.....	33
4.5.7 High Salt Extraction of Histones.....	34
4.5.8 Acidic Extraction of Histones.....	34
4.5.9 Extraction of Histones via Active Motif Histone Purification Kit	35
4.5.10 HPLC Purification of Histones.....	35
4.5.11 SDS-PAGE.....	36
4.5.12 Coomassie-staining.....	36
4.5.13 Quantitative Western Blot Analysis	36
4.6 MALDI-TOF	37
4.6.1 Preparation of Proteins for Linear Mode Mass Spectrometry.....	37
4.6.2 Preparation of Histones for Proteolysis	37
4.6.3 Digestion with Trypsin.....	37
4.6.4 Preparation of Peptides for Mass Spectrometry	38
4.6.4.1 Acid Extraction of Trypsin-digested Peptides.....	38
4.6.4.2 Zip Tip Extraction.....	39

4.6.4.3 <i>Carbon Tip Extraction</i>	39
4.6.5 <i>Analysis by MALDI-TOF</i>	40
4.6.5.1 <i>Linear Mode MALDI-TOF</i>	40
4.6.5.2 <i>Peptide Reflector Fingerprint Mode MALDI-TOF</i>	40
4.6.6 <i>Quantitation of MALDI-TOF Data</i>	41
4.7 Orbi-Trap	41
4.7.1 <i>Sample Preparation for Orbi-Trap</i>	41
4.7.2 <i>Analysis by Orbi-Trap</i>	42
4.7.3 <i>Identification of Histone PTMs by Orbi-Trap</i>	42
4.8 Heatmap Generation	43
5 Results	44
5.1 Histone Extraction from Stable Cell Lineages	44
5.1.1 <i>Direct Lysis of Nuclei</i>	44
5.1.2 <i>High Salt Extraction of Histones</i>	46
5.1.3 <i>Acid Extraction of Histones</i>	48
5.1.3.1 <i>Acid Histone Extraction with Perchloric Acid</i>	48
5.1.3.2 <i>Acid Histone Extraction with Hydrochloric Acid</i>	50
5.1.3.3 <i>Histone Extraction with Sulfuric Acid</i>	51
5.2 Histone Extraction of Embryos of Different Stages	52
5.2.1 <i>Nuclei Extraction with an Ultracentrifuge Step</i>	53
5.2.2 <i>Extraction of Nucleosomes with FPLC via Hydroxylapatite Column</i>	55
5.2.3 <i>Direct Lysis of Embryos and Further Purification via HPLC</i>	58
5.2.4 <i>Development of a New Protocol for Histone Extraction</i>	61
5.2.5 <i>Comparison to the Commercially Available Histone Purification Kit</i>	65
5.3 Technical Preparation for the Identification and Quantification of Histone Modifications by Mass Spectrometry	66
5.3.1 <i>Histone Sequences</i>	66
5.3.2 <i>Histone Preparation for Mass Spectrometric Analysis of Histone Modifications</i>	69
5.4 Identification and Quantification of Histone Modifications by Mass Spectrometry	71
5.4.1 <i>Post-Translational Modifications of Histones H2A and H2B</i>	74
5.4.2 <i>Post-Translational Modifications of Histone H4</i>	76

5.4.3 <i>Post-Translational Modifications of Histone H3</i>	79
5.4.4 <i>Post-Translational Modifications of murine ES Cells and MEFs</i>	87
6 Discussion	90
6.1 Technical Aspects of the Histone Extraction	91
6.2 Technical Aspects of the Histone Analysis by Mass Spectrometry	94
6.3 Global Alterations of Histone Modifications during <i>Xenopus laevis</i> Development	97
6.3.1 <i>Histone Modifications of Active Transcription</i>	98
6.3.2 <i>Repressive Histone Modifications</i>	99
6.3.3 <i>Clustering of Histone Modification Profiles</i>	103
6.4 Comparison of Histone Modification Profiles of Murine ES Cells and <i>Xenopus Blastulae</i>	105
6.5 Biological Implications of Histone Modification Profiles	108
6.5.1 <i>Reprogramming during <i>Xenopus laevis</i> Development</i>	108
6.5.2 <i>Biological Function of different Histone Modifications</i>	110
7 Abbreviations	113
8 References	123
9 Supplement	145

1 Summary

Vertebrate embryos are derived from a transitory pool of pluripotent embryonic cells. By the process of induction, these precursor cells are assigned to specific fates and differentiation programs. Histone post-translational modifications are thought to play a key role in the establishment and maintenance of stable gene expression patterns underlying these processes. While at gene level histone modifications are known to change during differentiation, very little is known about the quantitative fluctuations in bulk histone modifications during development. To investigate this issue histones isolated from four different developmental stages of *Xenopus laevis* were analysed by mass spectrometry.

Initially, a variety of different protocols for histone extraction from *Xenopus laevis* embryos and stable cell lines was tested and evaluated. Since non of the available methods worked sufficiently, a new reliable and effective protocol for nuclei preparation and histone extraction was established. Using mass spectrometry, core histone modifications were unambiguously determined. The techniques for identification and quantification of histone modifications by tandem mass spectrometry were improved as well.

In total, an average sequence coverage of 68% of modification sites for the four core histones was achieved by tryptic digestion after covalent modification of lysine residues with propionic anhydride. Using both LC-MS/MS and MALDI-TOF mass spectrometry, a total number of 2 modifications of H2A and 3 modifications H2B, 39 modifications of H3 and 20 modifications of H4 were identified and quantified. During this developmental period, an increase in the unmodified states, and a shift from histone modifications associated with transcriptionally active to transcriptionally repressive histone marks, was observed. Furthermore, these naturally occurring histone modifications were compared to the histone modifications of murine ES cells, detecting large differences in the methylation patterns of lysines 27 and 36 of histone H3 between pluripotent cells from *Xenopus blastulae* and murine ES cells. By combining all detected modification transitions, their patterns could be clustered according to their embryonic origin, defining specific histone modification profiles for each developmental stage. These specific histone modification profiles indicated a stepwise maturation of the embryonic

epigenome, which may be cause to the progressing restriction of cellular potency during development.

This thesis has revealed major quantitative shifts for several histone modifications known to be involved in gene regulation and furthermore enabled the definition of stage specific histone modification profiles accompanying and potentially regulating the transition from pluripotent to determined cell states using an antibody-independent method.

2 Zusammenfassung

Embryonen von Wirbeltieren entstehen aus Zellen, die sich für eine kurze Dauer in einem pluripotenten Zustand befinden. Durch den Prozess der embryonalen Induktion werden diese Vorläuferzellen ihrer spezifischen Differenzierung und der dafür notwendigen Regulierung zugewiesen. Man geht davon aus, dass der kovalenten Modifizierung der Histone eine Schlüsselrolle bei der Entstehung und Aufrechterhaltung von stabilen Genexpressionsmustern, die der Differenzierung zugrunde liegen, zukommt. Man hat heute bereits einige Informationen über die Veränderungen der Modifizierungen von Histonproteinen in Bezug auf einzelne Gene während der Differenzierung von Zellen erworben. Es ist jedoch wenig über das Ausmaß dieser Veränderungen auf globaler Ebene während der Entwicklung bekannt. Um diese Veränderungen beschreiben zu können wurden in dieser Studie Histonproteine aus vier verschiedenen Entwicklungsstufen des Krallenfrosches *Xenopus laevis* mittels Massenspektrometrie untersucht.

Zu Beginn der Versuche wurden verschiedene Methoden zur Aufreinigung von Histonproteinen aus den Embryonen sowie von Zelllinien des Krallenfrosches *Xenopus laevis* ausprobiert und bezüglich ihrer Qualität beurteilt. Da bei den Versuchen keine der verwendeten Methoden zufriedenstellende Ergebnisse erbrachte, wurde schrittweise eine neue Methode zur Isolierung von Zellkernen sowie Histonproteinen entwickelt. Die Charakterisierung und Identifizierung von verschiedenen Modifizierungen der vier Kern-Histonproteine konnte durch die Verwendung der Massenspektrometrie eindeutig erfolgen. Zudem konnten die Methoden zur Erkennung von Modifizierungen sowie deren Quantifizierung dieser weiter verbessert werden.

Durch die chemische Reaktion von Lysinseitenketten der Histonproteine mit

Propionsäureanhydrid und dem anschließenden Verdau der Histonproteine mit der Protease Trypsin konnten insgesamt 68% der Modifizierungsstellen der Aminosäuresequenz der vier Kern-Histonproteine beurteilt werden. Demnach konnten 2 Modifizierungen an dem Histon H2A, 3 an H2B, 39 an H3 und 20 an H4 beschrieben werden. In besonderem Maße wurde dies gewährleistet durch die Kombination der LC-MS/MS und der MALDI-TOF Massenspektrometrie.

Während der entwicklungsbiologischen Phasen, die untersucht wurden, konnte ein Abfall der Häufigkeit von Histon-Modifizierungen festgestellt werden. Zudem zeigte sich während der Entwicklung eine Verlagerung solcher Modifizierungen, die mit der Aktivierung von Gensequenzen in Zusammenhang gebracht werden können, zu solchen, die mit dem Abschalten von Gensequenzen in Verbindung gebracht werden. Darüber hinaus wurden die in dem Krallenfrosch vorkommenden Modifizierungsmuster mit solchen aus embryonalen Stammzellen, die von der Maus stammen, verglichen. Darin wurden besonders Unterschiede in der Methylierung von den Lysinresten 27 und 36 zwischen pluripotenten Embryonen des Frosches und den murinen Stammzellen gefunden. Mittels der Kombination aller beschriebener Modifizierungen der Histon-Proteine konnten für die jeweiligen analysierte Entwicklungsstufen charakteristische Muster identifiziert und in einer speziellen Graphik gruppiert werden. Diese stadienspezifischen Modifikationsmuster deuten auf eine schrittweise Reifung des embryonalen Epigenoms hin, die ursächlich an der zunehmenden Restriktion der zellulären Differenzierung und Potenz beteiligt sein könnte.

Diese Arbeit erbrachte einen bedeutenden Einblick in das Ausmaß der Bewegungen von zahlreichen Modifizierungen von Histonproteinen, die bei der Genregulierung beteiligt sind. Zudem ermöglichte diese Studie, Profile von charakteristischen Modifizierungsmustern der Histone zu definieren und diese entsprechend den Entwicklungsstadien zuzuordnen. Diese spezifischen Modifikationsmuster könnten bei der Regulierung des Übergangs von einem pluripotenten zu einem in der Differenzierung festgelegten Zustand der Zelle beteiligt sein.

3 Introduction

3.1 Development of Vertebrates

The embryonic development of a vertebrate organism is defined as the time between the fertilization of the egg and the end of organogenesis. The developmental process starts with only a single cell, the zygote, and leads over a period with fundamental steps to the formation of an adult organism, comprising of a huge variety of specialized and differentiated cell types. Among the events taking place during embryonic development are gastrulation, neurulation and organogenesis, the formation of the three germ layers - ectoderm, mesoderm and endoderm - and the stepwise specification of cell fates (Gilbert, 2006).

3.1.1 Morphological Differentiation

Embryonic development in *Xenopus* starts at fertilization, where male and female pronuclei fuse and the subsequent cleavage divisions begin. These cleavage divisions occur rapidly after each other and last only 20-30 minutes each. This short cell cycle is achieved by the lack of gap phases. The volume of the blastomeres is constantly reduced by each cell division and thus the nuclei to cytoplasm ratio increases constantly. After 12 cell divisions, the nuclei to cytoplasm ratio is assumed to reach a critical value leading to the subsequent activation of the zygotic genome. This time point is known as Mid-Blastula Transition (MBT; Newport *et al.*, 1982a/b). During this transition, the rate of cell division slows down and the cell cycle lengthens (Heasman, 2006).

The animal-vegetal axis is defined by the two hemispheres of the animal and vegetal pole of the oocyte. Coincidentally with the first cleavage division, cortical rotation takes place, which is a movement driven by microtubular organisation due to the entry of the sperm. The maximal movement, however, is directed into the opposite direction of the sperm entry point. The point of maximal cortical rotation becomes the future dorsal blastoporus and, thus, defines the anterior-posterior axis of the embryo (Heasman, 2006).

Shortly after the activation of the zygotic genome the embryo reaches with the gastrulation the next critical stage. During gastrulation, cells move to different positions within the embryo and establish the multilayered body plan. The amphibian gastrulation starts at the future dorsal side of the embryo by the formation of the blastopore lip. The

cells migrating from the outside to the inside of the embryo will form the endoderm and mesoderm; the cells remaining at the outer surface become the ectoderm. In order to cover the entire surface of the embryo, the ectodermal cells expand by epiboly, which is driven by the intercalation of cell layers in the animal hemisphere of the embryo. The vegetal cells will form the prospective endoderm (Beetschen, 2001). At the end of gastrulation, the three germ layers have been formed and the body axes have been determined. Furthermore, the segregation of the primordial germ cells, specific for reproduction, takes place during gastrulation (Gilbert, 2006).

By the end of gastrulation the three germ layers are differentiated and the formation of the organ precursors and the neurulation can begin. The ectodermal cells of the animal hemisphere show high BMP expression leading to epidermal differentiation. However, if the ectodermal cells are beneath the involuted mesodermal cells, the BMP signal is suppressed leading to proneural fate. Later, the neuroectoderm forms the neural plate, from which the brain and the neural tube arise. During the same time, organogenesis starts where the organ precursors - such as the heart, the archenteron and the kidney - are formed from derivatives of the three germ layers (Heasman, 2006).

Early embryonic development is ending on the way to the tadpole stage. Characteristic for amphibians - the tadpole continues its development with metamorphosis.

3.1.2 Differences concerning the Development of Amphibians and Mammals

In amphibians, such as the frog, development is extra-corporal of the female. On the other hand, in mammals, such as mice, both fertilization and development take place in the female's reproductive organs.

The mammalian embryo has no internal nutrient supply and, thus, the establishment of external nutrient and oxygen supply is crucial. Consequently, the earliest differentiation event of the mammalian embryo is the separation of trophoctodermal cells and the inner cell mass (ICM; Dyce *et al.*, 1987). The trophoctoderm will form the placenta together with the decidual cells of the female uterus. The amphibian embryo, in contrast, possesses a large yolk supply to nourish it until it starts feeding at day 7 of development.

Another difference between mammals and frog is the activation of the zygotic genome. As described above, the frog genome is activated after 12 rapid cell divisions

when the embryo consists of roughly 4000 cells at MBT. The mouse genome, however, becomes activated at the 2-cell stage (Moore *et al.*, 1993) and the human genome between the 4- and 8-cell stage (Braude *et al.*, 1988).

3.2 Regulatory Mechanisms of Early Development

As described above, major events take place during embryonic development of a vertebrate organism. All the morphological processes, the rapid cell divisions, the formation of the germ layers and the body axes during gastrulation as well as the following development, require a distinct regulation.

3.2.1 *Differential Gene Expression in Pluripotent and Somatic Cells*

Although the adult organism consists of a huge variety of differentiated and specialized cells, every cell - with the exception of lymphocytes - has the identical genomic information, the DNA. Pluripotent cells such as ES cells are transcriptional permissive and show a genome wide gene transcription at low levels (Efroni *et al.*, 2008), but the transcription is potentially unproductive due to stalled polymerase (Gunther *et al.*, 2007). Differentiated cells, however, have restricted transcriptional potential. These cells have both stably repressed genes as well as actively transcribed genes and thus use only a small portion of the genome present in the cell. This has been shown by experiments determining the mRNA and protein pool found in differentiated cells, which is highly specific for the cell type (Wetmur *et al.*, 1968). These results indicate that differentiated, somatic cells utilize differential gene expression.

During the embryonic development, the potency of cells is constantly reduced. At the beginning, the totipotent zygote is capable to produce all cell types of an adult organism as well as extra-embryonic tissues. During development, the cellular potency is reduced constantly, and there are pluripotent stem cells present, which are able to give rise to the embryonic tissues but have a more restricted ability to differentiate than the totipotent zygote. Adult organisms consist mostly of differentiated cells and have only small populations of multipotent stem cells remaining, like the haematopoietic stem cells, which give rise only to the different blood cells - erythrocytes, platelets and neutrophils.

Next to the totipotent zygote and the early blastomeres, there are embryonic stem cells (ESCs) derived from the ICM of the mouse blastocyst embryo (Evans *et al.*, 1981;

Martin, 1981). The cells of the ICM are still pluripotent. These pluripotent embryonic stem cells have characteristic features differing from somatic cell types. Embryonic stem cells cultured in vitro show permanent self-renewal, are able to form all embryonic cell types and, as determined in transplantation experiments, these cells can form teratocarcinomas and can give rise to derivatives of all the three germ layers as well as the germ line by implantation of the ES cells into blastulae embryos. On the other hand, these ES cell cultures require co-culture on a feeder layer or leukaemia-inducing factor (LIF) in the medium to prevent spontaneous differentiation (Prelle *et al.*, 2002).

Furthermore, mouse ES cells are characterized by the expression of the pluripotency core network. This network requires the expression of certain transcription factors such as Oct-4 and Sox-2 as well as the expression of either Klf-4 and c-Myc or Nanog and Lin-28 for reprogramming of somatic cells (Takahashi *et al.*, 2007; Yu *et al.*, 2007). These transcription factors share most of their target genes. Many targets of Oct-4, Sox-2 and Nanog encode key transcription factors for differentiation and development, and are transcriptionally inactive in ES cells.

3.2.2 *Inductive Events and Transcription Factor Cascades*

Inductive events and the activation of transcription factor cascades are necessary for the differential gene expression found in differentiating somatic cells.

Induction is the interaction of different cell types in close proximity, where one cell type is providing a signal called inducer to another cell type, which has to be competent for the signal and is responding to it accordingly. As for Conrad Waddington, competence is no passive state but an active ability to be responsive (Slack, 1993; Gilbert, 2006). Cellular competence can be explained best by the fact that a signal has different functions at different times, for example maternal Wnt dorsalizes, zygotic Wnt ventralizes the embryo. Competence may be the presence of a transcription factor in the cell, which is able to interact with the signalling molecule of the inducer.

Inductive events commit cells to a specific differentiation program. As the cells of the embryo stepwise differentiate into their somatic fate, they are first specified to become a certain cell type. This specification means that they are able to differentiate autonomously, but only in a neutral surrounding. However, they still can respond to environmental changes, which induce a change in cell fate. In the next step, these cells are

determined for their fate, meaning they are irreversibly committed to form one certain cell type (Slack, 1993).

On the other hand, transcription factors can bind to enhancer and promoter elements of genes from where they modulate gene transcription. The expression of certain transcription factors as well as their co-factors is very important for the establishment of gene expression. The presence of a transcription factor or transcription factor networks in the cell nucleus can establish differential gene expression, as it is found in differentiating cells. Among the transcription factor pathways are FGF, TGF- β , WNT- or BMP-signaling.

3.2.3 *Establishment and Maintenance of Differential Gene Expression*

Regarding the induction and the function of transcription factor networks, how is the early embryonic development regulated?

The zygote has a pool of maternal mRNAs and proteins, which become unevenly distributed during the rapid cell divisions at the beginning of embryonic development. Among those factors are transcriptional activators as well as repressors (Houston *et al.*, 2002, Kofron *et al.*, 2004). As mentioned above, the cortical rotation movement distributes the maternal Wnt11 mRNA. The Wnt pathway is crucial for the establishment of the organizer structure and the subsequent formation of dorsal structures in the embryo (Heasman, 2006).

By the start of zygotic transcription at MBT, a wave of inductive events is triggered. These signals lead to the activation of localized transcription factors. The activation of transcription factors becomes more and more stable through a process called self-enhancement.

At the beginning, the induction events lead to unstable expression of transcription factors. The initiation of transcription factor pathways, such as the FGF, TGF- β or Wnt, leads through the mechanisms of self-enhancement and feedback-loops to stable differentiated gene expression profiles. The consequence of the activation of specific pathways is the expression only of subsets of genes. This limited gene expression causes stepwise differentiation of the cell.

To establish and maintain the differential gene expression in different cells, the transcription factors are linked to mechanisms which affect the chromatin structure.

But how are these transcription profiles, established by the inductive signals,

maintained over several cell divisions?

Among the mechanisms underlying the establishment and maintenance of differential gene expression, the epigenetic regulation has increasingly gained attention during the last decades. Epigenetic mechanisms have shown to influence the regulation of the potency of cells such as pluripotency and differentiation, reprogramming of somatic nuclei as well as the transformation to malignant cancer cells (Jones *et al.*, 2007).

3.2.4 Epigenetics

Conrad Waddington defined the term “epigenetics” in 1957 as “a phenomenon that changes the final outcome of a locus or chromosome without changing the underlying DNA sequence. Thus, cellular differentiation may be considered [...] largely governed by changes in the “epigenetic landscape” rather than alterations in genetic inheritance” (Slack, 2002). A very attractive consequence of this hypothesis is the fact that these epigenetic modifications are stable enough to allow the faithful replication of gene expression patterns, but on the other hand have a higher level of variability than DNA sequence. The possibility to reverse the epigenetic modifications allows a return to the previous state of gene expression whenever necessary (Jaenisch *et al.*, 2003). This has led to the assumption, that through alterations in epigenetic modifications, somatic cells could be reversed to a pluripotent state and thus be differentiated *in vitro* in any cell type of interest (Campbell *et al.*, 1996; Gurdon *et al.*, 2000). This possibility would have tremendous impact on the medical treatment of diseases like strokes, myocardial infarction or dementia. Through the generation of replacement tissues from the patients’ own somatic cells, these illnesses could be cured in principle.

3.3 Chromatin

Epigenetic regulation involves mechanisms linked to the packaging of genes into chromatin, which regulate the accessibility of genes through different mechanisms like DNA methylation, histone post-translational modifications (PTMs) and histone variant incorporation, nucleosomal remodeling, non-coding RNA mediated silencing, chromatin compaction into higher order structures as well as its distribution in the nuclear space (Berger, 2007).

Chromatin, the template for gene transcription, DNA replication and repair,

consists of DNA wrapped around the nucleosomes (see Figure 1). It can be distinguished into the transcriptionally active euchromatin, the inactive facultative heterochromatin and the constitutive heterochromatin. The last contains only few genes and the centromeric and telomeric regions (Fischle *et al.*, 2003).

DNA can be methylated at CpGs, first described in 1948 and this methylation of CpGs is found in most eukaryotes (Hotchkiss, 1948). This modification is linked to gene regulation, genomic imprinting, X-chromosome inactivation and normal embryonic development in mammals. Three different DNA methyltransferases were found: Dnmt1, which is involved in maintenance of the methylation after DNA replication; and Dnmt3a as well as 3b, which both function during early development in establishing tissue specific DNA methylation signatures which are then transmitted to differentiating cells. DNA methylation can influence histone modifications via the MeCP2 complex (Shiraishi *et al.*, 2002), a reader of methylated DNA, which also functions as histone deacetylase. Thus DNA methylation could facilitate gene repression by removal of histone acetylations. In cancer cells, both hypomethylation and hypermethylation are found and may be linked to aberrant gene expression and cell function (Hermann *et al.*, 2004; Sulewska *et al.*, 2007).

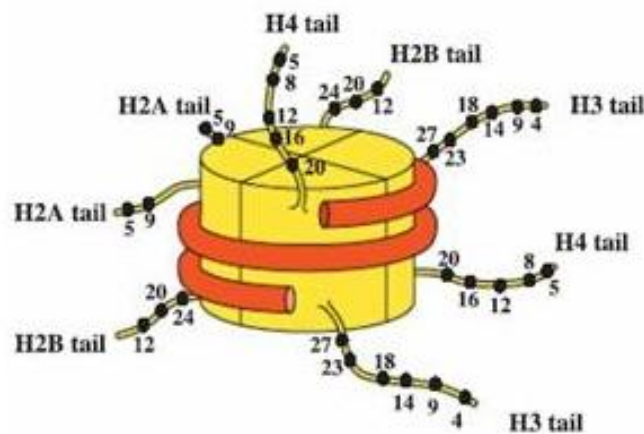


Figure 1: The Nucleosome

Depicted is the histone octamere with the DNA wrapped around (orange). The four core histones H2A, H2B, H3 and H4 (yellow) build the histone hetero-octamere in the center of the nucleosome, whereas the histone tails are protruding of the nucleosome. The histone tails are the sites of most histone modifications. Possible modifications are indicated by the dot and the number for the corresponding amino acid residue (modified after Alberts *et al.*, 2002).

During the last years, evidence accumulated that in addition to the chromatin modifications, non-coding RNA molecules are also involved in gene regulation among other functions. These interference RNAs (RNAi) are found in prokaryotes as well as in eukaryotes and their mechanism of action seems to be conserved among species as they function in degradation, modification and inhibition of the function of nuclei acids. The RNAi molecules base-pair with the nucleotide sequence of target RNAs mostly in the 3' untranslated region, and lead to degradation or translational inhibition of the target RNAs. There are three different small non coding RNAs known, small interfering (siRNA), micro RNA (miRNA) and PIWI interacting RNA (piRNA). All of them are derived from long double stranded RNA templates, that are processed by Dicer, a ribonuclease, to small-interfering RNAs, which are then loaded onto members of the Argonaute protein family such as the RISC complex (RNA-induced silencing complex; Peters *et al.*, 2007).

3.3.1 *The Nucleosome*

The fundamental and repeating unit of chromatin is the nucleosome, first detected by Kornberg in 1974 (Kornberg, 1974). The nucleosome consists of two copies of the four core histones - H2A, H2B, H3 and H4 - assembled to the histone octamere and 147bps of DNA which are wrapped around the histone octamere in a flat, left handed superhelix in 1.65 turns (Luger *et al.*, 1997; see Figure 1). This interaction is achieved by 14 mainly water mediated contact points between the minor groove of the DNA and the histone-octamere (see Figure 1). The water molecules arrange the interaction between the basic, positively charged amino acid side chains of the histones and the negatively charged phosphate backbone of the DNA and thus can overcome structural differences of the DNA sequence (Muthurajan *et al.*, 2003). The nucleosome is the most stable protein DNA complex under physiological conditions. At the entry of the DNA to the nucleosome, the linker histone H1 is situated to further compact and stabilize the complex. Two adjacent nucleosomes are connected through the linker DNA, which varies, tissue and species dependent, between 10 and 80 bps in length (Bednar *et al.*, 1998). The linker DNA together with the linker histone H1 and the nucleosome complex is referred to as the chromatosome.

3.3.2 *ATP-dependent Nucleosome Remodeling*

ATP-dependent nucleosome remodeling can alter the DNA histone interaction with the usage of ATP hydrolysis to overcome the non-covalent contacts between DNA and histone-octamers. Through their action, the underlying DNA sequence can become more accessible to enzymes such as the polymerases for gene transcription. Remodeling complexes can either slide the nucleosomes which either leads to dense package of the adjacent nucleosomes, opening of the DNA in between the nucleosomes, or total eviction of the nucleosome from the DNA. Chromatin remodeling factors can be grouped in four families, which are conserved from yeast to human. These complexes comprise of 2 to 12 different subunits with the common denominator being a single Swi2/Snf2 APTase subunit. The four families are: (i) Swi2/Snf2, (ii) ISWI/SNF2 type, (iii) CHD family and the (iv) INO80 complex (Becker *et al.*, 2002; Tsukiyama, 2002).

3.3.3 *The Higher Order Organization of the Chromatin*

The nucleosome condenses the primary DNA molecule in a five-fold manner and shapes the DNA molecule not just by bending, but also by facilitating higher order structures (Luger *et al.*, 1997). The unfolded “beads on a string” model (see Figure 2) consists of nucleosomes separated by 10-80 bps of linker DNA. This is then further compacted into the so-called 30 nm fiber, for which several models exist: (i) the “one start helix” with bent linker DNA followed immediately by the next nucleosome known as the solenoid (Thoma *et al.*, 1979), where the nucleosomes coil around a central cavity with eight nucleosomes per turn; (ii) the “two start helix” models consisting of straight linker DNA connecting two adjacent stacks of helically arranged nucleosome cores. The two start helical model can either form a helical ribbon or the crossed linker model (Chubb, 2009). The formation of higher order structures requires both the N-terminal domains of the histones as well as the linker histones. In the 30 nm fiber, DNA is compacted approximately 50 fold. Although the further levels of compaction are still under debate, there are models for a 60-130 nm chromonema fiber (see Figure 2) and chromatin-loop domains consisting of 100 kbps-1 Mbps (Figure 2; Cremer *et al.*, 2001). The overall compaction of DNA through this folding is roughly 400,000 fold in the eukaryotic nucleus. In the interphase nucleus, chromosomes are found unevenly distributed (see Figure 2), where gene rich territories are rather in the center of the nucleus and gene-poor regions

rather at the nuclear periphery. The interchromosome regions, which are thought to contain the complexes for transcription, and the chromosome territories build up an entire nuclear architecture (Cremer *et al.*, 2001). To further stabilize this nuclear organization, the chromosomes seem to be physically linked to the nuclear envelope via binding to specific proteins (Bartova *et al.*, 2008).

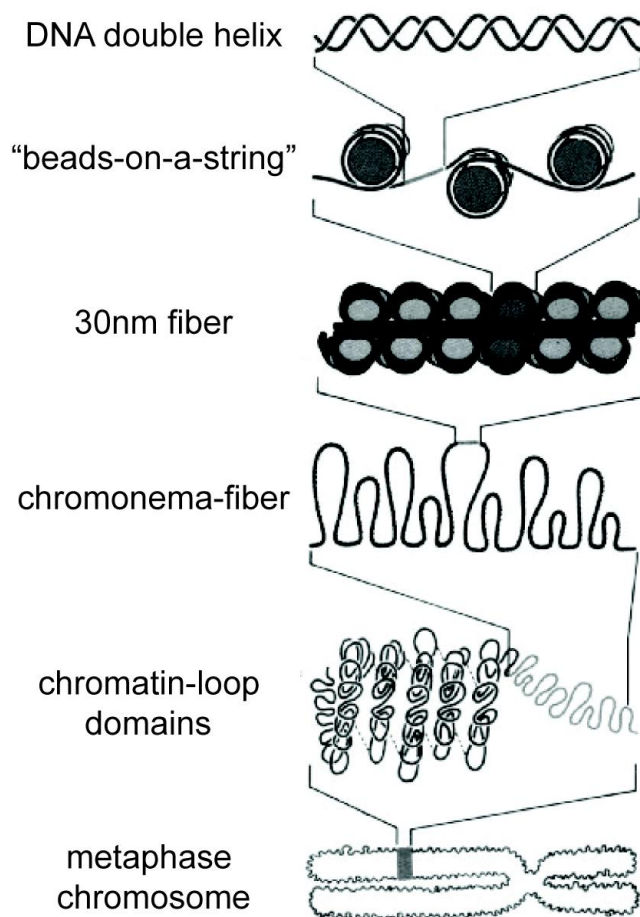


Figure 2: Levels of Chromatin Compaction

The DNA double helix is depicted in the uppermost panel. This DNA double helix is wrapped around the histone octamere to form the nucleosome. This conformation is often called “beads-on a string”, where histone octameres represent the beads and the DNA the string. Further compaction is achieved by the formation of the 30nm fiber, which is then folded into the chromonema fiber. The chromonema fiber is further compacted into the chromatin-loop domains. Finally, during mitosis, the chromatin loops are condensed to form the metaphase chromosomes (Picture taken from Felsenfeld *et al.*, 2003).

3.3.4 Histones and Histone-Variants

Histone proteins, first discovered in 1884 by Kossel (Kossel, 1884), are basic proteins grouped into the four core histones H2A, H2B, H3 and H4 and the linker histone H1 (Turner, 2005). The histone proteins are encoded in gene clusters, which are necessary for the allocation of high amounts of histones during DNA replication (Braastad *et al.*, 2004). The clustering of histone genes is well conserved among different species such as sea urchin, flies (Sierra *et al.*, 1982), mice (Wang *et al.*, 1996) and humans (Marzluff, 2002). Next to the conserved organization of the histone genes, the sequences of the histones are also among the most invariant proteins known (Kornberg, 1999). The four core histones are small proteins - their size ranging from 11 kDa and 16 kDa - consisting of a large globular domain and a small flexible N-terminal domain referred to as the histone tail. The globular domains are necessary for the interaction of the histone proteins among each other by the formation of the histone-octamere complex (Arents *et al.*, 1995; see Figure 1). The tail domains which are protruding from the nucleosome and thus are easily accessible for enzymes, contain high amounts of lysine and arginine residues, which are subjected to a variety of post-translational modifications. Removing the histone tails by genetic engineering in yeast is cell-lethal, which points out their importance.

For three of the four core histones - i.e. H2A, H2B and H3 - histone variants exist, which mostly differ only in few amino acids from their canonical counterparts. The so-called H3.1 is incorporated into the chromatin in a replication-dependent manner. The H3 variants H3.2 and H3.3, however, are incorporated in a replication-independent manner. Furthermore there are some specific variants, such as the testis specific H3.1t and CENP-A (centromeric protein A), which is found at centromeres and is required for correct kinetochore assembly (Henikoff *et al.*, 2005). H3.1 is associated with repressed gene loci and is mainly associated with repressive histone modifications. In contrast, the H3.3 variant is associated with actively transcribed genes and mostly carries active histone marks, such as K9, K18 and K23 acetylation. Nucleosomes containing the H3.3 variant are much less stable than those with H3.1 (Jin *et al.*, 2007). According to a common model, H3.1 is linked to constitutive heterochromatin, H3.2 to facultative heterochromatin and H3.3 to euchromatic regions (Hake *et al.*, 2006). However the proof for the function of H3.2 is still missing. Support for the idea that different variants have different functions may come from the finding that H3.1 is bound by the histone chaperone CAF-1 (chromatin assembly factor 1), whereas H3.3 is bound by Hira (Loyola *et al.*, 2006). However, this may, on the

other hand, simply reflect different modes of histone deposition.

Most histone variants are found for H2A, among the best characterized are H2A.X, H2A.Z, H2A.Bbd (Bar body dependent) and macroH2A (West *et al.*, 1980). H2A.X has high sequence similarity to H2A except the C-terminal domain, where it is specifically phosphorylated upon DNA damage (Rogakou *et al.*, 1998). This phosphorylated H2A.X, also referred to as γ -H2A.X, is directly linked to DNA double strand break repair processes. The function of H2A.Z is more controversial. H2A.Z is found ubiquitinated at the inactivated X-chromosome in mammals, suggesting its involvement in the formation or maintenance of heterochromatin structures (Sarcinella *et al.*, 2007). In yeast, however, H2A.Z is associated with transcriptionally active gene promoters (Jin *et al.*, 2007). The most diverged H2A variant is macroH2A. The molecular weight is roughly three times higher than canonical H2A, due to a large C-terminally located “macro” domain. MacroH2A is involved in X-chromosome inactivation of mammals by regulating ADP-ribosylation with its macro domain (Ladurner, 2003).

Next to the four core histones, there is the H1 linker histone family, which has the highest diversity among different species. Furthermore, there are many variants found of H1 proteins in humans, including five somatic subtypes, a spermatogenic subtype, and an oocyte subtype, suggesting diverse functions for linker histones (Wood *et al.*, 2009). Typically, they contain a globular domain and a highly basic C-terminal domain necessary for the maintenance of the nucleosome positioning (Zhang *et al.*, 2003). H1 hinders both spontaneous disruption of the nucleosome and ATP-dependent nucleosome remodeling. The stoichiometry of H1 to nucleosomes differs remarkably among different tissues and cell types, possibly reflecting the global accessibility of the genome (Rupp *et al.*, 2005).

3.3.5 *Post-Translational Histone Modifications*

Histone proteins are subjected to a variety of covalent post-translational modifications (PTMs). These modifications are mainly situated at the tails of the histones (see Figure 3), although there are also some modification sites found in the globular domains. As this thesis work only pertains to the four core histones, only the modifications of these histones will be discussed, with special focus on H3 and H4, that harbour most known modifications (Kouzarides, 2007; Lee *et al.*, 2009; Dambacher *et al.*, 2010).

Common modifications are the methylation of lysine (K) and arginine (R) residues (me; Allfrey *et al.*, 1964) and the acetylation of lysines (ac; Allfrey *et al.*, 1964). Next to these modifications there is also phosphorylation of serine (S) and threonine (T) residues (phos), ubiquitination (ubi, Shilatifard *et al.*, 2006) and SUMOylation of lysines (sumo, Nathan *et al.*, 2006), ADP-ribosylation of glutamines (G) and arginines (adp). Furthermore, core histones are known to be biotinylated and subjected to prolin isomerization (Nelson *et al.*, 2006; Kouzarides, 2007).

According to a recent rule of nomenclature, modifications are described by the core histone, the site and the kind of modification (Turner, 2005). For instance, the trimethylation state of lysine residue at amino acid position 4 of the histone H3 is referred to as H3K4me3 (Turner, 2005).

The different post-translational histone modifications differ in their functional readout and in the enzymes involved in the establishment and erasure of these modifications.

The histone modifications are established by specialized enzymes, also called the “writers”, and they get removed by another set of enzymes, also referred to as “erasers” (Berger, 2007). The steady state reflects the competing activities of the two sets of enzymes.

For the function of the histone modifications frequently other proteins and protein complexes are recruited, which are able to selectively “read” the histone modification via special domains. These enzymes lead to the establishment or erasure of other histone modifications (cross-talk), nucleosomal remodeling, initiation of transcription factor binding, DNA accessibility and repair, cell cycle progression or replication (Kouzarides, 2007).

3.3.5.1 Acetylation

Acetylation of lysines is found on K5, K9, K13 on histone H2A and of K5, K12, K15 and K20 on H2B, K4, K9, K14, K18, K23, K27, K36 and K56 of H3 as well as K5, K8, K12, K16 and K20 of H4 (Bhaumik *et al.*, 2007).

Acetylation of lysine residues of histones was first described in 1964 (Allfrey *et al.*, 1964). The enzymes establishing the acetyl marks are called histon acetyl transferase (HATs; Vaquero *et al.*, 2003). There are three classes of HATs known: GNAT, NYST and

CBP/p300. HATs are rather unspecific in their substrate recognition, however, the association of the HAT activity in different complexes provides substrate specificity (Lee *et al.*, 2007). Histone deacetylation is performed by histone deacetylases (HDAC). The acetylation and deacetylation of histones is a highly dynamic process.

In general, acetylation of lysines is linked to transcriptional activation. The acetylation of the core histones partially neutralizes the positive charge of the lysine residues consequently decreasing the binding efficiency of the histone to the DNA and, thus, facilitating transcription. Furthermore, acetylated histones are recognized by specific bromodomain containing proteins (Kouzarides, 2007).

Via the acetylation of histones and the interaction of protein-domains which bind to the acetylated histones, protein complexes such as nucleosome remodeling ATPases are subsequently recruited to chromatin and lead to transcriptional permissive chromatin regions. On the other hand, transcriptional silencing is associated with hypoacetylated nucleosomes (Braunstein *et al.*, 1993).

The histone acetylation cycle is probably best understood in yeast. For H4, acetylation occurs in a hierarchical fashion. The mono-acetylated histone is predominantly acetylated at K16. The diacetylated H4 is either acetylated at K8 or K12 together with K16, and K5ac is found only at the tri- or tetra-acetylated histones (Clarke *et al.*, 1993). In general, H4K5 and K8 are frequently found acetylated in euchromatic regions with active transcription. In contrast, H4K12ac is increased in heterochromatic regions (Kimura *et al.*, 2005). High acetylation levels of H3K9ac and K14ac are found in active genes (Liang *et al.*, 2004). Acetylation of the N-terminus of H3, nominally at H3K4, K9, K14, K18, K23, K27 and K36 is linked to active transcription (Taverna *et al.*, 2007). H3K56ac is found on newly synthesized histones. H3K56 is acetylated by Rtt109 in yeast, however, in humans, this mark is established by CBP and p300 and removed by SIRT1/2 (Das *et al.*, 2009). H3K56ac shows only a weak correlation to gene expression in genome wide ChIP analysis (Xie *et al.*, 2009).

3.3.5.2 Methylation

Methylation is found on both lysine as well as arginine residues of histones. Lysines can be mono-, di- or tri-methylated, in contrast to arginines, which can only be mono- and di-methylated (Paik *et al.*, 1969). Methylation of histones is always interpreted

by enzymes, which recognize the methylation by specialized domains like the chromodomain, the WD40 repeat, the Tudor domain, the MBT domain or the PHD finger domain (De la Cruz *et al.*, 2005).

Lysine methylation is established by very specific histone methyltransferases (KMTs) and for arginines by arginine methyltransferases (PRMTs). Methylation of both arginines and lysines is reversible. Methylated arginine can be deiminated to citrulline by PADI4 (peptidylarginine deiminase 4; Wang *et al.*, 2004). This does not totally reverse the modification but at least removes the methyl-group. Whether the citrulline then gets converted to an arginine residue again or the histone gets replaced is currently unknown. For the removal of methylation of lysines there are two classes of lysine demethylases known, the LSD1/BHC110 (Shi *et al.*, 2004) class and the jumonji class (Klose *et al.*, 2006).

Arginine methylation is linked to transcriptional active genes. Methylation of arginines occurs at H3R2, H3R8, H3R17, H3R26 and H4R3 (Kouzarides, 2007).

In contrast, methylation of lysine residues is involved in diverse function. The methylation of H4K20 was first described in 1969 (DeLange *et al.*, 1969). H4K20me1 is involved in the inactivation of the X-chromosome in mammals. The mono-methylation is established by PR-Set7 (Nishioka *et al.*, 2002). Lack of PR-Set7 in mouse embryos is lethal at the eight-cell stage. These embryos show depletion of H4K20me1 and reduced levels of H4K20me2 and me3 (Oda *et al.*, 2009). H4K20me2 and me3 is associated with constitutive heterochromatin and silencing of imprinted gene clusters (Schotta *et al.*, 2004). The di- and tri-methylation is established by the Su(var) 4-20h1/h2 enzymes, which use the K20me1 as their substrate. In addition to its function in transcription, H4K20me also functions in cell cycle progression and DNA repair (Yang *et al.*, 2008).

H3K4me is established by MLL HMTs, especially SETD1A/B and Ash1 (Gregory *et al.*, 2007). These enzymes belong to the Trithorax group, which is involved in the segmental activation of the Hox genes. Demethylases of both the LSD1/BHC110 class and the Jumonji class are able to remove the methyl mark of H3K4. H3K4me3 is found at the 5' end of ORFs of genes and is associated with actively transcribed genes (Bernstein *et al.*, 2002). H3K4me is able to recruit a variety of different factors such as nucleosome remodeling enzymes (Flanagan *et al.*, 2005), HATs (Taverna *et al.*, 2006) and HDACs (Shi *et al.*, 2006). Furthermore H3K4me is involved in the binding and initiation of Polymerase II (Pol II; Fuda *et al.*, 2009).

H3K9me2 and me3 is established by the SETDB1 enzymes. The Jumonji class of

histone demethylases is able to remove the me2 and me3 again. H3K9me2 and me3 was shown to be repressive and associated with heterochromatin (Rea *et al.*, 2000). However, H3K9me1 has also been found at active transcribed regions (Barski *et al.*, 2007). H3K9me2 and me3 is found in heterochromatin where it is established by Su(var)3-9 (Rea *et al.*, 2000). The methylation of H3K9 is a prerequisite for the subsequent recruitment of other factors such as HP1 and the compaction of heterochromatin. In euchromatic regions, H3K9me2 and me3 are established mainly by G9a and Glp (Tachibana *et al.*, 2002). Recently K9me2 was found in large regions in differentiated tissues by ChiP analysis. These regions are conserved between mice and humans and the H3K9me2 occupancy has low levels. H3K9me2 was shown to correlate inversely to the expression of the underlying genes. Wen and coworkers termed these regions as large organized chromatin K9 modifications (LOCKS; Wen *et al.*, 2009).

H3K27me2 and me3 are established by the Enhancer of Zeste (Ezh2) HMT (Cao *et al.*, 2002). Ezh together with Suz12 and Eed belongs to the Polycomb repression complex 2 (PRC2). H3K27me2 and me3 are hardly found at active promoters and genes, thus being strongly correlated with silenced genes. At euchromatic regions, H3K27me3 is necessary for gene silencing and it functions in heterochromatin maintenance. The H3K27me3 mark recruits the PRC1 complex. PRC1, together with the associated protein Cbx2, establishes the H2AK119ubi mark via the Ring1A subunit (Wang *et al.*, 2004). Those genes being occupied by the PRC2 and PRC1 machinery are mainly developmental regulators (Boyer *et al.*, 2006). H3K27me1, however, is found at regions with active transcription (Barski *et al.*, 2007).

H3K36me is established by SETD2 of the Set2 class of the HMTs and associated with active RNA-Pol II. H3K36me2 and me3 are predominantly found near the 3' end of genes (Bannister *et al.*, 2005). The Jumonji class of histone lysine demethylases is capable to remove H3K36me2 and me3 again. H3K36me3 can recruit HDAC to the gene body to prevent the aberrant initiation of the Pol II in the gene body (Carroza *et al.*, 2005).

Lysine 79 is located in the globular domain of H3 and can be methylated in most eukaryotic species. The mark is established by DOT1, the only HMT lacking the Set domain (Ng, *et al.*, 2002). H3K79me3 is linked to constitutive heterochromatin and the silencing of telomers, as the name of the enzyme DOT1 implies (disruptor of temoleric silencing). Furthermore, methylation is described for H2BK5, H3K14 and H3K56 (Bhaumik *et al.*, 2007).

3.3.5.3 Phosphorylation

All core histones have been shown to be phosphorylated (see Figure 3). The modification is established by kinases and removed by specific histone phosphatases again. Phosphorylation of histones is associated with chromosome condensation during mitosis, as well as with transcriptional activation. Sites of phosphorylation are H2AS1 and H2AT120, H2BS13, H3T3, H3S10, H3T11, H3S28 and H4S1 (Kouzarides, 2007). The best characterized site of phosphorylation is H3S10, which inhibits the modification of the adjacent H3K9 residue (Fischle *et al.*, 2005). H3K10phos is established by the Aurora B kinase, which is involved in mitotic regulation (Vaquero *et al.*, 2003).

3.3.5.4 Ubiquitination

Ubiquitination of histones is known for all core histones except H4, whereas H2A and H2B show higher levels than H3 (see Figure 3). This modification occurs as mono- ubiquitination and involves the transfer of the 76-amino-acid peptide ubiquitin to a lysine residue (Shilatifard *et al.*, 2006). De-ubiquitinating enzymes can remove this modification again. Poly-ubiquitination is linked to the degradation of the histone by the proteasome, whereas the mono-ubiquitination functions in transcriptional regulation. Ubiquitinated lysines are found at position 119 at H2A and 120 at H2B. H2AK119ubi is linked to transcriptional repression as it is established by Bmi/Ring 1a. This ubiquitin-ligase, a subunit of the polycomb repression complex 1 (PRC1), is supposed to be involved in inhibition of the elongating RNA-Pol II. Recent data, however, questions the silencing function of H2AK119ubi as the protein ZRF1, which specifically binds to the histone modification is also involved in gene activation (Richly *et al.*, 2011).

H2BK120ubi is associated with transcriptional activation in yeast (Zhu *et al.*, 2005). However, deubiquitination of H2BK120 also seems to represent an activating stimulus. According to a model of sequential modification, H2BK120 is deubiquitinated first by Ubp8, a deubiquitination factor and component of the SAGA complex. This leads to the methylation of H3K4. Then the recruitment of Set2 is possible, which methylates H3K36, a mark of elongating Pol II (Kouzarides, 2007).

3.3.5.5 SUMOylation

SUMOylation (Small Ubiquitin like Modifier), related to ubiquitination, is also a

very large modification. It has been found at all the four core histones (see Figure 3; Kouzarides, 2007). It is linked to transcriptional repression as it correlates inversely to the level of acetylation (Lee *et al.*, 2007a). SUMOylation is found at H4K5, K8 and K16 (Bhaumik *et al.*, 2007).

3.3.5.6 ADP-ribosylation

ADP-ribosylation (Adenosine-diphosphate) of histone proteins involves the enzymatic transfer of ADP-ribose from nicotinamide adenine dinucleotide (NAD⁺) to glutamine or arginine residues and it occurs as mono- or poly-ribosylation (Hassa *et al.*, 2006). All four core histones have been shown to bear possible modification sites. The highest prevalence of ADP-ribosylation is found at H2B. As the establishment of the ADP-ribosylation mark is NAD⁺ dependent, it may be connected to cell metabolism. Only little is known about the function of ADP-ribosylation. Because it is enriched at hyperacetylated H4, it might be linked to transcriptional activation as it is found together with hyperacetylated histone H4 (Vaquero *et al.*, 2003).

3.3.5.7 Biotinylation

Biotinylation of H4K12 is linked to heterochromatin (Latham *et al.*, 2007). It is found at H2AK9 and K13, H3K18 and H4K12 (see Figure 3). So far, little is known about the function of Biotinylation (Bhaumik *et al.*, 2007).

3.3.5.8 Interactions among Histone Modifications

As mentioned above, the establishment or removal of one histone PTM can influence other histone modifications. In that model, different histone modifications are either elusive or concomitant. An example of histone PTM cross-talk is the methylated H3K4 mark. The active H3K4me₃ mark is found on hyperacetylated histones, but never coexists with the repressive H3K9me₃ mark. Thus, H3K4me₃ has been found to coexist with up to penta-acetylated histones at H3K9, K14, K18, K23 and K27 (Taverna *et al.*, 2007). Furthermore, the methylation of H3K4 by MLL1 complex links the acetylation of H4K16 by the MOF acetyltransferase, because both enzymatic activities are found in one complex. Both these marks are required for optimal transcription *in vitro* and at the Hox a9 gene in

vivo (Dou *et al.*, 2005). It is also known, that acetylation of H3K18 and K23 promotes the methylation of H3R17 by CARM1 in the course of estrogen-dependent gene activation (Daujat *et al.*, 2002). Interestingly, demethylation of H3R17 is enzymatically linked to histone deacetylation (Lee *et al.*, 2006).

Chromatin may be divided in eu- and heterochromatin. Euchromatin, which is less condensed, accessible and easily transcribed, contains the genes, which are either active or silent. Active genes are associated with high levels of H3K4me1, me2 and me3, K9me1, K36me3, K27me1 and H4K20me1. Silent genes in euchromatin are marked by high levels of K27me2 and me3, K79me3 and lowest levels of K4me3, K36me3, K27me1, K9me1 and H4K20me1 (Barski *et al.*, 2007).

The genome consists to a large extent of non-coding DNA sequences such as satellite repeats, telomeric repeats, mobile elements and interspersed elements. To prevent their transcription and thus to maintain genome stability, it is necessary to have these elements under tight control. This constitutive heterochromatin can be divided into pericentric and telomeric heterochromatin. Pericentric heterochromatin consists of major satellite repeats. It is marked by H3K9me3 and H4K20me3. In yeast, this histone PTM pattern is established by the Su(var) 3-9 homologue, which is recruited by RNAi. The Su(var) 3-9 enzymes establish H3K9me3, which in turn recruits HP1 (Lachner *et al.*, 2001). Via HP1 Su(var) 4-20 enzymes are recruited and the H4K20me3 mark is established (Schotta *et al.*, 2004). Telomeric heterochromatin is very similar to pericentric heterochromatin. Chromatin fragments enriched in H3K9me3 and H4K20me3 often contain elevated levels of H3K79me3 established by Dot1.

3.4 Mass Spectrometry

The soft-ionization methods in mass spectrometry such as Matrix Assisted Laser Desorption/ Ionization – Time of Flight (MALDI-TOF; Karas *et al.*, 1988) and Electro Spray Ionization (ESI; Whitehouse *et al.*, 1985) led to the identification of new sites of histone modifications (Zhang *et al.*, 2003) as well as description of bulk chromatin modifications (Strahl *et al.*, 2000). Histone modifications have been described using mass spectrometry in several species like *Drosophila* (Bonaldi *et al.*, 2004), *Xenopus* oocytes (Nicklay *et al.*, 2009), chicken erythrocytes (Zhang *et al.*, 2002) and human cells (Strahl *et al.*, 2001).

Common to all mass spectrometric methods is that the sample is first ionized and then accelerated in an electrical field. In case of MALDI-TOF, the electrical energy is equal the charge of a molecule q multiplied with the electrical voltage of the electrical field U (see Figure 4A). Thus the energy a single molecule gains in the electrical field is the same for all molecules given that the charge is identical. The electrical energy, however, is equal to the kinetic energy of the molecule according to the law of conservation of energy. The kinetic energy depends on the mass and the velocity of the molecule (see Figure 4A). The velocity is defined as the distance s traveled during the time t (see Figure 4B). By measuring the time of flight (t) a molecule needs for a defined distance (s), one is able to determine the mass of the molecule (see Figure 4C; Lottspeich *et al.*, 2006).

$$E_{\text{elec}} = q * U = \frac{1}{2} m * v^2 = E_{\text{kin}} \quad (\text{A})$$

$$v = \frac{s}{t} \quad (\text{B})$$

$$m = \frac{2 q * U * t^2}{s^2} \quad (\text{C})$$

Figure 4: Physical Equations

As mentioned above, there exist several techniques for the ionization of the analyte. Using MALDI (matrix assisted LASER desorption/ionization) the analyte is embedded in a solid matrix, consisting of the co-crystallized matrix and analyte molecules. As the solid matrix absorbs the energy of the LASER beam, the solid matrix is desorbed and matrix and analyte molecules are transferred to the gas phase. Simultaneously, the analyte molecules are ionized via proton transfer (Lottspeich *et al.*, 2006).

Mass spectrometry allows the exact measurement of the mass of the analyte. As proteins consist of only twenty different amino-acids, the mass of different proteins can be identical. To increase the specificity for the identification, the proteins are digested into peptides with proteinases. The combination of different peptides of one protein - like a fingerprint - is much more characteristic for a protein than its single mass. Consequently

this peptide mass fingerprint technique is used to identify proteins.

Another difficulty concerning the detection of modifications of proteins is that modifications lead to a mass shift, but do not necessarily point to the exact localization of the modification. To circumvent this problem, the digested peptides are further disassociated into smaller fragments. This fragmentation occurs theoretically at all bonds of the molecule, but peptide bonds are favoured. One ends up with a series of different fragments of one peptide. The mass shift of a protein modification can then be exactly localized due to the altered mass of the corresponding fragments (Lottspeich *et al.*, 2006).

Knowing the sequence of a protein, covalent modifications can easily be detected in peptide mass fingerprint without the need for peptide sequencing. To be able to distinguish between the mass shift for tri-methylation (42.0470Da) and the one for acetylation (42.0106Da) of an amino acid, ultra-high resolution technology is necessary (Freitas *et al.*, 2004; Gropengiesser *et al.*, 2009). The separation of different histone modification states in a reversed-phase HPLC chromatography prior to mass spectrometry allows the precise quantification of different modifications (Freitas *et al.*, 2004). Striking advantages of mass spectrometry are the possibility to quantify different proteins or modifications and to possibly identify new proteins or protein modifications (Gropengiesser *et al.*, 2009).

3.5 Objectives

The aim of this thesis was to establish a broad compendium of covalent histone modifications and their quantitative flux during normogenesis in a vertebrate model organism.

Recent scientific studies were conducted in *Drosophila* revealing alterations in histone modifications during *Drosophila* development (Bonaldi *et al.*, 2004) and the mammalian systems such as mouse and human (Bernstein *et al.*, 2002; Phantsiel *et al.*, 2007). As experimental data is difficult to achieve during mammalian development, several studies were performed in the *in vitro* system using embryonic stem cells (Bernstein *et al.*, 2002; Phantsiel *et al.*, 2007). Analysis of histone modifications of the vertebrate organism *Xenopus laevis* allows to fill the gap between the results of *Drosophila* studies and the *in vitro* data sets of murine ES cells. To allow comparison to recent data sets of ES cells (Bernstein *et al.*, 2002; Phantsiel *et al.*, 2007), analysis of primary embryonic fibroblasts and murine ES cells was included. Furthermore, a considerable amount of developmental studies in the field has been conducted in *Xenopus laevis* (Heaseman, 2006) and our laboratory has exquisite expertise in *Xenopus* development.

During normogenetic development of *Xenopus laevis*, precursor cells are assigned to specific fates and differentiation programs. Histone post-translational modifications are thought to play a key role in the establishment and maintenance of stable gene expression patterns underlying these processes. Thus the aim was to achieve a detailed and quantitative description of histone modifications and their alterations during vertebrate development.

One goal was to define a broadly applicable method for histone purification for diverse samples such as *Xenopus* embryos as well as cultured cells. In order to obtain quantitative information on variety of different histone modifications and modification sites by an antibody-independent approach, mass spectrometry is the method of choice. The information from ChIP-based studies is invariably dictated by many technical parameters, most notably the antibody quality (Minard *et al.*, 2009; Trelle *et al.*, 2007). Mass spectrometry, on the other hand, has the advantage of quantification, the possibility to identify new modifications, no limitations in the availability of specific antibodies and no limit in the detection of combinatorial modifications (Gropengiesser *et al.*, 2009).

4 Material and Methods

4.1 Laboratory Equipment

CCD camera: Q Imaging (Retiga) Gel-documentation System

Centrifuges: Eppendorf centrifuge 5417C (Eppendorf), Sorvall RC-6 Plus (Thermo Scientific), Optima MAX-E Ultracentrifuge (Beckman), PicoFuge (Stratagene), miniSpin (Eppendorf), miniCentrifuge (neoLab), 3-18 (Sigma)

Columns: 5ml FPLC self preparable column (Pharmacia Biotech), Jupiter C4 5ml 300Å (15061.0 Phenomex)

FPLC: Äkta FPLC (Amersham Pharmacia Biotech)

Gel Documentation: (peqLab), G-Box (Syngene)

Homogenizer: Glass-Glass Homogenizer 5ml narrow (Braun, Melsungen), LSC Homogenizer LH-21 (Yamato)

HPLC: Ultimate 3000 (Dionex), Ettan microLC (Amersham Pharmacia Biotech)

MALDI-TOF: Voyager-DE™ STR, BioSpectrometry™ Workstation (Applied Biosystems)

Microscopes: Stereomicroscope Stemi SV6, Axiophot (Zeiss),

Orbitrap: LTQ-XL (Thermo Scientific)

Software: Photoshop CS5 (Adobe), Illustrator CS5 (Adobe), MacVector 7.1 (Oxford Molecular Group), Office 2004 for Mac (Microsoft), QCapture Imaging (Zeiss/Axiophot), Voyager for MALDI-TOF (Applied Biosystems), Data Explorer for MALDI Analysis (Applied Biosystems), Bioworks for Orbitrap Analysis (Thermo Scientific), Xcalibur for Orbitrap Analysis (Thermo Scientific), Manuelito (open source: <http://sourceforge.net/projects/manuelito>), Vision Capture (Vilber Lourmat), Unicorn 5.1 (G.E. Healthcare), GPMW 5.5 (Lighthouse Data), Gene Snap (Synoptics), R (open source: <http://www.r-project.org>)

Sonicator: Bioruptor™ (Diagenode), Ultrasonic Bath (Merck)

Target Plate: Hydrophobic Target Plate (Applied Biosystems)

Shaker: Thermomixer comfort (Eppendorf)

Western-Blot Quantification: Li-cor Odyssey (Bioscience)

4.2 Reagents

4.2.1 Chemicals

Acetonitril HPLC-grade (Sigma); Ammoniumbicarbonate (Sigma); Ammoniumpersulfate (Merck); Aprotinin (Sigma); Chicken Serum (Invitrogen); Coomassie brilliant blue G250 (Serva); DMSO (Sigma); DTT (Sigma); Enhanced Chemiluminescence solution (GE); Fetal Calf Serum (GIBCO BRL); Gentamycin (GIBCO BRL); Human Chorionic Gonadotropin (Sigma, Amsa); ddH₂O HPLC-grade (Merck); Dulbecco's Modified Eagle Medium, DMEM (Sigma); Leupeptin (Sigma); L-Glutamine (GIBCO BRL); Mitomycin C (GIBCO BRL); MNase (Sigma); Nonessential Amino Acids (GIBCO BRL); NP-40 (Fluka); Monothioglycerol (Sigma); Page Ruler Prestained Protein Ladder (Fermentas); Penicillin (GIBCO BRL); Pepstatin (Sigma); PMSF (Sigma); Propionic Anhydride (Sigma); Proteinase K (Sigma); Precision Plus Protein Prestained Standard (Biorad); Recombinant Mouse Leukemia Inhibitory Factor (ESGRO); Sinapinic Acid (Fluka); Sodium Butyrate (Sigma); Spermine (Sigma); Spermidine (Sigma); Streptomycin (GIBCO BRL); Triton X100 (Roth); Trypsin (Promega); Tween (Sigma); α -Cyano-4-Hydroxycinnamic Acid (Sigma), β -Mercaptoethanol (Sigma).

Other fine- and biochemicals were ordered at the following companies: Fluka, Merck, Sigma-Aldrich, Roth and Biomol.

4.2.2 Consumables

0.2ml tubes, stripes of 8 (Nunc); 0.5ml tubes LoBind (Eppendorf); 10cm dishes (Falcon); 15cm dishes (Greiner Cellstar); 15mL, 50mL tubes (Sarstedt); Dialysis Membrane, MWCO 6-8 (SpectraPor); Histone Purification Kit (Active Motif, Cat.No 40025); Carbon Tips (Glygen); Mirocloth® (CalBiochem); ZipTips μ -C18 (Millipore)

4.3 Biological Material

4.3.1 *Xenopus laevis* (Nasco, *Xenopus Express*)

Adult wild-type *Xenopus laevis* frogs were used. The frogs were kept in tap water with a temperature of 17-19°C and a population density of one frog per 5l water. The animals were fed three times per week with Pondsticks Premium brittle (Interquell

GmbH, Wehringen; Sieve *et al.*, 1989).

4.3.2 XTC and A6 Culture Cells (ATCC, LGC-Promochem)

Cells were cultured in 75% DMEM medium, supplemented with 15% fetal bovine serum, 100U/ml penicillin/100ng/ml streptomycin, at 26°C in 26mm Petri dishes in a humidified atmosphere with 5% CO₂. A6 and XTC cells were subcultured with 0.25% trypsin and 0.03% EDTA for cell detachment. Medium was renewed twice per week. Cells were split in a dilution 1:2 every 3-4 days (Rafferty 1969; Pudney *et al.*, 1972). A6 and XTC cells were kindly provided by Barbara Hölscher.

4.3.3 ES Cell and Primary Embryonic Fibroblast Culture

Electroporation and isolation of stable clones using the murine ES cell line GSES was performed according to standard protocols with minor modifications (Muller *et al.*, 2000). Transgenic ES cells were grown on inactivated MEFs (3 hrs. 8µg/ml mitomycin C) in high glucose DMEM supplemented with 10% heat-inactivated ES qualified fetal calf serum, 2mM L-glutamine, 50U/ml penicillin, 50µg/ml streptomycin, 1x nonessential amino acids and 0.1mM β-mercaptoethanol. They were kept undifferentiated under feeder free conditions by addition of 1000U/ml recombinant mouse LIF. Cells were maintained at 37°C in a humidified atmosphere of 5% CO₂/95% air. Monolayers were passaged by trypsinization at 70–80% confluency. The growth medium for the attached differentiation cultures was changed every day. ES cells were kindly provided by Christiane Groß and Robert David, Grosshadern.

4.4 Embryological Methods

4.4.1 Solutions

Modified Barth's Saline (MBS): 5mM HEPES-KOH, 88mM NaCl, 1mM KCl, 0.7mM CaCl₂ (added freshly), 1mM MgSO₄, 2.5mM NaHCO₃ (pH 7.6 at 23°C)

Cystein: 2% L-Cystein in 0.1xMBS (pH7.8 at 23°C, adjusted with 5M NaOH)

Human Chorionic Gonadotropin (HCG): 1000U/ml HCG in ddH₂O

MBS/high salt: MBS with 50mM NaCl

0.1xMBS/Gentamycin: 0.1x MBS, 10 μ g/ml Gentamycin

MBS/CS: 0.8x MBS high salt with 20% chicken serum, 200U/ml penicillin and 200 μ g /ml streptomycin

4.4.2 Superovulation of Female *Xenopus laevis*

Xenopus laevis females were stimulated to lay eggs by injection of 500-800 units of human chorionic gonadotropin into the dorsal lymph sac and incubation at 18-20°C over night. Egg laying started about 10-18h later (Sieve *et al.*, 1989).

4.4.3 Preparation of Testis

A male frog was anaesthetized in 0.1% 3-Aminobenzoic acid-ethyl-ester in ddH₂O for 30min, cooled down in ice-cold water, and killed by decapitation. The two testes were taken from the abdominal cavity. Until use, the testes were stored in MBS/CS for maximal 7 days (Sieve *et al.*, 1989).

4.4.4 In vitro Fertilization of Eggs and Culture of the Embryos

For *in vitro* fertilization, a piece of testis was minced in 1x MBS and mixed with freshly laid eggs. Afterwards the embryos were cultured in 0.1x MBS at 16-23°C in 110mm Petri dishes.

4.4.5 Removal of the Egg Jelly Coat

One hour after fertilization or later, the egg jelly coat was removed in 2% cysteine solution pH 7.8 for about 5min with gentle agitation in an Erlenmeyer flask. Embryos were washed three times with 0.1x MBS and cultured further in 0.1x MBS/Gentamycin at 16-23°C.

4.5 Biochemical Methods

4.5.1 Solutions

3x Lämmli buffer: 150mM Tris/HCl (pH6.8 at 23°C), 300mM DTT, 4% SDS, 30% glycerol;

E1-buffer: 110mM KCl, 50mM Tris/HCl (pH7.4 at 23°C), 5mM MgCl₂, 0.1mM spermine, 0.1mM EDTA, 2mM DTT, 0.4mM PMSF;

PBS buffer: 140mM NaCl, 27mM KCl, 8mM Na₂HPO₄, 1.5mM KH₂PO₄ (pH7.4 at 23°C);

Glycine buffer: 15mM HEPES-KOH, pH 7.6, 10mM KCl, 5mM MgCl₂, 0.05mM EDTA, 0.25mM EGTA, 1mM DTT, 10% Glycerol, 0.4mM PMSF;

Sucrose buffer: 15mM HEPES-KOH, pH 7.6, 10mM KCl, 5mM MgCl₂, 0.05mM EDTA, 0.25mM EGTA, 1mM DTT, 30mM Surose, 0.4mM PMSF;

TE-buffer: 10mM Tris-HCl, pH 7.6, 1mM EDTA;

No-Salt Buffer: 3mM EDTA, 0.2mM EGTA;

High-Salt Solubilization Buffer: 50mM Tris-HCl, pH 8.0, 2.5M NaCl, 0.05% NP-40

High Salt Dialysis Solution: 10mM Tris-HCl, pH 8.0;

Acid Extraction Dialysis Solution: 100mM Acetic Acid, 1mM DTT;

HPLC solvent A: 0.065% TFA in H₂O;

HPLC solvent B: 84% ACN with 0.05% TFA;

4.5.2 *Nuclear Preparation from Cell Cultures*

Approximately 10⁷ cells of A6, XTC, ES cells or murine embryonic fibroblasts were collected by trypsin treatment. The cells were washed twice with PBS and then resuspended in 1ml PBS buffer containing 0.3% Triton; with freshly added Spermine, Spermidine, Sodiumbutyrate and Pepstatin, Leupeptin, Aprotinin and PMSF. The cells were put on a rotating wheel at 4°C for 20min and then centrifuged in an Eppendorf Centrifuge 5417C with top speed. The supernatant containing the cytoplasmatic fraction was discarded and the nuclear pellet was used either directly for SDS-PAGE (for protocol see 4.5.11), or for extraction with high salt (for protocol see 4.5.7) or acidic conditions (for protocol see 4.5.8).

4.5.3 *Nuclear Preparation from Xenopus laevis Embryos*

Around 50 to 100 embryos were collected at the required developmental stage defined by the normal table of Nieuwkoop and Faber (Nieuwkoop *et al.*, 1967) in 1.5ml Eppendorf tubes and washed with 1ml E-1 buffer containing 0.2% NP-40 and 0.25M sucrose as well as with freshly added Spermine, Spermidine, Sodiumbutyrate and

Pepstatin, Leupeptin, Aprotinin and PMSF (=complete protease inhibitors). Embryos were homogenized by pipetting in E-1/0.25M sucrose as well as complete protease inhibitors (20µl per 5 embryos) with a 200µl tip until the suspension appeared homogenous. Additional E-1/0.25M sucrose buffer was added to a final volume of 8µl per embryo, followed by vortexing for 10s on top speed (Vortex, Bachofer) with another 200µl E-1 with 0.05% NP-40 and 2.2M sucrose as well as complete protease inhibitors. The homogenate was layered on a 10µl cushion of E-1/2.2M sucrose as well as complete protease inhibitors in a 5x20mm centrifuge tube (Beckman) and centrifuged at 130 000g for 2h at 4°C in an Beckman ultra-centrifuge Optima MAX-E (TLA-55 rotor). Yolk and lipids were on top of the supernatant, which contained the cytoplasmatic fraction. The nuclear pellet was used for acidic extraction of core histone proteins (for protocol see 4.5.8).

4.5.4 *Cell Lysate from Xenopus laevis Embryos*

Around 200 embryos were collected at the appropriate stage according to Nieuwkoop and Faber (Nieuwkoop *et al.*, 1967) in 15ml Falcon tubes and washed in 3ml E-1 buffer containing 0.2% NP-40 and 0.25M sucrose as well as complete protease inhibitors. The embryos were transferred to a narrow 5ml glass-glass homogenizer together with 1ml E-1/0.25M sucrose as well as complete protease inhibitors and homogenized until the embryos were completely disrupted. The embryo lysate was used for acidic extraction of core histone proteins (for protocol see 4.5.8) and then further purified using a reversed-phase HPLC (see 4.5.10).

4.5.5 *Histone Purification by FPLC*

Between 800-1200 embryos were collected at the appropriate stage according to Nieuwkoop and Faber (Nieuwkoop *et al.*, 1967) in 15ml Falcon tubes. The embryos were washed 3x in glycine buffer and then re-suspended in 10ml glycine buffer, containing complete protease inhibitors. The solution was homogenized with a Homogenizer (Yamato) with 6 strokes at 1000rpm at 4°C. Afterwards, the homogenate was filtered through Mirocloth® to remove larger un-minced particles. The homogenate was centrifuged in a Sorvall RC-6+ centrifuge at 8000rpm for 10 min with the HP-4 rotor. The supernatant containing the cytoplasmatic fraction was discarded and the nuclear pellet was washed twice in Sucrose buffer and finally resuspended in 2ml Sucrose buffer. CaCl₂

was added to a final concentration of 1M, necessary for the correct function of the MNase, and the solution was incubated in a water bath at 27°C for 10min. 3µl MNase (50U/l) were added for the digestion of DNA and the reaction was stopped after 10 min at 27°C in the water bath by adding 20µl 0.5M EDTA. The suspension was centrifuged in an Eppendorf 5417C centrifuge with 8000rpm for 10min. The supernatant was discarded and the pellet was re-suspended in 1.5ml of TE buffer pH 7.6 and put on a rotating wheel for 30min at 4°C. The solution was centrifuged in an Eppendorf 5417C centrifuge with top speed for 10min. The supernatant, containing the fragmented chromatin, was used for further purification via FPLC. 8ml hydroxyapatite suspension was washed several times with 0.63M KCl/100mM KPO₄ with a total volume of 150ml. The hydroxyapatite was put in a 5ml FPLC column. The column was equilibrated with 0.63M KCl/100mM KH₂PO₄/K₂HPO₄ buffer for 120min with a flow rate of 0.5ml/min. The column was loaded with the supernatant containing the fragmented chromatin and washed with 0.5ml/min 0.63M KCl/100mM KH₂PO₄/K₂HPO₄ buffer for 100min. The core histones were eluted with a step elution to 2M KCl/100mM 100mM KH₂PO₄/K₂HPO₄ buffer. The histones were collected in 500µl fractions to separate them from DNA and other contaminating proteins. The eluted histones of the fractions were precipitated with TCA and the entire pellet was loaded onto SDS-PAGE (Simon *et al.*, 1979).

4.5.6 Nuclear Preparation from *Xenopus laevis* Embryos by Density Gradient Centrifugation

Around 50 to 100 embryos were collected at the required developmental stage according to the normal table of Nieuwkoop and Faber (Nieuwkoop *et al.*, 1967) in 1.5ml Eppendorf tubes and washed 3x with E-1/0.25M sucrose as well as complete protease inhibitors. The embryos were incubated in 1ml of E-1/0.25M sucrose buffer with complete protease inhibitors at room temperature for 20min and then transferred to a narrow 5ml glass-glass homogenizer. Another 2ml of E-1/0.25M sucrose buffer was added and the embryos were disintegrated according to Table 1. The correct number of strokes is crucial to minimize disruption of the nuclei and subsequent loss of histone proteins. The homogenate was transferred to a 15ml Falcon tube and centrifuged in a 3-18 Sigma centrifuge with 1000rpm for 10min at 4°C. The supernatant containing the cytoplasmic fraction was discarded and the pellet containing the nuclei was re-suspended in 3ml

E1/0.25M sucrose buffer containing 0.5% triton and 0.5% NP-40. The solution was incubated on ice for 20min. 5ml E1/1.25M sucrose buffer containing complete protease inhibitors was put in a 50ml Falcon tube as cushion and the nuclei suspension was carefully layered on top and centrifuged in a 3-18 Sigma centrifuge with 1000rpm for 30min at 4°C. The pellet was washed in 1ml E1 buffer and used for acidic extraction (for protocol see 4.5.8).

Stage	NF9	NF12	NF18	NF37
Amount of Embryos	200	100	100	50
Strokes for Disruption	10	20	20	20

Table 1: Specification for Nuclear Preparation

Amount of *Xenopus laevis* embryos for nuclear preparation by density gradient centrifugation and number of strokes for complete homogenization

4.5.7 High Salt Extraction of Histones

Nuclear pellets from cell lines (see 4.5.2) were re-suspended in 1ml of no-salt buffer and vortexed for 10s on top speed (Vortex, Bachofer) discontinuously. Afterwards the solution was incubated on a rotating wheel for 30min at 4°C and centrifuged in an Eppendorf 5417C centrifuge with top speed for 10min. The supernatant was discarded and the pellet was re-suspended in 400µl High-Salt Solubilization buffer. After vortexing, the solution was incubated on a rotating wheel for 30min at 4°C and centrifuged in an Eppendorf 5417C centrifuge with top speed for 10min. The supernatant was transferred to dialysis tubing and put in a beaker with 1l of high salt dialysis solution for about 3h. The dialysed histones were precipitated by adding TCA to a final concentration of 30%. The fractions were centrifuged in an Eppendorf 5417C centrifuge with top speed for 10min and the pellet was washed twice with 4% perchloric acid, twice with acidified acetone and twice with pure acetone.

4.5.8 Acidic Extraction of Histones

Nuclear pellets from embryonic material (see 4.5.6) were re-suspended in 1ml 0.4M HCl containing complete protease inhibitors and incubated on a rotating wheel overnight at 4°C. The next day, the solution was centrifuged in an Eppendorf 5417C centrifuge

with top speed for 10min, after which the supernatant was transferred to dialysis tubing and put in a beaker with 1l of acid extraction dialysis solution for about 3h. The dialysed histone solution was put in the Concentrator (Eppendorf) for 4-5h at room temperature and then stored at -20°C.

4.5.9 Extraction of Histones via Active Motif Histone Purification Kit

Around 800-1200 embryos were collected at the appropriate stage according to Nieuwkoop and Faber (Nieuwkoop *et al.*, 1967) in 15ml Falcon tubes. The histone extraction was performed as described in the instructions for tissue material according to the manufacturer's manual. 5ml Extraction buffer was added to the embryos and then completely homogenized in a narrow 5ml glass-glass homogenizer. The suspension was transferred to a 15ml Falcon tube and put on a rotating wheel over night at 4°C. The homogenate was centrifuged in a 3-18 Sigma centrifuge at top speed and the supernatant was transferred to a 15ml Falcon tube. Neutralization buffer was added until pH reached 7.0. The column was prepared as explained in the instruction manual, including pre-wetting with H₂O of the filter and equilibration with EB-buffer. The filter disc was put in the column tip and washed twice with water. The entire homogenate was passed through the column by gravity flow and the flow-through fraction was discarded. The column was washed with 10ml of Wash buffer and 5ml H2A/H2B Elution buffer was added on the column. The eluate was collected in 1ml fractions. Histones H3 and H4 were eluted with 5ml of H3/H4 Elution buffer and collected in 1ml fractions. 70% Perchloric acid was added to a final concentration of 4% to all sample fractions and put on a rotating wheel over night. The fractions were spun in an Eppendorf 5417C centrifuge with top speed for 10min and the supernatant was kept. Histones were precipitated by adding TCA to a final concentration of 30%. The fractions were cleaned in an Eppendorf 5417C centrifuge with top speed for 10min and the pellet was washed twice with 4% perchloric acid, twice with acidified acetone and twice with pure acetone. The pellet of the entire input was air dried and used for SDS-PAGE (for protocol see 4.5.11).

4.5.10 HPLC Purification of Histones

The histones, obtained by acidic extraction (see 4.5.8) and stored in 0.1% TFA, were loaded directly onto a C4 reversed-phase HPLC column (Phenomenex). The histone

proteins were eluted via HPLC (Ettan microLC), using 0.065% TFA in H₂O (solvent A) and 84% Acetonitril with 0.05% TFA (solvent B). The flow rate was set to 40µl/min after 10 min and solvent B was increased from 35% to 55% over a 15min period. After another 45min, the concentration of solvent B was increased to 65% over 30min and maintained for 15min. Then the solvent B concentration was set to 100% for 10min. 50µl fractions were collected. Fractions were lyophilized and subsequently used for SDS-PAGE and mass spectrometry.

4.5.11 SDS-PAGE

SDS-PAGE (SDS-polyacrylamide gel electrophoresis) was carried out according to standard protocols (Sambrook *et al.*, 1989), with 15% to 18% PAA gels to separate the small core histone proteins (40mA, 130-190V). The lyophilized protein was dissolved in 3x Laemmli and boiled up to 95°C for 5min in order to denature the proteins before loading into the stacking gel pockets. Approximately 5-20 embryo equivalents were loaded per slot for Coomassie staining and approximately 1-3 embryo-equivalents for Western-blotting.

4.5.12 Coomassie-staining

SDS-PAA gels were incubated over night at room temperature in 0.4g of Coomassie brilliant blue G250 dissolved in 200ml of 40% (v/v) methanol in water. Afterwards the SDS-PAA gels were destained in 60% ddH₂O, 30% methanol and 10% acetic acid for roughly 8h, with exchange of the destain-solution every 2 hours, stopped as soon as the background color had vanished.

4.5.13 Quantitative Western Blot Analysis

Western blot analysis was carried out according to standard protocols (Sambrook *et al.*, 1989). After SDS-PAGE proteins were blotted with 1.5mA/cm² for 1h to nitrocellulose membranes, which were blocked using PBS containing 3% BSA for 1h. Primary H4K20me3 antibody (1:500; Schotta *et al.*, 2008) and core region-H3 (1:25 000; Abcam) were incubated using PBS containing 3% BSA. Secondary fluorophore linked antibody (Li-cor, IRDye 700Dx goat anti-rabbit, 1: 5 000) was incubated in PBS with 5%

milk powder, 0.1% Tween 20. Quantification was done using Li-cor Odysse wet membrane technique for 700nm absorbance.

4.6 MALDI-TOF

4.6.1 *Preparation of Proteins for Linear Mode Mass Spectrometry*

For mass spectrometric preparations HPLC-grade fine chemicals and water were used. The proteins in the 50 μ l fractions from reversed-phase HPLC purification (for protocol see 4.5.10) were lyophilized. Then the proteins were re-dissolved in 1-2 μ l of 50% ACN/0.1% TFA saturated with sinapinic acid and eluted onto a hydrophobic target plate. The sinapinic acid forms crystals with the proteins and allows a soft-ionization technique.

4.6.2 *Preparation of Histones for Proteolysis*

The protein bands were cut out of the Coomassie brilliant blue stained SDS-PAGE and minced into small pieces (1mm x 1mm x 1mm) with a clean scalpel. Gel pieces were covered in water to prevent them from drying out. The gel pieces were washed twice with 200 μ l H₂O by shaking them in a Shaker (Thermomixer comfort, Eppendorf) at 650rpm at 37°C for 5min. For neutralization of acidic components, 100 μ l of 10mM Ammoniumbicarbonate (Ambic) was added to the gel pieces. The subsequent destaining of the gel pieces from the Coomassie stain was done with 100 μ l of 0.1M Ambic and 100 μ l of HPLC-grade Acetonitrile (ACN). Afterwards the sample was shaken in a Shaker (Thermomixer comfort, Eppendorf) at 650rpm 37°C for 30-90 min. After destaining, gel pieces were washed twice with 200 μ l H₂O and shaken in a Shaker (Thermomixer comfort, Eppendorf) at 650 rpm 37°C for 1min. The gel was dehydrated by application of 100 μ l ACN until the gel pieces appeared white (2-3 times). Excess ACN was removed with a small gel loader tip.

4.6.3 *Digestion with Trypsin*

For detection of the core histones in MALDI-TOF the peptide reflector mode was used. In order to apply this method the proteins had to be digested prior to analysis. The core histones were propionylated to chemically modify lysines before digestion with

trypsin. As histones are very basic proteins, they contain a high number of lysines and arginines. The endoprotease trypsin cleaves after every lysine and arginine side chain, except if a proline is following. Thus, untreated histones are digested into many peptides too small for convenient detection with MALDI-TOF. Therefore, lysines were blocked for trypsin digestion with propionic anhydride in a reaction called propionylation.

Lysines in histones can either be unmodified, acetylated or mono-, di- or trimethylated. Propionic anhydride can react only with unmodified or mono-methylated lysine residues; other modifications of lysines are not recognized by the proteinase trypsin. Thus cleavages occur only after unmodified arginine residues. Propionylation also happens at the N-terminus of each peptide and gives a mass shift of 56.0262Da that has to be included in the calculation of to the expected peptide masses (Villar-Garea *et al.*, 2008). For the propionylation reaction, 1 μ l propionic anhydride and 49 μ l of 1M Ambic were added to the gel pieces. The propionylation mixture was incubated in a Shaker (Thermomixer comfort, Eppendorf) at 37°C and 650rpm for one hour. Afterwards, the gel pieces were washed with 50 μ l 0.1M Ambic thoroughly. 100 μ l of ACN was added and removed again until the gel pieces appeared white. When ACN was completely removed, the gel pieces were air dried in room air.

All core histones were digested with the endoproteinase trypsin. Trypsin was added to the reaction on ice to avoid self digestion. 1 μ l of 0.2 μ g/ μ l trypsin in 50 μ l 0.1M Ambic was pipetted on the dried gel pieces until complete absorption. The digestion mix was incubated in a Shaker (Thermomixer comfort, Eppendorf) for at least 3h or over night at 37°C and 650rpm. The reaction was stopped by freezing the solution.

4.6.4 Preparation of Peptides for Mass Spectrometry

4.6.4.1 Acid Extraction of Trypsin-digested Peptides

To maximize the recovery, an acidic extraction was performed after digestion. The supernatants of the different steps were pooled in a 0.5ml LoBind tube. First, the supernatant of the digestion was put in the LoBind tube. 30 μ l 25mM Ambic was applied on the gel pieces and incubated in a Shaker (Thermomixer comfort, Eppendorf) for 15min at 37°C and 650rpm. 40 μ l ACN were added and the gel pieces were shaken in a Shaker (Thermomixer comfort, Eppendorf) for 15min at 37°C and 650rpm. The supernatants were pooled in the LoBind tubes with the trypsin mixture. 50 μ l of 5 % formic acid were applied

to the gel pieces and incubated for 15min at 37°C 650rpm then 50 μ l of ACN was added and incubated in a Shaker (Thermomixer comfort, Eppendorf) for 15min at 37°C and 650rpm. The supernatants were pooled and the gel pieces were incubated a second time with 30 μ l ACN, which was then transferred to the other supernatants. The LoBind tubes were then lyophilized in a speed vac until complete evaporation. The pellets were thoroughly re-dissolved in 10 μ l 0.1% TFA and stored at -20°C until further analysis.

4.6.4.2 Zip Tip Extraction

As a preparative step for MALDI-TOF analysis, peptides were desalted and concentrated using ZipTips. The ZipTip contains a stationary phase which binds the peptides. The bound peptides can then be washed and eluted onto a hydrophobic target plate. The peptides of interest required a hydrophobic stationary phase, such as a μ -C18 column. The ZipTip volume is 10 μ l. Solutions for wetting, equilibrating and washing were prepared in 1.5ml tubes before starting. Wetting buffer consisted of 50 % ACN, the stationary phase was equilibrated and washed with 0.1% TFA. The sample was diluted in 0.1% TFA and eluted using 50% ACN/0.1% TFA. For the α -4-hydroxy-cinnamic acid matrix, 50% ACN/0.3% TFA together with a small spatula of the α -4-hydroxy-cinnamic acid powder were vortexed for 60 seconds to achieve a saturated solution. The tip was first wetted three times in 50% ACN. To equilibrate the tip, 0.1% TFA was applied to the tip. Then, the sample was loaded to the tip-column. To remove contamination, the column was washed with 0.1% TFA. Finally the peptides were eluted using 2 μ l elution buffer and spotted onto the target plate. After complete drying, 1 μ l of matrix solution was applied on the spot.

4.6.4.3 Carbon Tip Extraction

Carbon Tips were used to enrich the peptide aa 3-8 of the histone H3, which was hardly detected in MALDI-TOF because of its low ionization efficacy. After digestion with trypsin, the peptide solution was used for Carbon Tip enrichment. The Carbon Tip was wetted with 20 μ l 40% ACN, 30% 0.6% TFA and 20% H₂O. The peptides were loaded onto the Carbon Tip matrix. To remove contamination, the Carbon Tip was washed with 0.1% TFA. The peptides were eluted from the Carbon Tip matrix with 20 μ l 0.1% TFA followed by a second elution step with 20 μ l 40% ACN, 30% 0.6% TFA and 20% H₂O. The

eluted peptides were lyophilized and reconstituted in 10 μ l of 0.1% TFA and spotted as described above.

4.6.5 Analysis by MALDI-TOF

MALDI-TOF is a standard technique in protein and peptide analysis, which uses the assistance of a matrix for softened and efficient laser desorption and ionization. The peptides or proteins are embedded in a matrix, which absorbs the energy of the laser shot. The energy of the laser is then transferred to the peptides/proteins via the matrix, leading to ionization. Moreover, the matrix increases the intensity of single peptides/proteins in the spectra and facilitated better analysis. The matrix is selected depending on the mode in which the mass spectrometer is operated and the composition of the analyte (Lottspeich *et al.*, 2006).

4.6.5.1 Linear Mode MALDI-TOF

Undigested proteins were analysed using the linear mode MALDI-TOF. Depending on the instrument settings, a mass up to 50 000Da can be detected. The core histones have a molecular mass between 11kDa and 16kDa, these settings were sufficient for histone detection. As matrix for the soft-ionization we used sinapinic acid. The detailed instrument settings of the instrument are provided in Table 2. Analyzed spectra resulted from 500 laser shots.

Program Mode: Linear Mode	
Acceleration Voltage:	25 000 V
Grid Voltage:	90%
Mass Range:	1000-50000 Da
LASER Intensity:	800-1200 W

Table 2: Settings of Linear Mode

4.6.5.2 Peptide Reflector Mode MALDI-TOF

The peptide reflector mode is especially suited for the detection of peptides because of its higher resolution. The unique combination of peptides originating from the

digested protein like a unique fingerprint allows the identification of the protein as well as detection of peptide modifications of known proteins. The instrument settings are displayed in Table 3.

Program Mode: Peptide Reflector Mode	
Acceleration Voltage:	20 000 V
Grid Voltage:	66%
Mass Range:	500-3500 Da
LASER Intensity:	800-1200 W

Table 3: Settings of MALDI-TOF for Peptide Reflector Mode

4.6.6 *Quantitation of MALDI-TOF Data*

Spectra were analyzed using the Data Explorer software. First, spectra were de-isotoped. For internal calibration the two peaks 842.51Da and 2211.10Da of trypsin were used. The advantage of this calibration is that the two peaks are adjacent to the detection limits of the spectra (500-3500Da) and, thus, lead to adequate calibration in between the range of most of the peptides of interest. For quantification of the different post-translational modifications of the various peptides obtained after digestion, the relative intensities of each PTM were taken into account. For this purpose, the software integrates the area under the curve reflecting the number of extracted ion counts (XIC), which correspond to the relative intensities. The areas of all modifications of one peptide were summarized and from this, the relative abundance of each modification was calculated in percentage.

4.7 Orbi-Trap

4.7.1 *Sample Preparation for Orbi-Trap*

Samples for Orbi-Trap analysis were digested with the proteinase trypsin followed by Zip-Tip treatment to remove the salts.

4.7.2 *Analysis by Orbi-Trap*

The sample was lyophilized and stored or reconstituted in 15 μ l 0.1% TFA. 10 - 15 μ l of the sample were loaded to the reversed-phase HPLC in an analytical C18 micro column (75 μ m packed with C18 PepMapTM, 3 μ m, 100 \AA by LC Packings) and with a 80min gradient from 5 to 60% ACN in 0.1% FA. The eluate from the LC was directly electro-sprayed into the LTQ-Orbitrap mass spectrometer. The MS instrument was operated in data dependent mode to automatically switch between full scan MS and MS/MS acquisition. Survey full scan MS spectra (from m/z 300 - 2000) were acquired in the orbitrap with resolution $R=60,000$ at m/z 400. The six most intense peptide ions with charge states between 2 and 4 were sequentially isolated and fragmented in the linear ion trap by collision induced dissociation (CID). All fragment ion spectra were recorded in the LTQ part of the instrument. For all measurements with the orbitrap detector, 3 lock-mass ions from ambient air ($m/z=371.10123$, 445.12002, 519.13882) were used for internal calibration as described. Typical mass spectrometric conditions were: spray voltage, 1.5 kV; no sheath and auxiliary gas flow; heated capillary temperature, 200°C.

4.7.3 *Identification of Histone PTMs by Orbi-Trap*

For the analysis of different histone modifications the resulting raw files (.raw) were converted into data files (.dta) and searched against the NCBI database using the SEQUEST search algorithm to identify the corresponding histones using the BLOWORKS 3.3 software package (Thermo). In order to determine the post-translational modifications on the histone peptides, the data files (.dta) were then re-searched against a targeted histone database containing all histone molecules from the NCBI database with the following modification settings (static modification on lysine: 56.0262; variable modifications: me1: 14.01565; me2: -27.99490; me3: -13.97925; ac: -14.10565 on K and phos: 79.96633 on S). The resulting SEQUEST files (.srf) were filtered for a peptide probability score of 0.0005 and X correlation values of 1.5, 2.0, 2.5 and 3 for charge states 1, 2, 3 and 4 accordingly. In the cases where a MS/MS spectrum could be interpreted by two different modifications with both having a probability score of lower than 0.0005, only the one with the lowest score (top hit) was counted. A similar search against a reversed histone decoy database did not lead to an identification of peptides that matched the filter criteria. Therefore the false discovery rate was estimated to be less than 1%. For all peptides that

were quantified an extracted ion chromatogram (XIC) was obtained from the raw file using the Xcalibur 2.0.7 software (Thermo) extracting doubly and triply charged ions with a user defined mass tolerance of 5-10ppm and a mass precision of 4 decimals to be able to distinguish between modifications such as acetylation or tri-methylation with a mass difference of 0.043Da. For quantification particular attention was paid to the retention times of the corresponding peptides from the pep map column. The retention time for each specific peptide varied less than one minute when the same LC-MS/MS method was used. Therefore, only these XIC were used for quantification, which eluted at time points where at least two MS/MS spectra that unambiguously identified the identity of the peptide and the position of the modification within the peptide. For a relative quantification of the modifications the same method as described above was used (4.6.6).

4.8 Heatmap Generation

The Heatmap shown in Figure 24 was generated using R (<http://www.r-project.org>) and the *gplots* package (<http://CRAN.R-project.org/package=gplots>). All functions were called using default parameters if not indicated otherwise. First, the mean and the standard deviation of the quantifications obtained by mass spectrometry across the four developmental cell stages were calculated (each stage containing two biological replicates). Histones with small mean values contain low mass spectrometry quantifications that could have obscured the clustering analysis that were observed in the heatmap. Thus, as a preprocessing step of the data, modifications with mean values less than 0.5% (22 modifications) were discarded. In order to compare appropriately the contribution of each modification to the the specific cell stages, the mass spectrometry quantifications were standardized using individual Z-scores. Using these Z-scores, the Heatmap was produced using the Ward's minimum variance method to perform the hierarchical clustering analysis.

5 Results

5.1 Histone Extraction from Stable Cell Lineages

Several methods have been described in the literature for enrichment and purification of histone proteins from various sources. The objective of this thesis was to identify histone modifications using mass spectrometry and thus purity as well as enrichment of histone proteins was important for this analysis. As different sources for histone purification such as early embryos or cell cultures were used, which all have different features considering the abundance of the histones and the amount and abundance of contaminating moieties such as fat or yolk, the extraction of histones was considered to be a major topic. To test different common extraction methods, two different *Xenopus laevis* cell lines, XTC and A6, were used. These cells are easily grown under standard culture conditions and furthermore have only little contaminants such as yolk and lipids compared to differently staged embryos. Moreover, a protocol was established which purifies nuclei of the two cell lines. To compare different extraction protocols, nuclei were all obtained using this protocol. Following the nuclei extraction, different protocols for histone extraction were applied and their quality was evaluated by Coomassie stained SDS-PAGE and MALDI-TOF spectra of these histones. After all, four different protocols were tested and evaluated with extraction of entire nuclei without further extraction, high salt extraction and three different protocols for acidic extraction.

5.1.1 Direct Lysis of Nuclei

As histones are a major fraction of nuclear proteins, boiling of entire nuclei in 3x Laemmli buffer is expected to yield high amounts of histone proteins (Rodriguez-Collazo *et al.*, 2009). Direct lysis is done by preparation of nuclei of the cell lines followed by lysis of the nuclei via boiling in 3x Laemmli buffer. The nuclear and cellular proteins were then used for Coomassie stained SDS-PAGE (see Figure 5A).

First, the reliability of the nuclear preparation protocol was evaluated using phase contrast microscopy and DAPI (4',6-Diamidino-2-phenylindol) staining of the nuclei (see Figure 5B). Nuclei were clearly visible in both phase contrast and DAPI stain pointing to successful nuclear preparation. The theoretical masses of the core histone proteins range from 11kDa for Histone H4 and 16kDa for Histone H3. In Coomassie stained SDS-PAGE,

the four core histone give a characteristic pattern, which was visible for the direct lysed nuclei in Figure 5C lane 1 and 3. The direct lysis method was compared to acidic extraction in Figure 5C lanes 2 and 4. The four core histone proteins were detected in the Coomassie stained SDS-PAGE, but the concentration of histone proteins was much lower than in acidic extraction for the same input (Compare lanes 1 and 3 with 2 and 4 in Figure 5C).

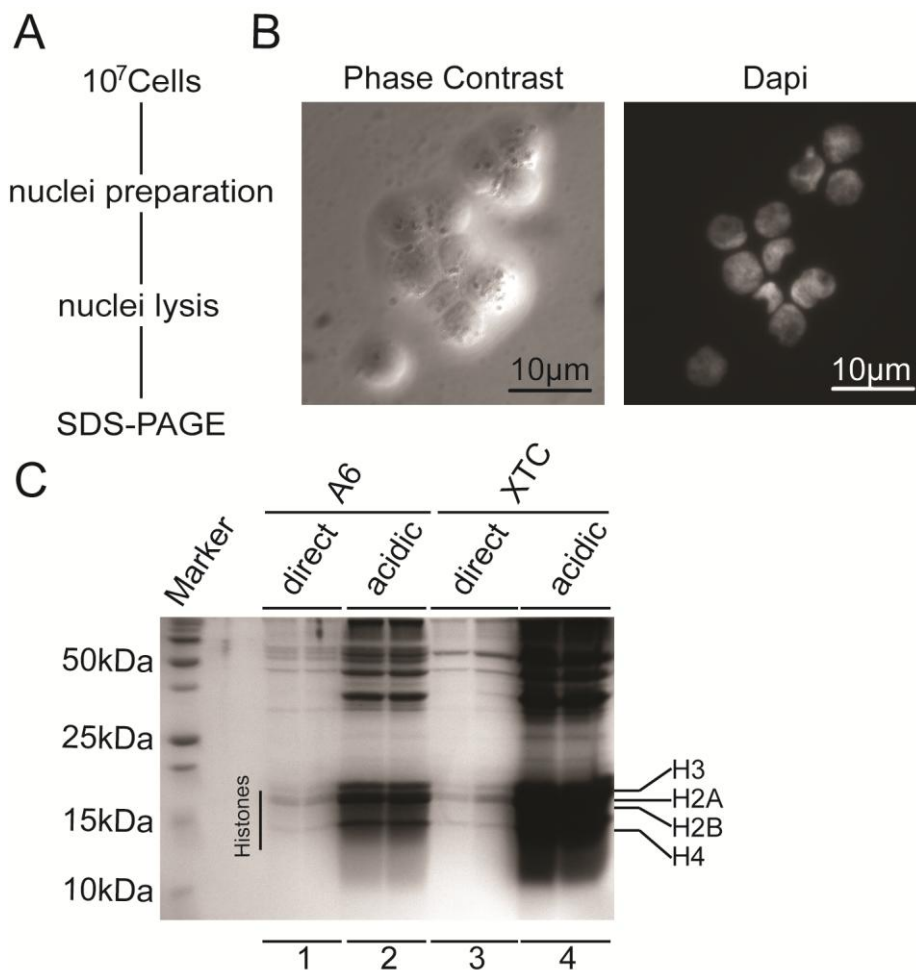


Figure 5: Preparation of Nuclei from Stable Cell Lineages and Direct Lysis for Histone Purification

A) Flowchart of nuclei preparation from stable cell lineages and histone purification with SDS-PAGE. **B)** Microscopic pictures of nuclei after nuclei preparation of A6 cell line in phase contrast and stained with DAPI. The preparation was successful as the nuclei are clearly visible. **C)** Coomassie stained SDS-PAGE of one biological repeats of histones of A6 and XTC in direct lysis compared to hydrochloric acid extracted histones. The yield of core histone proteins of the same amount of cells (each lane 10⁶ cells) was much higher in the hydrochloric acid extracted samples compared to the directly lysed ones.

Mass spectrometry was used to identify the trypsin digested proteins from Coomassie stained SDS-PAGE with the use of the MASCOT database search. The sequence of the four proteins of the SDS-PAGE matched the core histone sequences significantly. The reproducibility, signal to noise ratio and the quality of the internal calibration of the MALDI-TOF spectra were evaluated. For the spectra of histones extracted with direct lysis the quality of internal calibration was low and only 50% of the spectra had sufficient signal to noise ratio for analysis. Furthermore the recovery of modifications of technical replicates was low with a correlation coefficient of 0.702.

In summary, direct lysis was sufficient for histone extraction of cell lineages. These results have also been confirmed by mass spectrometry and subsequent Database search resulting in significant identification of histone proteins. However, the histones purified with this protocol could not be used to quantify histone modifications as the spectra quality was not sufficient. A considerable advantage of direct lysis is its ease of use and the quick protocol. On the other hand, comparing the yield of histone proteins in the SDS-PAGE of this method to the one of acidic extraction, it is approximately 5-10 times lower with equal amount of input (Compare Figure 5C, lane 1 and 3 to lane 2 and 4).

5.1.2 High Salt Extraction of Histones

The method of high salt extraction for histone proteins, first described by Burton in 1978, uses buffers containing high salt concentrations to dissolve the DNA-histone interaction. The use of buffers containing high concentrations of salt results in precipitation of the contaminants, whereas the four core histones remain soluble and thus can be enriched and purified (Burton *et al.*, 1978).

For high salt extraction, nuclei were prepared and disrupted as described above followed by histone extraction with the high salt buffers and Coomassie stained SDS-PAGE (see Figure 6A). SDS-PAGE with Coomassie staining revealed the characteristic core histone pattern visible between 11kDa and 16kDa (Figure 6B).

Mass spectrometry of tryptic digested proteins was performed. Using the MASCOT Database Search, the sequences of the four proteins were significantly matched to the sequences of the core histones. Compared to the amount of histones present in the SDS-PAGE, this extraction method yielded around twice as much compared to direct lysis (see Figure 5C).

The reproducibility, signal to noise ratio and the quality of the internal calibration of the MALDI-TOF spectra was evaluated for the aberration of the expected mass to the measured peaks. For the spectra of histones extracted with high salt extraction, the quality of internal calibration was low, but around 70% of the spectra had sufficient signal to noise ratio for analysis. The recovery of modifications of technical replicates was low, but better than in direct lysis with a correlation coefficient of 0.79.

In summary, high salt extraction as described by Burton yielded more core histone proteins compared to direct lysis (Burton *et al.*, 1978). The quantity of histone proteins and the quality of spectra, however, was not sufficient for analysis of post-translational histone modifications.

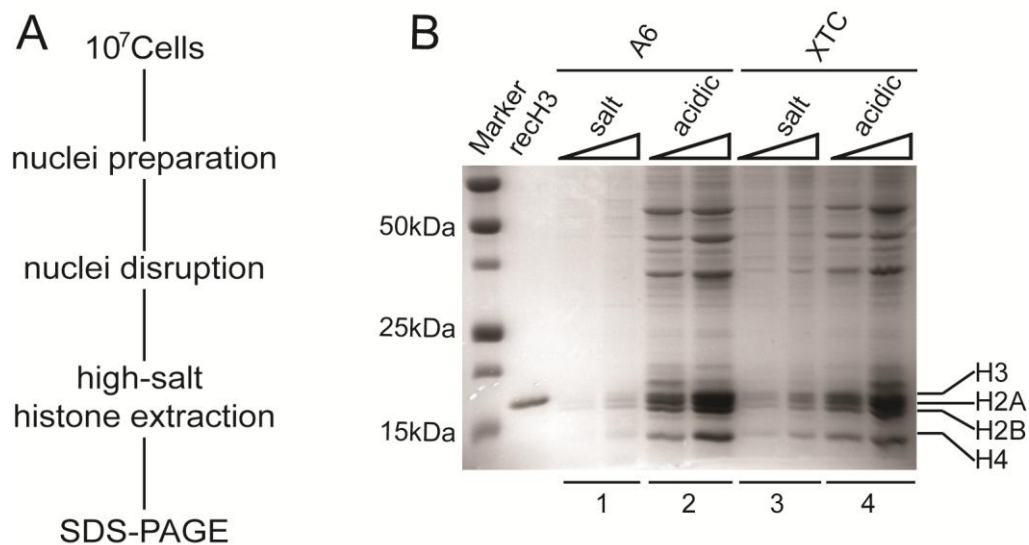


Figure 6: High Salt Extraction of Histones of Stable Cell Lineages

A) Flowchart of nuclei preparation and histone purification with buffers containing high salt concentrations and SDS-PAGE. **B)** Coomassie stained SDS-PAGE of high salt extracted histones of A6 and XTC versus hydrochloric acid extracted histones in two concentrations (1×10^6 cells and 4×10^6 cells). The yield of core histone proteins of same amount of cells was higher in the hydrochloric acid extracted samples compared to the high salt extracted ones.

5.1.3 Acid Extraction of Histones

Acid extraction is used frequently for histone proteins and is the method of choice according to Lindner (Lindner *et al.*, 2008). This method is described to be easy to use and to have a high yield. Acidic extraction relies on the high amount of basic amino residues mainly in the N-termini of the histone proteins as this feature leads to solubility despite the presence of high concentrations of acid. Furthermore it is described to be especially well suited for samples containing only little amounts of chromatin and thus might be useful for histone extractions of low cell numbers or young embryos such as blastulae or 10^6 cells. In the literature, the use of three different acids is described for the extraction. These are hydrochloric acid and sulphuric acid for core histone purification (Ausio *et al.*, 1998) and perchloric acid for separate enrichment of linker histones (Lindner *et al.*, 1998). Perchloric and hydrochloric acid extraction in series was evaluated first, to find out whether the two step approach led to satisfying results.

5.1.3.1 Acid Histone Extraction with Perchloric Acid

Perchloric linker histone extraction is performed by preparation and disruption of nuclei. As perchloric acid extraction is supposed to extract only the linker histone fraction, it is followed by either hydrochloric or sulphuric acid extraction of core histones. Histone proteins were separated by Coomassie stained SDS-PAGE (see Figure 7A; Lindner *et al.*, 1998).

In the Coomassie stained SDS-PAGE, there was a protein band visible correlating to the 40kDa Marker band. Mass spectrometric analysis of the tryptic digested putative linker histone identified the protein sequence as the linker histone H1 by MASCOT Database Search. Furthermore, the characteristic pattern of the four core histones was present in much lower concentration than the linker histones in the perchloric extraction. Thus linker histones were enriched in the perchloric acid extraction. The pellet after perchloric linker histone extraction and hydrochloric core histone extraction showed few core histones and no linker histones. As the focus of this thesis was mainly on the four core histones, sequential elution with first perchloric linker histone extraction followed by hydrochloric core histone extraction led to a separation of the core histones in fractions.

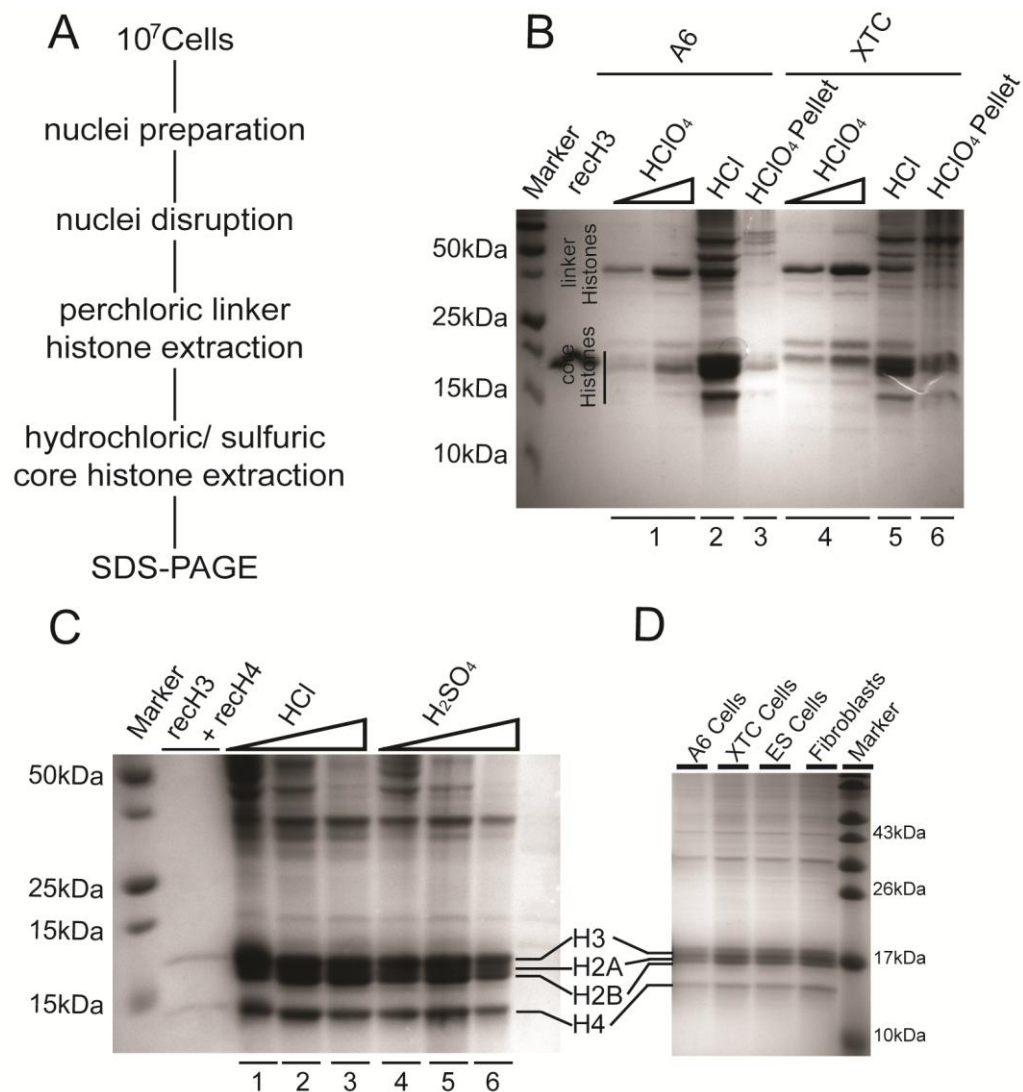


Figure 7: Acidic Extraction of Histones from Somatic Cell Lines

A) Flowchart of nuclei preparation and subsequent linker histone extraction with perchloric acid and core histone extraction via hydrochloric or sulfuric acid. **B)** Coomassie stained SDS-PAGE of increasing concentration (1×10^6 cells and 4×10^6 cells) of linker histones extracted with perchloric acid (lanes 1, 4) and core histones extracted with hydrochloric acid (lanes 2, 5) of A6 and XTC nuclei. The nuclear pellet after both extractions (lanes 3, 6) had core histones left, but no linker histones. The linker histone extraction (lanes 1, 4) enriched core histones. The hydrochloric extraction (lanes 2, 5) did still contain linker histones, despite of the separate linker histone extraction step and yielded a high amount of core histone proteins. **C)** Coomassie stained SDS-PAGE of histone proteins extracted with hydrochloric (lanes 1-3) and sulfuric acid (lanes 4-6). The concentration of acid used for extraction increased in the lanes 1-3 and 4-6 from 0.4M over 0.8M to 1.2M acid. As the concentration of the acid used for the extraction increased so did the purity of the histone extraction, mainly seen at the proteins present above the 25kDa marker band. **D)** Coomassie stained SDS-PAGE of the core histones extracted by acidic extraction with 0.4M HCl of the four different cells used.

Potentially, specific post-translational histone modifications might be extracted together with the linker histones in perchloric extraction, leading to biased post-translational histone modification analysis.

In summary, the method of linker histone extraction with perchloric acid in series with core histone extraction with hydrochloric acid yielded linker histones as described by Lindner (Lindner *et al.*, 1998). No linker histones were lost in the pellet, which did not contain any linker histone in the Coomassie stained SDS-PAGE. However, as linker histones were also present in the extraction with hydrochloric acid, the linker histones were separated in several fractions. As the focus of the thesis was on core histones, the two step extraction was not superior to a one step hydrochloric histone extraction (Figure 7B and C).

5.1.3.2 *Acid Histone Extraction with Hydrochloric Acid*

As described above, the sequential extraction of histones with perchloric acid, followed by hydrochloric acid has led to a low yield of core histones. In a next experiment the yield of hydrochloric extraction on its own was compared to sequential acidic histone extraction.

As depicted in Figure 7C, the characteristic pattern of the four core histone proteins was visible in the Coomassie stained SDS-PAGE. The tryptic digested proteins were analysed using mass spectrometry and the sequences matched significantly to the core histones by MASCOT Database Search. The yield of the core histones was much higher compared to directly lysed nuclei (see Figure 5C, lane 1 and 3), high-salt extraction (see Figure 6B, lane 1 and 3) as well as in comparison to the two-step perchloric-hydrochloric extraction (see Figure 7B). However, there were more other protein bands visible in the Coomassie stained SDS-PAGE compared to other extraction protocols (compare Figures 5C, 6B and 7C). To possibly reduce the amount of other proteins, different concentrations of hydrochloric acid, 0.4M, 0.8M and 1.2M respectively, were tested. The results are depicted in the Coomassie stained SDS-PAGE (Figure 7C, lanes 1-3). The protein bands visible between 40kDa and 60kDa of the Marker band were reduced as the concentration of hydrochloric acid was increased. The amount of the four core histone proteins was not changed by the increased hydrochloric acid concentration in the Coomassie stained SDS-PAGE. Tryptic digested histone proteins were analysed using

mass spectrometry. MASCOT database search matched the sequences significantly to the four core histone proteins, confirming the enrichment of core histones at higher acid concentrations.

The reproducibility, signal to noise ratio and the quality of the internal calibration of the MALDI-TOF spectra was evaluated for the different hydrochloric acid concentrations. As the concentration of hydrochloric acid was increased, the reproducibility of histone spectra, the quality of internal calibration and the signal to noise ratio for analysis was reduced. The recovery of modifications of technical replicates of 0.4M hydrochloric extraction had a correlation coefficient of 0.96 and thus was much better than of all other extraction methods tested.

Summarizing, hydrochloric acid extracted core histones in high amounts and resulted in high quality of MALDI-TOF spectra for post-translational histone modification analysis. The use of 0.4M hydrochloric acid was on the one hand sufficient for enrichment and purity of histones and on the other hand resulted in best mass spectrometric analysis conditions.

5.1.3.3 *Histone Extraction with Sulfuric Acid*

Sulfuric acid can also be used to extract core histones, although there are fewer remarks on this compared to hydrochloric acid (H. Gould, 1998). To evaluate differences in the two acidic core histone protocols, the method using sulphuric acid was performed in the same way as hydrochloric acid extraction (see Figure 7A).

Coomassie stained SDS-PAGE of three different concentrations of sulfuric acid, 0.4M, 0.8M and 1.2M revealed the characteristic core histone pattern (see Figure 7C). Mass spectrometric analysis of tryptic digested proteins followed by MASCOT Database search confirmed the sequence of the core histone proteins. The different concentrations of sulfuric acid, 0.4M, 0.8M and 1.2M resulted in reduced protein bands visible between 40kDa and 60kDa of the Marker band as the concentration of hydrochloric acid was increased in the Coomassie stained SDS-PAGE (Figure 7C). The amount of the four core histone proteins was not changed by the increased sulfuric acid concentration in the Coomassie stained SDS-PAGE.

The reproducibility, signal to noise ratio and the quality of the internal calibration of the MALDI-TOF spectra were evaluated for the different sulfuric acid concentrations. As the

concentration of sulfuric acid was increased, the reproducibility of histone spectra, the quality of internal calibration and the signal to noise ratio for analysis was reduced. The recovery of modifications in technical replicates was comparable to hydrochloric extraction. However the correlation coefficient of sulphuric acid extracted core histone was 0.89 and thus was slightly lower as of hydrochloric extracted histones.

Certainly, both hydrochloric and sulphuric acidic extraction methods for core histones were comparable in regard to yield and purity of histones, as well as reliability and internal calibration of mass spectra. However the reproducibility of hydrochloric acid extracted histones was superior to sulphuric extracted histones.

In summary, all tested protocols were able to yield core histones in Coomassie stained SDS-PAGE. However, the four different protocols produced very different amounts of histones with lowest yield in directly lysed nuclei and the highest yield in both hydrochloric and sulphuric acid extractions, differing at least by an order of magnitude. Using tryptic digested proteins for mass spectrometry followed by MASCOT Database Search the four core histones were identified significantly. However there were significant differences between the reproducibility of the spectra acquisition and the analysis of post-translational histone modifications using mass spectrometry. The highest quality as well as best correlation of different histone modifications of spectra was achieved with the method of 0.4M hydrochloric acid core histone extraction.

For further core histone extraction from cell nuclei, 0.4M hydrochloric acid was used (Figure 7D). One additional advantage of hydrochloric acid extraction is that this method is frequently used for histone preparations and thus is best suited to compare our results to published data (Zhang *et al.*, 2003, Bonaldi *et al.*, 2004).

5.2 Histone Extraction of Embryos of Different Stages

Aiming to establish a histone modification profile of *Xenopus laevis* embryos of four different stages - blastula, gastrula, neurula and tadpole stage - there was the need for an easy, reliable and efficient protocol to obtain histones of these embryos (see Figure 8). The embryos consist of much more proteins compared to cultured cells as well as of yolk and fat. Thus, histone extraction of embryos requires a higher degree of purification and enrichment compared to histone extraction of stable cell lines. As there was hydrochloric

extraction from nuclei of cell lines described previously to be a method with high yield, high purity and sufficient quality in mass spectrometry, the first attempt was to apply this method to embryos.

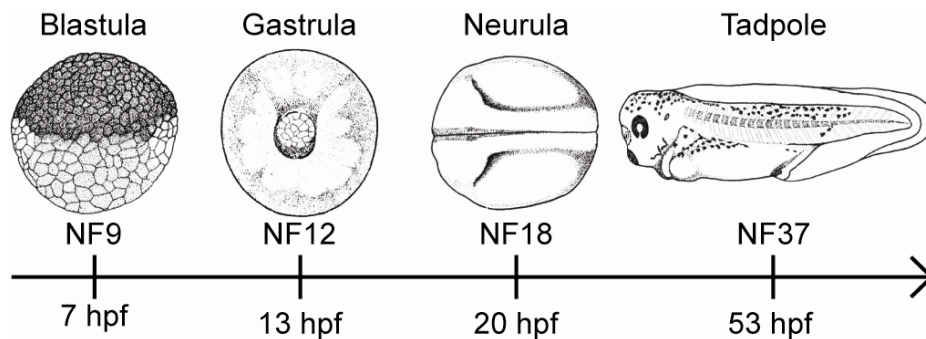


Figure 8: Xenopus Developmental Time Line

Time line of *Xenopus laevis* embryonic development – NF stages according to Nieuwkoop and Faber and by hours after fertilization (hpf). The stages selected for mass spectrometric analysis are characterized by the following embryonic and cellular features: blastula (NF9) naïve/multipotent cells; gastrula (NF12) germ layers specified; neurula (NF18) germ layer patterning and differentiation; tadpole (NF37) embryonic development completed, larvae hatched. (Pictures taken from <http://www.xenbase.org/xenwiki/index.php>)

5.2.1 Nuclei Extraction with an Ultracentrifuge Step

In our lab, protocols established by Blobel were used previously for nuclei preparation of different embryonic stages of *Xenopus laevis* (Blobel *et al.*, 1966). The method uses staged cohorts of embryos which were then homogenized by pipetting. The nuclei were then separated from the cell debris by ultracentrifugation on a density gradient of sucrose (Figure 9A). Before applying hydrochloric acid extraction to the nuclei, the nuclei were visualized in phase contrast microscopy and by DAPI staining (see Figure 9B). Phase contrast microscopy revealed the presence of many vesicles and cell debris in addition to nuclei, with the DAPI fluorescent images identifying the nuclei. These results indicate that the method quite effectively disrupted embryos, tissues and cells, but failed to separate nuclei from cellular debris. The pellet of the centrifugation was used for hydrochloric acid extraction to enrich and purify the histone proteins (Murray *et al.*, 1966). The Coomassie stained SDS-PAGE of the acidic extraction showed the characteristic core histone (see Figure 9C). However, there was a wide spread of other protein bands visible in the SDS-PAGE. Thus, the histone extraction of embryos is not as pure as for cell lines.

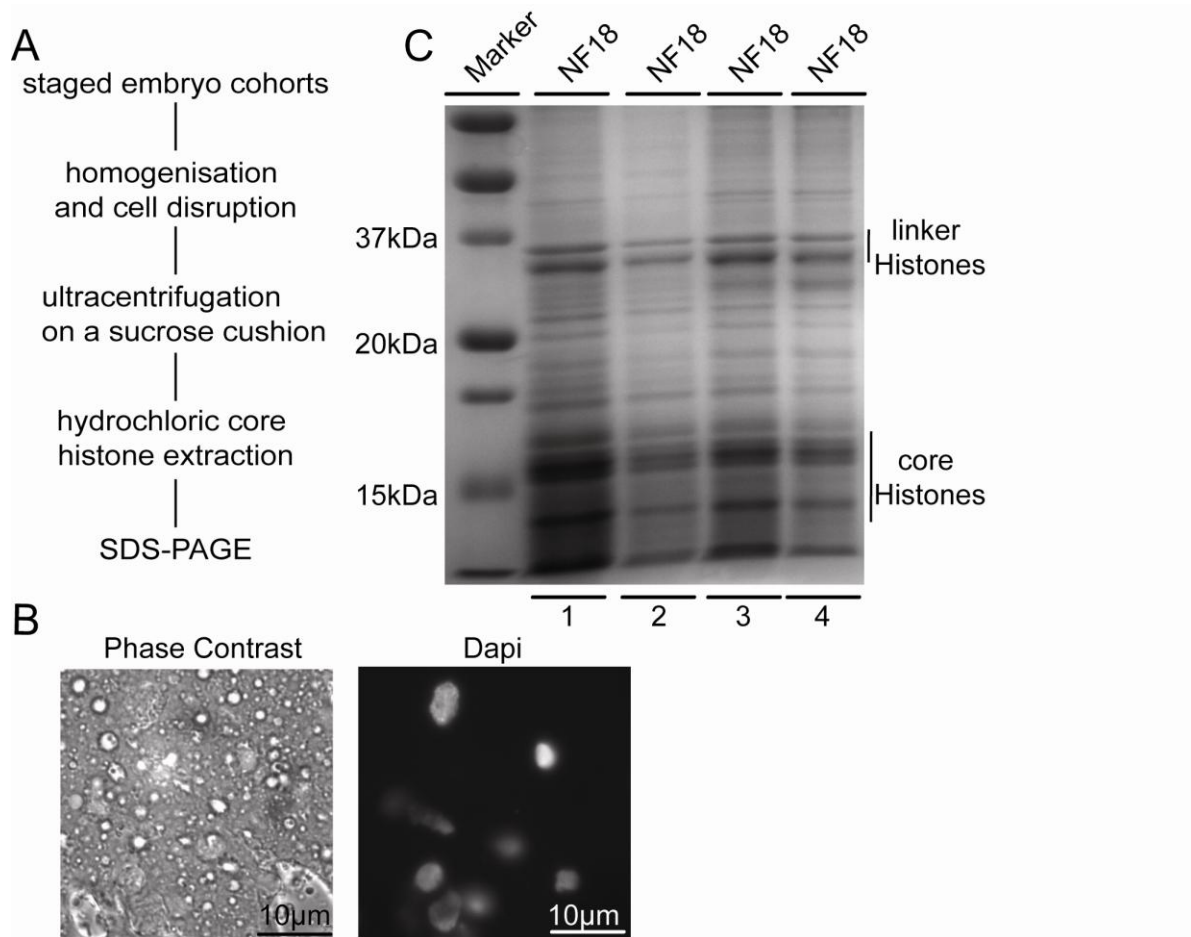


Figure 9: Nuclei Preparation of *Xenopus laevis* Embryos and Acidic Histone Extraction

A) Flowchart of nuclei preparation of *Xenopus laevis* and hydrochloric acid histone extraction. **B)** Microscopic pictures of nuclei after nuclei preparation of neurula stage *Xenopus laevis* embryos in phase contrast and stained with DAPI. The preparation was successful as the nuclei were visible, but together with the nuclei a huge amount of debris was present in the sample. **C)** Coomassie stained SDS-PAGE of four biological repeats of histones extracted with hydrochloric acid after nuclei preparation of neurula staged *Xenopus laevis* embryos. The four core histone proteins were visible in the gel but the yield differed significantly among the different samples reflecting the heterogeneity of the nuclear preparation.

As the different amount of histone proteins and other proteins in the four biological repeats in the Coomassie stained SDS-PAGE indicated, there is some variation between extracted samples of the same amount of input. This reflected the heterogeneity of the nuclear preparation and thus the hydrochloric acid histone extraction. Mass spectrometry of tryptic digested proteins was performed. The subsequent MASCOT database search did

not identify sequences of core histone proteins. The purity and the amount of extracted histone proteins were not sufficient for significant identification. The quality of spectra was low and other proteins were present in the same samples (External contamination was ruled out as cause for the reduced quality of the histone spectra).

Homogenization of embryos followed by density gradient ultracentrifugation with sucrose yielded no intact nuclei in phase contrast microscopy and DAPI staining. Furthermore the destructed nuclei were not separated from cell debris. By Coomassie stained SDS-PAGE putative core histone proteins were detected next to high amounts of other proteins. Thus, applying hydrochloric acid to the ultracentrifuge pellet did not result in rather pure histones. By the use of mass spectrometry, no histones were identified and the quality of histone spectra was not sufficient for post-translational histone modification analysis.

5.2.2 *Extraction of Nucleosomes with FPLC via Hydroxylapatite Column*

As the nuclei preparation of embryos followed by acidic extraction did not result in histones suited for post-translational histone modification analysis, the need for another extraction method was obvious. Another frequently used technique for histone purification is the use of chromatography via a hydroxyapatite resin (Brand *et al.*, 2008). This method uses staged cohorts of embryos followed by nuclei preparation (see Figure 10A). To increase the yield, the nuclei were digested with Micrococcal Nuclease (MNase) producing small chromatin fragments consisting of 1-3 nucleosomes (see Figure 10B). MNase predominantly cleaves DNA in the internucleosomal regions and the aim was to obtain 1-3 nucleosome fractions. Figure 10B shows an Etidiumbromide stained agarose gel of separated DNA purified after different digestion times for MNase. Chromatin fragments containing 1-3 nucleosomes were visible after 5 and 10 minutes of MNase digestion. This digestion was sufficient for loading soluble chromatin onto the hydroxyapatite matrix. The DNA-nucleosome complex was bound tightly to the hydroxyapatite matrix. The hydroxyapatite resin was then washed to reduce the amount of contaminants. The method relies on the stable DNA-nucleosome complex, which remains bound to the resin by the use of low salt buffers, whereas other proteins are eluted.

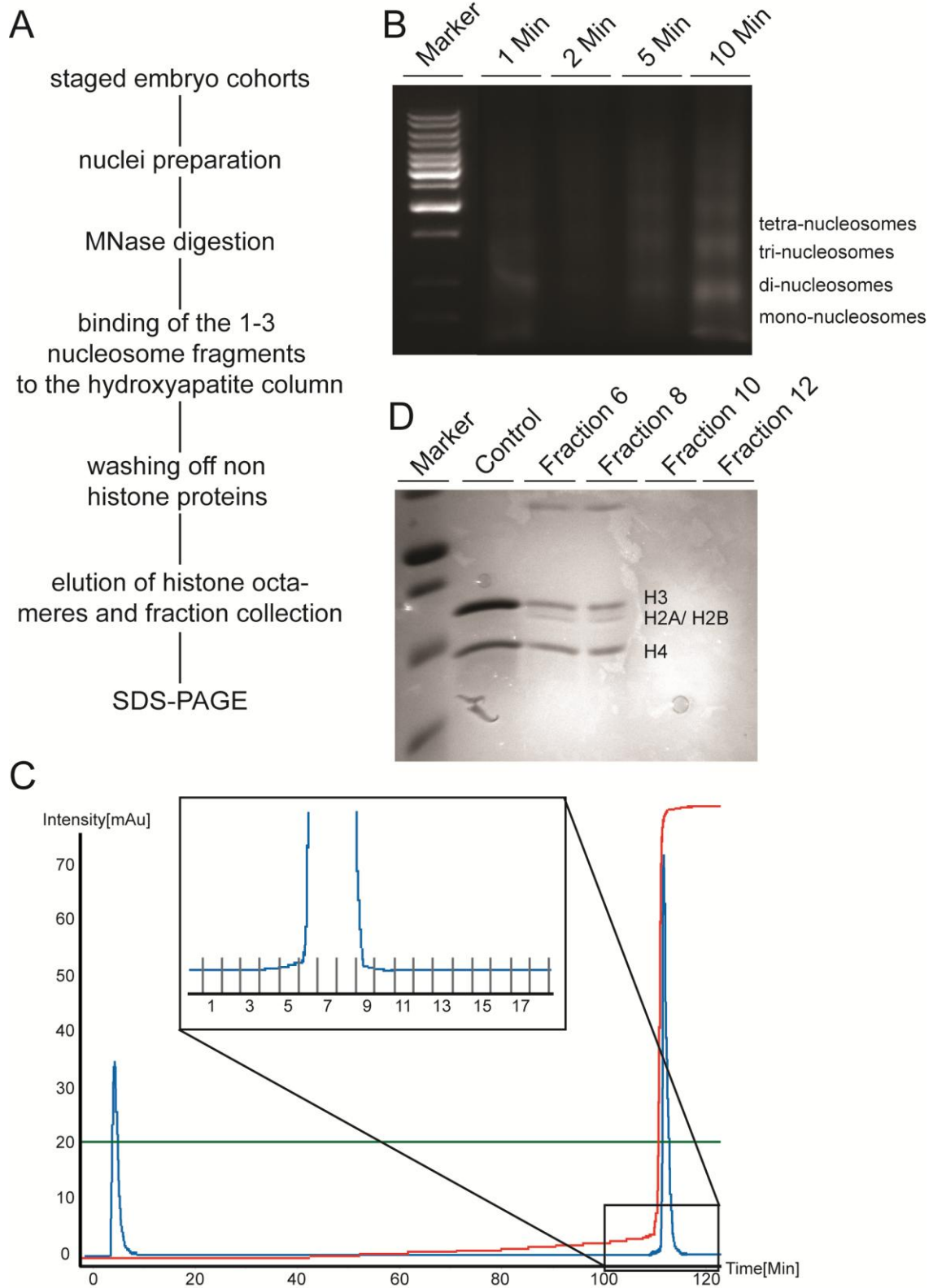


Figure 10: Hydroxyapatite Histone Octamere Purification after Nuclei Preparation of *Xenopus laevis* Embryos

A) Flowchart of nuclei preparation and hydroxyapatite purification of histone

octameres of *Xenopus laevis* embryos. **B)** Etidiumbromide stained Agarose gel of MNase digestion of chromatin after nuclei purification and lysis of different digestion duration. After 10 Min of MNase digestion the number of nucleosomes per DNA segment was between 1 and 3 and best suited for subsequent hydroxyapatite purification. **C)** Chromatogram of histone octamere purification with a hydroxyapatite column. Abscissa displays the time, ordinate displays the UV absorption. In blue is the UV absorption, in green the flow rate and in red the H_2PO_4 concentration. A step gradient of H_2PO_4 after 110 minutes elutes the histone octameres indicated by the UV absorption (box). **D)** Coomassie stained SDS-PAGE of precipitated fractions as indicated after the FPLC run. Histone H3 and H4 are present in aequimolar amounts whereas H2A and H2B are present in much less quantity.

After washing, a sharp increase in the salt concentration led to the elution of the core histones whereas the DNA remained bound to the resin (see Figure 10C). One big advantage of this method is the high purity of histones and the preservation of the heterooctamer conformation in buffers containing high salt concentrations (Bloom *et al.*, 1977). On the other hand, disadvantages of this method are the high amount of starting material necessary for the resin together with low recovery of histones from the resin (Ausio *et al.*, 1998). The fractions eluting from the resin were precipitated with TCA followed by Coomassie stained SDS-PAGE (see Figure 10D). The recovery of the elution of the FPLC in the SDS-PAGE was small (see Figure 10D). Furthermore, the four core histones did not elute in equal amounts, but H3 and H4 were present in higher concentration compared to H2A and H2B. Elution of histone proteins in other fractions was checked and ruled out by Western blot (data not shown). Mass spectrometric analysis of tryptic digested proteins revealed the presence of H3 and H4, however the quality of spectra was neither sufficient for significant identification of the histones nor for analysis of post-translational histone modifications. Furthermore, the recovery of histones for early embryonic stages such as blastula and gastrula was low yielding no histone proteins in Coomassie stained SDS-PAGE. Therefore, the material input became limiting, at least for the younger embryonic stages. These findings pointed to either insufficient nuclei preparation or high loss of histone proteins during the chromatography. To check the nuclei preparation, phase contrast microscopy and DAPI staining was performed. Independent from the embryonic stage analysed, no intact nuclei were detected by microscopic means (data not shown).

In the end, the histone fractions from hydroxyapatite purification were pure, but not enriched. The nuclear preparation was not efficient as nuclei were not detectable by microscopy. Furthermore, the recovery of histones was very low regarding the high amount of starting material and the size of the hydroxyapatite resin. The amount of

embryos necessary for extraction exceeded our possibilities. For early embryonic stages (NF9, NF12), no histones were present most likely caused by the low abundance of nuclei in these stages. The quality and reproducibility of mass spectra was not sufficient for significant histone identification or analysis of post-translational histone modifications. Consequently this method did not meet the expectations for an easy, efficient and reliable method for histone extraction.

5.2.3 *Direct Lysis of Embryos and Further Purification via HPLC*

As described above, two independent methods for nuclei extraction were used followed by two different techniques for histone extraction. The yield of one of these methods was low as nuclei preparation was not possible as the nuclei were not separated from cytoplasmatic and yolk components. Furthermore, both methods did not yield histones with purity sufficient for subsequent mass spectrometric analysis. As the common denominator of both approaches was the low quality of the nuclear preparation, a different technique starting from whole embryo lysates that are further purified by reversed-phase chromatography. Using reversed-phase HPLC, the histone proteins were separated from other proteins according to their chemical and physical properties. The different fractions of the RP-HPLC were then analysed by SDS-PAGE or mass spectrometry (see Figure 11A).

Reversed-phase HPLC for histone purification, first introduced by Certa and Ehrenstiel in 1981, is a quick and efficient method especially suited for small-scale purification. This method allows to separate even histone variants under optimal conditions (Schechter *et al.*, 2007).

In the method described by Schechter, H3 typically eluates as a broad peak and is well separated from H2A, H2B and H4, which elute in close proximity (Schechter *et al.*, 2007). Regarding these advantages of RP-HPLC together with whole embryo lysates to circumvent the difficulties with nuclei preparation, this method seemed to offer an easy and quick solution. After embryo disruption, the samples were evaluated by SDS-PAGE prior to the RP-HPLC runs (data not shown). There was a variety of different protein bands visible in the Coomassie staining together with putative histone proteins.

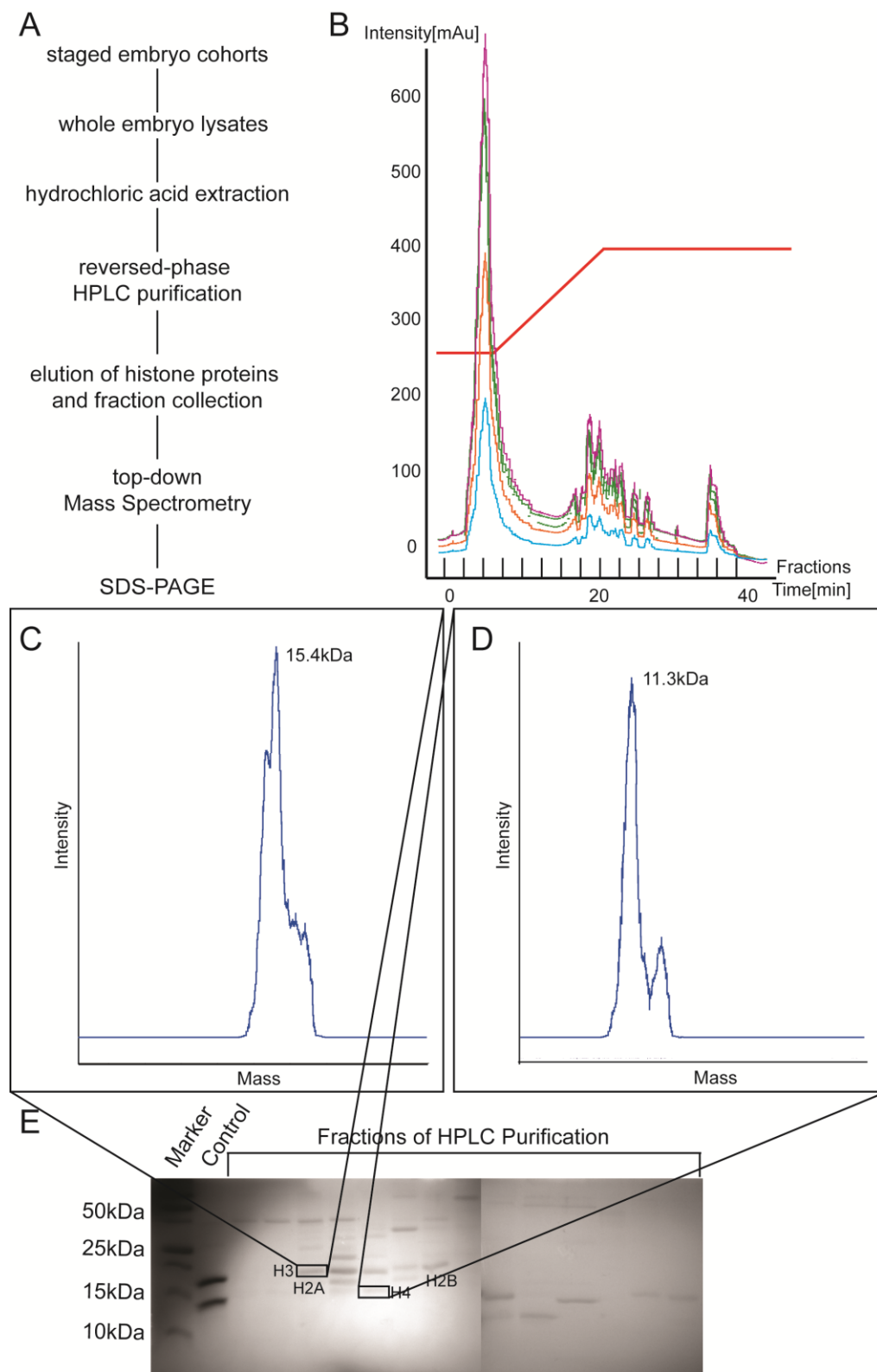


Figure 11: Reversed-phase HPLC Purification of Histones of Xenopus Whole Embryo Lysates

A) Flowchart of reversed-phase HPLC histone protein purification. **B)** reversed-

phase HPL-Chromatogram of whole lysates of different stages of *Xenopus laevis* embryos. Abscissa displays the time, ordinate displays the UV absorption. Light blue was the UV absorption of blastula staged embryos, in orange gastrula, in green neurula and in purple tadpole. The red line indicates the gradient of the hydrophobic buffer. The four different embryonic stages showed a similar UV absorption profile, but different composition and intensity according to the different stages. C) MALDI-TOF spectrum of undigested H3 protein in linear mode of reversed-phase HPLC fractions. Abscissa displays mass, ordinate displays the intensity. H3 protein was clearly detectable in the fractions. D) MALDI-TOF spectrum of undigested H4 protein in linear mode of reversed-phase HPLC fractions. Abscissa displays mass, ordinate displays the intensity. H4 protein was clearly detectable in the fractions. E) Coomassie stained SDS-PAGE of pooled fractions of reversed-phase HPLC runs of neurula staged *Xenopus laevis* embryos. The histone proteins were visible in several fractions and were eluting at different time points as expected. Nevertheless there were other proteins eluting at the same time points as histone proteins.

Then RP-HPLC runs were performed using an increasing gradient of ACN to stepwise elute the core histones (see Figure 11B). The eluting proteins were collected in fractions of 200 μ l. The chromatogram shows the increasing gradient of ACN and the optical density measurements of several runs of differently staged embryos (see Figure 11B). After 5-10 minutes, the proteins which did not bind to the resin were eluted. According to the optical density measurement, several different peaks were eluting at around 18-25 minutes and 35 minutes. The pattern of the RP-HPLC runs was very similar comparing different runs of the differently staged embryos. To check for histone proteins in those fractions with high UV absorption, mass spectrometry in linear mode was performed to visualize undigested protein (see Figure 11C and D). The use of the linear mode allows detection of undigested protein, however, significant identification of the protein is not possible. The molecular weight of the proteins found in mass spectrometry correlated to H3 (see Figure 11C) and H4 (see Figure 11D). In the Coomassie stained SDS-PAGE of different fractions from the RP-HPLC runs, putative histone proteins matched the mass of recombinant histones H3 and H4. Thus, putative core histones were visible in several fractions. The histone H3 was eluting earlier than H4, which was consistent with published data (Ausio et al., 1998; Schechter *et al.*, 2007). Mass spectrometric analysis of tryptic digested proteins revealed low reproducibility and signal to noise ratio. Histone H4 was significantly identified by the use of MASCOT database search, but not H3 for several technical and biological repeats. The recovery of modifications of histone H4 of

technical replicates was low, but better than in direct lysis with a correlation coefficient of 0.67.

Summarizing, RP-HPLC purification of histones derived from whole lysates of different staged embryos did not result in histones of sufficient quantity and quality for mass spectrometric analysis. Based on these results, further investigation with this method was not made.

5.2.4 *Development of a New Protocol for Histone Extraction*

Having tested three different approaches for histone purification and enrichment, which did not yield enough pure core histones, the idea arose that rather the nuclei preparation is the cause of the problem, than the histone extraction methods. This idea was further supported by the findings using the somatic cell lines A6 and XTC, where even direct lysis of entire nuclei led to histone enrichment (see Figure 5C). The use of whole embryo lysates with RP-HPLC purification did not improve the histone extraction (see Figure 11E). As the frequently used methods for nuclear preparation did not result in nuclei, the need to develop a new protocol for nuclei extraction of *Xenopus laevis* embryos was obvious. A variety of methods to extract histone proteins from pure nuclei, like high-salt (Burton *et al.*, 1978), acid extractions (Murray *et al.*, 1966) or chromatography methods (Simon *et al.*, 1979; Certa *et al.*, 1981) is available.

There are several methods described in the literature how to extract nuclei of embryos or cells (Gould, 1998). The use of aqueous buffers is described to better preserve the structure of the nucleus as compared to other extraction methods. Thus, three different buffer conditions on the basis of aqueous buffers were tested (see Figure 12A).

After homogenization of the embryos of stage NF18, the quality of the nuclei was checked by phase contrast microscopy and DAPI staining (see Figure 12B). Best results were obtained using buffer E1 and further extractions were done with this buffer. In the next step, homogenization was tested with two 5ml-glass-glass homogenizers differing in the gap width. Nuclei structure was best conserved using the loose fit homogenizer compared to the tight fitting homogenizer (data not shown). The best quantity of strokes was determined to yield as much intact nuclei as possible together with high efficient cell disruption. Best results were obtained by 20 strokes (data not shown).

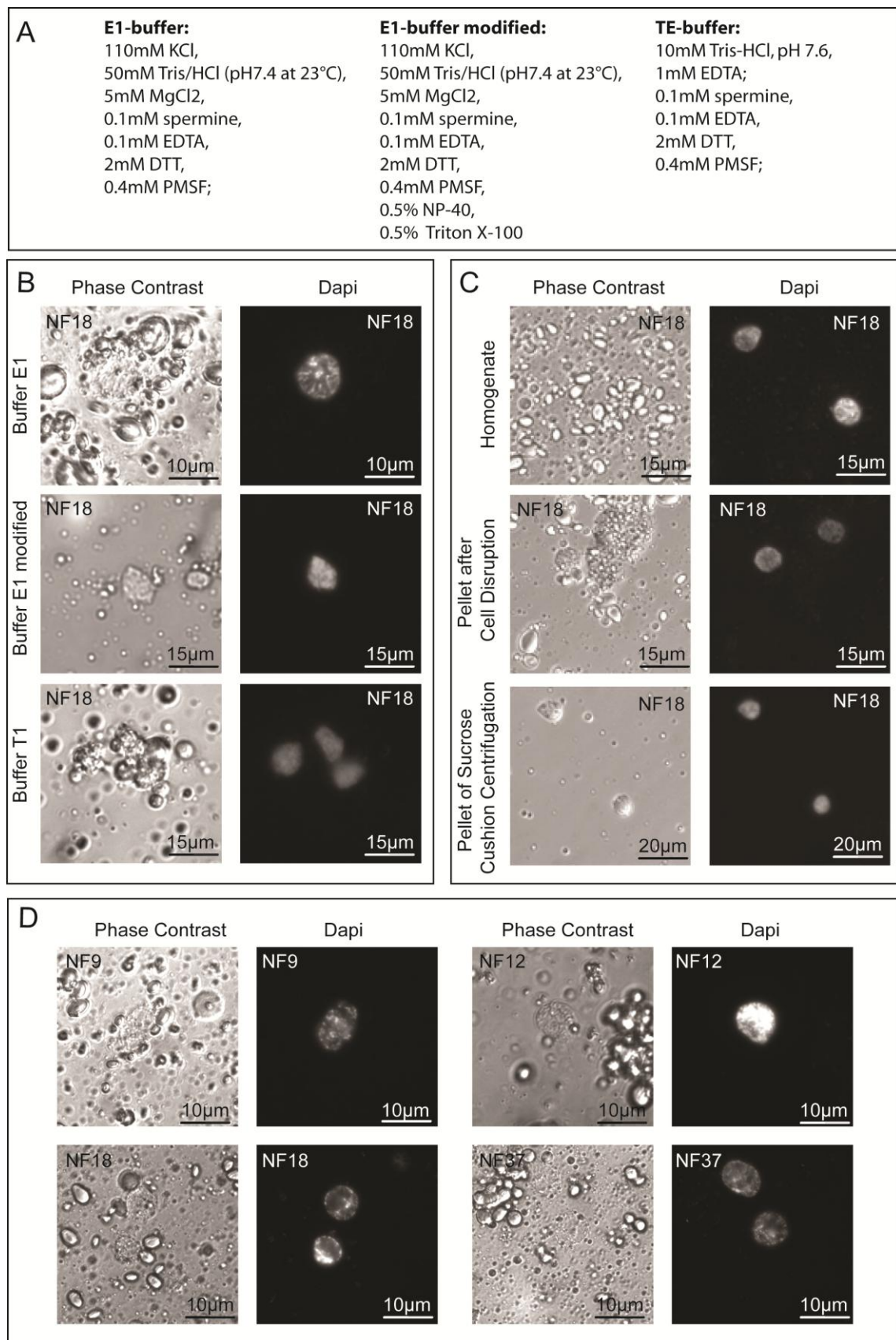


Figure 12: Optimizing the Preparation of Embryonic Cell Nuclei for *Xenopus laevis* Embryos

A) Different buffers used for the extraction. **B)** Microscopic pictures of neurula (NF18) staged *Xenopus laevis* embryos after homogenization in different buffers in DAPI stain and phase contrast. Nuclei were visible in all buffers tested but the morphology of the nuclei extracted with modified buffer E1 and T1 appeared irregular whereas the nuclei extracted with buffer E1 appeared normal. **C)** Microscopic pictures of neurula (NF18) staged *Xenopus laevis* embryos after homogenization, after centrifugation of the homogenate and after a second centrifugation step in phase contrast and DAPI stain. The nuclei in DAPI stain were clearly visible and the debris and contaminants were highly reduced by the two centrifugation steps, as seen in the phase contrast pictures. **D)** Microscopic pictures of blastula (NF9), gastrula (NF12), neurula (NF18) and tadpole (NF37) staged *Xenopus laevis* embryos after homogenization and first centrifugation in phase contrast and DAPI stain. Nuclei were visible for all the four embryonic stages.

In the first steps of the method development, stable buffer and homogenizing conditions were established. Then, a method to separate the nuclei from the cytoplasmic components was developed.

Different centrifugation velocities were tested from 100rpm to 3000rpm in a swing-out rotor. At a centrifugation velocity of 1000rpm, high amounts of nuclei were detectable by phase contrast microscopy and DAPI staining, whereas no nuclei were present in the supernatant. Centrifugation with more than 1000rpm resulted in more cell contaminants in the pellet, whereas the yield of nuclei was not increased. Thus the centrifugation velocity was set to 1000rpm. In the pellet cell debris and other contaminants were present despite the centrifugation (Figure 12C). In the next step, the nuclei were washed in buffer E1 containing the non-ionic detergents Triton X-100 and NP-40 to get rid of lipoproteins and yolk. A second centrifugation on a gradient of sucrose with 1000rpm to further purify the nuclei was performed. The density of the nuclei is usually higher than the density of cell debris and yolk. Consequently, the centrifugation on a density gradient such as sucrose was expected to purify the nuclei. As optimal sedimentation of nuclei was already determined at 1000rpm, different concentrations for the sucrose gradient were tested. Best results were obtained using 2.0M sucrose in buffer E1 (Figure 12C). As shown in Figure 12B and C, the newly developed method was able to deliver rather pure nuclei of roughly 50 neurula staged embryos.

Finally, this protocol had to be adopted to the other developmental stages with slight changes (see Figure 12D and Figure 13A).

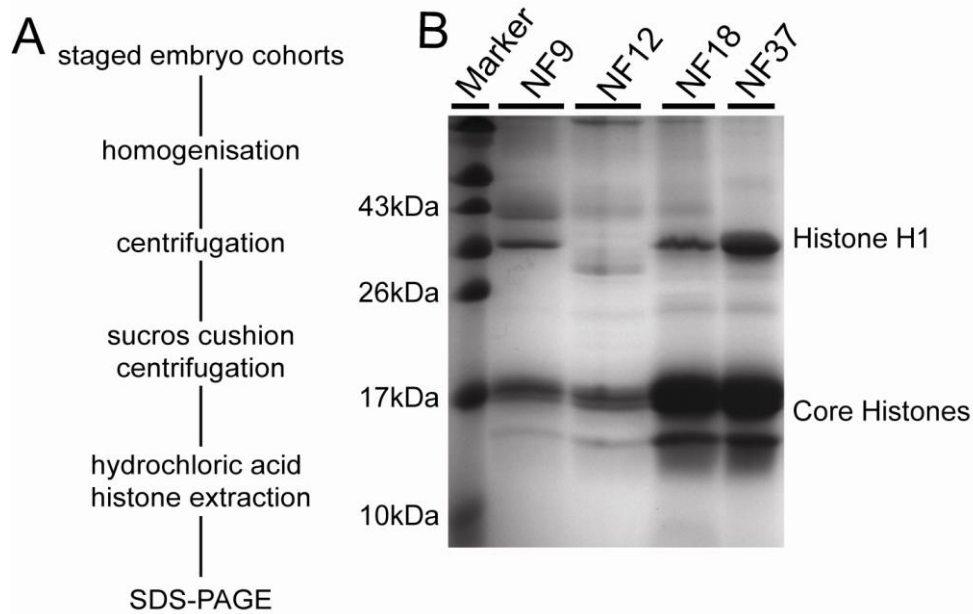


Figure 13: Purification of Core Histone Proteins from *Xenopus laevis* Embryos

A) Flowchart of nuclei extraction and purification and acidic histone extraction. **B)** Coomassie stained SDS-PAGE of blastula (NF9), gastrula (NF12), neurula (NF18) and tadpole (NF37) staged *Xenopus laevis* embryos. The four core histones were clearly visible and enriched in the SDS-PAGE and only little other proteins were present in the extracts.

Trypsin digested putative histone proteins were analysed using mass spectrometry. MASCOT database search matched the sequences significantly to the four core histone proteins. The reproducibility of histone spectra, the quality of internal calibration and the signal to noise ratio was sufficient for analysis. The recovery of modifications of technical replicates of 0.4M hydrochloric acid extraction had a correlation coefficient of 0.93 and thus was comparable to the correlation coefficient found for histones from cell lines.

In summary, a new, efficient and reliable method for preparation of rather pure nuclei of maximal 50 embryos for one biological repeat of different staged *Xenopus laevis* embryos was developed. Using acid extraction of these nuclei resulted in high yield of pure core histones in SDS-PAGE (see Figure 13B). The extraction conditions of the four core histones met the requirements for analysis of post-translational histone modifications in mass spectrometry.

5.2.5 Comparison to the Commercially Available Histone Purification Kit

During the time when the new method for nuclei purification of *Xenopus laevis* embryos was developed, a commercially available histone purification Kit was placed on the market by Active Motif. This kit claimed to purify histones of all kinds of tissues and organisms.

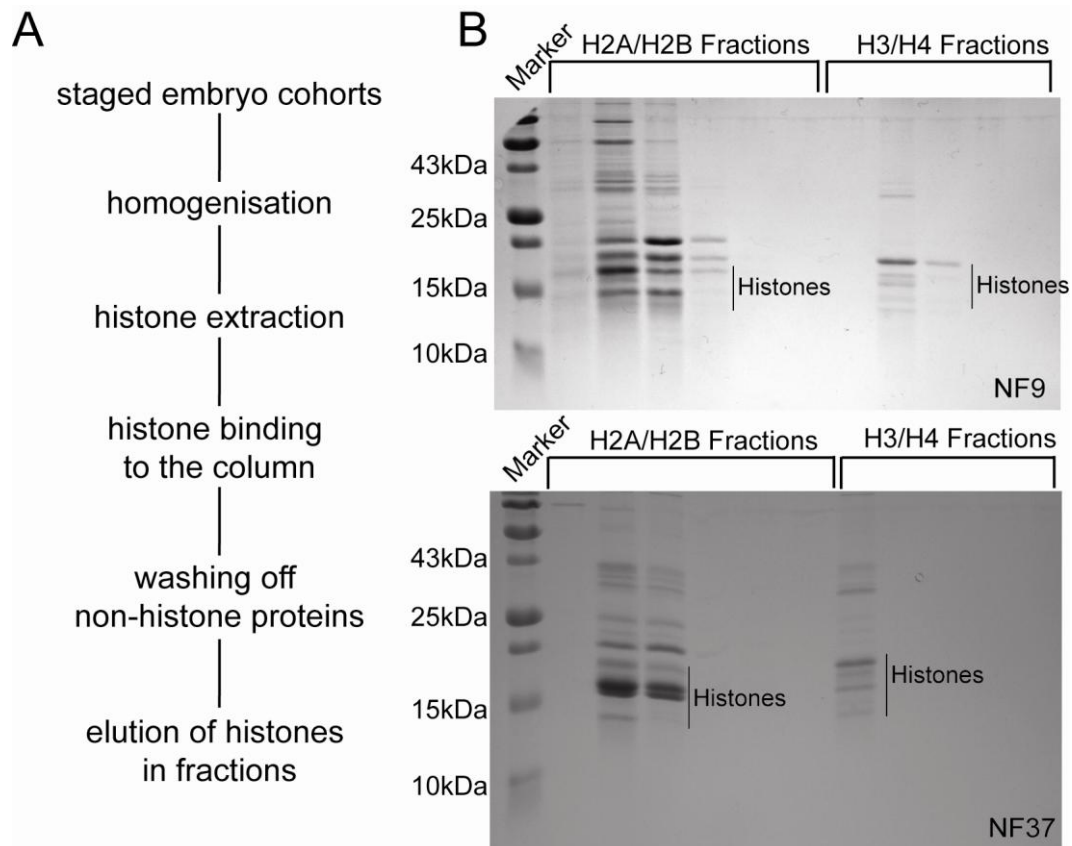


Figure 14: Histone Purification via the Active Motif Histone Purification Kit

A) Flowchart of the purification as described in the users manual. **B)** Coomassie stained SDS-PAGE of histones purified with the kit of blastula (NF9) and tadpole (NF37) staged *Xenopus laevis* embryos collected in H2A/H2B and H3/H4 fractions. Histone proteins were present in the fractions although the two fractions were not able to elute H2A/H2B separate from H3/H4 and lots of other proteins were present in the extracts as well.

To check the efficiency of the kit, the protocol (see Figure 14A) for purification of core histones was applied as mentioned in the manufacturer's manual. Staged cohorts of embryos were used for homogenization followed by histone extraction using a histone binding resin provided with the kit. After loading and washing, the histones were eluted

in fractions (Figure 14A and B). The Coomassie stained SDS-PAGE of different fractions revealed a variety of proteins. The amount of the four histones was not equally distributed in Coomassie stained SDS-PAGE (see Figure 14B). Tryptic digested proteins analysed by mass spectrometry revealed high reproducibility but low signal to noise ratio. MASCOT database search identified the core histones significantly, however, the spectra had a low correlation coefficient of 0.80 likely caused by the low recovery from the column. The amount of embryos needed for the kit was very high with at least 800-1200 embryos for one technical repeat.

With the use of the histone purification Kit as described by Active Motif, histone proteins could be isolated. The kit is very expensive and together with the high input of embryonic material especially the use in preparative scale histone purification is not suitable. The histones were significantly identified in mass spectrometry, but the quality of the spectra was lower compared to the ones of acidic extracted histones. The kit was considered as an alternative histone extraction method for *Xenopus laevis* embryos next to the protocol previously developed.

In summary, hydrochloric acid extraction of histones of embryonic cell nuclei turned out to be the method of choice, based on yield and purity of histones for the quantitative analysis by mass spectrometry. Thus core histones of the eight different conditions of interest, four embryonic stages (blastula, gastrula, neurula and tadpole; see Figure 13B) as well as of the four cell lines (A6, XTC, ES cells and Feeder cells; see Figure 7D) were prepared according to this new protocol.

5.3 Technical Preparation for the Identification and Quantification of Histone Modifications by Mass Spectrometry

5.3.1 Histone Sequences

In order to investigate histone post-translational modifications during vertebrate embryogenesis, core histone proteins were purified from unmanipulated *Xenopus laevis* embryos of four different developmental stages (see Figure 8). The stages included blastula (NF9), gastrula (NF12), neurula (NF18) and tadpole (NF37) embryos, representing key steps in vertebrate development (see Figure 8).


```

variant Histone H3 (H3.3 - replication-independent)

H3.3 Xenopus      1 ARTKQTARKSTGGKAPRKQLATKAARKSAPSTGGVKKPHRYRPGTVLREIRRYQKSTEL  60
H3.3 Human        1 ARTKQTARKSTGGKAPRKQLATKAARKSAPSTGGVKKPHRYRPGTVLREIRRYQKSTEL  60
H3.3 Mus          1 ARTKQTARKSTGGKAPRKQLATKAARKSAPSTGGVKKPHRYRPGTVLREIRRYQKSTEL  60
H3.3 Drosophila   1 ARTKQTARKSTGGKAPRKQLATKAARKSAPSTGGVKKPHRYRPGTVLREIRRYQKSTEL  60
*****

H3.3 Xenopus      61 LIRKLPFQRLVREIAQDFKTDLRFQSAAIIGALQEASEAYLVGLFEDTNLCAIHAKRVTIM 120
H3.3 Human        61 LIRKLPFQRLVREIAQDFKTDLRFQSAAIIGALQEASEAYLVGLFEDTNLCAIHAKRVTIM 120
H3.3 Mus          61 LIRKLPFQRLVREIAQDFKTDLRFQSAAIIGALQEASEAYLVGLFEDTNLCAIHAKRVTIM 120
H3.3 Drosophila   61 LIRKLPFQRLVREIAQDFKTDLRFQSAAIIGALQEASEAYLVGLFEDTNLCAIHAKRVTIM 120
*****

H3.3 Xenopus      121 PKDIQLARRIRGERA 135
H3.3 Human        121 PKDIQLARRIRGERA 135
H3.3 Mus          121 PKDIQLARRIRGERA 135
H3.3 Drosophila   121 PKDIQLARRIRGERA 135
*****

Histone H4

H4 Xenopus        1 SGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLKV  60
H4 Human          1 SGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLKV  60
H4 Mus            1 SGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLKV  60
H4 Drosophila     1 TGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLKV  60
*****

H4 Xenopus        61 FLENVIRDAVITYTEHAKRKTIVTAMDVVYALKRQGRTLYGFGG 102
H4 Human          61 FLENVIRDAVITYTEHAKRKTIVTAMDVVYALKRQGRTLYGFGG 102
H4 Mus            61 FLENVIRDAVITYTEHAKRKTIVTAMDVVYALKRQGRTLYGFGG 102
H4 Drosophila     61 FLENVIRDAVITYTEHAKRKTIVTAMDVVYALKRQGRTLYGFGG 102
*****

```

Figure 15: Conservation of Core Histone Sequences

Clustal alignment of core histone proteins from frog, human, mouse and fruitfly showing conservation of the histone sequences.

At late blastula, shortly after the onset of zygotic transcription, embryos consist mostly of uncommitted, pluripotent cells (Nieuwkoop et al., 1967; Nicetto and Rupp, unpublished data). By the gastrula stage, the germ layers have been induced, and embryonic patterning increases the cellular diversity of the embryos during neurulation (Heaseman, 2006; De Robertis, 2009). After hatching, tadpoles are composed largely of differentiated, although premetamorphic, somatic cells. To evaluate the results from the embryonic samples, histones from *Xenopus* A6 and XTC cell lines, as well as from murine germline-transmission competent ES cells (Maisonneuve et al., 2009) and primary

embryonic fibroblasts (MEF) were prepared.

Essential for the identification of post-translational modifications of proteins is detailed knowledge of the protein sequence. In order to align the sequences of *Xenopus laevis* core histone with histones of other species, a NCBI database search for *Xenopus laevis* core histone sequences was done (see Supplement Figure 1). Alignment of the primary sequences of *Xenopus* histones with core histones of other species such as human, mouse and fruitfly revealed some variability of the H2A and H2B histones and high interspecies conservation of the H3 and H4 (see Figure 15). For histone H2A, there are differences in single amino acid residues at several positions of the histone sequences among different species. Noteworthy is the extended N-terminus of the histone H2A of mouse, which has another seven amino acids more than the other three species. The amino acid exchanges do not cluster in one domain but are distributed over the entire sequence. The sequence of Histone H2B has amino acid exchanges in different species. The differences of the amino acid sequence cluster in the N-terminus. In *Drosophila* H2B, insertions of amino acids are found. For histone H2B, the vertebrate sequences show higher similarity to the sequence of *Drosophila*. The histone variant H3.1 is only found in mammals and thus not present in *Xenopus laevis*. The alignment of the replication-dependent histone variants H3.2 and the replication-independent H3.3 shows no sequence differences among different species. The amino acid sequence of histone H4 is well conserved among species and has one amino acid exchange at the first position.

5.3.2 *Histone Preparation for Mass Spectrometric Analysis of Histone Modifications*

In a next step, *in silico* digestion of the four core histones was done using the endoproteinase Trypsin. Trypsin, which cleaves N-terminally of lysine and arginine residues, is frequently used for digestion of proteins in mass spectrometry. The resulting peptides of tryptic digestion of the four core histones are depicted in Table 4. The mass spectrometer MALDI-TOF requires a peptide mass between 500 and 3500 Da, other peptides cannot be analysed. The four core histone proteins have a high number of arginine and lysine residues in close proximity resulting in frequent cleavage sites. The peptides of tryptic digestion are often too small for detection in MALDI-TOF. To prevent trypsin from cleavage and thus to increase the size of the peptides, a covalent modification of lysine residues with propionic anhydride was introduced.

H2A Trypsin		H2B Trypsin		H3 Trypsin		H4 Trypsin	
aa	mass [Da]	aa	mass [Da]	aa	mass [Da]	aa	mass [Da]
1-3	319.17	1-5	541.30	1-3	377.20	1-4	450.21
4-5	204.13	6-11	570.33	4-5	248.16	5-6	204.13
6-9	389.21	12	147.11	6-9	475.26	7-9	261.16
10-11	276.17	13-15	291.17	10	147.11	10-13	374.24
12-13	218.15	16	147.11	11-15	688.41	14-17	332.19
14-15	218.15	17-20	418.27	16-18	343.21	18	175.12
16-17	276.17	21-23	376.22	19	147.11	19-20	312.18
18-20	349.18	24	147.11	20-24	560.34	21	147.11
21-30	1015.57	25-27	319.16	25-27	317.19	22-24	387.27
31-33	411.25	28	147.11	28	147.11	25-36	1325.75
34-36	401.29	29	175.12	29-37	787.43	37	175.12
37	147.11	30	175.12	38-41	537.33	38-40	359.24
38-43	709.33	31	147.11	42-50	1032.60	41	175.12
44-72	2945.60	32-33	262.15	51-53	417.25	42-45	360.22
73-75	376.18	34	147.11	54	175.12	46	175.12
76	147.11	35-43	1135.57	55-57	438.24	47-56	1180.62
77-78	276.17	44-46	359.27	58-64	831.49	57-60	416.29
79-82	498.34	47-57	1168.60	65	147.11	61-68	989.58
83-89	836.51	58-72	1759.81	66-70	660.38	69-78	1134.54
90-96	861.40	73-79	703.37	71-73	387.27	79	175.12
97-119	2287.38	80-85	745.40	74-80	850.43	80	147.11
120	147.11	86	175.12	81-84	504.28	81-92	1310.07
121-125	551.27	87-92	664.36	85-116	3456.68	93	175.12
126-128	335.19	93-99	816.46	117	175.12	94-96	360.2
129-130	234.15	100-108	953.60	118-123	688.41	97-103	714.35
		109-116	828.42	124-129	715.41		
		117-120	418.27	130	175.12		
		121-125	569.29	131-132	288.2		
				133-135	361.18		
				136	90.06		

H2A Propionylation-Trypsin		H2B Propionylation-Trypsin		H3 Propionylation-Trypsin		H4 Propionylation-Trypsin	
aa	mass [Da]	aa	mass [Da]	aa	mass [Da]	aa	mass [Da]
1-4	506.24	1-30	3837.16	1-2	302.18	1-3	375.20
5-12	943.53	31	175.12	3-8	760.43	4-17	1494.88
13-18	768.48	32-34	446.27	9-17	1013.57	18-19	312.18
19-21	349.18	35-73	4718.39	18-26	1098.66	20-23	571.39
22-30	944.53	74-80	703.37	27-40	1601.91	24-35	1381.78
31-33	411.25	81-87	957.53	41-49	1032.60	36	175.12
34-36	401.29	88-93	664.36	50-52	417.25	37-39	359.24
37-43	893.45	94-100	816.46	53	175.12	40	175.12
44-72	2915.59	101-126	2936.63	54-63	1306.74	41-45	572.35
73-78	873.48			64-69	844.50	46-55	1180.62
79-82	498.34			70-72	387.27	56-67	1442.87
83-89	836.51			73-83	1391.72	68-78	1346.67
90-130	4611.55			84-116	3668.81	79-92	1706.95
				117-128	1440.82	93-95	360.20
				129	175.12	96-102	714.35
				130-131	288.20		
				132-134	361.18		
				135	90.06		

Table 4: Tryptic Peptides of Core Histone Proteins

Displayed are the peptides of H2A, H2B, H3 and H4 after digestion with Trypsin (upper panel) and after Propionylation of lysine residues and Trypsin digestion (lower panel). Peptides highlighted in grey are not detectable by MALDI-TOF analysis. Abr. aa = amino acid.

The propionic anhydride only reacts with unmodified or mono-methylated lysine residues, however, if the lysine residue is di- or tri-methylated, trypsin will also not cleave after the lysine residue. The peptides resulting from these digestions are displayed in Table 4. For all the core histones, the sequence coverage was higher with tryptic digestion with covalent lysine modification with propionic anhydride (see Table 4). In total, the trypsin digestion of the four core histones had sequence coverage of 51.5% for H2A, 68.8% for H2B, 68.9% for H3 and 87.3% for H4.

Notably, most of the best known sites of modifications were covered. After the acquisition of MALDI-TOF spectra they were evaluated for their quality. Only spectra fulfilling the quality characteristics such as high signal to noise ratio and sufficient internal calibration were used for the quantification of post-translational histone modifications. The MALDI-TOF spectra display the mass at the x-axis and the intensity at the y-axis. The sum of the intensities of different modified states of a single histone peptide is 100%. The relative abundance of a modification is the intensity of the modification divided by the summarized intensity being equal to 100% (Bondarenko *et al.*, 2002).

5.4 Identification and Quantification of Histone Modifications by Mass Spectrometry

Histone post-translational modifications were analysed in triplicate biological repeats for MALDI-TOF and duplicate for Orbi-Trap and each biological sample was analysed in at least two technical repeats. The average correlation coefficient between biological replicates was 0.9962, indicating high robustness of the analysis for the reported epigenetic states.

For the identification of histone post-translational modifications, preferentially high mass accuracy LC-MS/MS mass spectrometry was used (Orbi-Trap, Thermo Scientific). This technique combines the advantages of reversed-phase HPLC and of high-resolution tandem mass spectrometry. In the reversed-phase LC, the different peptides and their modified counterparts were separated according to their hydrophobic interaction with the micro-column in the liquid chromatography. The histone proteins of the Coomassie stained SDS-PAGE were propionylated and then digested using the endoproteinase trypsin (see Figure 16A).

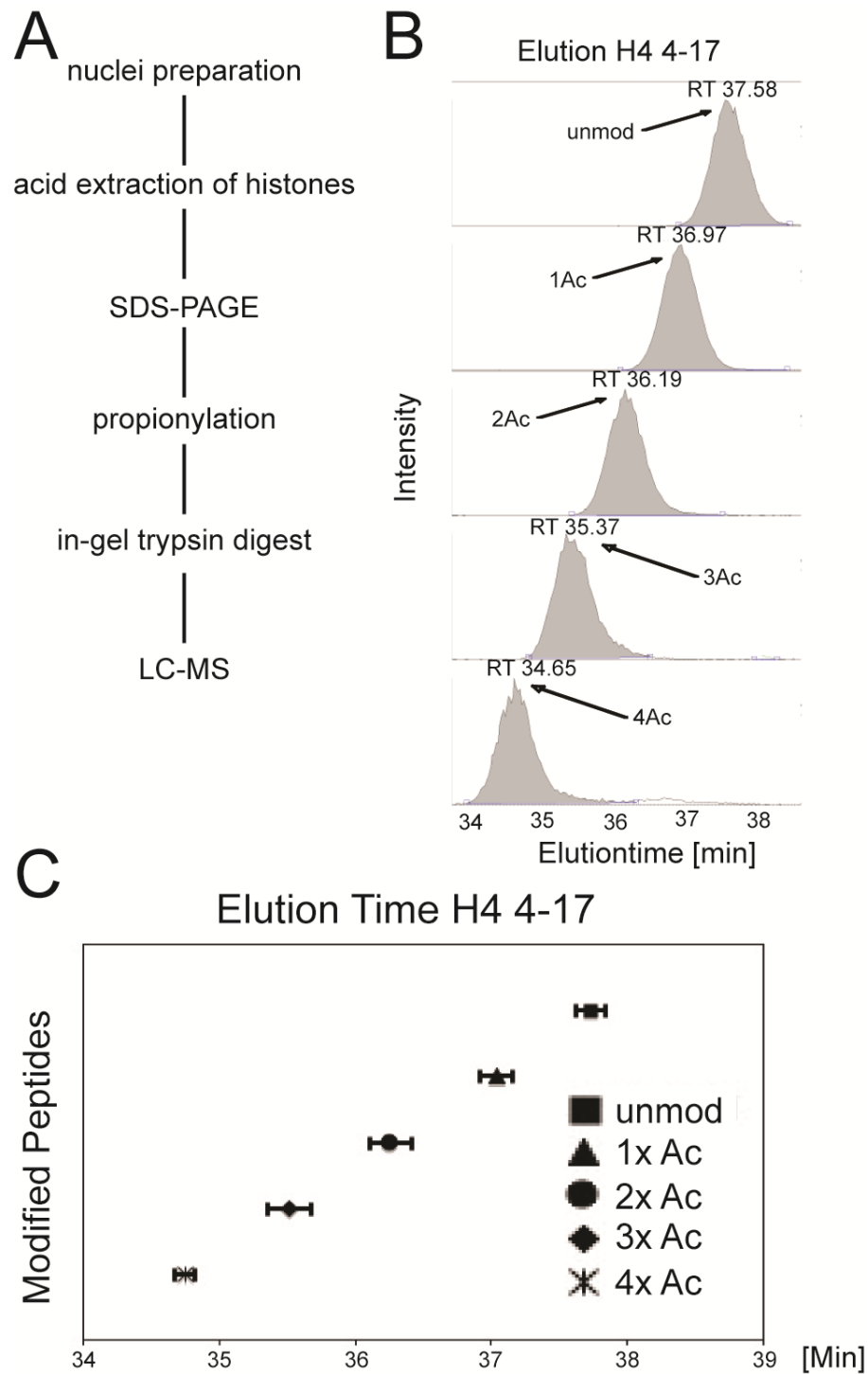


Figure 16: LC-MS/MS Quantification

A) Flow chart of sample preparation for LC-MS/MS **B)** Elution time of different acetylation states of H4 4-17 peptides in LC. The elution on the LC shifted to earlier time points as the amount of acetyl groups increased. **C)** Averaged elution times of different acetylation states of H4 4-17 peptides in LC. The error bars indicate standard deviation of several independent LC runs of different samples. The elution time was a conserved feature of a modification.

According to the hydrophobic binding properties of the reversed-phase column, different modified peptides of one core histone could be separated. The different modified peptides were eluting from the column at different time points (see Figure 16B). Specifically, peptides with different acetylation states of the H4 4-17 peptide led to an earlier elution from the column in the LC run.

Higher acetylation states further polarized this effect by even eluting earlier. Histones of different sources showed a similar distribution on the LC and consequently the peak elution time of different modifications of the peptides did not spread. Exemplary shown for the peptide 4-17 of H4 in Figure 16C are the averaged elution times of several independent LC runs of differently staged *Xenopus laevis* embryos, and stable cell lines A6 and XTC. The addition of one acetyl-group shifted the elution time by 0.7 Min to an earlier elution. These findings indicated that in addition to the mass of a histone peptide and its fragmentation spectrum, the elution time was also a unique parameter to characterise the modification of a peptide. In the following analysis, the elution time was used as a third parameter to determine the identity of a given peptide.

The feature of the elution time to be unique together with the mass of a peptide was also used for precise quantification of different modifications of histone peptides. The area under the peak of the LC run was determined by the extracted ion counts (XICs) of the ions with this mass and was proportional to its abundance in the sample of the corresponding doubly and triply charged ions. As different modified histone peptides were separated by the reversed-phases chromatography, different modification states could be quantitated individually.

Another advantage of the tandem MS approach is the possibility to fragment peptides of interest with collision by small gas molecules, called collision induced dissociation (CID). The collision of the histone peptides with Argon gas molecules leads to a random fragmentation of mainly the peptide backbone. The consequence of the fragmentation is the generation of series of smaller ionized fragments. Depending whether the ion is situated at the N-terminus or the C-terminus the fragment is referred to as b- or y-fragment. The two fragments originating of one peptide are complementary. The series of fragments are able to cover the entire sequence of the peptide and thus a modification at an amino acid residue can be traced to its exact localization in the peptide.

Using LC-MS/MS, a total number of 36 modifications on 5 peptides of H3 and 16 modifications on 4 peptides of H4 were detected (see Supplement Table 2-5). Specific

biochemical characteristics of some peptides precluded analysis by LC-MS/MS and, therefore, were analysed by conventional MALDI-TOF mass spectrometry to investigate their modification states. Thus, for H2A, one modification state was detected, for H2B 3 modifications, for H3 4 modifications and for H4 3 modifications were detected using MALDI-TOF (see Supplement Table 1).

5.4.1 *Post-Translational Modifications of Histones H2A and H2B*

A total of 6 peptides of histone H2A and 12 peptides of H2B were detected using MALDI-TOF mass spectrometry. For the peptide 83-89 with the sequence HLQLAVR of histone H2A mass shifts corresponding to mono- and di-methylation were detected identifying the methylation of the arginine residue 89. The level of mono-methylation was 3.31% in NF9 and was reduced to 0.8% in NF12, 0.98% in NF18 and 0.68% in NF37 (see Figure 17A and B). However, the reduction of the levels of R89me1 was not significant. With mono-methylation of 3.61% and 4.49% for A6 and XTC, respectively, the levels of mono-methylation were in the same range as the embryonic levels. Di-methylation of R89 was generally less abundant than the levels of mono-methylation. R89me2 was found with 0.66% in NF9, 0.43% in NF12 and 2.96% in NF37 and 0.34% in A6, but it was not detected in NF18 and XTC (see Figure 17A and B).

The peptide 93-99 of histone H2B has the sequence EIQTAVR. Mass shifts corresponding to di-methylation, but not to mono-methylation, were detected. For NF12 R99me2 was low with 2.92% (see Figure 17C and D). Next to di-methylation, phosphorylation was detected on the same peptide 93-99. The level of T99phos was 10.36% in NF37, 5.19% in A6 and 1.08% in XTC (see Figure 17C and D).

Phosphorylation was also detected for the peptide 87-92 of H2B. As the sequence, STITSR, of this peptide revealed, there are several amino-acids which may be phosphorylated. As this peptide was analysed using MALDI-TOF mass spectrometry which did not have the possibility for CID, precise localization of the phosphorylation was not possible. Furthermore, the phosphorylation was only detected for A6 with a level of 10% (see Figure 17E).

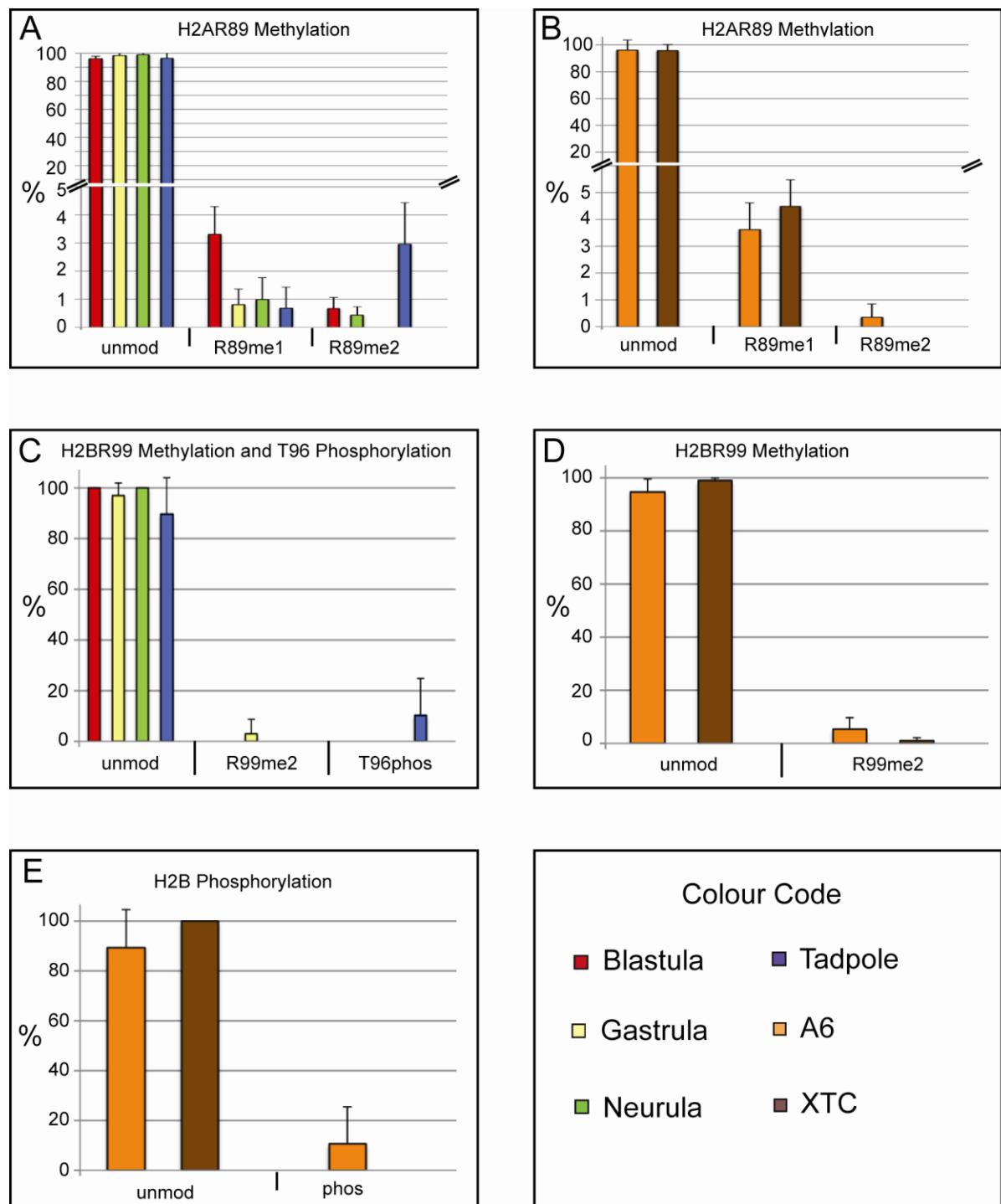


Figure 17: Histone H2A and H2B Modifications in Xenopus Embryos

Bar-Charts showing the relative abundance of histone post-translational modifications, identified by MALDI-TOF mass spectrometry. Abbreviations: unmod = unmodified peptide, Kac = acetylated lysine residue, Rme1 = mono-methylated, Rme2 = di-methylated arginine residue, phos = phosphorylation.

5.4.2 *Post-Translational Modifications of Histone H4*

For the analysis of post-translational modifications of histone H4, preferentially high mass accuracy LC-MS/MS was used. The small mass and specific biochemical characteristics of the peptide 20-23 precluded analysis by LC-MS/MS and, therefore, this peptide was analysed by conventional MALDI-TOF mass spectrometry to investigate its modification states. A total of 7 peptides of H4 were detected and in five of these peptides, post-translational histone modifications were identified (see Supplement Table 1). The peptide 4-17 with the sequence GKKGKGLGKGGAKR has four lysine residues, which according to previous findings are subjected to multiple acetylations (Kouzarides, 2007). The mono-acetylated peptides were always modified at the lysine K16. The blastula staged embryo was in 38% of the total H4 4-17 peptide acetylated at K16. The level rised significantly in later developmental stages to 44% in NF12 and remained at high levels in NF18 (43%) and NF37 (42%; see Figure 18A). The percentage of K16 mono-acetylation was lower in A6 with 38% and 34% in XTC (see Figure 18B). In this case, the amount of unmodified peptide was much higher than at any embryonic stage tested with 54% in A6 and 58% in XTC. Thus, the embryos had a higher amount of multiply acetylated H4 peptides. The highest prevalence was found in the di-acetylated N-terminus. The blastula staged embryo had 27% of the total H4 protein acetylated at K5 plus K12. This modification, however, was not present at any other condition tested, neither in the embryos nor in the two cell lines (see Figure 18A and B). The gastrula, though, had 11% of K8 plus K16 acetylation. This combination of acetylation marks peaked at the neurula stage with 17% and decreased again to 11% in the tadpole and even further to 6.6% and 6.8% in the two cell lines. The blastula staged embryo, however, had none of this di-acetylation pattern. Considering tri-acetylation, there was one combination detected, which was K5ac plus K8ac and K16ac. The global trend of this modification was a reduction from 9% in NF9 to 3% in tadpole stage with a transient peak of 7% in the neurula stage. The two stable cell lines showed lower levels of this combinatorial mark with both times 1.3% compared to embryos. The developmental reduction in tetra-acetylation was similar to the tri-acetylation mark. Globally, there was a reduction in the levels of tetra-acetylation with an exception at the neurula stage embryo, which had the highest levels with 3.4%. The amount of tetra-acetylated H4 peptide decreased from 1.9% in NF9 to 1.15% in NF37 and to 0.3% in both of the cell lines.

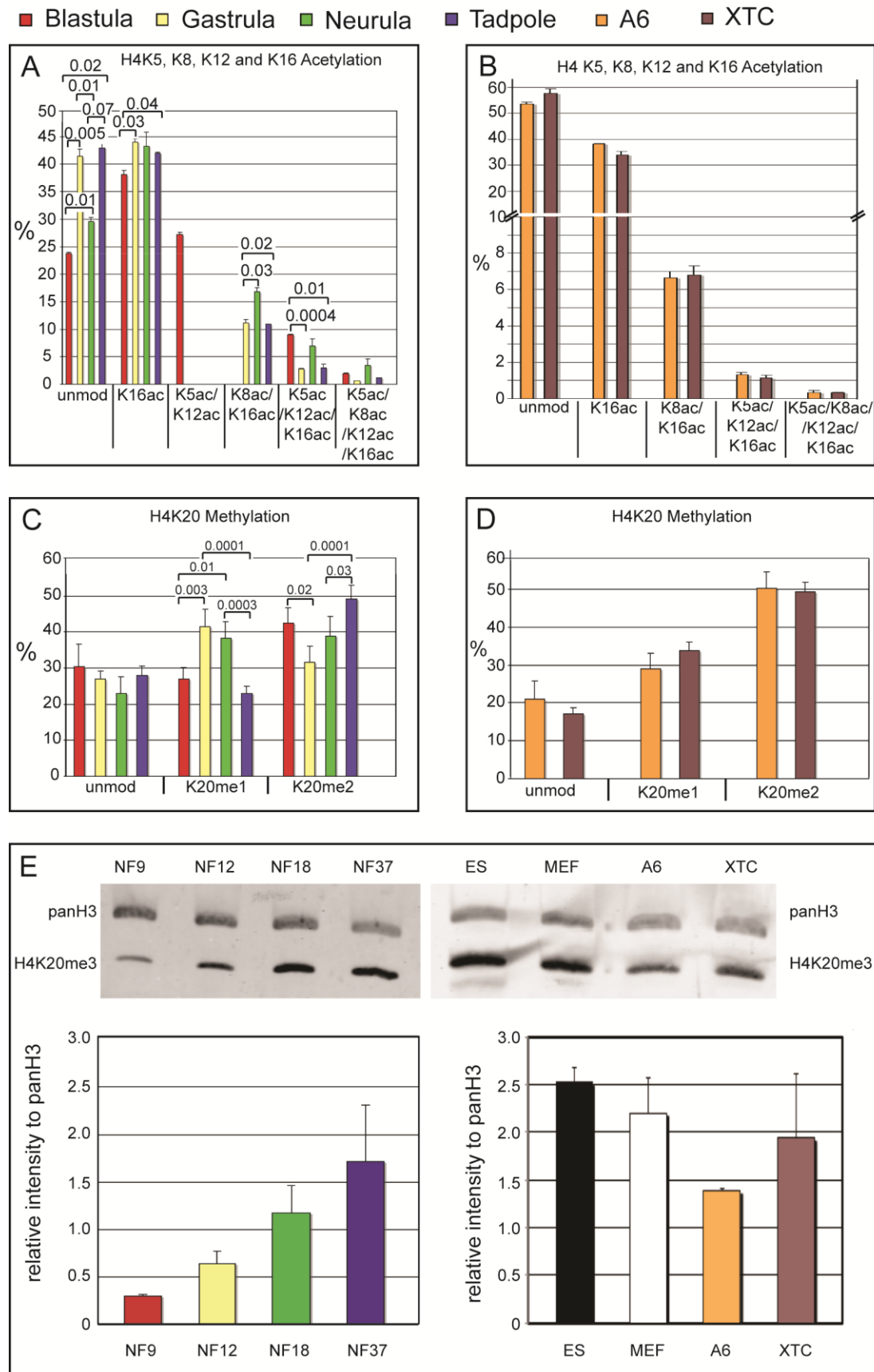


Figure 18: Histone H4 Modifications in Xenopus Embryos and Cell Lines

Bar-Charts show the relative abundance of histone modifications, identified either by Orbi-Trap LC-MS/MS mass spectrometry (**A+B**; error bars indicate SD of two independent biological replicates) or MALDI-TOF mass spectrometry (**C+D**; error bars indicate SD of three independent biological replicates). **E**) Immunoblotting for total histone H3 and histone H4 tri-methylated at lysine 20. Upper panel - Western blot Odyssey infrared imaging signals; lower panel - bar chart showing increase of H4K20me3 levels relative to total histone H3 protein during development. Abbreviations: unmod = unmodified peptide, Kac = acetylated lysine residue, Kme1 = mono-methylated, Kme2 = di-methylated, Kme3 = tri-methylated lysine residue. Where applicable, p values are given by numbers above brackets to indicate significant differences in the abundance of a histone modification between samples.

Overall, this data indicated a general trend towards reduction of multiple acetylated lysine residues in the tail of H4 as the embryo developed from blastula to tadpole stage. At blastula stage, there was a total level of multiple acetylation marks of 38% decreasing to 17% in tadpole stage and 8% in XTC cells.

The tryptic peptide 20-23 from propionylated histone H4 with the sequence KVLRL has one lysine residue at the position 20, which is known to be methylated (Schotta *et al.*, 2004). Its different methylation states have been linked to diverse epigenetic functions. The H4K20me3 marks are correlated with the formation of pericentromeric heterochromatin, maintenance of genome integrity and transcriptional repression, while the role of H4K20me1 is controversial (Dambacher *et al.*, 2010). Using MALDI-TOF, peptides corresponding to unmodified, mono-methylated and di-methylated states were detected in all samples (see Figure 18C and D). The amount of the unmodified peptide decreased slightly from blastula to tadpole stage, and was even lower in the A6 (20%) and XTC (17%) cell lines. It is known that the specific properties of the tri-methylated H4 peptide 20-23 makes it very difficult to quantify this PTM by mass spectrometry, and reproducible detection in the samples was not possible. To determine the relative abundance of H4K20me3 in frog embryos, Western blot analysis with antibodies against tri-methylated H4K20 as well as pan-histone H3 was performed. This confirmed the consistent presence of the H4K20me3 mark during embryonic development (Figure 18E), and revealed also that its abundance increased approximal five-fold from blastula to tadpole stages.

Next to the two peptides described above, mass shifts corresponding to post-translational histone modifications were detected by MALDI-TOF in the peptides 24-35, 56-67 and 79-92 (see Supplement Table 1). The peptide 24-35 has the sequence

DNIQGITKPAIR. A mass shift corresponding to acetylation of lysine 31 was detected. The overall abundance of this modification ranged between 0.17% in NF12 and 1.49% in XTC. Furthermore, a mass shift corresponding to a mono-methylation of the peptide 56-67 was found in NF12-NF37. The level of K59me1 ranged between 0.67% in NF18 and 6.6% in NF37. The H4 peptide 79-92 with the sequence KTVTAMDVVYALKR has two lysine residues at position 79 and 91, respectively, explaining the two mass shifts corresponding to mono- and di-acetylation. The di-acetylated peptide was found in the *Xenopus laevis* embryonic stages with the lowest abundance in NF37 with 0.79% and the highest abundance in NF18 with 3.34%. The level of di-acetylation was 0.53% in XTC.

5.4.3 *Post-Translational Modifications of Histone H3*

For the analysis of post-translational modifications of histone H3 preferentially high mass accuracy LC-MS/MS was used. However, the small mass and specific biochemical characteristics of the peptide 3-8 precluded analysis by LC-MS/MS, and, therefore, this peptide was analysed by conventional MALDI-TOF mass spectrometry to investigate its modification states. 8 out of 9 peptides from H3 were found to be modified.

The peptide 3-8 with the sequence TKQTAR has one lysine residue which is known to be methylated (Zhang *et al.*, 2003). As for specific biochemical characteristics, this peptide was difficult to detect in MALDI-TOF (see Figure 19A). To improve the detectability of the peptide 3-8, a special enrichment technique with charcoal mini-tips (Carbon-Tips, Glygen) was used. According to Chin and coworkers, the peptide 3-8 is bound selectively to the charcoal and thus is enriched (Chin *et al.*, 1998). Our data indicated an enrichment of the peptide as detection was improved as well as the signal to noise ratio was increased (compare Figure 19A and B). By the use of this enrichment technique, analysis of modification of the H3 peptide 3-8 was possible.

The amount of the unmodified peptide 3-8 increased significantly during embryonic development from 27% in NF9 to 44% in NF37. This was largely due to a decrease in tri-methylation from 28% in NF9 to 16% in NF37. The levels of the mono- and di-methylated isoforms remained relatively constant. In comparison, the two cultured frog cell lines had much higher levels of unmodified (67% in A6 and 62% in XTC), and much lower levels of di- and tri-methylated H3K4 (10% each; Figure 20A and B). This data indicated an unusual and transcriptionally permissive state of the uncommitted blastula

epigenome, in which half of the histone H3 tails appeared to be di- or tri-methylated at Lysine 4.

One of the best characterized modifications that is associated with heterochromatin and transcriptionally repressed regions is di-/tri-methylation of lysine 9 on the H3 tail.

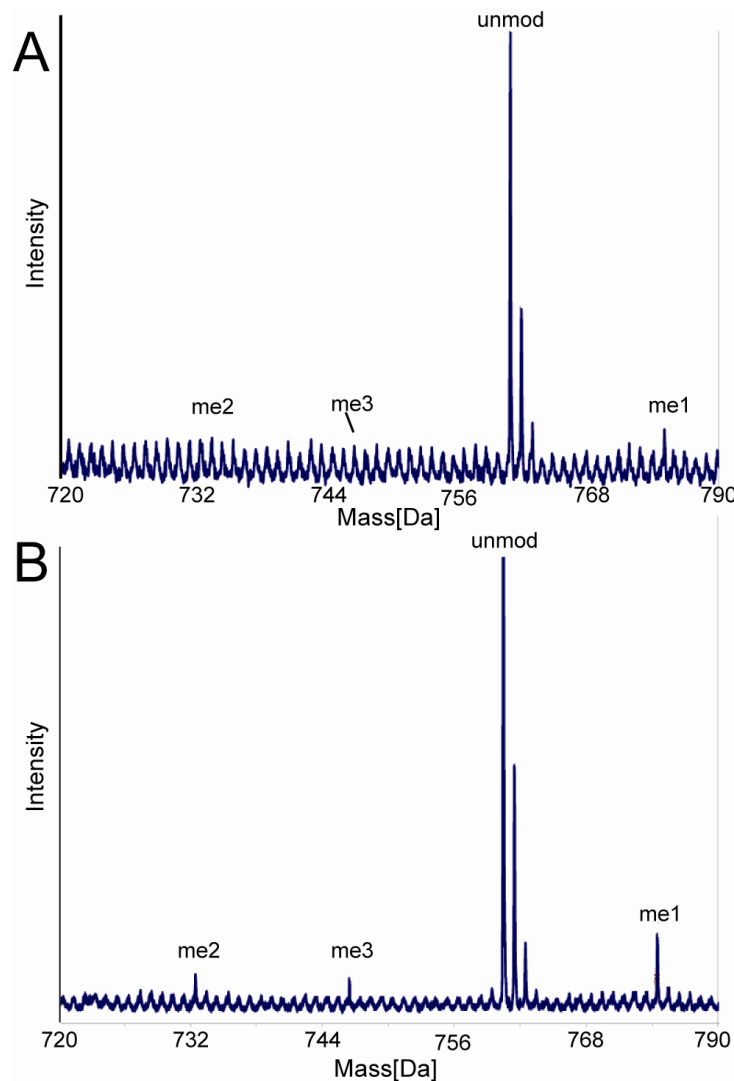


Figure 19: MALDI-TOF Spectra of Porous Graphite Carbon Enrichment of H3 Peptide 3-8

A) MALDI-TOF spectra of H3-3-8 peptide **B)** MALDI-TOF spectra of enriched H3 3-8 peptide with porous graphite carbon. MALDI-TOF spectra display mass on x-axis and intensity on y-axis. Abbreviations: unmod = unmodified peptide, me1 = mono-methylated, me2 = di-methylated, me3 = tri-methylated lysine residue.

In contrast, K9 mono-methylation has been associated with regulatory DNA elements and maintenance of activation potential during differentiation. K9 mono-, di- and tri-

methylation was present in all the samples, although at very different levels (Figure 20C). The different modification states of K9 fluctuated less than two-fold during frog development, except for K9me1, which showed a small but significant transient peak at gastrula, when it accounts for 20% of the total H3 9-17 peptides. A similar distribution of mono-, di- and tri-methylated K9 was found also in the two cultured *Xenopus* cell lines (Figure 20D).

Combined, the K9me2/me3 marks made up for less than 4% of the total H3 tails in the different *Xenopus* samples, suggesting a rather low density of this mark in both constitutive and facultative heterochromatin.

The peptide 9-17 with the sequence KSTGGKAPR has two lysine residues at the position 9 and 14 and a serine residue at the position 10 which is known to be phosphorylated (Kouzarides, 2007). At low abundance, di-acetylated peptides (K9ac + K14ac) were detected in all stages, however, acetylation of K9 without the presence of K14c was not found (Figure 20E and F). The level of mono-acetylation of this peptide decreased significantly as development proceeded from 45% in blastulae to 31% in gastrulae, after which it increased again gradually up to 37% in tadpoles. As it is the case for H4 mono-acetylation, the tissue culture cells contained less acetylated H3 (Figure 20F). Di-acetylated H3 K9/K14 molecules had low abundance and their levels did not change significantly during development. However, the levels were slightly lower again in the two frog cell lines, compared to blastulae, which contained the lowest value of the four embryonic stages (1.17%).

In addition to the K9 and K14 residues, the nearby lysines 18 and 23 on the H3 tail can also be acetylated. A similar pattern for both mono-acetylated states with a significant shift between NF9 and the three other embryonic stages was found (Figure 20G). Like for the K9/K14 pair, K18 acetylation was found only in combination with acetylated K23. These levels were between 4% in NF12 and 9% in NF37. Mono-acetylated K23 ranged from 57% in blastula to 45% in the other embryonic stages. In the two *Xenopus* cell lines, the level of K23 mono-acetylation was around 20% and the di-acetylation at 2% (see Figure 20H). Thus, in the same manner as for H3K9 and K14, the acetylation of K18 and K23, was very different comparing the four embryonic samples and the somatic cell lines from *Xenopus*.

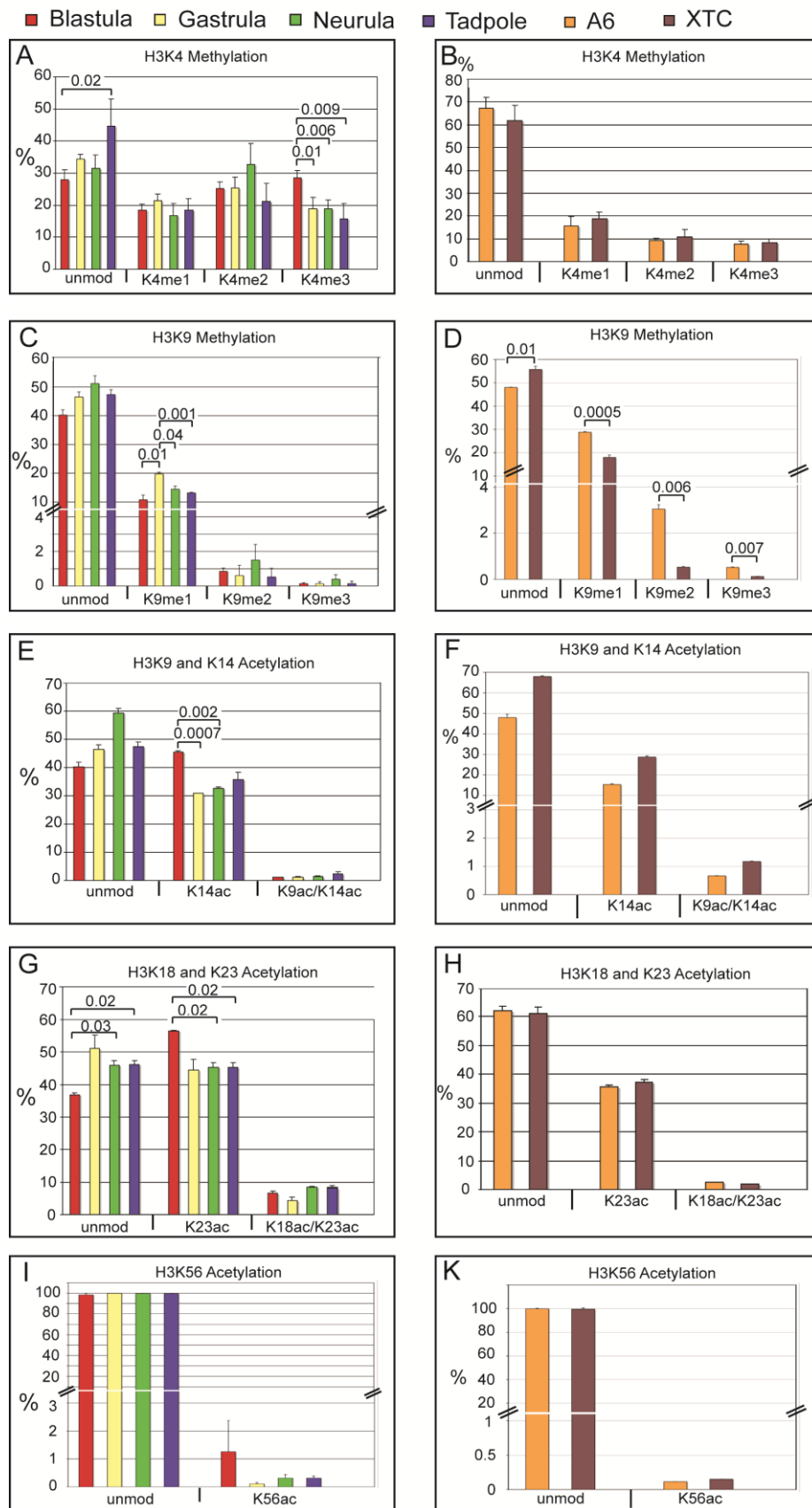


Figure 20: Histone H3 Modifications in Xenopus Embryos and Cell Lines

Bar-Charts showing the relative abundance of histone modifications, identified either by MALDI-TOF mass spectrometry (**A+B**; error bars indicate SD of three independent biological replicates) or Orbi-Trap LC-MS/MS mass spectrometry (**C-K**; error bars indicate SD of two independent biological replicates). Abbreviations: unmod = unmodified peptide, Kac = acetylated lysine residue, Kme1 = mono-methylated, Kme2 = di-methylated, Kme3 = tri-methylated lysine residue. Where applicable, p values are given by numbers above brackets to indicate significant differences in the abundance of a histone modification between samples.

Besides N-terminal acetylation marks, acetylation in the H3 histone-fold domain at Lysine 56 was also detected (H3K56ac; Figure 20I and K). The frog blastula contained 1.3% of acetylated K56, and the level of this mark was decreasing to 0.11% in the gastrula stage, however, not significantly (see Figure 20I and K). After this, the abundance of acetylated K56 remained low with 0.32% in the tadpole embryo. Notably, both cultured frog cell lines contained K56ac levels around 0.1%, which were comparable to the lowest level found in the gastrula stage (Figure 20I). Thus, the overall abundance of K56ac mark was quite low in frog embryos, but decreased rapidly during the blastula/gastrula transition.

Furthermore, the peptide 64-69 was found to be di-methylated at levels of 0.74% in NF9 to 3.56% in MEFs (see Supplement Table 1). According to Zhang the peptide 73-83 is methylated at lysine 79 (Zhang *et al.*, 2003). By detection of mass shifts corresponding to mono-, di- and tri-methylation of lysine 79, these findings were confirmed. The mono-methylated peptide's abundance increased as the embryo developed further from 2.3% in NF9 to 7.2% in NF37. In the cell lines A6 and XTC, the mono-methylation level varied between 4.4% in XTC and 5.94% in A6. On the other hand, the tri-methylation was decreasing from 1.44% at NF9 to 0% at NF37. The level of di-methylated peptide was low in all embryonic samples. For the H3 peptide 117-128 a mass shift corresponding to a di-methylation with an abundance of 1.68% was detected only for NF37 (see Supplement Table 1).

Methylation of lysines 27 and 36 within the histone H3 tail has been mapped to mutually exclusive regions within the genome of most eukaryotes. The H3K27me2/me3 marks were almost exclusively found at inactive genomic regions, whereas the H3K36me2/me3 marks were localized predominantly in actively transcribed genes (Barski *et al.*, 2007). Both modifications reside on the same tryptic peptide (H3, 27-40; see Figure 21A), which provided the possibility to determine combinations of modification states on these two lysines by mass spectrometry.

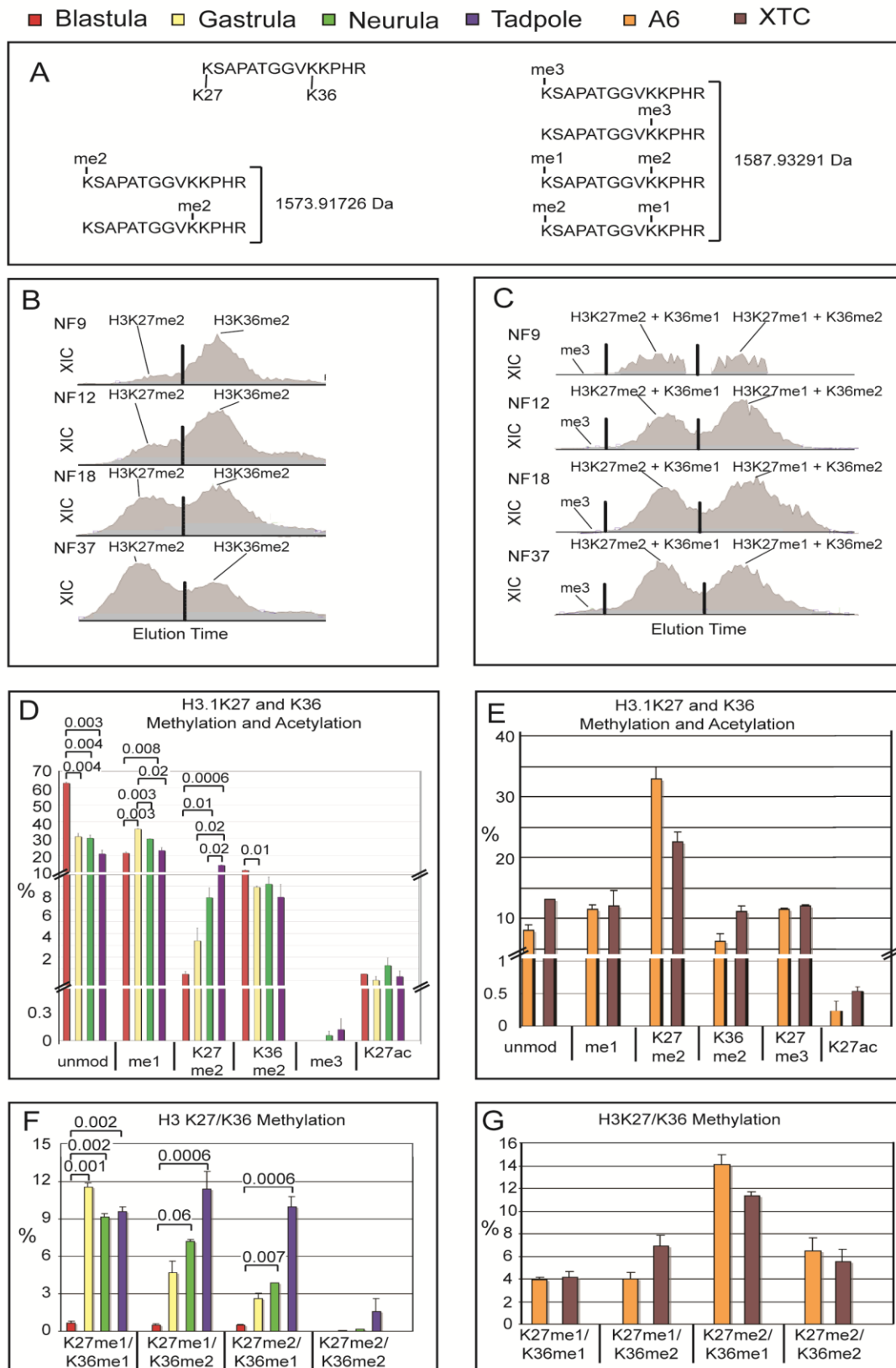


Figure 21: Histone Modifications on the H3 27-40 Peptide in Xenopus Embryos and Cell Lines

A) Amino-acid sequence of the H3 27-40 peptide and its isobaric modification forms, which have identical mass although being differently modified at K27 and K36 residues. Panels **B+C**; Extracted ion chromatograms (XICs) showing separation of isobaric H3 27-40 peptides from the four *Xenopus* embryonic stages by differential elution from a C18 micro-column on the reversed-phase HPLC of on-line mass spectrometry. X-axis represents retention time, y-axis the intensity of ion currents in the quadrupole Orbi-Trap mass spectrometer. Peak separation was called by the ICIS peak detection algorithm program (Thermo), indicated here by the vertical black lines. **B**) XICs of isobaric di-methylated peptides, modified at either K27 or K36. **C**) XICs of isobaric tri-methylated peptides representing K27me3/K36me3, K27me2+K36me1 and K27me1+K36me2. **D+E**) Bar-Chart of H3K27 and K36 modification states. Note that isobaric mono- (data not shown) and tri-methylated peptides, which are methylated either at K27 or K36, elute simultaneously and cannot be distinguished. **F+G**) Bar-Chart of combinatorial K27/K36-methylated peptides. Abbreviations: unmod = unmodified peptide, me1 = single mono-methylated lysine at position 27 or 36, K27me2 = di-methylated lysine 27, K36me2 = di-methylated lysine 36, Kme3 = single tri-methylated lysine at position 27 or 36, K27me1/K36me1 = double mono-methylated, K27me2/K36me2 = double di-methylated, K27me1/K36me2 = combinatorial tri-methylated peptide with dimethylated K36, K27me2/K36me1 = combinatorial tri-methylated peptide with dimethylated K27. Error bars represent SD. Where applicable, p values are given by numbers above brackets to indicate significant differences in the abundance of a histone modification between samples.

However, as peptides carrying a single tri-methyl group at either K27 or K36, or combinations of a mono-methyl and a di-methyl group at these two lysines, are isobaric (i.e. they have the same mass), they cannot be distinguished solely by this parameter (see Figure 21A). Although tandem MS/MS analysis can facilitate the assignment of the modifications, the relative quantification based on extracted ion chromatograms was only possible, when the differentially modified isobaric peptides were physically separated.

Indeed, tandem mass spectra revealed that most peptides that have a higher methylation degree at position 27 eluted earlier from a C18 reversed phase column than the corresponding ones that carry the same number of methyl groups on K36 (Figure 21B and C). With the exception of the isobaric mono-methylated peptides (i.e. K27me1 or K36me1), which were not separated on the C18 micro-column, 12 additional, distinct methylation states for the K27 and K36 positions on single H3 27-40 peptides were identified.

In genome wide studies, tri-methylation of H3K36 is found at strongly expressed genes, which are also preferentially enriched for the replication-independent histone variant H3.3. On canonical histone H3 (i.e. replication-dependent H3.2) that is highly

abundant in early embryos, H3K36me₃ was detected only in frog tadpoles at very low abundance by its fragmentation spectrum (see Figure 21D and E; Supplement Table 4). The ratio of H3.3 to H3.2, as measured by comparing all spectra derived from the peptides unique for each isoform, decreased during development from 20% in blastula to approximately 7% in tadpoles (see Supplement Figure 2). However, a very similar distribution of modifications within these peptides was found in all cases, and therefore the analysis was focused on the peptide derived from the more abundant, canonical H3 molecule. Overall, this data indicates that H3K36me₃ levels are low during frog development.

Consequently, these findings implied that practically all the single tri-methylated 27-40 peptides of H3 represent K27me₃, a conclusion that was supported by the fragmentation spectra (see Supplement Table 4). H3K27me₃ was readily identified in many, but not all samples. Notably, it was not detected in blastula and gastrula, and appeared only at very low levels in neurula (0.05%) and tadpole stages (0.1%; Figure 21B and D). However, in *Xenopus* cell lines (Figure 21E), as well as in mammalian cells (see below), H3K27me₃ levels were at least 100-fold higher than in bulk embryonic chromatin up to neurula stages. It was concluded that the repressive K27me₃ modification is present in the histone extracts, but extremely underrepresented in the early frog embryo compared to cultured frog and mammalian cell lines.

Overall, the abundance of methylated H3K27 rose during development (Figure 21B-G). This trend became particularly obvious, when the XIC profiles of the dimethylated peptide species were compared to each other. While the K36me₂ peak stayed more or less constant, the K27me₂ peak increased gradually from 1.6% in the blastula to 14% in the tadpole stage (Figure 21D). During this time, the ratio of H3K27me₂ to H3K36me₂ changed from 0.14 in the blastula to 1.6 in the neurula stage, i.e. by more than 10-fold. In comparison, the cultured cell lines A6 and XTC showed the highest percentage of H3K27me₂ peptides (33% and 22% respectively), and contained more than 10% of H3K27me₃ (Figure 21E).

The abundance of peptides with combinatorial methylation on both K27 and K36 residues rose significantly during development. In the frog blastula, less than 2% of the H3 tails were simultaneously modified on these sites, while one third is combinatorially modified in tadpoles (Figure 21F and G). Among these, low levels of unexpected combinations such as K27me₃ paired with K36me₁/me₂ in late embryonic and tissue

culture cell samples were found, too. In general, the level of modifications within this peptide showed a major increase during development. This trend may be connected functionally to the establishment of stable gene expression profiles as cells differentiate.

5.4.4 *Post-Translational Histone Modifications of murine ES Cells and MEFs*

Most cells in the blastula stage of *Xenopus* embryos are uncommitted and capable to differentiate into derivatives of all three germ layers (Heaseman *et al.*, 1984). This raised the issue, whether the observed histone modification profiles in the frog blastula represented common features of pluripotent cells. To address this point, the histone PTMs of murine ES cells and of primary embryonic fibroblasts (MEFs), which were used as feeder cells in the ES cell culture and represent a murine somatic cell type for comparison, were also analysed.

The degree of acetylation on the H3 18-26 and H4 4-17 peptides of MEFs and ES cells indicated that both cell types were separated successfully from each other (see Supplement Figure 3). In depth analysis confirmed then that the two murine cell types contained very specific profiles of histone modifications, and that the histone modifications of pluripotent cells from *Xenopus* blastulae were related to, but also clearly distinguishable from the profiles of pluripotent murine ES cells. Bivalent chromatin fragments, which simultaneously carry the active H3K4me3 and the repressive H3K27me3 marks, are considered a hallmark of pluripotent ES cells. With this technical approach it was not determined, whether histone H3 tails were modified on both Lysine 4 and 27, since these residues were located on different tryptic peptides. Nevertheless, the overall abundance of modifications at these sites was observed. For H3K4, this comparison revealed a high similarity between *Xenopus* blastula and murine ES cells (Figure 22A). Of all samples, these two contained the lowest amount of unmodified K4 (below 30%), but maximal amounts of H3K4me2/me3 (>50% together). In contrast, H3K4me1 levels were not only similar between blastula, ES cells and MEFs, but almost constant in all samples (Figure 22A). In contrast to H3K4, the methylation profiles of the H3 peptide 27-40, including Lysines 27 and 36, were very different between blastulae and ES cells. As described earlier, the majority of the di-methyl marks were situated on K36 in frog blastulae.

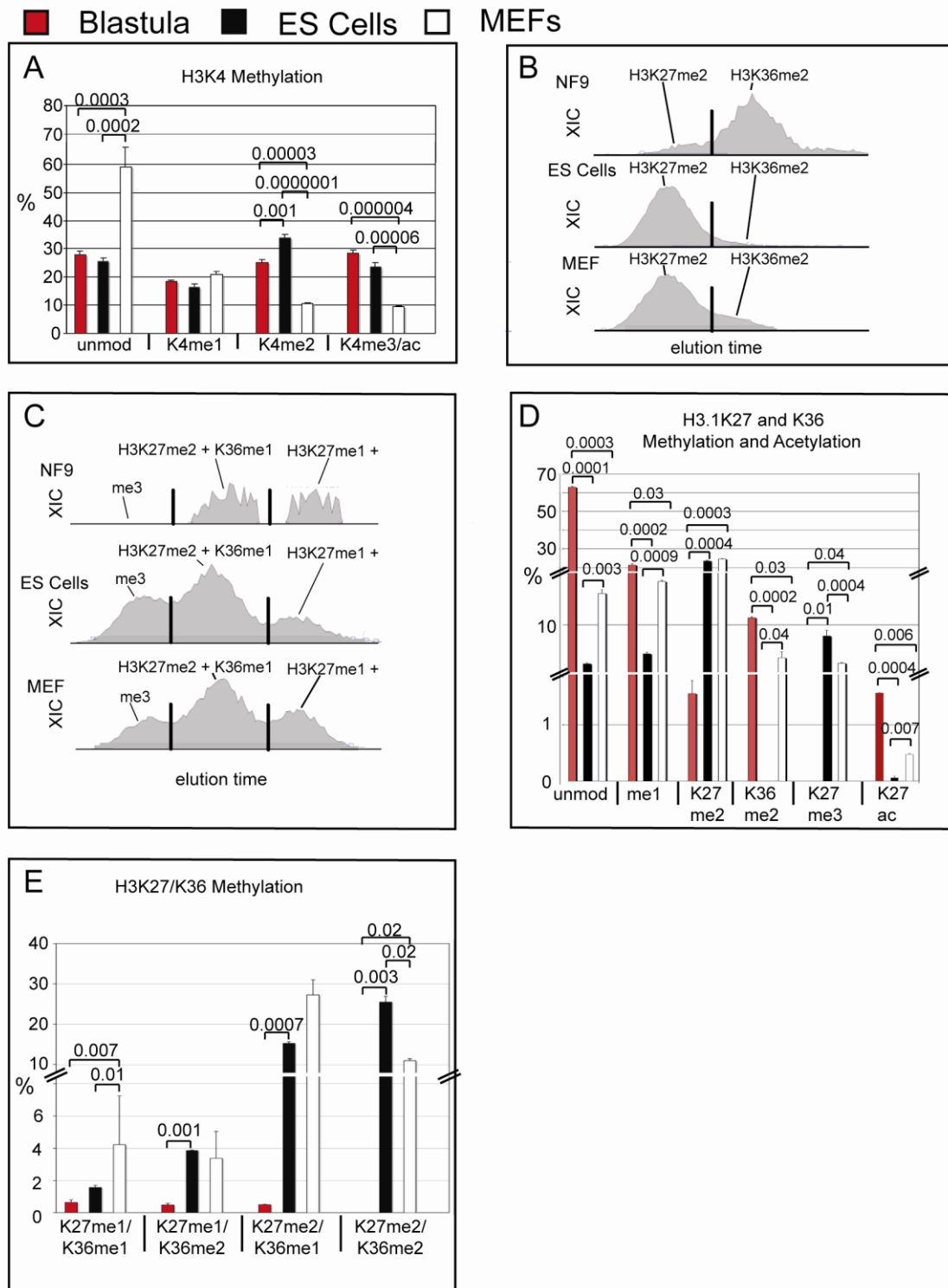


Figure 22: Opposing Histone Modifications on the H3 27-40 Peptide of *Xenopus* Blastulae Compared to Murine ES Cells and MEFs.

A) Bar-Charts showing H3K4 methylation states of *Xenopus laevis* blastulae, ES cells and mouse embryo fibroblasts. Data obtained by MALDI-TOF mass spectrometry, error bars indicate SD of three independent biological replicates. B) XIC profiles of K27me2 and K36me2 modification states of the H3 27-40 peptide of *Xenopus laevis* blastulae, murine ES Cells and MEFs on a C18 micro-column on a reversed-phase.

HPLC during Orbi-Trap on-line mass spectrometry. C) XIC profiles for K27/K36me3, K27me2+K36me1 and K27me1+K36me2 of the H3 27-40 peptides from the same samples. In B and C, x-axis gives relative elution times of peptides, the y-axis shows their intensity according to the ion current of the quadrupole Mass Spectrometer. Peak separation was called by the ICIS peak detection algorithm program (Thermo), indicated here by vertical black lines. D) Bar-Chart of mutually exclusive H3K27 and K36 methyl states. E) Bar-Chart of peptides with combinatorial H3K27 and K36 methylation. D and E - data from Orbi-Trap Mass Spectrometry, error bars indicate SD of two independent biological replicates. Abbr.: unmod = unmodified peptide, Kme1 = mono-methylated lysine residue, Kme2 = di-methylated lysine residue, Kme3 = tri-methylated lysine residue. Where applicable, p values are given by numbers above brackets to indicate significant differences in histone modifications between samples.

In ES cells, however, they were found predominantly on H3K27 (Figure 22C and D). MEFs had a similar di-methyl-distribution like ES cells, but contained significantly more K36me2 than these. A dichotomy was also found for the higher modification states of this peptide, including three to five methyl groups (Figure 22E). Whereas in frog blastulae, H3K27me3 was below the detection limit, mouse ES cells had with 8.3% the third-highest K27me3 level of all samples tested, and twice as much as MEFs (Figure 22D). Combinatorial tri-methylated states, such as K27me1/K36me2 and K27me2/K36me1 were also significantly more abundant in ES cells than in frog blastulae, and comparable to the levels found in MEFs (Figure 22E). Notably, the highest levels of tetra- (K27me2+K36me2) and penta-methylated (K27me3+K36me2) states of the H3 27-40 peptides were also found in ES cells and to a less extent in MEFs, but were not detected in frog blastulae (Figure 22E). Therefore, with the exceptions of the mono-methyl (K27 or K36) and the K36me2 states, frog blastulae were characterized by the lowest methylation levels for the H3 27-40 peptide of all the samples tested.

Overall, these results indicate that frog blastulae and murine ES cells share high H3K4me2/me3 levels, but the former were dramatically undermethylated at the H3 27-40 peptide. In fact, uncommitted frog cells had a 15-fold higher level of the unmodified state of this peptide than pluripotent ES cells.

6 Discussion

In this study, a variety of different protocols for histone extraction from *Xenopus laevis* embryos was tested and evaluated. In concern of this test, a reliable and efficient protocol for nuclei preparation and histone extraction from *Xenopus laevis* embryos was established. By the use of both MALDI-TOF and the Orbi-Trap as a tandem mass spectrometer, many core histone modifications were quantitatively determined. The techniques for both identification and quantification of histone modifications in mass spectrometry were improved.

A total number of 63 modification states was detected on the four core histones. Specifically, using LC-MS/MS a total number of 36 modifications on 5 peptides of H3 and 16 modifications on 4 peptides of H4 were detected. Analysis by conventional MALDI-TOF mass spectrometry revealed one modification state of H2A, 3 modifications of H2B, 4 modifications of H3 and 3 modifications of H4. The histone modification states at important developmental time points of *Xenopus laevis* were analysed and compared to the histone modification profiles of somatic cell lines. Furthermore, the quantitative shifts, which were observed during development, were clustered into stage-specific histone modification profiles accompanying and potentially regulating the transition from pluripotent to determined cell states. As a characteristic histone modification profile of the pluripotent blastula staged *Xenopus* embryos was determined, the *in vivo* data of the blastulae were compared to the profile of murine pluripotent ES cells. Despite the pluripotent nature of both *Xenopus* blastulae and murine ES cells, there were significant differences in the histone modification profiles of both samples.

This mass spectrometric analysis of histone modifications present in bulk embryonic chromatin through *Xenopus laevis* development revealed major quantitative shifts for several histone modifications known to be involved in gene regulation, and has also identified specific differences between pluripotent cells from blastula staged *Xenopus* embryos and later developmental stages such as tadpole or even the somatic cell lines A6 and XTC. Thus, the study complemented the data of Bonaldi and colleagues who did analysis of histone PTMs at three distinct time points of *Drosophila* development (Bonaldi *et al.*, 2004) and Nicklay and Schechter, who analysed predeposition histone modifications of *Xenopus* sperm and oocytes (Nicklay *et al.*, 2009; Schechter *et al.*, 2009).

6.1 Technical Aspects of the Histone Extraction

The extraction and purification of histone proteins is frequently described in the literature as a non-demanding method (Murray *et al.*, 1966; Burton *et al.*, 1978). For the subsequent analysis of the histone modifications by mass spectrometry, however, a sufficient purity as well as quantity of histones is required. In order to determine the yield as well as the suitability for mass spectrometric modification analysis, different frequently used histone extraction protocols were tested.

Common protocols for histone extraction rely on acidic histone extraction (Murray *et al.*, 1966) or high salt extraction (Burton *et al.*, 1978). Next to these protocols, a variety of chromatographic methods has been applied to purify histones, such as the selective absorption of histone-DNA complex to the hydroxyapatite matrix (Simon *et al.*, 1979) and the separation of histones from contaminants by reversed-phase HPLC (Certa *et al.*, 1981; Lindner *et al.*, 1989). While these methods provide histones for Western blot analysis of post-translational modifications, it has been very difficult to use them for mass spectrometric analysis. Consequently, mass spectrometric analysis of histone modifications from embryos has rarely been published (Bonaldi *et al.*, 2004; Nicklay *et al.*, 2009).

As the histone proteins have high abundance of basic amino acid residues, the pI ranges around 12. Consequently the histone proteins can be enriched and purified using 0.4M hydrochloric acid, because the high acid concentrations lead to the precipitation of most proteins whereas the histones remain soluble because of the high pI. Although the method was the first to be introduced, it remains the methods of choice for many studies (Rodriguez-Collazo *et al.*, 2009; Schechter *et al.*, 2007).

The experiments with the cell lines A6 and XTC as well as with *Xenopus* embryos revealed that acidic extraction of nuclei with both hydrochloric and sulfuric acid had the highest yield of all protocols tested (see Figure 7C). Important for the high yield and purity of the extraction was the proper isolation of nuclei in advance to the histone extraction. The development of an efficient and reliable method to isolate intact nuclei from embryos of different developmental stages was the most critical prerequisite for subsequent isolation of histones by acidic extraction. The conclusion of these findings is, that acidic histone extraction is the method of choice as it had the highest yield of all protocols tested, but proper nuclei extraction is the determinant of success.

High salt extraction is described as a useful alternative to acidic histone

extraction. Salt extraction relies on the different affinities of non-histone proteins and histones to DNA. Non-histone proteins dissolve from the DNA at low salt concentrations, followed by the linker histones. The four core histones remain bound to the DNA until 2M salt buffers are applied. Thus this method is described to purify and enrich histone proteins by collection of eluting fractions (Burton *et al.*, 1978). High salt extraction for *Xenopus* cell lines was performed and the core histone proteins were purified as described (see Figure 6). However, the yield of histone proteins was very low compared to acidic extraction and the purity was not sufficient for subsequent mass spectrometric analysis.

Summarizing, these findings suggested that high-salt extraction is a useful alternative for histone purification, however it seemed not to be suited for mass spectrometry analysis of post-translational histone modifications. Despite these results, this method might still be useful for the analysis of post-translational histone modifications of the linker histones.

Hydroxyapatite purification of proteins was first described by Bloom in 1977 (Bloom *et al.*, 1977). Simon and Felsenfeld applied this technique the first time for the purification of histone proteins (Simon *et al.*, 1979). However, the method is also described to require high amounts of starting material together with a low yield of only 20-30% (Simon *et al.*, 1979; Kurokawa *et al.*, 1984).

On the other hand, reversed-phase HPLC purification of histones is described as a powerful technique on an analytical as well as preparative scale. The recovery is around 80-90% of input material and the method is described to be efficient and quick (Lindner, 2008). The purification is achieved by a separation of the mixture of proteins by elution over time. Depending on the use of mobile and stationary phase and the gradient, the method is able to even separate differently modified histone proteins (Schechter *et al.*, 2007).

The experiments of both chromatographic techniques used for the separation and thus purification of histones yielded histone proteins in the hydroxyapatite purification, but not in the reversed-phase HPLC. The amount of histone proteins after the purification was not sufficient for further mass spectrometric analysis in both techniques (see Figures 10 and 11).

Considering these findings, chromatographic techniques may offer a convenient approach for large cell numbers, but the yield is too low for limiting starting material such

as embryos.

In summary, among the different protocols described for histone extraction in the literature, there were major differences concerning both quantity and quality of both purification and enrichment. Successful histone extraction of *Xenopus* embryos was finally achieved by proper nuclei preparation prior to hydrochloric acidic extraction.

As pointed out before, the preparation of nuclei from *Xenopus* embryos was crucial for the histone extraction. In the literature, this step is not frequently described. Histone analysis by Dimitrov and colleagues was done by autoradiography, because histone visualization using silver staining in 2D gels was not possible due to low abundance of histones after the extraction (Dimitrov *et al.*, 1993).

Notably, a procedure for the isolation of nuclei from *Xenopus* embryos together with the subsequent histone extraction sufficient for mass spectrometry has not been described so far.

One of the first protocols published describes the preparation of nuclei by the use of an ultracentrifugation step. This method was described to yield about 90% of intact nuclei (Blobel *et al.*, 1966). One difficulty for the application of protocols to *Xenopus laevis* embryos is the nuclear size, being around $2000\mu\text{m}^3$ at the blastula stage. The nuclear size constantly decreases to $300\mu\text{m}^3$ at the tadpole stage (Heaseman *et al.*, 1984).

Applying the protocol described by Blobel and colleagues did neither yield intact nuclei nor histones (see Figure 9). Thus, the need to develop a new protocol for the nuclei preparation of *Xenopus laevis* was obvious. The stepwise approach to the new method revealed that as the size of the nuclei varies, the homogenization conditions had to be adapted to the nuclei volume. Consequently, the method was optimized for each embryonic stage (see Figure 12D). The efficiency of the method, yielding high purity as well as enrichment, was able to decrease the amount of starting material necessary for histone extraction. Furthermore this protocol solved the technical problems with the embryo specific contaminants such as yolk and lipids. The subsequent acidic extraction of the nuclei then yielded enough material for several technical repeats as well as of enough purity for SDS-PAGE and mass spectrometric analysis. The high efficiency of the protocol reduced also the amount of material required for Western blot analysis to less than one embryo equivalents per lane. This amount was totally sufficient for reliable detection of rare histone modifications reducing the amount of starting material by one to two orders

of magnitude (Dario Nicetto, personal communication).

In brief, the only protocol for nuclei preparation of *Xenopus* embryos described by Blobel yielded only insufficient amounts of histones (Blobel *et al.*, 1966). Other nuclei preparation protocols - such as hydroxyapatite and reversed-phase HPLC purification - did not result in histones sufficient for subsequent analysis. By the step-wise establishment of a method for nuclei preparation of *Xenopus laevis* embryos, histone extraction was possible in a quick, efficient and reliable manner. Furthermore this protocol was well suited for both mass spectrometric and Western blot analyses.

6.2 Technical Aspects of the Histone Modification Analysis by Mass Spectrometry

Mass spectrometry is a highly useful tool for the investigation of histone post-translational modifications, as it allows quantification over a high dynamic range and of several different modifications as well as identification of novel histone modifications (Neilson *et al.*, 2011). For these reasons, mass spectrometry has frequently been used for detection of histone modifications of different sources (Zhang *et al.*, 2003; Bonaldi *et al.*, 2004; Phanstiel *et al.*, 2007; Nicklay *et al.*, 2009). However, some histone modifications require special techniques for proper detection using mass spectrometry. For example, ubiquitination still poses a challenge as this modification is often heterogenous, and on SDS-PAGE the modification leads to a mass shift due to the ubiquitination. Nevertheless, improvements with ubiquitination detection have been made by special chromatographic means and the use of ultra high resolution tandem MS (Kirkpatrick *et al.*, 2005). SUMOylation and ADP-ribosylation of histones also require special technical means for mass spectrometric detection. While SUMOylation is still difficult to detect, a first report describes the fragmentation spectrum for ADP-ribosylation. A specific cleavage of ADP-ribosyl-arginine into N-ADP-ribosyl-carbodiimide and ornithine was described as a specific fragmentation pattern for this modification (Osago *et al.*, 2009). Arginine methylation has low abundance and other modifications, such as phosphorylation, are unstable and thus difficult to detect in mass spectrometry (Thomas *et al.*, 2006; Buszczak *et al.*, 2009).

Next to the necessity of specific enrichment techniques for some modifications, the ionization efficiency of the mass spectrometer has an impact on the identification of different histone modifications. Methylation of both lysine and arginine residues as well

as acetylation of lysines is known to have high ionization efficiency and thus might be slightly overrepresented (Pesavento *et al.*, 2006; Gropengiesser *et al.*, 2009). Over the last years, reliable and effective methods for label free quantification have been developed to overcome these technical difficulties. Among these methods is the integration of peptide peak areas of the chromatographic runs, also called extracted ion chromatograms (XICs). This approach is expected to be quite accurate and to decrease possible overrepresentation of histone modifications (Bondarenko *et al.*, 2002; Eberl *et al.*, 2011).

For the analysis of post-translational histone modifications, tryptic digestion of propionic anhydride treated histones was used (see Table 4). Some peptides of the digestion were difficult to detect due to low ionization efficiencies or loss during preparative procedures.

This problem was overcome by absorption to porous graphite carbon (TipTop, Gylgen) which retained and enriched the small hydrophilic peptides, which would otherwise be lost (see Figure 19; Chin *et al.*, 1998).

For the identification of known and novel post-translational histone modifications, the most reliable technique is to use fragmentation of the modified peptides. The fragmentation allows sequencing of the peptide and thus the determination of the exact localization of the modification (see Supplement Tables 4 and 5). Furthermore, biological variability of the peptide sequences might pretend possible post-translational modifications which can be ruled out by sequencing due to fragmentation spectra (Eberl *et al.*, 2011). The use of search algorithms, such as SEQUEST, for the identification of peptide modifications together with decoy searches against reversed or scrambled databases reduces the false negative rate and helps to improve the significant identification of histone modifications (Kapp *et al.*, 2005).

This study was focused on the identification of methylation and acetylation of lysine residues, however, phosphorylation, methylation of lysine and arginine residues of H2A and H2B was detected, too (see Figure 23). As phosphorylation is often underrepresented due to reduced elution times of phosphorylated peptides in the RP-HPLC, phosphorylation of histones was not considered for significant quantification (Eberl *et al.*, 2011).

Using RP-HPLC as well as MALDI-TOF, the H3 3-8 peptide was not detected. The porous graphite carbon is especially well suited for polar molecules such as the methylated H3 3-8 peptide.

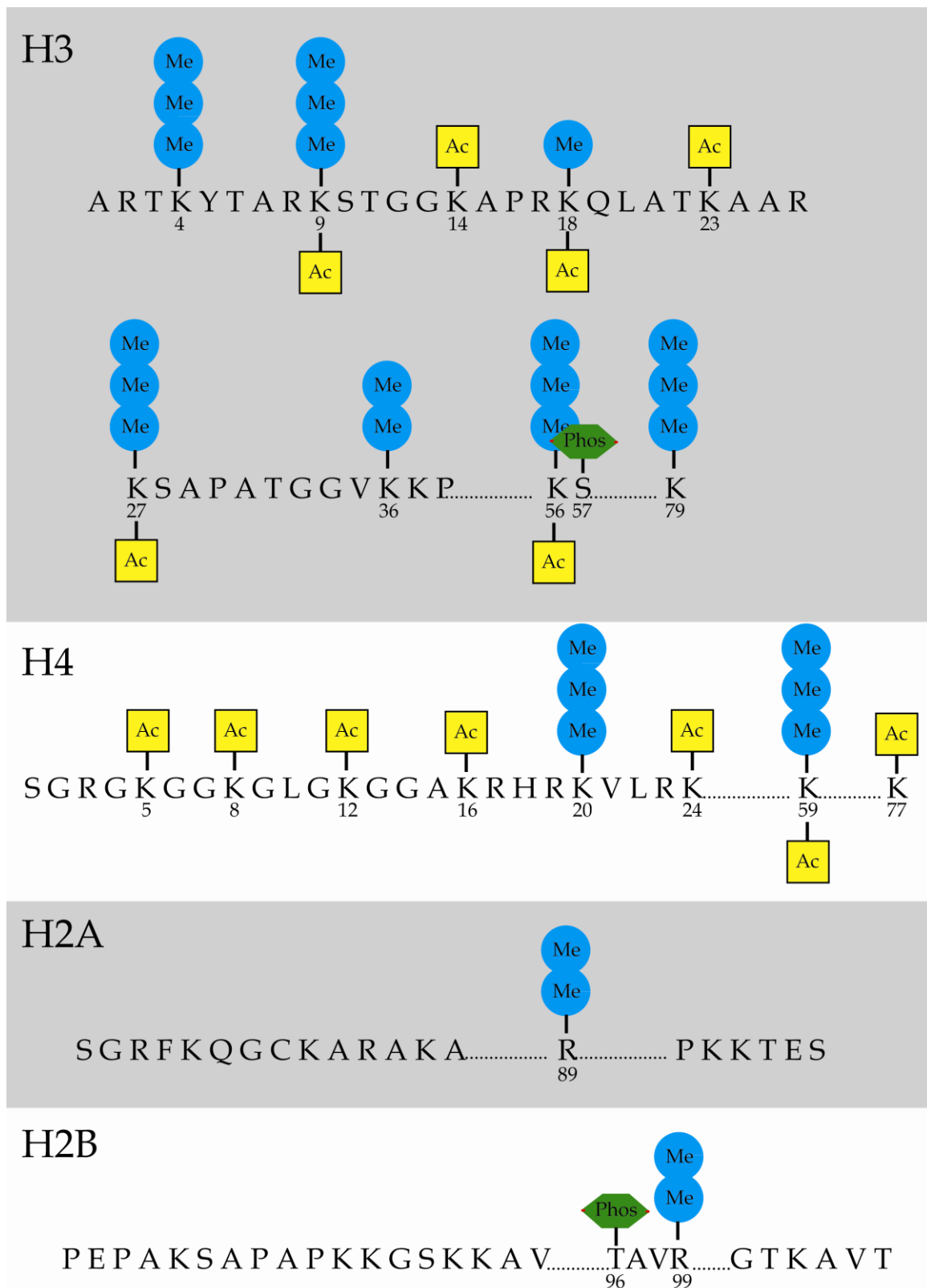


Figure 23: Detected Histone Modifications of the Core Histones

The number displays the modified amino acid residue. In blue is methylation, in yellow acetylation, in green phosphorylation of Histones. Abr.: Me-methylation, Ac-acetylation, Phos-phosphorylation.

In some cases, the retention of small polar components in RP-HPLC is not sufficient and the use of porous graphite carbon can improve the retention (West *et al.*, 2010). Thus, for detection and quantification of the methylation H3K4, the special enrichment technique with tips (TipTop, Glygen) containing a small stationary porous graphite carbon phase was used (see Figure 19).

Although the unmodified, mono- and di-methylated H4 20-23 peptides were detected using MALDI-TOF, the tri-methylated isoform was not detectable. The use of the porous graphite carbon did not improve the detection of the H4K20me3 mark (data not shown). Special chemical methods to derivatize the small peptide consisting of only 4 amino acids such as the reaction of malondialdehyde with arginine residues can improve the detectability in LC-MS. However, as the unmodified arginine residues were already modified by propionic anhydride to prevent digestion by trypsin, this technique was not applicable. Recent publications of histone modification analysis using mass spectrometry did either not detect H4K20me3 (Bonaldi *et al.*, 2004) or detected it in very low overall abundance of under 1% (Zhang *et al.*, 2003; Phanstiel *et al.*, 2007; Nicklay *et al.*, 2009). This implies a general technical problem with this small peptide, which together with its low abundance precluded the detection of the H4K20me3 mark by MALDI-TOF. To determine the course of the tri-methylation mark during development, quantitative Western blot analysis had to be performed (see Figure 18E).

Summarizing, mass spectrometry is a highly reliable technique for histone modification analysis (Eberl *et al.*, 2011). The method is frequently used, and comparison to other datasets is possible (Zhang *et al.*, 2003; Bonaldi *et al.*, 2004; Nicklay *et al.*, 2009).

6.3 Global Alterations of Histone Modifications during *Xenopus laevis* Development

Embryonic development, comprising of phases such as reprogramming and differentiation, is guided by changes in gene expression. Both transcription factors and epigenetic regulation such as histone modifications establish cellular memories and lineage specificity (Reik *et al.*, 2007). Thus, differences in the epigenetic states between different stages of *Xenopus* development were expected for histone modifications to reflect active transcription or repressive epigenetic states.

6.3.1 *Histone Modifications of Active Transcription*

Several histone modifications have been shown to be linked to active genes, either by initiation or elongation of transcription (Lee *et al.*, 2009). During *Xenopus* development, a general shift from highly abundant histone modifications of active transcription to a lower abundance of active marks in later developmental stages was obvious (see Figures 18 and 20). There was a drop in the level of methylation - H3K4me₃, H3K36me₂ - and acetylation of active histone modifications - H3K14ac, H3K23ac, and H4K5ac/K12ac/16ac as well as H4K5ac/K12ac - from the blastula to the gastrula stage. These modifications, which all function in gene activation, showed a similar course over development. During gastrulation, there were highest levels found for the active marks H4K16ac, H4K20unmod and me₁, H3K9me₁, H3K4me₁, which again had a similar time course over development.

These findings implied a general trend towards active transcription in the blastula stage with a stepwise reduction of transcriptional activity as development of the embryo proceeded as well as the establishment of repressed gene loci.

Recent studies of the ICM of the mouse embryo as well as later developmental stages revealed high levels of H3K9ac, H3K4me₂ and me₃ as marks of active transcription. As the mouse embryo proceeds its development, these histone modifications of active transcription were reduced in their abundance to lower levels (Hemberger *et al.*, 2009). These findings are consistent with this data set for *Xenopus* development (see Figures 18, 20 and 24).

Previous ChIP studies showed, that H3K4me₃, which functions as a binding platform for example for NURF, was found at promoters of active genes (Wysocka *et al.*, 2006). Methylation of some lysine residues, such as H3K4, is linked to gene activation (Bernstein *et al.*, 2002). H3K4me₃ is found at the 5' end of ORFs of genes and is often associated with actively transcribed genes (Bernstein *et al.*, 2002). However, recent human ChIP data suggested that H3K4me₃ marks promoter and enhancer regions irrespective of production of full length transcripts (Guenther *et al.*, 2007). Next to the recruitment of a great variety of different factors, H3K4me₁ can also activate HATs (Taverna *et al.*, 2006). This links the active transcription of genes, marked by H3K4me₁, to the presence of multiple acetylated histones.

Commitantly, during *Xenopus* development, the overall level of H3K4 methylation is reduced with a concomitant increase in unmodified K4 residues (see Figure 20A).

Furthermore, high levels of acetylation of the H3 N-terminus were linked to active

transcription (Liang *et al.*, 2004; Taverna *et al.*, 2007). As mentioned above, H3K4me3 is linked to gene activation and was found in high levels in *Xenopus* blastulae. H3K4me3, however, is recognized by the TFIID complex, a member of the pre-initiation complex for transcription, which also recognizes the acetylation of H3K9 and H3K14 by its bromodomain (Dambacher *et al.*, 2010). Thus the active mark H3K4me3 was linked to the acetylation of the H3 N-terminus. The clusters of the *Xenopus* embryo histone modifications revealed a comparable course for both the marks H3K4me3 as well as for the N-terminal acetylation of H3K9, K14, K18 and K23 (see Figures 20 and 24).

H3K36me2 and me3 is found on genes with elongating Pol II and is highly enriched at the 3' end of the ORF (Bannister *et al.*, 2005). Again, H3K4me3 and H3K36me3 have been linked and both marks have been implicated in transcriptional elongation in yeast. The methylation of H3K36 was shown to recruit deacetylase complexes to the coding regions and inhibits inappropriate initiation of transcription from cryptic start sites (Carrozza *et al.*, 2005).

The *Xenopus* embryo histone modification clusters showed a comparable course for H3K4me3 and H3K36me2 (see Figure 24). For unknown reasons, H3K36me3 was rarely detected in our data set. In a ChIP study of human T cells, acetylation of H3K9 and H3K27, and H3K9me1 and H3K4me1/me2/me3 were found on active transcribed genes (Wang *et al.*, 2008). H3K4me3, H3K9ac and H3K36me1 are found at promoter regions of actively transcribed genes (Barski *et al.*, 2007).

Summarizing, the blastula stage had much higher levels for numerous histone modifications associated with active transcription compared to the other embryonic stages. Furthermore, several key modifications of active transcription were clustered together implying comparable regulation during development. There is a general trend of high abundance of active histone marks in blastula to a gradual reduction in gastrula to the neurula and tadpole stage and consistent with findings of other studies (Bernstein *et al.*, 2002; Barski *et al.*, 2007; Hemberger *et al.* 2009).

6.3.2 Repressive Histone Modifications

Gene silencing, as determined by ChIP profiling across silenced genes, is linked to H3K27me3 as a prominent modification for gene repression, and H3K9me3 and H4K20me3. Surprisingly, the early *Xenopus* embryo has very low abundance of histone modifications associated with gene silencing (see Figure 21C and D). Heatmap cluster

analysis grouped H4K20me2 and me3, H3K27me2 together with highest levels in tadpole stage (see Figure 24). These findings implied that the early embryos such as blastulae and gastrulae have lowest levels of repressive histone modifications, which only become established during later development.

A recent study showed that differentiated cells carry large domains, up to several megabases long, with high levels of H3K9me2, and genes within these domains are repressed (Wen *et al.*, 2009). This mode of repression involves the recruitment of KMTases and HP1 to the promoter of repressed genes.

The level of H3K9me2 and me3 did not change significantly during development of *Xenopus laevis* (see Figure 20C). This unexpected finding might be explained by recent studies showing that a large proportion of silent genes did not carry any of the tested epigenetic modifications. It is possible that these genes are passively repressed and that their silent state is just due to the lack of activating factors (Bas van Steensel, 2011).

Major satellite repeats, which are also repressed, have a distinct H3K9me3 and H4K20me3 chromatin signature, which is found in almost all cell types and developmental stages suggesting that these modifications have a general function in heterochromatin. H3K9me3 is established by Suv3-9h1 and Suv3-9h2 enzymes (Rea *et al.*, 2000). Two other HMTases, Suv4-20h1 and Suv4-20h2 establish H4K20me3 (Schotta *et al.*, 2004).

Both the levels of mono- and di-methylation of H4K20 remained unaltered during development (see Figure 18C). However, Western blot quantification revealed a steady and linear increase in the tri-methylation of H4K20 between NF9 and NF37 (see Figure 18E). These results point to a steady increase of the level of repressive H4K20me3 during development.

The histone mark with the strongest association with gene repression is H3K27me3. Both H3K27me2 and me3 are linked to the repression of genes via the Polycomb machinery. H3K27me2 and me3 is hardly found at active promoters. At euchromatic regions, H3K27me3 is necessary for gene silencing and it functions in heterochromatin maintenance. H3K27me3 recruits the PRC1 complex, which further compacts the chromatin structure via the establishment of H2AK119ubi. H3K27me1, however, is found at regions with active transcription (Barski *et al.*, 2007).

Both levels of H3K27me2 as well as H3K27me3 were strikingly low at blastula and gastrula stages but, then, increased during later development to reach high levels in the

two cell lines A6 and XTC (see Figure 21B-E). Remarkably, H3K27me3 levels in blastula and gastrula stages were below the detection limit. Technical problems are unlikely to explain this finding, as high levels of H3K27me3 were found in other samples (see Figures 18 and 20). Moreover, these findings of low repressive histone modifications in early *Xenopus* embryos (NF9 and NF12), are in agreement with recent ChIP-Seq results from *Xenopus tropicalis* (Akkers *et al.*, 2009).

It is not known, why the chromatin of the frog blastula is so devoid of the H3K27me2/me3 marks. Recently, an antagonistic relationship between methylation and acetylation of Lysine 27 has been reported that reflects a competition between PcG and trithorax/MLL complexes, which recruit H3K27 methyltransferase or acetyltransferase activities, respectively (Pasini *et al.*, 2010). Since frog embryos contained the highest relative K27ac levels in this data set (see Figure 21D and E), one might speculate that some regulatory mechanism in blastula embryos favors MLL-mediated recruitment of HATs. However, a direct competition between trx and PcG proteins on target genes is unlikely to explain entirely the initial depletion in H3K27 methylation, because K27ac levels remain high during subsequent development despite the dramatic increase in K27me2/me3 from gastrula stages onwards (see Figure 21D and E). Furthermore, K27me3 is as high as 10% in ES cells whereas the level of H3K27ac reaches only 1.5% in blastulae. This makes it very unlikely that the acetylation of H3K27 accounts for the low levels of K27me3 (Bernstein *et al.*, 2002).

Chromatin fragments have been identified that simultaneously carry active as well as repressive histone modifications, specifically the H3K4me3 and H3K27me3 marks. This chromatin state has been termed bivalent to indicate its ambiguous regulatory nature (Bernstein *et al.*, 2006, Azura *et al.*, 2006). Bivalent domains first have been described in embryonic stem (ES) cells, where they may keep developmental regulatory genes in a state poised for subsequent activation during cellular differentiation. However, bivalent domains are not restricted to pluripotent cells and most likely have functions beyond priming genes for activation (Sharov *et al.*, 2007). In contrast to mammalian ES cells, bivalent chromatin domains are practically absent from *Xenopus* blastulae (Akkers *et al.*, 2009). Lysine tri-methylation of H3K4 and H3K27 appears in the zebrafish epigenome only after the maternal-zygotic transition and in the same sequence as in frogs.

Although the co-existence of the H3K4me3 and H3K27me3 mark on one histone tail cannot completely be ruled out, both levels for H3K4me3 and H3K27me3 can be

described separately in this study. As no H3K27me3 was found in *Xenopus* blastulae, it appears very unlikely that bivalent histone modifications exist (see Figure 21C and D).

In general, our study described the presence of tri-methylation of other modification sites as well as for the K27 site in other stages, limiting the possibility of technical problems. However, it may be that mass spectrometry is not as sensitive as antibody based approaches for the detection of H3K27me3.

According to the recent publication of Akkers and coworkers, the presence of H3K27me3 was ruled out at the early stages of the zebrafish embryo (Akkers *et al.*, 2009). This finding suggests that the formation of somatic cell lines in *Xenopus* species occurs from an epigenetic state practically devoid of bivalent domains. Future studies in *Xenopus* will have to address, whether the derepressed state at blastula exists also in earlier pre-MBT stages such as the morula. General changes in histone modifications occur in mouse embryos between fertilisation and the blastocyst stage (Hemberger *et al.*, 2009; Young, 2011). However, histone PTMs have not been investigated from post-implantation embryos, except for one pioneering study using carrier ChIP (Rugg-Gunn *et al.*, 2009). Here, both K4me3 and K27me3 marks were found on several developmental loci in epiblast tissue, but the bivalent status of these loci was not rigorously established and the opposing marks may derive from distinct cell populations. Nevertheless, this study demonstrates that K27me3 exists in the mouse embryo at the time of germ layer formation, unlike in *Xenopus* embryos.

Nicklay and Schechter have incubated sperm nuclei with *Xenopus* egg extract *in vitro* to generate chromatin of an “early embryo equivalent” that may represent the chromatin state before or at MBT (Nicklay *et al.*, 2009). This sample contains quite abundant K27me3 levels (estimated 10-50% by Western blot and mass spectrometry), confirming the maternal expression of biologically active PRC2 components in early embryos (Aldiri *et al.*, 2009). However, neither this study nor others (Akkers *et al.*, 2009) have found evidence for this in blastulae. If chromatin of the “early embryo equivalent” corresponds to a physiological state, one would have to postulate that practically all K27me3 marks become erased before midblastula. The presence of maternal and potentially active PRC2 complexes in turn suggests for the unmanipulated frog embryo that PRC2 activity is either negatively controlled or efficiently antagonized by demethylation around the onset of zygotic transcription. The sudden increase of K27me2 at gastrula suggests a regulatory switch that quite rapidly brings a significant portion of

the epigenome under the influence of the polycomb system. This switch may include the upregulation of Suz12 mRNA levels that has been described during gastrulation (Aldiri *et al.*, 2009).

6.3.3 Clustering of Histone Modification Profiles

As mentioned above, significant quantitative changes in histone modifications of both active and repressive function were observed during frog embryogenesis (see Figures 18, 20 and 21). In order to investigate, whether particular developmental stages could be described by characteristic groups of histone modification states, hierarchical clustering analysis was performed to identify potential epigenetic patterns. A heatmap of the obtained clusters was generated to visualize the results (see Figure 24, kindly provided by Jose Arteaga-Salas). The dendrogram on the left axis of the heatmap shows four well-defined clusters between the second and third levels. The most obvious features of the heatmap are the four red-colored areas, each representing a cluster of histone modifications at their individual maximal abundance, which describe the four developmental stages that were analysed. The separation is very clear between blastula and tadpole clusters, where the red groups generally indicate a global shift from active to repressive histone marks as the embryonic cells undergo differentiation. It is less clear between gastrula and neurula stages, in particular at the bottom third of the heatmap, where several histone modification states persist at similar, intermediate abundances until the tadpole stage. As gastrula and neurula stages are biologically characterized by ongoing cellular diversification, these modifications may be associated with transitory features of germ-layer committed precursor cell populations. Importantly, the apparent clustering provides unique evidence for a global, stepwise maturation of the embryonic epigenome in vertebrate embryos. This finding extends the hypothesis of individualized epigenomes for different cell types (Schechter *et al.*, 2009) to whole embryos. This surprising result does not necessarily imply that all cells of an embryo at a given stage share the same histone PTM profile, but rather reveals the existence of stage-specific constraints that shape the histone modifications according to global cellular transitions, such as the development from pluripotency to germ layer precursor to committed and differentiated cell states.

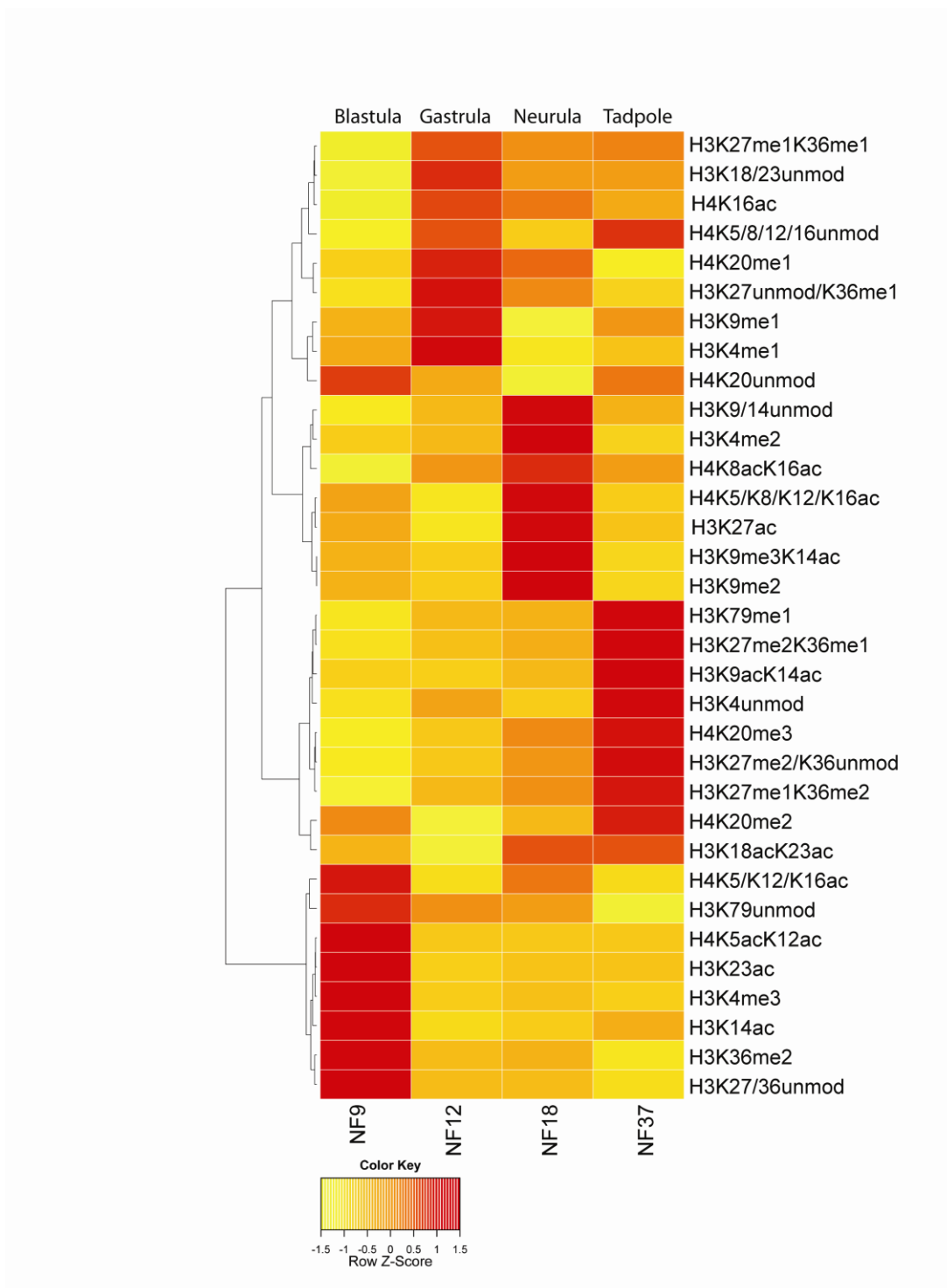


Figure 24: Stage-Specific Histone Modification Profiles

The heatmap visualizes clusters of histone modifications according to their relative abundance at the four developmental stages, which have been investigated. The hierarchical clustering analysis produced a dendrogram, shown on the left side, with four major branches that correspond to specific developmental stages. The four clusters define histone modifications profiles (HMPs), which reflect the gradual transition from uncommitted to determined cell fates.

6.4 Comparison of Histone Modification Profiles of Murine ES Cells and *Xenopus* Blastulae

During the amphibian blastula-gastrula transition, pluripotent embryonic cells are programmed by induction to germ-layer related cell fates (De Robertis *et al.*, 2009). This phase corresponds in mammalian embryos to the specification of the epiblast at implantation, from when on pluripotency is gradually lost (Silva, 2008). This raised the issue, whether the histone modification profiles of *Xenopus* blastulae represent a common feature of pluripotent cells. To address this point, the analysis was extended to murine ES cells and of primary embryonic fibroblasts (MEFs), which were used as feeder cells in the ES cell culture and represent a murine somatic cell type for comparison.

As the primary embryonic fibroblasts were used as feeder cells in the ES cell culture, their successful separation had to be evaluated first. The analysis confirmed that the two murine cell types contained very specific profiles of histone modifications (see Supplement Figure 3).

Despite its low abundance, acetylation of H3K56 is specifically enriched at promoters bound by the core pluripotency factors Oct-4/Sox-2/Nanog. The acetylation of H3K56 is situated at the major groove of the DNA histone complex. The chromatin remodeling complex SWI/SNF can be recruited to genes recognizing H3K56ac and thus transcription is facilitated at certain genes in yeast (Xu *et al.*, 2006). Upon cellular differentiation, H3K56ac is found on genes involved in the differentiation process (Xie *et al.*, 2009).

K56ac levels stay rather constant between blastula and gastrula (Figure 20I). The analysis of murine ESC displayed higher levels of H3K56ac than in MEFs and comparable levels to *Xenopus* embryos (see Supplement Table 3). At the same time, the embryonic K9me1 level rises to a transient maximum of 20% abundance.

H3K9me1 has been recognized as a stable mark of cis-regulatory elements for both active and inactive genes (Lee *et al.*, 2009). Together, the observed changes in the H3K56ac and H3K9me1 profiles may thus be important for the programming of cell fates.

Consequently, between the *in vivo* pluripotency model of *Xenopus laevis* blastulae and the *in vitro* model for pluripotent mouse ES cells, significant differences in the levels of H3K56ac and H3K9me1 were present.

The bivalent pattern of two modifications, so far known to be involved in opposing functions, might be characteristic for pluripotency as it had first been found in

ES cells, but resolved upon differentiation. The hypothesis evolved, that the dual marking of histones at developmental genes are poised for quick activation (Bernstein *et al.*, 2006).

With the bottom up approach, it was not possible to determine, whether histone H3 tails are modified on both Lysine 4 and 27, since these residues were located on different tryptic peptides (see Table 4). Nevertheless, comparison of the overall abundance of modifications at these sites revealed high similarity between *Xenopus* blastula and murine ES cells for H3K4me3 (see Figure 21). However, for H3K27me3, there were dramatic differences in the abundance. The murine ESCs had the highest levels of H3K27me3 of all samples tested, whereas the level of the K27 tri-methylation was not detectable in *Xenopus* blastulae (see Figure 21D). Thus, it was expected that the blastulae had no bivalent histone modifications.

Furthermore, the combinatorial methylation profiles of the H3 peptide 27-40, including Lysines 27 and 36, were very different between blastulae and ES cells. As described earlier, the majority of the di-methyl marks resided on K36 in frog blastulae. In ES cells, however, they were found predominantly on K27 (Figure 20C and D). MEFs had a similar di-methyl-distribution as ES cells, but contained significantly more K36me2 than these. A dichotomy was also found for the higher modification states of this peptide, including three to five methyl groups. Whereas in frog blastulae H3K27me3 was below the detection limit, mouse ES cells had with 8.3% the third-highest K27me3 level of all samples tested, and twice as much as MEFs (Figure 21E). The combinatorial tri-methylated states, such as K27me1/K36me2 and K27me2/K36me1 were also significantly more abundant in ES cells than in frog blastulae, and comparable to the levels found in MEFs. Notably, the highest levels of tetra- (K27me2+K36me2) and penta-methylated (K27me3+K36me2) states of the H3 27-40 peptides were also found in ES cells and to a less extent in MEFs, but were not detected in frog blastulae (Figure 21E).

Overall, these results indicated that frog blastulae and murine ES cells shared high H3K4me2/me3 levels, but the former were dramatically undermethylated at the H3 27-40 peptide. Based on this, the polycomb repression complex 2 (PRC2) activity seems to have a very different impact on the epigenomes of frog embryos and *in vitro* cultured murine ES cells. This dramatic increase in K27 methylation was consistent with the results of recent ChIP-studies in *Xenopus tropicalis* and *Dario rerio*, which reported a hierarchical acquisition of H3K4 and K27 methyl marks (Akkers *et al.*, 2009; Vastenhouw *et al.*, 2010). In extension to these studies, however, we have found that the K27me3 mark is at least

100-fold less abundant in *Xenopus* early embryos than in cultured frog or mammalian cell lines, including murine ES cells. As a consequence, the K27me3 density in chromatin may be rather low throughout the embryo. Alternatively, it could be high within a subfraction of either cells or genomic regions, and consequently being depleted from the rest. Whatever may be the case, this finding is very unexpected, given current models on the epigenetic nature of pluripotency. Functional tests are required to clarify, whether at this low abundance (<0.1%) H3K27me3 represses transcription in *Xenopus* blastulae as reported for human and murine ES cells, in which H3K27me3 is much more abundant (Dai *et al.*, 2007; and our data here).

Furthermore, it seems possible that different histone modifications, although opposing in their single function, act totally different in concert in order to fine-tune the current activation state of a locus. Thus, co-existence of H3K4me3 and H3K27me3, although opposing in their single readout, may have completely different function for the underlying genomic region.

Another possible explanation for this finding might be the fact that the epigenetic regulation during mouse and frog development is different. Differences in the histone modifications of mouse development have been described for imprinting of genes, which does not exist in *Xenopus* but is present in mouse as well as for global alterations in DNA methylation during development. Genetic imprinting is frequently found in gametes, where it especially marks genes differing in their expression from parental and maternal origin. Whereas the mouse genome becomes demethylated during early development the frog genome remains hypermethylated during this time (Veenstra *et al.*, 2001). However the Polycomb complex, which is required for the establishment of H3K27me3, is shown to be present in *Xenopus* embryos. The two PcG proteins EED and YY1 of *Xenopus laevis* are highly expressed in the fertilized egg and throughout the blastula stage. The EED transcripts decline during gastrula stages, the YY1 during neurula stages (Satijn *et al.*, 2001). Moreover the *Xenopus* polycomb XPcg and Xbmi-1 are also expressed in early *Xenopus* embryos with similar kinetics (Reijnen *et al.*, 1995). However, in *Xenopus laevis*, both genes are predominantly expressed in the ectoderm, especially in the central nervous system compared to the mouse embryos, where both genes have a much broader distribution. Even though bivalency may be important for mouse development, it seems not to be required in non-mammalian vertebrates such as the frog.

Given the reprogramming of the genome after fertilization, the removal of

repressive marks and re-establishment of active marks is required. It might be that the high abundance of the bivalent histone modification in the in vitro system of ES cells just reflects a short time window during the epigenetic reprogramming of the gamet's precursors. As the repressed state of both maternal and parental pronuclei are reprogrammed to an active epigenetic state, a short period of the co-existence of the repressed and active histone modifications might be present. Because development is a highly dynamic process and pluripotency of the zygote is no enduring state, the bivalent modifications are present only for a very short time window in vivo. The presence of the bivalent state of histone modifications in ES cells, however, might simply reflect an artificial state of pluripotency in vitro.

Either way, it remains to be determined whether the bivalent histone modifications are of biological importance for pluripotency.

6.5 Biological Implications of Histone Modification Profiles

6.5.1 *Reprogramming during Xenopus laevis Development*

As mentioned above, the fertilization of the zygote is followed by epigenetic reprogramming (Mayer *et al.*, 2000). Even before fertilization, characteristic histone modification patterns are established during germ cell differentiation (Hajkova *et al.*, 2008), and nucleosomes, which are retained in low numbers in human sperm, carry histone modifications on particular developmental loci (Hammoud *et al.*, 2009). Following the fertilization of the zygote, epigenetic reprogramming and at least partial erasure of marks such as imprints occurs in the embryo (Mayer *et al.*, 2000). Reprogramming during development comprises hypermethylation at H3K4 and differential H3K4 di/tri-methylation at tissue-specific genes during haematopoiesis (Orford *et al.*, 2008), establishment of differentiation-specific H3K9me2 regions associated with silencing (Wen *et al.*, 2009) and the modulation of bivalent domains (Bernstein *et al.*, 2006; Mikkelsen *et al.*, 2007).

The vertebrate model organism *Xenopus* has recently provided key information on epigenetic changes in early development. Histone modifications from various *Xenopus laevis* cell types including oocytes, sperm and somatic cells, have been characterized by immunoblotting and mass spectrometry, indicating unique histone modification signatures for each cell type (Schechter *et al.*, 2009; Nicklay *et al.*, 2009). The histone modification profile of sperm showed very low levels of H3K4me3, medium levels of

H3K9me3, low levels of H3K27me3, but a variety of combinations of H3K27 and K36 methylation. The level of H4K20me3 was high in sperm, however, it was not detected for the early embryo equivalent. Furthermore, the early embryo equivalent had no H3K4me3, high levels of H3K27me3 but no combinations of the methylation of both K27 and K36 (Nicklay *et al.*, 2009). The *in vivo* relevance of the early embryo equivalent, which is generated *in vitro* by incubation of pronuclei with *Xenopus* egg extract is unclear. However, the early embryo equivalent displays a totally different histone modification profile compared to sperm. Furthermore, the profile of the early embryo equivalent dramatically differs from the findings of blastula staged embryos from this data set.

Considering both our data and the findings from Nicklay and Schechter, reprogramming between the early embryo equivalent and the blastula stage includes the reduction of H3K27me3 levels to not detectable as well as the levels of H4K20me3, which have very low abundance by quantitative Western blotting. On the other hand, the blastula stage has very high levels of H3K4me3 indicating that the blastula stage has a transcriptionally permissive epigenome. This abundance of H3K4me3 (almost 30%) in frog blastulae is a puzzling observation. In mammals, the H3K4me3 mark is found on gene promoters and enhancers. Given that frog and human genomes are comparable in size and gene numbers, this level exceeds the expected genome proportion represented by these transcriptional regulatory elements. Notably, murine ES cells have similarly high levels, and even in somatic cells (A6, XTC and MEFs) its abundance is around 10%. Although H3K4 acetylation levels are generally low (Garcia *et al.*, 2007), it is possible that some fraction of the H3K4me3 peak represents acetylated peptides. Nevertheless, one may speculate that H3K4me3 could be deposited differently in embryonic cells compared to somatic cells.

In addition to the typical promoter-proximal peaks, which have been observed in recent genome-wide ChIP experiments both in frog and zebrafish, this mark could exist at higher base-levels throughout the embryonic epigenome, and/or be enriched significantly in still unannotated areas. The high abundance of H3K4me3 in embryos certainly merits further investigation.

Another example for the transition from repressed to activated chromatin states is given by the androgen receptor (AR), which is part of the steroid hormone receptor family of ligand-activated transcription factors. The cytoplasmic androgen receptor dissociates upon hormone binding from chaperones, and translocates to the nucleus,

where it binds to androgen response elements (AREs) of target genes (Metzger *et al.*, 2006). The activation of genes through the action of the androgen receptor requires a distinct histone modification profile of H3K4 and H3K9. An important first step in the transition from a repressed state to an activated state is the demethylation of the repressive H3K9me3. Following the removal of H3K9me3, the acetylation of both H3K9 and K14 and tri-methylation of H3K4 at several AR-receptor-dependent genes leads to transcriptional gene activation (Metzger *et al.*, 2008). The level of H3K4me3, H3K9ac and K14ac rises and the level of H3K9me3 decreases between the early embryo equivalent and blastula stage (Nicklay *et al.*, 2009). These findings indicate a general shift of the chromatin template from repressive to active during reprogramming.

6.5.2 *Biological Function of different Histone Modifications*

Different modes exist how histone modifications influence chromatin and genomic activation. Among these mechanisms is the ability of a post-translational histone modification to cause changes in chromatin dynamics such as heterochromatin or euchromatin formation. H4K16ac is involved in the inhibition of the formation of compacted 30-nanometer structures (Shogren-Knaak *et al.*, 2006). A second mechanism involves the ability of a histone mark to serve as docking sites for non-histone proteins. Through the recruitment of proteins and complexes, diverse outcomes of the chromatin template such as establishment or removal of histone modifications involved in transcription, DNA synthesis and repair or gene silencing may be the consequence.

The notion that certain histone modifications seem to carry out consistent functions on the chromatin template led to the hypothesis of a distinct histone code (Strahl *et al.*, 2000). Unlike the genetic triplet code, where a triplet codon always results in the same amino acid in the mRNA, one single histone post-translational modification mark does not necessarily correlate with one outcome. However, the single histone modifications may act as the single bases of the triplet and only function in concert with presence or absence of other histone modifications or trans-acting enzymes (Oliver *et al.*, 2011). The interaction of several histone modifications situated in close proximity, however, not necessarily on the same nucleosome, can function in a (i) cooperative manner – several histone marks act together – in (ii) independent manner – several modifications co-exist but do not interfere – and in (iii) antagonistical manner – presence of one modification blocks the modification of an adjacent residue (Seet *et al.*, 2006).

An example for the cooperative manner of histone modifications is H3K4me₃, which is recognized by the NuA3 histone acetyltransferase complex which then recruits a histone acetyltransferase complex for acetylation of H3K14 (Taverna *et al.*, 2006). Additionally, recent data revealed that a single point mutation in histone H3K14, hindering its acetylation, resulted in specific loss of H3K4 tri-methylation, but not mono- or di-methylation in yeast (Nakanishi *et al.*, 2008). During *Xenopus* development, the level of H3K4me₃ and H3K14ac showed a sharp decrease from blastula to gastrula stage, indicating at least compatible behavior during early development (see Figure 20).

Antagonistic behavior of histone modifications has been shown for H3K9me₃ recruiting heterochromatin protein 1 (HP1). However, when the adjacent serine residue is phosphorylated during mitosis or during gene activation, the binding of HP1 is prevented (Fischle *et al.*, 2005).

Although a huge amount of data exists for the function of single histone modifications (Kouzarides, 2007; Lee *et al.*, 2009; Dambacher *et al.*, 2010), the exact meaning of the co-existence of different histone modifications is only at the beginning to be discovered. This may reflect the difficulties to study the function of several histone modifications in close proximity. For mass spectrometry, the bottom-up approach with digestion of histone proteins prior to identification and quantification is used most frequently (Zhang *et al.*, 2003, Bonaldi *et al.*, 2004, Nicklay *et al.*, 2009). The potential complexity of co-existence of up to 60 or more histone modifications at one histone tail poses a challenge to the scientist and is rarely addressed (Phanstiel *et al.*, 2007). However, by the use of digested histone peptides, the description of the co-existence of different histone modifications at one histone tail is mostly impossible. In this study, the bottom-up technique was used for this reason and the co-existence of different histone modifications was described for H3K9me/ac and H3K14ac, H3K18ac and H3K23ac, H3K27me/ac and H3K36me as well as H4K5ac/8ac/12ac/16ac. On the other hand, the use of antibody based techniques such as ChIP to identify histone modifications in the genome have frequently been used (Bernstein *et al.*, 2002; Barski *et al.*, 2007). However, to be able to describe the presence of several histone modifications at one histone tail, the use of special and technically challenging Re-ChIP experiments has to be performed. Sequential ChIP experiments indicated, that only a small proportion of the two bivalent histone modifications, K4me₃ and K27me₃, co-exist on the same nucleosome (Bernstein *et al.*, 2006). Studies at higher resolution rather favored the localization of the opposing histone

marks at different nucleosomes (Barski *et al.*, 2007; Akkers *et al.*, 2009).

One has to keep in mind, that histone modifications also function in DNA repair, cell cycle progression and may have currently unknown functions. Consequently, the co-existence of several marks may simply reflect the coincidence of different chromatin influences such as DNA repair, gene silencing, transcription, cell cycle progression as well as nuclear spacing. Furthermore, histone modification crosstalk might also be dictated by a time dependent manner by which these modifications are introduced and, thus, just looking at the pattern of chromatin modifications at a locus at an arbitrary time point is not sufficient to determine its gene expression status (Lee *et al.*, 2010). Summarizing, further studies are needed to provide better insight into the language of modification crosstalk and the function of co-existing histone modifications.

7 Abbreviations

A	Alanin
Å	Angström
aa	Amino acid
Abr	Abreviation
ac	Acetylation
ACN	Acetonitril
ADP	Adenosin-diphosphate
adp	ADPribosylation
Ambic	Ammoniumbicarbonate
AR	Androgen Receptor
ARE	Androgen Response Element
Asf-1	Anti Silencing Function-1
Ash1	Histone Methyltransferase
ATCC	Biological Resource Center
ATP	Adenosin-triphosphate
ATPase	Adenosin-triphosphate Hydrolase
AuroraB	Kinase involved in Centromere Formation during Mitosis
A6	Adult kidney Cell Line of <i>Xenopus laevis</i>
b-fragment	Peptide Fragment with the Ion situated at the N-terminus of the Peptide after CID in MS/MS
BHC110	Histone Demethylase
Bioworks	Mass Spectrometric Software
BMI	Component of the PRC1 complex
BMP	Bone-Morphogenetic Protein
bps	Base Pairs
BSA	Bovine Serum Albumine
C	Cystein
CaCl ₂	Calcium Chloride
CAF-1	Chromatin-Assembly Factor-1
Carbon Tip	Pipette Tip with Porous Graphite Carbon

CARM1	Arginine Methyltransferase
CBP/p300	Histone Acetyltransferase
Cbx2	Chromobox Protein Homolog 2
CENP-A	Centromeric Protein-A
CHD	Chromodomain-Helicase-DNA-Binding Protein
ChIP	Chromatin Immuno-Precipitation
CID	Collision Induced Dissociation
cm	Centimeter
cm ²	Square Centimeter
c-Myc	Protooncogene
CO ₂	Carbondioxide
CpG	Cytosin followed by a Guanosin in the DNA, which is possibly methylated
C-terminus	Carboxy-Terminus of a Protein/Peptide
C4	Length of the Fatty Acid Chains of the Resin in the HPLC
C18	Length of the Fatty Acid Chains of the Resin in the HPLC
D	Aspartart
d	Days
Da	Dalton
DAPI	4',6-Diamidin-2-phenylindol
ddH ₂ O	Double distilled Water
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethylsulfoxide
DNA	Desoxyribonucleic Acid
Dnmt	DNA Methyltransferase
Dot1	Disruptor of Telomeric Silencing 1
dta	Data Filtering File of MS/MS
DTT	Di-Thio-Threitol
Dx	Wavelength for measurement of the quantitative Western Blot
E	Glutamate
EB	Buffer used for Histone Extraction via Histone Purification Kit
EBs	Embryoid Bodies
ECL	Enhanced Chemiluminescence

EDTA	Ethylendiamin-Tetra Acetat
EED	Subunit of Polycomb Repression Complex
EGTA	Ethylene Glycol Tetraacetic Acid
ESC	Embryonic Stem Cell
ES	Embryonic Stem
ESI	Electro-Spray Ionization
Ezh2	Enhancer of Zeste
E1	Buffer used for Nuclei Preparation of Xenopus embryos
F	Phenylalanine
FA	Formic Acid
FGF	Fibroblast Growth Factor
FPLC	Fast Protein Liquid Chromatography
FRAP	Fluorescence Recovery after Photobleaching
G	Glycine
g	Gravity
G9a	Histone Methyltransferase
GLp	Histone Methyltransferase
GNAT	Histone Acetyltransferase
GSES	Murine Embryonic Stem Cell Line
H	Histidine
h	Hour
HAT	Histone Acetyl-Transferase
HCG	Human Chorionic Gonadotropine
HCl	Hydrochloric Acid
HClO ₄	Perchloric Acid
HCO ₃	Hydrogencarbonate
HDAC	Histone Deacetylase
HEPES	2-(4-(2-Hydroxyethyl)-1-Piperazinyl)-Ethansulfon Acid
high salt	High-salt Extraction
Hira	Histone Chaperone
HMT	Histone Lysine Methyltransferase
Hox	Homeobox
HP1	Heterochromatic Protein-1

HPLC	High-Pressure-Liquid-Chromatography
H1	Histone 1, linker Histone
H2A	Histone 2A
H2A.Bbd	Histone 2A Bar body dependent, Histone Variant of H2A
H2A.Z	Histone 2A.Z, Histone Variant of H2A
H2B	Histone 2B
H ₂ O	Water
H ₂ PO ₄	Dihydrogen Phosphoric Acid
H ₂ SO ₄	Sulfuric Acid
H3.1	Replication-dependent Histone Variant of H3
H3.1t	Testis specific Histone Variant of H3.1
H3.2	Replication-independent Histone Variant of H3
H3.3	Replication-independent Histone Variant of H3
H4	Histone 4
I	Isoleucin
ICM	Inner Cell Mass
INO80	ATPase subunit
in vitro	Studies isolated from an Organism or their usual Biological Context
in vivo	Studies performed in an Organism or in their usual Biological Context
ISCOVES	Media for ES Cell Culture
ISWI/SNF2	Chromatin Remodeling Complex
I.U.	International Units
K	Lysine
kbp	Kilo basepairs
KCl	Potassium Chloride
kDa	Kilo Dalton
KH ₂ PO ₄	Potassiumdihydrogenphosphate
K ₂ HPO ₄	Potassiumhydrogenphosphate
Klf-4	Transcription Factor involved in Pluripotency
KMT	Histone Lysine Methyltransferase
KOH	Potassiumhydroxide
KPO ₄	Potassiumphosphate

kV	KiloVolts
L	Leucine
l	Liter
LASER	Light Amplification of Stimulated Emission or Radiation
LC	Liquid Chromatography
LC-MS/MS	Liquid Chromatography-Tandem Mass Spectrometry
LIF	Leukemia Induced Factor
Lin-28	Transcription Factor involved in Pluripotency
LoBind	Special coated tube for minimal Protein Binding
LOCK	Large Organized Chromatin K9 Modifications
LSD1	Lysine Demethylase-1
LTQ	Linear Trap Quadrupole, a mass spectrometer
M	Methionine
M	Molar
mA	Milli Ampere
macroH2A	Histone 2A Variant, contains the Macro Domain
MALDI-TOF	Matrix Assisted LASER Desorption Time of Flight Mass Spetrometer
MASCOT	Search Engine for Peptide Analysis
mAU	Milli Absorption Units
Mbps	Mega Base Pairs
MBS	Modified Barth Saline
MBT	Mid-Blastula Transition
MBT	Protein Domain Recognizing Methylation
me	Methylation
me1	Mono-methylation
me2	Di-methylation
me3	Tri-methylation
MeCP2	Methyl CpG binding Protein 2
MEF	Mouse Embryonic Fibroblast
MgCl ₂	Magnesiumchloride
microLC	Liquid Chromatography using small Sample and Buffer Volumes
min	Minute
miRNA	MicroRNA

ml	Milliliter
MLL	Histone Methylase
mM	Milli Molar
mm	Millimeter
MNase	Micrococcus Nuclease
MOF	Histone Acetylase
mRNA	Massanger Ribonucleic Acid
MS	Mass Spectrometry
MS-1	No sequential Mass Spectrometry
MS/MS	Sequential or Tandem Mass Spectrometry
m/z	Mass over electrical Charge
N	Asparagine
NaCl	Sodium Chloride
NAD ⁺	Nicotine Amide Adenine Dinucleotide
NaHCO ₃	Sodiumhydrogencarbonate
NaH ₂ PO ₄	Sodiumdihydrogenphosphate
Nanog	Transcription Factor involved in Pluripotency
NCBI	National Center for Biotechnology Information
NF	Nieuwkoop-Faber
ng	Nano Gram
NH ₂	Amino Group
nm	Nano Meter
Non-coding RNA	RNA Molecule which is not translated into a Protein
NP-40	Normidet P-40
N-terminus	Amino-Terminus of a protein/peptide
NuA3	Histone Acetyltransferase Complex
NURF	Chromatin Remodeling Complex
NYST	Histone Acetyltransferase Complex
Oct-4	Transcription Factor involved in Pluripotency
OD	Optical Density
Orbi-Trap	Mass Spectrometer containing a Trap for the Ions for subsequent CID
ORF	Open Reading Frame

P	Proline
PAA	Poly Acryl Amide
PADI4	Peptidyl Arginine Deiminase 4
pan	Total
PBS	Phosphate Buffered Saline
PcG	Polycomb
pH	Measurement of the Acidity or Basicity of an Aqueous Solution
PHD	Protein Domain Recognizing Methylation
Phos	Phosphorylation
piRNA	Piwi interacting RNA
PIWI	P-element induced Wimpy Testis Protein, involved in piRNA Formation
pmf	Peptide Mass Fingerprint
PMSF	Phenylmethylsulfonylfluorid
Pol II	Polymerase II
ppm	Parts per Million
PRC1/2	Polycomb Repression Complex 1/2
PRE/TRE	Polycomb/ Thrithorax Repression Element
PRMT	Arginine Methyltransferase
PR-SET7	Histone Methyltransferase
PTM	Post-Translational Modification
Q	Glutamine
R	Arginine
R	Open Source Software for Statistical Analysis
RAW	Raw Data File of the MS/MS
rec	Recombinant
Re-ChIP	Sequential Chromatin Immunoprecipitation
Ring1A	Polycomb Subunit, Ubiquitin Ligase
RISC	RNA-induced Silencing Complex
RNA	Ribonucleic Acid
RNAi	Ribonucleotide Inhibition
RP-HPLC	Reversed-Phase High-Preassure Liquid Chromatography
rpm	Rounds per Minute

Rtt109	Histone Acetyltransferase
S	Serine
s	Seconds
SAGA	Acetyltransferase Complex
SAM	S-Adenosyl Methionine
SEM	Standard Error of the Mean
SDS-PAGE	Sodiumduodecosylsulfate-Polyacrylamid-Gelelectrophorese
SET	Protein Domain of Histone Methyltransferases
SETD1A/B	Histone Methyltransferases
SETD2	Histone Methyltransferase
SEQUEST	Software for Mass Spectrometry
siRNA	Small Interfering Ribonucleotic Acid
SIRT1/2	Histone Deacetylases
Sox-2	Transcription Factor involved in Pluripotency
srf	SEQUEST result file
SUMO	Small Ubiquitin like Modifier
sumo	SUMOylation
SU(VAR)3-9	Supressor of Variegation, H3K9
SU(VAR)4-20	Supressor of Variegation, H4K20
Suz12	Subunit of Polycomb
SWI/SNF	Chromatin Remodeling Complex
Swi2/Snf2	Remodeling Complex
T	Threonine
T-box	Gene Family involved in Embryonic Development
TE	Tris and EDTA Buffer
TFA	Trifluoroacetic Acid
TGF- β	Transfroming Growth Factor- β
Tip Top	Pipette Tip used for Prouse Graphie Carbon enrichment
TM	Copyright
Tris/HCl	Tris(hydroxymethyl)-aminomethan Buffer
trx	Trithorax
Tudor	Protein Domain recognizing Di-methylated Arginine
U	Unit

ubi	Ubiquitination
Ubp8	Ubiquitin-specific Protease, Component of the SAGA Complex
unmod	Unmodified Peptide
UV	Ultraviolet
V	Valine
V	Volts
VegT	Transcription Factor
W	Watt
W	Tryptophan
WD40	Structural Motif forming a Propeller-Structure for Protein-Protein Interactions
Wnt	Signalling Molecule important for Development
X	X-chromosome
x	Multiplied by
X-chr	X-chromosome
Xbmi	Xenopus Homologue of the Polycomb Group Gene bmi
Xbra	Transcription Factor
XIC	Extracted Ion Count
XPcg	Xenopus Homologue of the Polycomb Gene
XTC	Xenopus Tissue Cell Culture
Y	Tyrosine
y-fagment	Peptide Fragment with the Ion situated at the C-terminus of the Peptide after CID in MS/MS
YY1	Xenopus Homologue of the Polycomb Group Gene Ying-Yang-1
ZipTip	Pipette Tip with a C18 RP Column for sample desalting
2D	Two-Dimensional
3′	C3 Atom of the Phosphate Backbone of the DNA
3-8	Peptide consisting of for instance amino acid 3 to 8
5′	C5 Atom of the Phosphate Backbone of the DNA
°C	Degree Celsius
%	Percent
®	Copyright
γ-H2A.X	Histone 2A Variant

μg	Micro Gram
μl	Micro Liter
μm	Micro Meter
μM	Micro Molar
μm^3	Cubic Micrometer

8 References

Akkers RC, van Heeringen SJ, Jacobi UG, Janssen-Megens EM, Francoijs KJ, Stunnenberg HG, Veenstra GJ, A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in *Xenopus* embryos, *Developmental Cell* 17: 425-434 (2009)

Alberts B, Johnson A, Lewis J, *Molecular biology of the cell*, Cell, NY (2002)

Aldiri I, Vetter ML, Characterization of the expression pattern of the PRC2 core subunit Suz12 during embryonic development of *Xenopus laevis*, *Dev Dyn* 238: 3185-3192 (2009)

Allfrey VG, Faulkner R, Mirsky AE, Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis, *Biochemistry* 51: 786-794 (1964)

Arents G, Moudrianakis EN, The histone fold: A ubiquitous architectural motif utilized in DNA compaction and protein dimerization, *Proc. Natl. Academic Science* 92: 11170-11174 (1995)

Ausio J, Moore SC, Reconstitution of chromatin complexes from High-Performance Liquid Chromatography-purified histones, *A Companion to Methods in Enzymology* 15: 333-342 (1998)

Azura V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merckenschlager M, Fisher AG, Chromatin signatures of pluripotent cell lines, *Nature Cell Biology* (2006)

Bannister AJ, Schneider R, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T, Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes, *Journal of Biological Chemistry* 280: 17732-17736 (2005)

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K, High-resolution profiling of histone methylations in the human genome, *Cell* 129: 823-837

(2007)

Bartova E, Krejci J, Harnicarova A, Galiova G, Kozubeck S, Histone modifications and nuclear architecture, *Journal of Histochemistry and Cytochemistry* 56: 711-721 (2008)

Bas van Steensel, B, Constructing the big picture, *EMBO Journal* 30: 1885-1895 (2011)

Becker PB, Hörz W, ATP-dependent nucleosome remodeling, *Annual Review Biochemistry* 71: 247-273 (2002)

Bednar J, Horowitz RA, Grigoryev SA, Carruthers LM, Hansen JC, Koster AJ, Woodcock CL, Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin, *Proc. Natl. Academic Science* 95: 14173-14178 (1998)

Beetschen JC, Amphibian gastrulation: history and evolution of a 125 year-old concept, *International Journal of Developmental Biology* 45: 771-795 (2001)

Berger SL, The complex language of chromatin regulation during transcription, *Nature* 447: 407-412 (2007)

Bernstein BE, Humphrey EL, Erlich RL, Schneider R, Bouman P, Liu JS, Kouzarides T, Schreiber SL, Methylation of histone H3 Lys 4 in coding regions of active genes, *PNAS* 99: 8695-8700 (2002)

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Path K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES, A bivalent chromatin structure marks key developmental genes in embryonic stem cells, *Cell* 125: 315-326 (2006)

Bhaumik SR, Smith E, Shilatifard A, Covalent modifications of histones during development and disease pathogenesis, *Nature Structural and Molecular Biology* 14: 1008-1016 (2007)

Blobel G, Potter VR, Nuclei from rat liver: an isolation method that combines purity with high yield, *Science* 154: 1662-1665 (1966)

Bloom KS Anderson JN, Fractionation and characterization of chromosomal proteins by the hydroxyapatite dissociation method, *Journal of Biological Chemistry* 258: 4446-4450 (1977)

Bonaldi T, Imhof A, Regula JT, A combination of different mass spectrometric techniques for the analysis of dynamic changes of histone modifications, *Proteomics* 5: 1382-1396 (2004)

Bondarenko PV, Chelius D, Shaler TA, *Anal Chem* 74:4741-4749 (2002)

Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, Bell GW, Otte AP, Vidal M, Gifford DK, Young RA, Jaenisch R, Polycomb complexes repress developmental regulators in murine embryonic stem cells, *Nature* 441: 349-353 (2006)

Braastad CD, Hovhannisyan H, van Wijnen AJ, Stein JL, Stein GS, Functional characterization of a human histone gene cluster duplication, *Genes* 342: 35-40 (2004)

Brand M, Rampalli S, Chaturvedi CP, Dilworth FJ, Analysis of epigenetic modifications of chromatin at specific gene loci by native chromatin immunoprecipitation of nucleosomes isolated using hydroxyapatite chromatography, *Nature Protocols* 3: 398-409 (2008)

Braude P, Bolton V, Moore S, Human gene expression first occurs between the four- and eight-cell stages of preimplantation development, *Nature* 332: 459-461 (1988)

Braunstein M, Rose AB, Holmes SG, Allis CD, Broach JR, Transcriptional silencing in yeast is associated with reduced nucleosome acetylation, *Genes and Development* 7: 592-604 (1993)

Burton DR, Butler MJ, Hyde JE, Phillips D, Skidmore CJ, Walker IO, The interaction of core histones with DNA: equilibrium binding studies, *Nuclei Acid Research* 5: 3643-3664 (1978)

Buszczak M, Paterno S, Spradling AC, *Drosophila* stem cells share a common requirement for the histone H2B ubiquitin protease scrawny, *Science* 323: 248- 251 (2009)

Campbell KHS, McWhir J, Ritchie WA, Wilmut I, Sheep cloned by nuclear transfer from a cultured cell line, *Nature* 380: 64-66, (1996)

Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y, Role of histone H3 lysine 27 methylation in Polycomb group silencing, *Science* 298: 1039-1043 (2002)

Certa U, Ehrenstein G, Reversed-phase high-performance liquid chromatography of histones, *Anal Biochemistry*, 118: 147-154 (1981)

Chin ET, Papac DI, The use of a porous graphitic carbon column for desalting hydrophilic peptides prior to matrix-assisted LASER desorption/ Ionization Time of Flight mass spectrometry, *Analytical Biochemistry* 273: 179-185 (1998)

Chubb JR, Gene activation at the edge of the nucleus, *EMBO Journal* 28: 2145-2146 (2009)

Clarke DJ, O'Neill LP, Turner BM, Selective use of H4 acetylation sites in the yeast *Saccharomyces cerevisiae*, *Biochemistry Journal* 294: 557-561 (1993)

Cremer T, Cremer C, Chromosome territories, nuclear architecture and gene regulation in mammalian cells, *Nature Reviews Genetics* 4: 292-301(2001)

Dai B, Rasmussen TP, Global epiproteomic signatures distinguish embryonic stem cells from differentiated cells, *Stem Cells* 25: 2567-2574 (2007)

Dambacher S, Hahn M, Schotta G, Epigenetic regulation of development by histone

lysine methylation, *Heredity* 105: 24-37 (2010)

Das C, Lucia MS, Hansen KC, Tyler JK, CBP/p300-mediated acetylation of histone H3 on lysine 56, *Nature* 459: 113-117 (2009)

Daujat S, Weiss T, Mohn F, Lange UC, Ziegler-Birling C, Zeissler U, Lappe M, Schübler D, Torres-Padilla ME, Schneider R, H3K64 tri-methylation marks heterochromatin and is dynamically remodeled during developmental reprogramming, *Nature Structural and Molecular Biology* 16: 777-782 (2002)

De la Cruz X, Lois S, Sanchez-Molina S, Martinez-Balbas MA, Do protein motifs read the histone code? *Bioassays* 27: 164-175 (2005)

DeLange RJ, Fambrough DM, Smith EL, Bonner J, Calf and pea histone IV, *Journal of Biological Chemistry* 244: 5569-5679 (1969)

De Robertis EM, Spemann's organizer and the self-regulation of embryonic fields, *Mech Dev* 126: 925-941 (2009)

Dimitrov S, Almouzni G, Dasso M, Wolffe AP, Chromatin transitions during early *Xenopus* embryogenesis: changes in histone H4 acetylation and in linker histone type, *Developmental Biology* 160: 214-227 (1993)

Dou Y, Milne TA, Tackett AJ, Smith ER, Fukuda A, Woyssocka J, Allis CD, Chait BT, Hess JL, Roeder RG, Physical association and coordinate function of the H3K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF, *Cell* 121: 873-885 (2005)

Dyce J, George M, Goodall H, Fleming TP, Do trophectoderm and inner cell mass cells in the mouse blastocyst maintain discrete lineages?, *Development* 100: 685-698 (1987)

Eberl HC, Mann M, Vermeulen M, Quantitative Proteomics for Epigenetics, *Chem Bio Chem* 12: 224-234 (2011)

Efroni S, Duttagupta R, Cheng J, Dehghani H, Hoepfner DJ, Dash C, Bazett-Jones DP, Le Grice S, McKay RD, Buetow KH, Gingeras TR, Misteli T, Meshorer E, Global transcription in pluripotent embryonic stem cells, *Cell Stem Cell*. 5: 437-447 (2008)

Evans MJ, Kaufman MH, Establishment in culture of pluripotent cells from mouse embryos, *Nature* 292: 154-156 (1981)

Felsenfeld G, Groudine M, Controlling the double helix, *Nature* 421: 448-453 (2003)

Fischle W, Wang Y, Allis CD, Histone and chromatin cross-talk, *Current Opinion in Cell Biology* 15: 172-183 (2003)

Fischle W, Tseng BS, Dormann HL, Ueberheide BM, Garcia BA, Shabanowitz J, Hunt DF, Funabiki H, Allis CD, Regulation of HP1 chromatin binding by histone H3 methylation and phosphorylation, *Nature* 438: 1116-1122 (2005)

Flanagan JF, Mi LZ, Chruszcz M, Cymborowski M, Clines KL, Minor W, Rastinejad F, Khorasanizadeh, Double chromodomains cooperate to recognize the methylated histone H3 tail, *Nature* 438: 1181-1185 (2005)

Freitas MA, Sklenar AR, Parthun MR, Application of mass spectrometry to the identification and quantification of histone post-translational modifications, *Journal of Cellular Biochemistry* 92: 691-700 (2004)

Fuda NJ, Ardehali MB, Lis JT, Defining mechanisms that regulate RNA polymerase II transcription in vivo, *Nature* 461: 186-192 (2009)

Garcia BA, Hake SB, Diaz RL, Kauer M, Morris SA, Organismal differences in post-translational modifications in histones H3 and H4, *Journal of Biological Chemistry* 282: 7641-7655 (2007)

Gilbert SF, *Developmental Biology*, Sinauer Associates, Incorporation, Sunderland, Massachusetts, USA, Eighth Edition, (2006)

Gregory GD, Vakoc CR, Rozovskaia T, Zheng X, Patel S, Nakamura T, Canaani E, Blobel GA, Mammalian ASH1 is a histone methyltransferase that occupies the transcribed region of active genes, *Molecular and Cellular Biology* 27: 8466-8479 (2007)

Gropengiesser J, Varadarajan BT, Stephanowitz H, Kraus E, The relative influence of phosphorylation and methylation on responsiveness of peptides to MALDI and ESI mass spectrometry, *Journal of Mass Spectrometry* (2009)

Gould H, Chromatin: A practical approach, The practical Approach series by Rickwood D, Hames BD, Oxford University Press, Oxford, New York, Tokyo, (1998)

Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA, A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells, *Cell* 130: 77-88 (2007)

Gurdon JB, A community effect in animal development, *Nature* 336: 772-774 (1988)

Gurdon JB, Hopwood N, The introduction of *Xenopus laevis* into developmental biology: of empire, pregnancy testing and ribosomal genes, *International Journal of Developmental Biology* 44: 43-50 (2000)

Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Chromatin dynamics during epigenetic reprogramming in the mouse germ line, *Nature* 452: 877-881 (2008)

Hake SB, Allis CD, Histone H3 variants and their potential role in indexing mammalian genomes: the H3 barcode hypothesis, *PNAS* 103: 6428-6435 (2006)

Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Distinctive chromatin in human sperm packages genes for embryo development, *Nature* 460: 473-478 (2009)

Hassa PO, Haenni SS, Elser M, Hottiger MO, Nuclear ADP-ribosylation reacts in mammalian cells: where are we today and where are we going?, *Microbiology and Molecular Biology Reviews* 70: 789-829 (2006)

Heasman J, Wylie CC, Hausen P, Smith JC, Fates and stages of determination of single vegetal pole blastomeres of *Xenopus laevis*, *Cell* 37: 185-194 (1984)

Heasman J, Patterning the early *Xenopus* embryo, *Development* 133: 1205-1217 (2006)

Hemberger M, Dean W, Reik W, Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal, *Chromatin Dynamics* 10: 526-537 (2009)

Henikoff S, Ahmad K, Assembly of variant histones into chromatin, *Annual Review in Cell Developmental Biology* 21: 133-153 (2005)

Hermann A, Gowher H, Jeltsch A, Biochemistry and biology of mammalian DNA methyltransferases, *Cellular Molecular Life Science* 61: 2571-2587 (2004)

Hotchkiss RD, The quantitative separation of purines, pyrimidines and nucleosides by paper chromatography. *J. Biol. Chem.* 175: 315-332 (1948)

Houston DW, Kofron M, Resnik E, Langland R, Destree O, Wylie C, Heasman J, Repression of organized genes in dorsal and ventral *Xenopus* cells mediated by maternal XTcf3, *Development* 129: 4015-4025 (2002)

Jaenisch R, Bird A, Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals, *Nature Genetics Supplement* 33: 245-254 (2003)

Jin C, Felsenfeld G, Nucleosome stability mediated by histone variants H3.3 and H2A.Z, *Genes & Development*, 21: 1519-1529 (2007)

Jones PA, Baylin SB, The epigenome of cancer, *Cell* 128: 683-692 (2007)

Kapp EA, Schütz F, Conolly LM, Chakel JA, Meza JE, Miller CA, Fenyo D, Eng JK, Adkins JA, Omenn GS, Simpson RJ, An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: sensitivity and specificity analysis, *Proteomics* 5: 3475-3490 (2005)

Karas M, Hillenkamp F, LASER desorption ionization of proteins with molecular masses exceeding 10,000 daltons, *Analytical Chemistry* 60: 2299-2301 (1988)

Kimura A, Matsubara K, Horikoshi M, A decade of histone acetylation: marking eukaryotic chromosomes with specific codes, *Journal Biochemistry* 138: 647-662 (2005)

Kirkpatrick DS, Denison C, Gygi SP, Weighing in on ubiquitin: the expansion role of mass spectrometry-based proteomics, *Nature Cell Biology* 7: 750-757 (2005)

Klose RJ, Yamane K, Bae Y, Zhang D, Erdjument-Bromage H, Tempst P, Wong J, Zhang Y, The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36, *Nature* 442: 312-316 (2006)

Kofron M, Wylie C, Heaseman J, The role of Mixer in patterning the early *Xenopus* embryo, *Development* 131: 2431-2441 (2004)

Kornberg RD, Chromatin structure: a repeating unit of histones and DNA, *Science* 184: 868-871 (1974)

Kornberg RD, Lorch Y, Twenty-five years of the nucleosome, fundamental particle of the eucaryote chromosome, *Cell* 98: 285-294 (1999)

Kossel A, *Z. Physiologisch Chemie* 8: 511-515 (1884)

Kouzarides T, Chromatin modifications and their function, *Cell* 128: 693-705 (2007)

Kurokawa M, MacLeod MC, Separation of histones by reversed-phase high-performance liquid chromatography: analysis of the binding of carcinogens to histones, *Analytical Biochemistry* 144: 47-54 (1984)

Kurth T, Meissner S, Schackel S, Steinbeisser H, Establishment of mesodermal gene expression patterns in early *Xenopus* embryos: the role of repression, *Dev Dyn* 233: 418-

429 (2005)

Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T, Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins, *Nature* 410: 116-120 (2001)

Ladurner AG, Tick-tock goes the acetylation clock, *Nature Structural Biology* 10: 83 (2003)

Latham JA, Dent SRY, Cross-regulation of histone modifications, *Nature Structural and Molecular Biology* 14: 1017-1024 (2007)

Lee MG, Wynder C, Bochar DA, Hakimi MA, Cooch N, Shiekhattar R, Functional interplay between histone demethylase and deacetylase enzymes, *Molecular and Cellular Biology* 26: 6395-6402 (2006)

Lee KK, Workman JL, Histone acetyltransferase complexes: one size doesn't fit all, *Nature Reviews* 8: 284-295 (2007)

Lee JS, Shilatifard A, A site to remember: H3K36 methylation a mark for histone deacetylation, *Mutation Research* 618: 130-134 (2007a)

Lee BM, Mahadevan LC, Stability of Histone Modifications Across Mammalian Genomes: Implications for „Epigenetic“ Marking, *Journal of Cellular Biochemistry* 108: 22-34 (2009)

Lee JS, Smith E, Shilatifard A, The Language of Histone Crosstalk, *Cell* 142: 682- 685 (2010)
Minireview

Liang G, Lin JC, Wei V, Yoo C, Cheng JC, Nguyen CT, Weisenberger DJ, Egger G, Takai D, Gonzales FA, Jones PA, Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome, *Proc Natl Acad Sci* 101: 7357-7362 (2004)

Lindner HH, Analysis of histones, histone variants and their post-translationally modified forms, *Electrophoresis* 29: 2516-2532 (2008) Review

-
- Lindner H**, Sarg B, Hoertnagl B, Helliger W, The microheterogeneity of the mammalian H1° histone, *Journal of Biological Chemistry* 273: 13324-13330 (1998)
- Lindner H**, Wesierska-Gadek J, Helliger W, Puschendorf B, Sauermann G, *Journal of Chromatography* 472: 243-249 (1989)
- Lottspeich F**, Engels JW, *Bioanalytik*, Spektrum Verlag, Munich, (2006)
- Loyola A**, Bonaldi T, Roche D, Imhof A, Almouznis G, PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state, *Molecular Cell* 24: 309-16 (2006)
- Luger K**, Mäder AW, Richmond RK, Sargent DF, Richmond TJ, Crystal structure of the nucleosome core particle at 2.8Å resolution, *Nature* 389: 251-260 (1997)
- Maisonneuve C**, Guilleret I, Vick P, Weber T, Andre P, Bicaudal C, a novel regulator of Dvl signaling abutting RNA-processing bodies, controls cilia orientation and leftward flow, *Development* 136: 3019-3030 (2009)
- Martin GR**, Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells, *Proc. Natl. Acad. Science* 78: 7634-7638 (1981)
- Marzluff WF**, Gongidi P, Woods KR, Jin J, Maltais LJ, The human and mouse replication dependent histone genes, *Genomics* 80: 487-498 (2002)
- Mayer W**, Niveleau A, Walter J, Fundele R, Haaf T, Demethylation of the zygotic paternal genome, *Nature* 403: 501-502 (2000).
- Meehan RR**, Dunican DS, Ruzov A, Pennings S, Epigenetic silencing in embryogenesis, *Experimental Cell Research* 309: 241-249 (2005)
- Meshorer E**, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T, Hyperdynamic

plasticity of chromatin proteins in pluripotent embryonic stem cells, *Developmental Cell* 10: 105-116 (2006)

Metzger E, Wissmann M, Schüle R, Histone demethylation and androgen-dependent transcription, *Current Opinion in Genetics and Development* 16: 513-517 (2006)

Metzger E, Yin N, Wissmann M, Kunowska N, Fischer K, Friedrichs N, Patnaik D, Higgins JM, Potier N, Scheidtmann KH, Buettner R, Schüle R, Phosphorylation of histone H3 at threonine 11 establishes a novel chromatin mark for transcriptional regulation, *Nature Cellular Biology* 10: 53-60 (2008)

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE, Genome-wide maps of chromatin state in pluripotent and lineage-committed cells, *Nature* 448: 553-560 (2007)

Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Dissecting direct reprogramming through integrative genomic analysis, *Nature* 454: 49-55 (2008)

Minard ME, Jain AK, Barton MC, Analysis of epigenetic alterations to chromatin during development, *Genesis* 47: 559-572 (2009)

Moore GD, Kopf GS, Schultz RM, Complete mouse egg activation in the absence of sperm by stimulation of an exogenous G protein-coupled receptor, *Developmental Biology* 159: 669-678 (1993)

Muller M, Fleischmann K, Selbert S, Ji GJ, Endl E, Middeler G, Müller OJ, Schlenke P, Wobus AM, Hescheler J, Katus HA, Franz WM, Selection of ventricular-like cardiomyocytes from ES cells in vitro, *FASEB* 14: 2540-2548 (2000)

Murray K, The acid extraction of histones from calf thymus desoxyribonucleoproteins, *Journal of Molecular Biology* 15: 409-419 (1966)

-
- Muthurajan** UM, Park YJ, Edayathumangalam RS, Suto RK, Chakravarthy S, Dyer PN, Luger K, Structure and dynamics of nucleosomal DNA, *Biopolymers* 68: 547-556 (2003)
- Nathan** D, Ingvarsdottir K, Sterner DE, Bylebyl GR, Dokmanociv M, Dorsey JA, Whelan KA, Krsmanovic M, Lane WS, Meluh PB, Histone SUMOylation is a negative regulator in *Saccharomyces cerevisiae* and shows dynamic interplay with positive acting histone modifications, *Genes and Development* 20: 966-976 (2006)
- Neilson** AK, Ali NA, Muralidharan S, Mirzaei M, Mariani M, Assadourian G, Lee A, van Sluyter ST, Haynes PA, Less label, more free: Approaches in label-free quantitative mass spectrometry, *Proteomics* 11: 535-553 (2011)
- Nelson** CJ, Santos-Rosa H, Kouzarides T, Proline isomerization of histone H3 regulates lysine methylation and gene expression, *Cell* 126: 905-916 (2006)
- Newport** J, Kirschner M, A major developmental transition in early *Xenopus* embryos: I. characterization and timing of cellular changes at the midblastula stage, *Cell* 30: 675-686 (1982a)
- Newport** J, Kirschner M, A major developmental transition in early *Xenopus* embryos: II. Control of the onset of transcription, *Cell* 30: 687-696 (1982b)
- Ng** HH, Feng Q, Wang H, Erdjument-Bromage H, Tempst P, Zhang Y, Struhl K, Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association, *Genes and Development* 16: 1518-1527 (2002)
- Nicklay** JJ, Schechter D, Chitta RK, Garcia BA, Shabanowitz J, Allis CD, Hunt DF, Analysis of histones in *Xenopus laevis* part II: mass spectrometry reveals an index of cell-type specific modifications on H3 and H4, *PNAS* (2009)
- Nieuwkoop** PD, Faber J, Normal table of *Xenopus laevis* (Daudin). Amsterdam: North-Holland Publishing Company (1967)

Nishioka K, Rice JC, Sarma K, Erdjument-Bromage H, Werner J, Wang Y, Chuikov S, Valenzuela P, Tempst P, Steward R, Lis JT, Allis CD, Reinberg D, PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin, *Molecular Cell* 9: 1201-1213 (2002)

Oda H, Okamoto I, Murphy N, Chu J, Price SM, Shen MM, Torres-Padilla ME, Heard E, Reinberg D, Monomethylation of histone H4-lysine 20 is involved in chromosome structure and stability and is essential for mouse development, *Molecular and Cellular Biology* 29: 2278-2295 (2009)

Olvier SS, Denu JM, Dynamic Interplay between Histone H3 Modifications and Protein Interpreters: Emerging Evidence for a "Histone Language", *ChemBioChem* 12: 299-307 (2011), Minireview

Orford K, Kharchenko P, Lai W, Dao MC, Worhunsky DJ, Ferro A, Janzen V, Park PJ, Scadden DT, Differential H3K4 methylation identifies developmentally poised hematopoietic genes, *Developmental Cell* 14: 798-809 (2008)

Osago H, Yamada K, Shibata T, Yoshino K, Hara N, Tsuchiya M, Precursor ion scanning and sequencing of arginine-ADP-ribosylated peptide by mass spectrometry, *Analytical Biochemistry* 393: 248-254 (2009)

Paik WK, Kim S Enzymatic methylation of histones, *Arch Biochemistry and Biophysics* 134: 632-637 (1969)

Pasini D, Malatesta M, Jung HR, Walfridsson J, Willer A, Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes, *Nucleic acids research* 38: 4958-4969 (2010)

Pesavento JJ, Mizzen CA, Kelleher NL, Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: human histone H4, *Analytical Chemistry* 78: 4271-480 (2006)

Peters L, Meister G, Argonaute proteins: mediators of RNA silencing, *Molecular Cell* 26: 611- 623 (2007)

Phanstiel D, Brumbaugh J, Berggren WT, Conard K, Feng X, Levenstein ME, McAllister GC, Thomson JA, Coon JJ, Mass spectrometry identifies and quantifies 74 unique histone H4 isoforms in differentiating human embryonic stem cells, *PNAS* 105: 4093-4098 (2007)

Prelle K, Zink N, Eckhard Wolf, Pluripotent stem cells- model of embryonic development, tool for gene targeting, and basis of cell therapy, *Anat. Histol. Embryology* 31: 169-186 (2002) Review

Pudney M, Varma MG, Leake CJ, Establishment of a cell line (XTC-2) from south African clawed toad, *Xenopus laevis*, *Experientia* 29: 466-467 (1972)

Rafferty KA, Mass culture of amphibian cells: methods and observations concerning stability of cell type, in: Mizell M, editor, *Biology of Amphibian Tumors*, Berlin: Springer Verlag, pp. 52-81 (1973)

Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD, Jenuwein T, Regulation of chromatin structure by site-specific histone H3 methyltransferases, *Nature* 406: 593-599 (2000)

Reijnen MJ, Hamer KH, den Blaauwen JL, Lambrechts C, Schoneveld I, van Driel R, Otte AP, Polycomb and bmi-1 homologs are expressed in overlapping patterns in *Xenopus* embryos and are able to interact with each other, *Mechanisms of Development* 53: 35-46 (1995)

Reik W, Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 447: 425-432 (2007).

Richly H, DiCroce L, The flip of the coin: the role of ZRF1 and histone H2A ubiquitination in transcriptional activation, *Cell Cycle* 10: 745-750 (2011)

Rodriguez-Collazo P, Leuba SH, Zlatanova J, Robust methods for purification of histones from cultured mammalian cells with the preservation of their native modifications, *Nucleic Acids Research* 1-15 (2009)

Rogakou EP, Pilch DR, Orr AH, Ivanova VS; Bonner WM, DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139, *Journal of Biological Chemistry* 273: 5858-5868 (1998)

Rugg-Gunn PJ, Cox BJ, Ralston A, Rossant J, Distinct histone modifications in stem cell lines and tissue lineages from the early mouse embryo. *Proc Natl Acad Sci* 107: 10783-10790 (2010)

Rupp RAW, Becker PB, Gene regulation by histone H1: new links to DNA methylation, *Cell* 123: 1178-1180 (2005)

Sambrook J, Fritsch EF, Maniatis T, *Molecular Cloning: A laboratory manual*, Cold Spring Harbour Laboratory Press, New York (1989)

Sanchez-Elsner T, Gou D, Kremmer E, Sauer F, Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax. *Science* 311: 1118-1123 (2006)

Sarcinella E, Zuzarte PC, Lau PN, Draker R, Cheung P, Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin, *Molecular Cell Biology* 27: 6457-6468 (2007)

Satijn DPE, Hamer KM, den Blaauwen J, Otte AP, The polycomb group proteins EEd interacts with YY1, and both proteins induce neural tissue in *Xenopus* embryos, *Molecular and Cellular Biology* 21: 1360-1369 (2001)

Schechter D, Dormann HL, Allis CD, Hake SB, Extraction, purification and analysis of histones, *Nature Protocols* 6: 1445-1457 (2007)

Schechter D, Nicklay JJ, Chitta RK, Shabanowitz J, Hunt DF, Allis CD, Analysis of histones in *Xenopus laevis* part I: A Distinct Index of Enriched Variants and Modifications exists in each cell Type and is removed During Developmental Transitions, *PNAS* (2009)

Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T, A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin, *Genes and Development* 18: 1251-1262 (2004)

Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, A chromatin wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse, *Genes Development* 22: 2048-2061 (2008)

Seet BT, Dikic I, Zhou MM, Pawson T, Reading protein modifications with interaction domains, *Nature Reviews Molecular and Cellular Biology* 7: 473-83 (2006)

Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y, Histone demethylation mediated by the nuclear amine oxidation homolog LSD1, *Cell* 119: 941-953 (2004)

Shi X, Hong T, Walter KL, Ewalt M, Michishita E, Hung T, Carney D, Pena P, Lan F, Kaadige MR, ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression, *Nature* 442: 96-99 (2006)

Shilatifard A, Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression, *Annual Review of Biochemistry* 75: 243-269 (2006)

Shiraishi M, Oates AJ, Sekiya T, An overview of the analysis of DNA methylation in mammalian genomes, *Biological Chemistry* 383: 893-906 (2002)

Shogren-Knaak M, Ishii H, Sun JM, Pazin MJ, Davie JR, Peterson CL, Histone H4-K16 acetylation controls chromatin structure and protein interactions, *Science* 311: 844-847 (2006)

Sierra F, Lichtler A, Marashi F, Rickles R, van Dyke T, Clark S, Wells J, Stein G, Stein J, Organization of human histone genes, *Proc. Natl. Science* 79: 1795-1799 (1982)

Sieve HL, Hattori K, Weintraub H, Progressive determination during formation of the anteroposterior axis in *Xenopus laevis*, *Cell* 58: 171-180 (1989)

Silva J, Smith A, Capturing pluripotency, *Cell* 132: 532-536 (2008)

Simon RH, Felsenfeld G, A new procedure for purifying histone pairs H2A + H2B and H3 + H4 from chromatin using hydroxyapatite, *Nuclei Acid Research* 6: 869-696 (1979)

Slack JM, Embryonic induction, *Mechanism in Development* 41: 91-107 (1993)

Slack JM, Conrad Hal Waddington: the last renaissance biologist?, *Nature Reviews Genetics* 3: 889-895 (2002)

Strahl BD, Allis CD, The language of covalent histone modifications, *Nature* 403: 41-45 (2000)

Strahl BD, Briggs SD, Brame CJ, Caldwell JA, Koh SS, Ma H, Cook RG, Shabanowitz J, Hunt DF, Stallcup MR, Allis CD, Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1, *Current Biology* 11: 996-1000 (2001)

Sulewska A, Niklinska W, Kozlowski M, Minarowski L, Naumnik W, Nikilinski J, Dabrowska K, Chyzewski L, DNA methylation in states of cell physiology and pathology, *Folia Histochemica et Cytobiologica* 45: 149-158 (2007)

Surani S, Hayashi K, Hajkova P, Genetic and epigenetic regulators of pluripotency, *Cell* 128: 747-762 (2007)

Tachibana M, Matsumura Y, Fukuda M, Kimura H, Shinkai Y, G9a/Glp complexes independently mediate H3K9 and DNA methylation to silence transcription, *EMBO* 27:

2681-2690 (2002)

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S, Induction of pluripotent stem cells from mouse embryonic and adult fibroblasts by defined factors, *Cell* 131: 861-872 (2007)

Taverna SD, Iling S, Rogers RS, Tanny JC, Lavender H, Li H, Baker L, Boyle J, Blair LP, Patel DJ, Aitchison JD, Tackett AJ, Allis CD, Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs, *Molecular Cell* 24: 785-796 (2006)

Taverna SD, Ueberheide BM, Liu Y, Tackett AJ, Diaz RL, Shabanowitz J, Chalt BT, Hunt DF, Allis CD, Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini, *PNAS* 104: 2086-2091 (2007)

Thoma F, Koller T, Klug A, Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin, *Journal of Cell Biology* 83: 403-427 (1979)

Thomas CE, Kelleher NL, Mizzen CA, Mass Spectrometric Characterization of Human Histone H3: A Bird's Eye View, *Journal of Proteome Research* 5, 240-247 (2006)

Trelle MB, Jensen ON, Functional proteomics in histone research and epigenetics, *Expert Reviews Proteomics* 4: 491-503 (2007)

Tsukiyama T, The in vivo functions of ATP-dependent chromatin remodeling factors, *Nature Reviews*, 3: 422-429 (2002)

Turner BM, Reading signals on the nucleosome with a new nomenclature for modified histones, *Nature Structural and Molecular Biology* 12: 110-112 (2005)

Vaquero A, Loyola A, Reinberg D, The constantly changing face of chromatin, *Science of Aging Knowledge Environment* 1- 16 (2003)

Vastenhouw NL, Zhang Y, Woods IG, Imam F, Regev A, Chromatin signature of embryonic pluripotency is established during genome activation, *Nature* 464: 922-926 (2010)

Veenstra GJ, Wolffe AP, Gene-selective developmental roles of general transcription factors, *Trends in Biochemistry Science* 26: 665-671 (2001)

Villar-Garea A, Israel L, Imhof A, Analysis of histone modifications by mass spectrometry, *Current Protocols in Protein Science*, 14.10.1- 14.10.14 (2008)

Wang Y, Woysocka J, Sayegh J, Lee YH, Perlin JR, Leonelli L, Sonbuchner LS, McDonald CH, Cook RG, Dou Y, Human PAD4 regulates histone arginine methylation levels via demethylination, *Science* 306: 279-283 (2004)

Wang ZF, Krasikov T, Frey MR, Wang J, Matera AG, Marzluff WF, Characterization of the mouse histone gene cluster on chromosome 13: 45 histone genes in three patches spread over 1 Mb, *Genome Research* 6: 688-701 (1996)

Wardle FC, Smith JC, Refinement of gene expression patterns in the early *Xenopus* embryo, *Development* 131: 4687-4696 (2004)

Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP, Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells, *Nature Genetics* 41: 246-350 (2009)

West MH, Bonner WM, Histone 2B can be modified by the attachment of ubiquitin, *Nuclei Acid Research* 8: 4671-4680 (1980)

West C, Elfakir C, Lafosse M, porous graphitic carbon: A versatile stationary phase for liquid chromatography, *Journal of Chromatography A* 1217: 3201-3216 (2010)

Wetmur JG, Davidson N, Kinetics of renaturation of DNA, *Journal of Molecular Biology*

31: 349-370 (1968)

Whitehouse CM, Dreyer RN, Yamashita M, Fenn JB, Electrospray interference for liquid chromatographs and mass spectrometry, *Analytical Chemistry* 57: 675-679 (1985)

Wood C, Snijders A, Willimason J, Reynolds C, Baldwin J, Dickman M, Post-translational modifications of the linker histone variants and their association with cell mechanisms, *FEBS Journal* 276: 3685-3697 (2009)

Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodeling, *Nature* 442: 86-90 (2006)

Xie W, Song C, Young NL, Sperling AS, Xu F, Sridharan R, Conway AE, Garcia BA, Plath K, Clark AT, Grunstein M, Histone H3 lysine 56 acetylation is linked to the core transcriptional network in human embryonic stem cells, *Molecular Cell* 33: 417-427 (2009)

Xu F, Zhang K, Grunstein M, Acetylation in Histone H3 Globular Domain Regulates Gene Expression in Yeast, *Cell* 121: 375-386 (2006)

Yang CA, Mizzen H, The multiple facets of histone H4 lysine 20 methylation, *Biochemistry Cell Biology* 87: 151-161 (2008)

Young RA, Control of the embryonic stem cell state, *Cell* 144: 940-954 (2011)

Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin II, Thomson JA, Induced pluripotent stem cell lines derived from human somatic cells, *Science* 318: 1917-1920 (2007)

Zhang K, Williams KE, Huang L, Yau P, Silino JS, Bradbury EM, Jones PR, Minch MJ, Burlingame AL, Histone acetylation and deacetylation, *Molecular and Cellular Proteomics* 7: 500-508 (2002)

Zhang L, Eugeni EE, Parthun MR, Freitas MA, Identification of noval histone post-translational modifications by peptide mass fingerprint, *Chromosoma* 112: 77-86 (2003)

Zhu B, Zheng Y, Pham AD, Mandal SS, Erdjument-Bromage H, Tempst P, Reinberg D, Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation, *Molecular Cell* 20: 601-611 (2005)

9 Supplement

Xenopus Histone H2A - Accession Number AAA49769

MSGRGKGGGK TRAKAKTRSS RAGLGFPVGR VHRLLRKGNV
AERVGAGAPV YLAHVLEYLT AEILELAGNA ARDNKKTRII
PRHLGLAVRN DEELNKLLGG VTIAGGGVLP NIGSTVLLPKK
TESAKSAKSK

Xenopus Histone H2B - Accession Number AAA49768

MPEPAKSAPA PKKGSKKAVT KTQKKDGKKR RKSARKESYAI YVYKVLKQVH
PDTGISSKAM SIMNSFVNDV FERIAGEASR LAHYNKRSTI TSREIQTAVR
LLLPGELAKH AVSEGTKAVT KYTSAK

Xenopus Histone H3 - Accession Number AAA49770

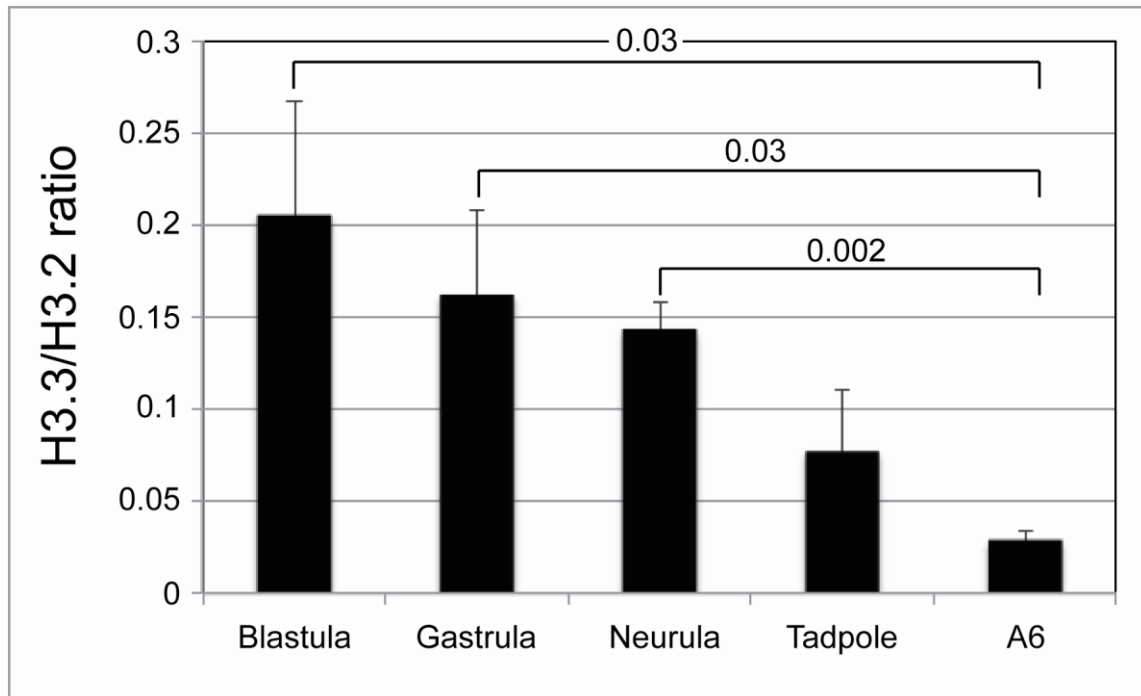
MARTKQTARK STGGKAPRKQ LATKAARKSA PATGGVKKPH RYRPGTVALR
EIRRYQKSTE LLIRKLPFQR LVREIAQDFK TDLRFQSSAV MALQEASEAY
LVGLFEDTNL CAIHAKRVTI MPKDIQLARR IRGERA

Xenopus Histone H4 - Accession Number AAA49771

MSGRGKGGGK LGKGGAKRHR KVLKRDNIGGI TKPAIRRLAR RDGVKRISGL
IYEETRGVLK VFLENVIRDA VTYTEHAKRK TVTAMDVVYA LKRGGRTLYG
FGG

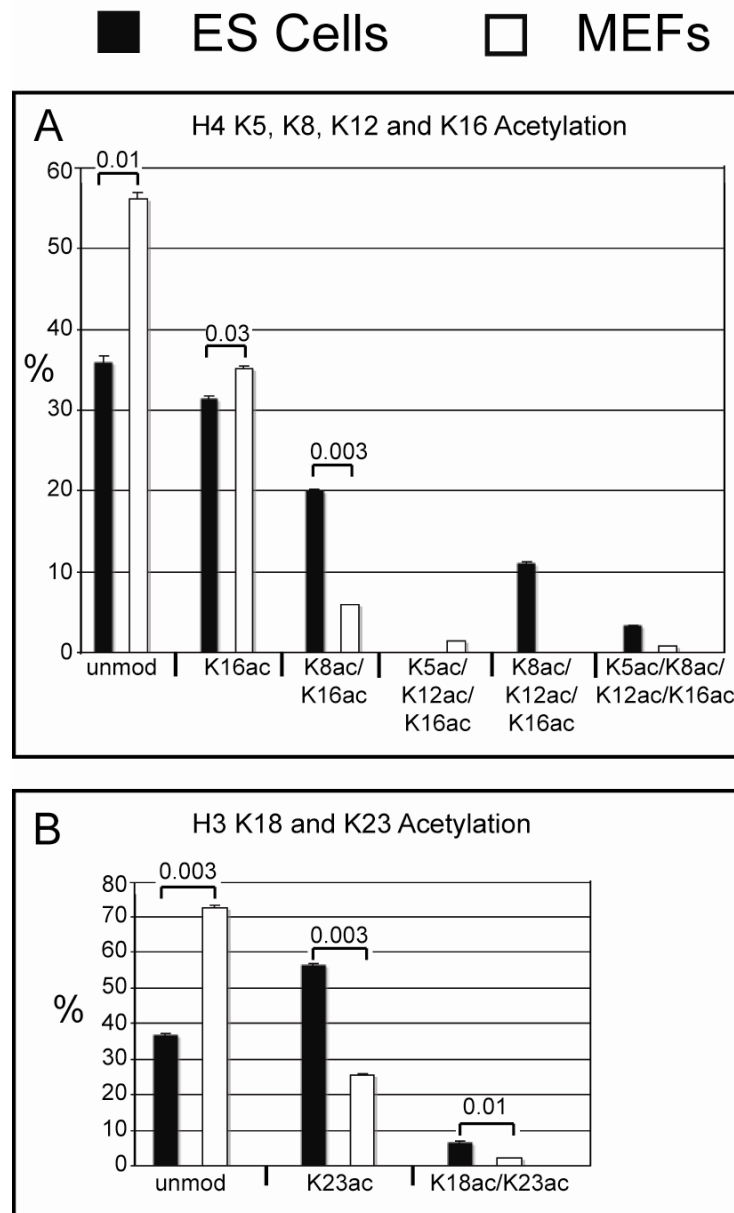
Supplement Figure 1: Histone Sequences of *Xenopus laevis*

Amino acid sequence of the four core histones of *Xenopus laevis*. Accession numbers of the histones according to NCBI database searches.



Supplement Figure 2: Quantification of H3.3/H3.2 Ratios in Xenopus Bulk Histones isolated from Different Developmental Stages

The ratios of MS/MS events for the unique H3.3 peptide KSAPSTGGVKKPHR and the corresponding H3.2 peptide KSAPATGGVKKPHR were calculated and used as a proxy to determine the ratio of H3.3/H3.2 in histones isolated from various embryo samples. Abbreviations: NF9 = blastula, NF12 = gastrula, NF18 = neurula, NF37 = tadpole, A6 - frog kidney cell line. Values represent the mean of at least two independent biological replicates. Error bars represent the calculated SD values.



Supplement Figure 3: Comparison of Representative Histone Modifications of Murine ES Cells and MEFs to illustrate adequate Separation by FACS

Bar-Charts of Histone Modifications. Data from Orbi-Trap Mass Spectrometry, Error bars indicate SD of two independent biological replicates (B,C). Abr. unmod = unmodified peptide, Kac = acetylated lysine residue, Kme1 = mono-methylated lysine residue, Kme2 = di-methylated lysine residue, Kme3 = tri-methylated lysine residue.

H2A											
Peptide	Modification	Mass [Da]	NF9	NF12	NF18	NF37	A6	XTC	ES	MEF	
83-89	unmod	836,51	96,03	98,45	99,02	96,36	95,96	95,51	100,00	100,00	
	me1	850,53	3,31	0,80	0,98	0,68	3,61	4,49	0,00	0,00	
	me2	864,54	0,66	0,43	0,00	2,96	0,34	0,00	0,00	0,00	
	phos	916,48	0,00	0,33	0,00	0,00	0,00	0,00	0,00	0,00	
H2B											
Peptide	Modification	Mass [Da]	NF9	NF12	NF18	NF37	A6	XTC	ES	MEF	
87-92	unmod	664,36	100,00	100,00	100,00	100,00	89,43	100,00	100,00	100,00	
	me1	678,38	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
	me2	692,39	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
	phos	744,33	0,00	0,00	0,00	0,00	10,57	0,00	0,00	0,00	
93-99	unmod	816,46	100,00	97,08	100,00	89,65	94,81	98,92	100,00	100,00	
	me1	830,47	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
	me2	844,49	0,00	2,92	0,00	0,00	0,00	0,00	0,00	0,00	
	phos	896,42	0,00	0,00	0,00	10,38	5,19	1,08	0,00	0,00	
H3											
Peptide	Modification	Mass [Da]	NF9	NF12	NF18	NF37	A6	XTC	ES	MEF	
3-8	unmod	760,43	27,88	34,41	31,37	44,78	67,34	61,97	25,66	59,15	
	me1	774,45	18,48	21,35	16,84	18,38	15,75	18,71	16,58	20,80	
	me2	732,44	25,17	25,44	32,83	21,15	9,19	10,92	34,13	10,68	
	me3	746,45	28,48	18,80	18,96	15,69	7,72	8,40	23,63	9,37	
	phos	840,40	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
64-69	unmod	844,50	98,63	98,87	99,25	98,99	97,02	98,88	98,67	96,44	
	me1	858,52	0,63	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
	me2	816,51	0,74	1,13	0,75	1,01	2,98	1,12	1,33	3,56	
	me3	830,53	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
73-83	unmod	1391,72	96,29	96,37	96,02	91,91	94,07	95,25	84,44	82,15	
	me1	1405,73	2,28	3,12	3,90	7,23	5,93	4,38	9,25	10,01	
	me2	1363,72	0,00	0,00	0,08	0,33	0,00	0,00	6,31	7,84	
	me3	1377,74	1,44	0,51	0,00	0,53	0,00	0,37	0,00	0,00	
117-128	unmod	1440,82	100,00	100,00	100,00	98,32	100,00	100,00	100,00	100,00	
	me1	1454,84	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
	me2	1412,83	0,00	0,00	0,00	1,68	0,00	0,00	0,00	0,00	
	me3	1426,85	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
	phos	1426,81	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
H4											
Peptide	Modification	Mass [Da]	NF9	NF12	NF18	NF37	A6	XTC	ES	MEF	
20-23	unmod	571,39	30,48	26,90	22,91	27,96	20,88	17,03	28,20	17,06	
	me1	585,41	26,97	41,51	38,24	22,91	28,98	33,69	29,72	22,41	
	me2	543,43	42,55	31,59	38,85	49,13	50,13	49,28	42,07	60,54	
	me3	557,41	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
56-67	unmod	1442,87	100,00	98,31	99,33	93,40	100,00	100,00	97,80	100,00	
	ac	1428,86	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
	me1	1456,89	0,00	1,69	0,67	6,60	0,00	0,00	2,20	0,00	
	me2	1414,88	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
	me3	1428,89	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
79-92	unmod	1706,95	98,43	97,82	96,39	99,21	100,00	99,47	100,00	100,00	
	ac1	1692,94	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
	ac2	1678,92	1,57	2,18	3,34	0,79	0,00	0,53	0,00	0,00	
	me1	1720,97	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	
	me2	1734,98	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	

Supplement Table 1: Modifications of the Four Core Histones Detected by MALDI-TOF

H3 peptides	M(cal) MH+	AV. RT	NF9 Mean	NF12 Mean	NF18 Mean	NF37 Mean	A6 Mean	XTC Mean	ES Cells Mean	MEFS Mean
Peptide 9-17										
unmod	1013,5738	31,9	40,19	46,48	51,17	47,33	48,03	55,70	49,00	42,61
K9me1	1027,5884	32,9	10,76	19,75	14,47	13,18	28,79	18,01	15,43	19,11
K9me2	985,5789	27,9	0,85	0,61	1,51	0,53	3,06	0,54	0,83	2,38
K9me3	989,5945	27,8	0,14	0,14	0,39	0,15	0,52	0,13	0,25	6,69
K14ac	989,5581	30,9	45,58	30,88	28,42	35,78	15,36	23,94	28,92	21,13
K8acK14ac	885,5425	30,0	1,17	1,16	1,20	2,41	0,67	0,88	4,63	1,89
K9me1K14ac	1013,5738	33,3	0,04	0,02	0,04	0,09	0,05	0,03	0,02	0,06
K9me2K14ac	971,5632	27,7	0,48	0,28	0,36	0,12	0,51	0,13	0,15	3,82
K9me3K14ac	985,5789	27,9	0,85	0,61	1,51	0,53	3,06	0,54	0,83	2,38
Peptide 18-26										
unmod	1089,6629	35,3	36,77	51,07	46,03	46,05	61,83	60,95	49,60	72,23
K23ac	1084,6473	34,3	56,51	44,50	45,38	45,39	35,59	37,08	44,01	25,45
K18me	1112,6786	36,4	0,02	0,05	0,04	0,00	0,09	0,13	0,18	0,27
K18acK23ac	1070,6316	33,3	6,71	4,40	8,56	8,57	2,41	1,86	6,22	2,07
peptide 27-40										
unmod	1601,9122	35,3	62,82	31,08	30,19	20,84	8,08	13,03	4,12	14,77
K27me1/K36me1	1615,9278	36,1	21,29	35,66	29,78	22,86	11,44	12,00	5,63	16,58
K27me2	1615,9642	32,3	1,55	4,39	8,07	14,05	32,89	22,48	23,60	24,56
K36me2	1573,9173	32,7	11,09	8,94	9,20	8,10	6,16	11,13	0,00	5,02
K27me3	1587,8985	32,4	0,00	0,00	0,05	0,10	11,48	11,91	8,33	4,21
K36me3	1587,8985	32,4	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
K27ac	1587,8985	34,5	1,56	1,04	2,29	1,35	0,23	0,54	0,08	0,47
K27me1K36me1	1629,9435	36,8	0,64	11,55	9,17	9,62	3,90	4,17	1,57	4,23
K27me1K36me2	1587,9329	34,0	0,48	4,70	7,23	11,40	4,02	6,83	3,87	3,38
K27me2K36me1	1587,9329	33,0	0,50	2,62	3,86	10,01	14,11	11,36	15,29	27,27
K27me2K36me2	1545,9224	30,7	0,00	0,04	0,18	1,59	6,45	5,63	25,52	10,96
K27me3K36me1	1601,9486	36,9	0,00	0,01	0,00	0,01	0,01	0,01	0,00	0,00
K27me3K36me2	1589,9380	30,8	0,00	0,00	0,00	0,11	1,24	0,93	12,03	2,09
peptide 54-63										
unmod	1306,7365	42,2	98,39	99,83	99,64	99,62	99,81	99,61	96,90	97,19
K56ac	1292,7209	41,5	1,26	0,11	0,18	0,32	0,12	0,15	1,86	1,50
K56meS57p	1400,7185	44,4	0,00	0,03	0,04	0,03	0,00	0,04	0,00	0,00
K56me1	1320,7522	43,5	0,36	0,02	0,09	0,03	0,07	0,05	1,24	1,32
K56me2	1278,7416	41,4	0,00	0,00	0,03	0,00	0,00	0,07	0,00	0,00
K56me3	1292,7573	41,5	0,00	0,00	0,04	0,00	0,01	0,09	0,00	0,00
peptide 73-83										
unmod	1391,7185	42,2	99,67	98,04	97,78	94,16	93,87	95,50	88,85	88,93
K79me1	1405,7322	42,5	0,39	1,85	2,15	5,33	6,01	4,38	5,98	9,57
K79me2	1393,7216	41,3	0,95	0,12	0,05	0,51	0,12	0,10	5,19	1,39
K79me3	1377,73728	41,4	0,00	0,00	0,02	0,00	0,01	0,03	0,01	0,00

Supplement Table 2: Tandem Mass Spectrometric Quantitative Results of Histone Modifications of Histone H3

H4 peptides	M(cal) MH+	Av. RT	NF9 Mean	NF12 Mean	NF18 Mean	NF37 Mean	A6 Mean	XTC Mean	ES Cells Mean	MEFs Mean
4-17										
GKGGKGLGKGG	1494,8750	37,7	23,82	41,45	29,57	42,92	53,56	57,58	35,90	56,20
unmod	1480,8594	37,0	38,08	43,96	43,26	41,97	38,18	34,01	31,46	35,21
K16ac	1466,8437	36,3	27,19	0,00	0,00	0,00	0,00	0,00	0,00	0,00
K5acK12ac	1466,8437	36,3	0,00	11,12	16,77	10,95	6,63	6,79	20,06	5,93
K8acK16ac	1452,8281	35,5	9,01	2,81	6,99	3,03	1,31	1,31	0,00	1,39
K5/K12/K16ac	1452,8281	35,5	0,00	0,00	0,00	0,00	0,00	0,00	11,07	0,00
K8/K12/K16ac	1438,8123	34,7	1,91	0,66	3,42	1,15	0,33	0,32	3,32	0,78
24-35										
DNIQGITKPAIR	1381,7798	40,6	99,89	99,90	99,79	99,94	99,90	99,85	99,90	99,91
unmod	1367,7641	39,8	0,11	0,10	0,21	0,07	0,11	0,15	0,11	0,09
K31ac										
58-67										
GVLKVFLENVIR	1442,8730	52,3	99,87	99,89	99,46	99,54	99,85	99,85	99,66	99,25
unmod	1428,8573	52,0	0,10	0,04	0,11	0,06	0,07	0,13	0,08	0,07
K59ac	1456,8886	53,4	0,00	0,08	0,34	0,21	0,04	0,02	0,23	0,69
K59me1	1414,8781	51,8	0,04	0,00	0,10	0,06	0,00	0,00	0,00	0,00
K59me2	1428,8937	51,7	0,00	0,00	0,01	0,06	0,05	0,00	0,04	0,00
K59me3										
68-78										
DAVTTYEHAQR	1346,6699	34,7	99,89	99,98	99,94	99,92	99,94	99,62	99,95	99,99
unmod	1332,6542	33,6	0,12	0,03	0,06	0,09	0,06	0,38	0,06	0,02
K77ac										

Supplement Table 3: Tandem Mass Spectrometric Quantitative Results of Histone Modifications of Histone H4

H3 peptides	Micall) MH+	Av. DeltaM	AV. RT	z	NF9	NF12	NF18	NF37	A6	XTC	ES Cells	MEFa
Peptide 3-8 TKOTAR												
unmod	760.4312	0.00028	27.6	2	1	2	1	2	1	2	1	2
K4me	774.4468	0.00004	29.7	2	1	1	3			1		1
Peptide 64-89 KLPFOR												
unmod	844.5039	0.00152	41.1	2	3	3	3	1	4	1	2	4
Peptide 9-17 KSTGGKAPR												
unmod	1013.5738	0.00090	31.9	2	2	9	4	3	3	3	8	20
K9me1	1027.5894	0.00188	32.9	2	3	3	3	2	1	3	3	1
K9me2	985.5789	0.00015	27.9	2	1	1	3	4	3	6	6	3
K9me3	999.5945	-0.00029	27.8	2	1	1	3	4	3	3	3	1
K14ac	999.5581	0.00044	30.9	2	4	7	6	5	2	5	3	4
K9acK14ac	985.5425	0.00004	30.0	2	1	3	3	4	2	3	3	3
K9me2K14ac	971.5632	0.00011	27.7	2	4	3	11	1	6	1	8	9
K9me3K14ac	985.5789	0.00005	27.9	2	2	2	1	1	1	1	1	1
Peptide 18-26 KQLATKAAR												
unmod	1098.6629	0.00072	35.3	2	2	4	69	35	10	3	7	6
K23ac	1084.6473	0.00138	34.3	2	2	3	39	9	5	2	4	10
K18me	1112.6786	-0.00001	36.4	2	2	2	2	2	2	2	2	3
K18acK23ac	1070.6316	0.00159	33.3	2	2	2	2	2	2	2	3	1
peptide 27-40 KSAPATGKPKPHR												
unmod	1601.9122	0.00261	35.3	2/3	6	5	28	11		4	7	
K27me1/K36me1	1615.9278	0.00311	36.1	2/3	4	5	4	4		4		
K27me2	1573.9173	0.00289	32.3	2/3	5	3	5	3	8	4	7	7
K36me2	1573.9173	0.00570	32.7	2/3	3	5	3		2	1	2	2
K27me3	1587.8965	0.00331	32.4	2/3								
K36me3	1587.8965	0.10720	32.4	2/3								
K27ac	1587.8965	0.00481	34.5	2/3								
K27me1K36me1	1629.9435	0.00452	36.8	2/3								
K27me1K36me2	1587.9329	0.00311	34.0	2/3								
K27me2K36me1	1587.9329	0.00523	33.0	2/3	2	2	1	2	3	3	1	1
K27me2K36me2	1545.9224	0.00058	30.7	2/3					1	5	7	2
K27me3K36me1	1601.9486	0.00796	36.9	2/3					1	1	4	1
K27me3K36me2	1559.9380	0.00092	30.8	2/3								2
Peptide 41-49 YRPGTVALR												
unmod	1032.5849	0.00216	41.5	2	3	2	9	6	2	2	4	2
peptide 54-63 YOKSTELLIR												
unmod	1306.7365	0.00143	42.2	2	23	23	81	82	52	8	67	43
K56ac	1292.7209	0.00172	41.5	2	1	1						
K56meS57p	1400.7185	-0.00309	44.4	3	4	2						
K56me	1320.7522	0.00033	43.5	2						1		1
peptide 73-83 EIAQDFKTLR												
unmod	1391.7165	0.00151	42.2	2	5	4	86	76	44	16	25	46
K79me1	1405.7322	0.00281	42.5	2								
K79me2	1363.7216	-0.00012	41.3	2								
peptide 117-128 VTMPKDIQLAR												
unmod	1440.8243	0.00165	46.2	2	3	53	6	9	4	5	7	5
Total Peptides					57	60	436	285	169	41	151	152
											87	115
												78
												169

Supplement Table 4: Tandem Mass Spectrometric MS2 Events after Collision Induced Dissociation (CID) of Histone H3

H4 peptides	Mic(a)l	MH+	Av. Delta M	Av. RT	z	NF9	NF12	NF18	NF37	A6	XTC	ES Cells	METFs							
4-17	GKGGKGLGGAKR					1	2	1	2	1	2	1	2							
unmod	1494,8750	0,00330	37,7	2	6	7	54	7	6	3	9	16	18	32	13	9	15			
K12ac	1480,8594	0,00611	37,0	2			1							1						
K16ac	1480,8594	0,00393	37,0	2	4	3	31	5	6	4	6	7	11	16	6	3	10	2		
K5acK12ac	1466,8437	-0,00109	36,3	2	1	1														
K8acK12ac	1466,8437	0,00050	36,3	2			1													
K8acK16ac	1466,8437	0,00609	36,3	2			3		3	3	2	4	1	3	1	3	1	2		
K12acK16ac	1466,8437	-0,00103	36,3	2																
K5/K12/K16ac	1452,8281	0,00324	35,5	2	2	2	3	3	2	1	1	1	1	4	3	2	3			
K8/K12/K16ac	1452,8281	0,00378	35,5	2			1		1					2		1				
K5/K8/K16ac	1452,8281	0,00555	35,5	2					2											
K5/K8/K12/K16ac	1438,8123	0,00199	34,7	2	3	3	4	6	4	3	3	4				6	5	2	2	
24-35	DNIQGITKPAIR																			
unmod	1381,7798	0,00226	40,6	2	4	10	89	7	10	32	15	43	20	52	25	41	8	16	23	
K31ac	1367,7641	0,00179	39,8	2			1										2			
46-55	ISGLIYEETR																			
unmod	1180,6208	0,00154	51,2	2	35	43	68	17	36	56	42	47	25	49	67	64	32	11	11	23
56-67	GVLKVFLENVIR																			
unmod	1442,8730	0,00286	52,3	2	9	19	36		9	65	39	47	13	58	42	61	14	14	18	
K59ac	1428,8573	0,00603	52,0	2																
K59me	1456,8886	0,00531	53,4	2					2											
68-78	DAVTYTEHAQR																			
unmod	1346,6699	0,00275	34,7	2	6	8	53	12	9	42	8	12	14	25	10	14	13	10	12	
K77ac	1332,6542	0,00078	33,6	2					2											
79-92	KTVTAMDVYVALKR																			
unmod	1706,9509	0,00299	48,6	3	6	10	26	1	6	12	15	10	14	33	11	18	22	8	10	10
Total Peptides						76	107	368	60	92	226	142	192	118	270	177	220	182	82	117

Supplement Table 5: Tandem Mass Spectrometric MS2 Events after Collision Induced Dissociation (CID) of Histone H4

Supplement Table 1: Modifications of H2A, H2A.Z and H2B detected by MALDI-TOF

Indicated is the Peptide with its amino acid sequence and the different modifications and their masses (M(cal)H⁺) and the percentile value of the modification according to the LC quantification. Abr. NF9 = blastula stage, NF12 = gastrula stage, NF18 = neurula stage, NF37 = tadpole stage, A6 and XTC are stable *Xenopus laevis* Cell Lines, ES Cells and MEFs are derived from mouse. M(cal)H⁺ is the mass of the positively charged peptide, Av. RT is the average Retention time of the peptide on the C18 micro column of the reversed-phase LC-MS, mean of two independent biological replicates.

Supplement Table 2: Tandem Mass Spectrometric Quantitative Results of Histone Modifications of Histone H3

Indicated is the Peptide with its amino acid sequence and the different modifications and their masses (M(cal)H⁺) and the percentile value of the modification according to the LC quantification. Abr. NF9 = blastula stage, NF12 = gastrula stage, NF18 = neurula stage, NF37 = tadpole stage, A6 and XTC are stable *Xenopus laevis* Cell Lines, ES Cells and MEFs are derived from mouse. M(cal)H⁺ is the mass of the positively charged peptide, Av. RT is the average Retention time of the peptide on the C18 micro column of the reversed-phase LC-MS, mean of two independent biological replicates.

Supplement Table 3: Tandem Mass Spectrometric Quantitative Results of Histone Modifications of Histone H4

Indicated is the Peptide with its amino acid sequence and the different modifications and their masses (M(cal)H⁺) and the percentile value of the modification according to the LC quantification. Abr. NF9 = blastula stage, NF12 = gastrula stage, NF18 = neurula stage, NF37 = tadpole stage, A6 and XTC are stable *Xenopus laevis* Cell Lines, ES Cells and MEFs are derived from mouse. M(cal)H⁺ is the mass of the positively charged peptide, Av. RT is the average Retention time of the peptide on the C18 micro column of the reversed-phase LC-MS, mean of two independent biological replicates.

Supplement Table 4: Tandem Mass Spectrometric MS2 Events after Collision Induced Dissociation (CID) of Histone H3

Indicated is the Peptide with its amino acid sequence and the different modifications and their masses (M(cal)H⁺) and the number of MS2 events for each modification. Numbers highlighted in grey are not used for quantifications as their peptide probability score was lower than other modifications of the same modification state. Abr. NF9 = blastula stage, NF12 = gastrula stage, NF18 = neurula stage, NF37 = tadpole stage, A6 and XTC are stable *Xenopus laevis* Cell Lines, ES Cells and MEFs are derived from mouse. M(cal)H⁺ is the mass of the positively charged peptide, Av. DeltaM is the average mass deviation of the theoretical mass and the measured values. Av. RT is the average Retention time of the peptide on the C18 micro column of the reversed-phase LC-MS. z is the load of the most abundant peptide, mean of two independent biological replicates.

Supplement Table 5: Tandem Mass Spectrometric MS2 Events after Collision Induced Dissociation (CID) of Histone H4

Indicated is the Peptide with its amino acid sequence and the different modifications and their masses (M(cal)H⁺) and the number of MS2 events for each modification. Numbers highlighted in grey are not used for quantifications as their peptide probability score was lower than other modifications of the same modification state. Abr. NF9 = blastula stage, NF12 = gastrula stage, NF18 = neurula stage, NF37 = tadpole stage, A6 and XTC are stable *Xenopus laevis* Cell Lines, ES Cells and MEFs are derived from mouse. M(cal)H⁺ is the mass of the positively charged peptide, Av. DeltaM is the average mass deviation of the theoretical mass and the measured values. Av. RT is the average Retention time of the peptide on the C18 micro column of the reversed-phase LC-MS. z is the load of the most abundant peptide, mean of two independent biological replicates.

