Universitat Autònoma de Barcelona

Facultat de Biociències

Departament de Bioquímica i Biologia Molecular

LINKER HISTONE POST-TRANSLATIONAL MODIFICATIONS AND EFFECTS OF PHOSPHORYLATION ON SECONDARY STRUCTURE AND CHROMATIN AGGREGATION

Report presented by Rita Lopez Ramos in order to earn the
PhD degree in Biochemistry and Molecular Biology

Studies performed in the Biosciences Unit of the Department of Biochemistry and Molecular Biology of the Autonomous University of Barcelona under the supervision of
Dr. Inma Ponte Marull and Dr. Alicia Roque Córdova

Bellaterra, October 2013

Abstract

Linker histones play an important role in establishing and maintaining chromatin higher-order structure[145,146] and in gene regulation. H1 histones can act in transcriptional regulation through the modulation of chromatin condensation or as a part of protein complexes responsible of either the activation or repression of certain genes[147,148,149]. H1 also participates in transcriptional regulation through nucleosome positioning[150], stabilization of DNA at the entry/exit of the core particle[149] and preferential binding to SAR regions[151].

In mammals, nine subtypes (H1a-e, $H1^0$, H1x, H1oo and H1t) have been identified. Subtypes H1a-e are usually referred to as the *somatic subtypes* since they are present in most somatic cells[166,167]. H1t and H1oo are germ line specific, and $H1^0$ and H1x are related to cell differentiation. H5 is a specific aian subtype which is equivalent to $H1^0$. Linker histones share a highly conserved globular domain sequence while exhibiting variation in the N and C-terminal domains[163].

It is thought that H1 subtypes differ in their functions. Functional differentiation is based in their timing of expression[152], moment and extent of phosphorylation[153], turnover rate[154,155] and in their ability to condense chromatin *in vitro*[155,156,157,158]. Developmental and gene expression studies also support the concept that the different subtypes play distinct roles in chromatin structure[160,161,162].

Histone H1 in vertebrates has a characteristic three-domain structure consisting of a short, flexible N-terminal domain (20-30 amino acids), a central globular domain (80 amino acids) and a long C-terminal domain (~100 amino acids)[213]. The amino- and carboxyl-terminal domains are highly basic and very lysine-rich, specially the C-terminal domain (CTD) which contains up to 40% of lysine residues[213]. Despite the sequence variability of the C-terminal domain among different histone subtypes, the charge distribution is quite uniform in all cases[224]. This domain binds to linker DNA, neutralizing the negative charge of phosphates and facilitating chromatin condensation into the 30 nm fibre and also the intermolecular aggregation[225,226,227]. Chromatin condensation is mediated through charge-neutralization of the negatively charged linker DNA.

Both terminal domains are mainly unstructured in aqueous solution[235,236,237]. Interaction with DNA induces the complete folding of the CTD under physiological conditions in a very stable manner[235]. The fact that the CTD is unstructured in aqueous solution but becomes completely folded when bound to DNA allowed including this domain in the group of intrinsically disordered proteins with coupled binding and folding.

Histone H1 is phosphorylated by CDK2 in a cell cycle-dependent manner: hyperphosphorylation is related to condensed chromatin while partial phosphorylation is associated with relaxed chromatin[249,254,266]. Linker histones are phosphorylated in the S/T-P-X-R/K motifs, mostly located in the CTD and involved in many structural and functional aspects of H1[232]. Previous studies had demonstrated that phosphorylation of the CTD of H1 has effects on secondary structure and DNA condensation and aggregation capacity[266].

In the present study, infrared spectroscopic analyses of the entire histone $H1^0$ were performed to determine the secondary structure of the protein in aqueous solution. In buffer solution, the amount of β-structure was significantly higher than the estimated from the sum of the β-structure content from the isolated C-terminal and globular domains, which seemed to indicate that the presence of the globular domain stabilizes the β-structure of the CTD in the entire protein. Furthermore, phosphorylation of the entire $H1^0$ induced an increase of the percentage of β-structure of histone $H1^0$ in aqueous solution (35%) with respect to the unphosphorylated protein free in solution (21%), indicating that phosphorylation led to a structural change in the protein that favoured the induction of β-structure in histone $H1^0$. This fact will be determined by a conformational change in the C-terminal domain, since all three phosphorylation sites of $H1^0$ are located in this domain.

The isolated CTD becomes completely folded when bound to DNA[235]. In this study, FTIR analysis of the entire histone $H1^0$ bound to DNA were performed at two different protein/DNA ratios, r(w/w), r=0.5 and r=0.7 with the aim of determining if this change also occurred in the entire $H1^0$. The binding of unphosphorylated histone $H1^0$ to DNA induced an increase in the proportion of α-helix from a 23% in the protein free in solution to a 34% when bound to DNA, which will be a consequence of the folding of the terminal domains upon DNA interaction[218,235], indicating that DNA binding to the unphosphorylated $H1^0$ induced its folding in α-helix conformation, as it happened with the isolated CTD. No effect of protein/DNA ratio was detected in this case.

Phosphorylated $H1^0$ bound to DNA was also studied by FTIR spectroscopy at physiological salt concentration. Under these conditions, the amount of β-structure increased considerably in contrast with the results obtained for $H1^00P$ bound to DNA. In this case, there was an effect of protein/DNA ratio, and the proportion of β-structure increased up to a 54% when r(w/w)=0.7. This structural change supposed a decrease of random coil percentage. This fact suggested that DNA binding increases the folding beyond phosphorylation in the protein in aqueous solution, which implies an important gain of β-structure, as it happened in the isolated CTD.

Folding of histone H1 may be determined by an increase of the hydrophobicity followed by charge compensation of the Lys positive charge in the CTD by the DNA phosphates. Because of that, the interaction with hydrophobic ligands could also trigger the folding of $H1^0$ as efficiently as DNA. In order to explore this idea, FTIR analyses of unphosphorylated and phosphorylated histone $H1^0$ in the presence of anionic detergent were performed. FTIR analysis of $H1^0$ in the presence of SDS demonstrated that the effect of this detergent mimicked the interaction with DNA. The amounts of secondary structure elements in both unphosphorylated and phosphorylated $H1^0$ in the presence of SDS at a molar ratio of 14:1 were nearly the same that the ones found in the complexes with DNA. In these conditions, the triphosphorylated protein had a 55% of β-structure, indicating that the CTD within histone $H1^0$ was also in an all-β conformation and formed amyloid fibres.

Mature chicken erythrocyte nuclei contain highly condensed and inert chromatin, mainly consisting of DNA and histone proteins, which is characteristic of this type of differentiated cells[354]. Six different subtypes of histone H1 (H1.01, H1.02, H1.03, H1.10, H1.1L and H1.1R) plus the avian specific subtype, histone H5[354,441], are present in chicken erythrocyte nuclei. Histone H5 replaces histone H1 in mature erythrocytes where the 30% of nucleosomes contain two linker histones[427].

In this study, chicken erythrocyte nuclei were treated with micrococcal nuclease and native chromatin was fractionated into soluble and insoluble fractions. Linker histones were analysed in order to observe their relative proportions in each chromatin fraction and to determine their post-translational modifications.

Transcriptional activity and mitosis are almost totally repressed in avian erythrocytes[1]. Because of that, avian linker histones are thought to be poorly phosphorylated. Separation of linker histones by HPCE allowed the identification of secondary peaks in the profiles of both soluble and insoluble chicken erythrocyte chromatin due to basal phosphorylation of histones H5 and H1 subtypes. No significant differences were observed between the two HPCE profiles. However, the relative amount of histone H1 subtypes with respect to histone H5 was higher in the insoluble chromatin fraction. This fact was confirmed by the calculation of the H1s/H5 ratio in a SDS electrophoresis. In the insoluble chromatin fraction, the H1s/H5 ratio was ~0.8 whereas it was ~0.3 in the soluble fraction.

Linker histones H5 and H1s were further separated and analysed by Tandem MS in order to map the post-translational modifications of all linker histones subtypes in each chromatin fraction.

Previous studies on post-translational modifications on chicken linker histones were carried out on bulk chromatin[432], considering that linker histones PTMs might be uniform along the chromatin molecule. In contrast to that, we fractionated native chromatin from chicken erythrocytes into soluble and insoluble fractions.

Several post-translational modifications in linker histones extracted from those two fractions were identified by ESI/LC-MS/MS, including N^{α}-acetylation, lysine acetylation and methylation and phosphorylation. From those, eight novel PTMs including lysine acetylation (H5K14ac, H1K34ac, H1K46ac, H1.02K17ac) and serine/threonine phosphorylation (H5T84phos, H1.02T3phos, H1S39phos, H1S104phos) were identified in chicken erythrocyte linker histones H5 (two novel PTMs) and H1 (six novel PTMs) in the Tandem MS analyses. One modification was determined to be specific of the soluble chromatin fraction, whereas five modifications were only identified in the insoluble fraction. These results indicated the existence of a different pattern of post-translational modifications between those two fractions.

Furthermore, comparison of our results with other recent mass spectrometry studies on linker histones PTMs in human, mouse, rat and chicken H1 variants[242,388,432] has allowed to identify some modifications in the N-terminal and globular domain that are conserved among different species, highlighting their importance in specific linker histone functions and epigenetics.

The effects of phosphorylation on chromatin structure may depend on the number and the position of the phosphate groups incorporated in H1; and these effects might be mediated by specific structural changes and not be a simple effect of the reduction of the net charge associated to phosphorylation[266]. For that reason, we decided to study the effects of *ex vivo* phosphorylation of histone H1 on chromatin aggregation.

In this study, native chromatin from chicken erythrocyte nuclei was purified and phosphorylated *ex vivo* with CDK2 kinase, which specifically phosphorylates the SPKK motifs present in linker histones. The aim was to study the changes in chromatin aggregation due to *ex vivo* phosphorylation.

The specificity of *ex vivo* phosphorylation of linker histones within chromatin was corroborated by radioactive CDK2-phosphorylation assay. Then, our purpose was to identify the different histone subtypes and to determine the level of phosphorylation and the position of the phosphate groups after *ex vivo* phosphorylation by different proteomic techniques.

HPCE analyses were performed for initial and qualitative identification of the phosphorylated species after *ex vivo* phosphorylation of chromatin. In the HPCE profiles, a progressive increase in the number of secondary peaks was observed with the increase of phosphorylation time.

MALDITOF-MS analyses allowed to observe a progressive increase of the phosphorylation levels with the increase of phosphorylation time, in both histones H5 and H1s. Mono-, di- and even tri-phosphorylated species could be identified. The results suggested that the increase in phosphorylation first occurred in the amount of monophosphorylated species and, then, lead to the appearance of diphosphorylated species. Calculation of the percentages of phosphorylated species of H5 CTD from the MALDITOF-MS analyses showed that, after overnight phosphorylation, the amount of phosphorylated species had increased up to 54%. However, this phosphorylation was partial, since only around 14% of potentially phosphorylated CDK2 sites of histone H5 were phosphorylated.

Tandem MS after proteolytic digestion revealed that, in all linker histones, the SPKK motifs were unphosphorylated in native chromatin, indicating that the phosphorylated peptides found at other times of reaction had been modified by *ex vivo* phosphorylation with CDK2. In H5, only S148 was identified in all the samples and was found phosphorylated after 1 hour. In the Tandem MS analysis of

histone H1 subtypes, all the CDK2 consensus sequences, with the exception of S171 (H1.01 numbering), were identified for H1.03, H1.1L and H1.1R. H1.03T16 was found phos`phorylated after 15 minutes; H1.1LS192 and H1.1RS186, after 1 hour; and H1.03S155, H1.1LS155 and H1.1RS153, after 3 hours.

We also studied the effect of *ex vivo* phosphorylation on chromatin aggregation. The addition of increasing milimolar concentrations of $MgCl_2$ leads to hierarchical condensation transitions of chromatin arrays[2]. Chromatin is self-repulsive in low ionic strength buffers, but that charge repulsion is reduced or eliminated in the presence of divalent and polyvalent cations[3]. The presence of Mg^{2+} affects the intramolecular folding of chromatin (condensation) and, increasing the amount of $MgCl_2$ above ~1.5mM concentration leads to intermolecular chromatin self-association (aggregation). In the present study, we analysed the induction of chromatin aggregation mediated by an increase in $MgCl_2$ concentration up to 1.6mM $MgCl_2$ of *ex vivo* phosphorylated chromatin by Dynamic Light Scattering (DLS).

Our analyses showed that the increase of the $MgCl_2$ concentration up to 1.6mM led to the appearance of a second peak of aggregated molecules in the DLS measurements, and that this aggregation not only depended on the presence of divalent cations (Mg^{2+}) in the medium but also of the temperature at which the previous incubation/phosphorylation reaction has been carried out.

DLS measurements of unphosphorylated and phosphorylated chromatin fragments showed that phosphorylation had a clear effect on chromatin aggregation of the samples. The most remarkable differences associated to *ex vivo* phosphorylation of linker histones within chromatin was a decrease in the hydrodynamic diameter of the aggregation peak measured by DLS at 1.6mM $MgCl_2$. The differences became greater with the increase of phosphorylation time, which conduced to the conclusion that, at 1.6mM $MgCl_2$, *ex vivo* phosphorylation of linker histones impaired chromatin aggregation.

Our results also showed that the size of the chromatin fragments had also an effect in the impairment of chromatin aggregation at 1.6mM $MgCl_2$ following *ex vivo* phosphorylation of chromatin. The difference in the size of the hydrodynamic diameter between the unphosphorylated and the phosphorylated samples after overnight incubation/phosphorylation was of ~69nm in the aggregation experiments of short chromatin fragments; whereas in larger chromatin fragments this difference increased up to ~654nm.

These results indicated tha histone H1 phosphorylation on its SPKK motifs has also a major influence on chromatin aggregation and compaction. In our case, partial phosphorylated chicken linker histones impaired chromatin aggregation induced by $MgCl_2$.

Abbreviations

ADP: adenosine di-phosphate

ATP: adenosine tri-phosphate

AU-PAGE: acetic acid/urea poliacrilamide gel electrophoresis

AUT-PAGE: AU-PAGE with added non-ionic detergent Triton X-100

BD: bromodomain

BPTF: Bromodomain PHD finger transcription factor

BRCA1: breast cancer DNA-repair protein

BRCT: C-terminal domain of BRCA1

CAP: catabolite gene activator protein

CD: circular dicroism

CDK: Cyclin-dependent kinase

CE: Capillary Electrophoresis

$CH1^0$/$cH1^0$/$H1^0$-CTD : C-terminal domain of histone $H1^0$

$CH1^0$0P/$cH1^0$0P/ $CH1^0$0p/$cH1^0$0p: unphosphorylated C-terminal domain of histone $H1^0$

$CH1^0$3P/$cH1^0$3P/ $CH1^0$3p/$cH1^0$3p: triphosphorylated C-terminal domain of histone $H1^0$

CHCA: *alpha*-cyano-4-hydroxycinnamic acid

ChIP on CHIP: Chromatin Immunoprecipitation on DNA chip

CHT-II: ceramic hydroxyapatite type II

CTD/Cter: C-terminal domain

DNA: deoxyribonucleic acid

DNA-PK: DNA-dependent protein kinase

DLS: Dynamic Light Scattering

ECD: electron-capture dissociation

EDTA: Ethylenediaminetetraacetic acid

EGTA: Ethylene glycol tetraacetic acid

ES cell: embryonic stem cell

ESI/ESI-MS: Electrospray Ionization Mass Spectrometry

ETD: electron-transfer dissociation

FRET: fluorescent resonance energy transfer

FT-ICR: Fourier transformed ion cyclotron resonance

FT-IR/ FTIR: Fourier transformed infrared spectroscopy

GD: globular domain

$GH1^0$/GH5: globular domain of histone $H1^0$/histone H5

$H1^0$0P/$H1^0$0p: unphosphorylated histone $H1^0$

$H1^0$3P/$H1^0$3p: triphosphorylated histone $H1^0$

HAT: histone acetyltransferase

HEPES: 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid

HDAC: Histone deacetylase

HILIC: Hydrophilic-Interaction Liquid Chromatography

HMT: Histone methyltransferase

HNF3γ: Hepatocite Nuclear Factor-3γ

HP1: Heterochromatin Protein 1

HPCE: High Performance Capillary Electrophoresis

HPLC: High Performance Liquid Chromatography

HPMC: Hydroxypropylmethylcellullose

HSH: helix-strand-helix

HTH: helix-turn-helix

IMAC: immobilized metal-ion affinity chromatography

ING2: Inhibitor of growth protein 2

IPTG: isopropyl-β-D-thiogalactoside

KDM: Histone lysine demethylase

KMT: Histone lysine methyltransferase

LC: Liquid Chromatography

LC-MS or LC-MS/MS: Liquid Chromatography coupled to tandem mass spectrometry

LTQ: quadrupole ion trap (mass spectrometer)

LX: DNA Ligase complexes

$NH1^0$: N-terminal domain of histone $H1^0$

Nter: N-terminal domain

NMR: Nuclear Magnetic Resonance

NMR-2D: two-dimension Nuclear Magnetic Resonance

NRL: Nucleosome-Repeat Length

NuRC: Nucleosome Remodeling Complex

MALDI/MALDI-MS: Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry

MALDITOF/MALDITOF-MS: Matrix-assisted laser-desorption ionization coupled with time-of-flight analyzer mass-spectrometry

MEL: Murine erythroleukemia

MMTV: mouse mammary tumour virus

MS: Mass Spectrometry

MS/MS: Tandem Mass Spectrometry

m/z: mass-to-charge ratio

on or o/n: overnight

PAD: Peptidylarginine deiminases

PARP: Poly-ADP-Ribosyltransferase

PCR: Polymerase Chain Reaction

PHD finger: Plant Homeo Domain finger

PEG: Polyethylene glycol

PELDOR: Electron paramagnetic resonance spin-labelling technique

PMF: Peptide Mass Fingerprint

PMSF: Phenylmethylsulfonyl fluoride

PMM: Peptide Mass Mapping

PRC1: polycomb-group repressive complex 1

PTM: post-translational modification

RP-HPLC/RPLC: Reversed Phase High Performance Liquid Chromatography

RT: Room Temperature

SAR: Scaffold-associated regions

SDS: Sodium Dodecyl Sulphate

SIC: Selected Ion Current

SUMO: Small Ubiquitin-like Modifier

TBE: Tris/Borate/EDTA

TCA: Trichloroacetic acid

TE: Tris/EDTA

TEA: Triethanolamine

TEM: Transmission Electron Microscopy

TEMED: N,N,N´,N´-tetramethylethylenediamine

TFE: Tetrafluoroethylene

ThT: Thioflavine T

TIC: Total Ion Current

TOF: time-of-flight

UV: ultraviolet

Introduction

An overview of chromatin structure

Chromosomes, from the Greek terms χρώμα, -τος *chroma*, colour, and σώμα, -τος *soma*, body; are the largest and most visible physical structures involved in the transfer of genetic information[4]**;** the essential unit for cellular division that must be replicated, divided, and passed successfully to their daughter cells to ensure genetic diversity and the survival of their progeny. Despite chromosomes may vary widely between different organisms, they are made up of a single DNA molecule, circular or linear, composed of 100,000 to 10,000,000,000 nucleotides[5]. In eukaryotes, nuclear chromosomes are packaged by proteins into a high-order structure called chromatin.

Chromatin is, according to the biochemists' operational definition, the DNA-protein-RNA complex extracted from eukaryotic lysed interphase nuclei. The exact composition of the extracted material will depend in part on the technique each researcher uses. Furthermore, the composition and properties of chromatin vary from one cell type to another, during development of a specific cell type, and at different stages in the cell cycle[6].

Close to two meters of DNA in each cell must be assembled into a small nucleus of a few microns in diameter and this fact is possible due to the packing of the genetic material into chromatin (Figure 1). This supramolecular structure is formed by the condensation of DNA-protein complexes up to levels of 5 orders of magnitude. Furthermore, despite its high degree of compaction, DNA must be rapidly accessible to protein complexes that regulate the different chromatin functions: replication, repair, recombination and gene expression. The dynamic organization of chromatin structure influences all functions of the genome. Chromatin is a highly dynamic structure from which nuclear proteins are continuously associating and dissociating. The accessibility of these proteins decreases progressively along the DNA length from the periphery to the centre of the nucleosome[7]. Therefore, the transcription of a certain gene depends on a dynamic sequence of events, each with a certain probability[8].

The nucleosome is the fundamental unit of chromatin. It is composed of histones and DNA, which are associated non-covalently in a precise stoichiometric ratio. Nucleosome organization in chromatin was first determined by partial digestion with micrococcal nuclease, which produced a series of discrete DNA fragments decreasing in length that showed a periodicity of about 200bp in gel electrophoresis[9].

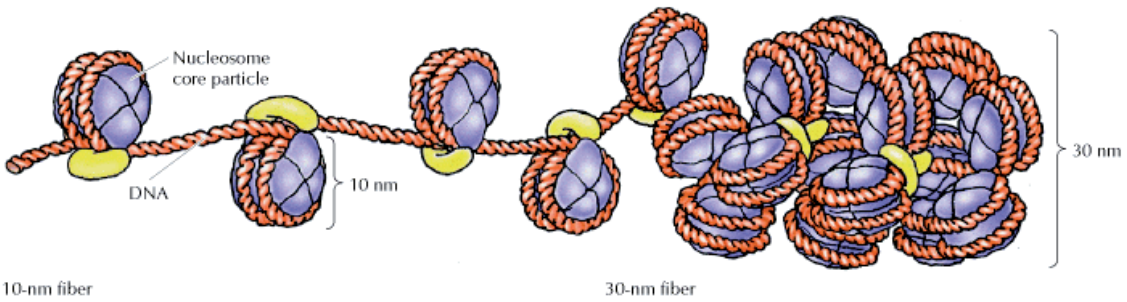


Figure 1. Molecular assembly of nucleosomes. The DNA (red) is wrapped around the histone octamer (blue) and both form the nucleosome core particle. This structure is locked by the linker histone H1 (yellow). The 10nm-chromatin fiber is further folded into a thicker fiber, the 30nm-solenoid.

The nucleosome is a nucleoprotein particle in which the DNA molecule is wrapped around an octameric core of histone proteins in about two left-handed superhelical turns and fastened externally by a single molecule of linker histone (histone H1). Extensive digestion with micrococcal nuclease led to the generation of the *core particle*, composed by the octamer, which contains two copies of each core histone $(H2A\text{-}H2B)_2$-$[(H3)_2$-$(H4)_2]$ and a DNA length of 147bp, that has the strongest contacts with the histone core. The DNA lost in this digestion, located in between consecutive nucleosome core particles, is called linker DNA (20-60bp). The overall architecture of the histone octamer, as well as all of the residues that are involved in direct protein– DNA interactions is highly conserved, and even a change up to 100 residues between frog (*X.laevis*) and yeast (*S.cerevisiae*) histones results in electron density maps with an overall root mean square deviation of only 1.6Å[10].

Nucleosome repeats with intervening linker DNA form the first level of DNA condensation, the 10nm-cromatin fibre, also well-known as the 'beads-on-a-string' configuration, due to its appearance in electron microscopy under low-salt conditions[11,12].

When H1 class histones bind to the internucleosomal DNA, the 10nm-chromatin fibre is packed further into the 30nm-fibre, which corresponds to the state of the great majority of chromatin during interphase. The precise structure of the chromatin fibre in the cell is not known in detail and there is still some debate about it[13].

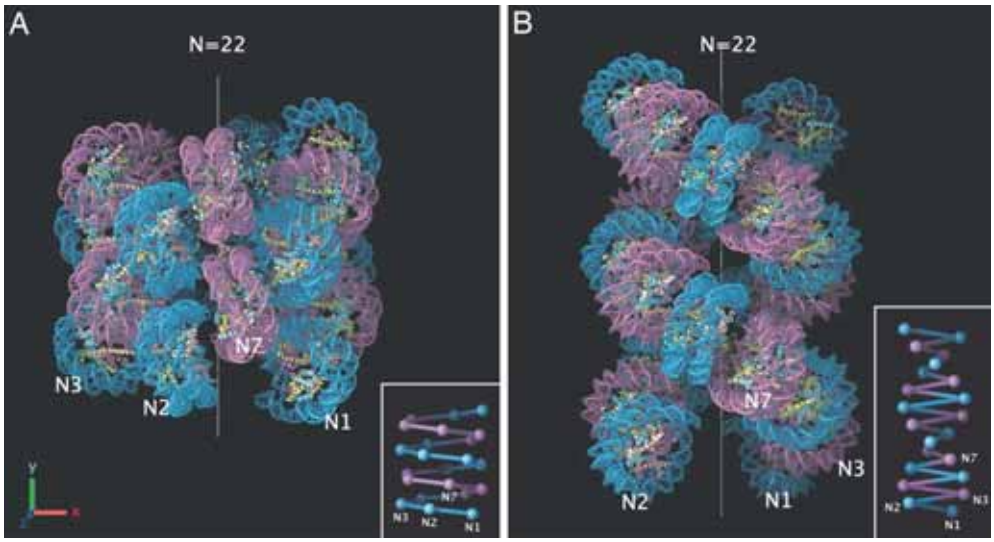


Figure 2. Chromatin organisation models. (A) One-start helix. (B) Two-start helix. From Robinson PJJ, *et. al.* (2006)[14]

There are several models that attempt to describe chromatin structure. These models can be classified into two classes[14]: the solenoid class in which a one-start helical stack of nucleosomes is linked together by bent DNA between adjacent octamers11,[15,16] and the two-start zigzag class, which states that two helical stacks of nucleosomes are connected by a relatively straight DNA linker (Figure 2). These models can further divide into the helical/twisted-ribbon model in which the linker DNA is oriented at variable angles along the length of the fibre[17,18]; and the crossed-linker model in which the linker DNA is oriented approximately perpendicular to the fibre axis[19,20]. A much more recent model is the one of the interdigitated solenoid[21], built up from experimental data of highly condensed chromatin fibres and that could represent a higher level of condensation than the ones described above.

Nevertheless, the existing models commonly accept that all experimental observations reflect different features of the same structure and agree in three important facts: the nucleosomes are located peripherally and lie perpendicular to the axis of the fibre, forming a zigzag chain in low ionic strength; the linker histone is arranged internally in the fibre and plays an essential role in chromatin condensation and nucleosome-nucleosome interactions either enable chromatin condensation or contribute to the stabilisation of the condensed chromatin fibre (Figure 3). It is important to point out that the limiting factor in chromatin condensation is the level of charge shielding along the phosphodiester backbone of DNA.

The 30 nm fibre is arranged into large loops of 30 to 100 Kb along a central protein scaffold to form transcriptionally active euchromatin. These loops would correspond to transcriptional dynamic domains, of about 30 to 100Kb[22]. Further compaction would lead to transcriptionally inactive heterochromatin.

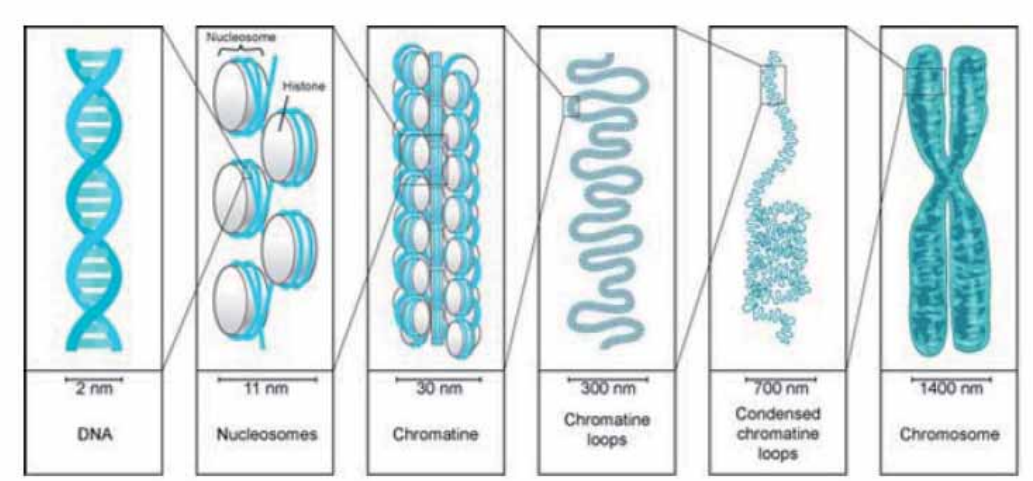


Figure 3. Chromatin compaction scheme

New experimental approaches, including chromatin conformation capture and cryo-electron microscopy, challenge the *in situ* evidence and universality of 30nm chromatin fibres as building blocks for chromosome folding *in vivo*, suggesting that the 10nm chromatin fibre is sufficient to describe the complexities of nuclear organization and gene regulation[23,24].

DNA structure in the nucleosome

First X-ray crystallographic studies of core particles at 7Å resolution demonstrated that DNA is located on the surface of the nucleosome, which has a disc-like shape 11nm in diameter and 5.6nm in height. The DNA is wrapped in 1.75 turns of left-handed super helix around the histone core[25]. Later crystallization studies of the nucleosome core particle up to 2.8Å[26] and 1.9Å[44] resolution have permitted a deeper knowledge of nucleosome morphology (Figure 4).

The structure of the nucleosome has a pseudodyad axis of symmetry that passes through the centre of the nucleosomal DNA. However, the winding of DNA around core histones is not uniform and the crystal structure suggests that the bending of DNA is sharper and preferentially distorted at ±1.5 helical turns from the dyad axis (the centre of the nucleosome)[27,28]. The helical periodicity changes when DNA is

wrapped around the octamer of histones from 10.5bp/turn in solution into an average of 10.2bp/turn, suggesting that it might be overwound[29]. This overwinding is not uniform over the whole DNA molecule in the core particle but results from distinct local regions with different local helical periodicities[30]. The central three turns of DNA in the core particle have 10.7bp/turn, in contrast with the rest of the nucleosomal DNA (10.0bp/turn). Crystallographic analysis of the octamer suggests that DNA at the dyad axis with 10.7bp/turn would make the most effective contacts with the histones, whereas out of this region, a helical periodicity of 10.0bp/turn would be optimal[31]. DNA at the dyad axis is relatively straight and it is mainly associated with histones H3 and H4, which create a central surface within their arginine residues remain exposed and spaced[32,33]. Differential helical periodicity in nucleosomal DNA leads to significant consequences in the association of other proteins that interact with DNA through sequence specificity.

Nucleosomal positioning

The local influence of DNA bending and stiffness has an effect in the translational and rotational positioning of the double-helix respect to the histone core. This phenomenon is known as *nucleosome positioning* and has significant biological consequences[34]. The translational position of the nucleosome refers to where the octamer begins and finishes being associated with DNA. The rotational position refers to which face of the double helix is in contact with or exposed away from the histone core4. Nucleosome positioning has been mostly studied in regulatory DNA.

The positioning of histone octamer appears to be primarily determined by the wrapping of DNA around the $(H3\text{-}H4)_2$ tetramer. However, linker histones can also influence the translational position of core histones with respect to DNA sequence[35,36,37].

Sequence-directed nucleosome positioning in DNA is accepted[38,39]. Nucleosome positioning sequences are DNA sequences that bias their own packing into nucleosomes by preferential positioning of the histone octamer. The sites of preferential positioning will be those having minimum free energy[40]. Thus, a particular DNA sequence that happens to be bent around the histone core to an appropriate degree for nucleosome packaging will have a higher affinity for the histone octamer and, thus, an increased nucleosome positioning power, since less free energy will be necessary to wrap the DNA on the histone surface[35,38,39,40].

Several studies have been performed to look for or create nucleosome positioning sequences, by both, *in vivo* and *in vitro* approaches. A "consensus DNA sequence" for nucleosome positioning has not been reported, but certain DNA sequence preferences or motifs for nucleosome positioning have been described and had been widely used in chromatin reconstitution experiments[40,41,42,43].

Core histones-DNA interactions

The nucleosome contains over 120 direct protein-DNA interactions and several hundred of water-mediated ones[44]. Direct protein - DNA interactions are not spread evenly over the octamer surface but rather located at discrete sites.

Although nucleosomes tend to prefer some DNA sequences over others[45], they are capable of binding to practically any sequence, which is thought to be due to the flexibility in the formation of these water-mediated interactions. The bulk of interactions between histones and DNA are salt bridges and hydrogen bonds established between side-chain groups (amine/hydroxyl) or main-chain amides and the DNA backbone phosphates. These interactions are non-sequence specific, so they play an important role in the limited-sequence dependence of nucleosome positioning. In addition, there are non-polar

interactions between protein side-chains and the deoxyribose groups, and an insertion of arginine side-chains that intercalate into the DNA minor groove at all 14 sites where the arginines face the octamer surface.

Electron paramagnetic resonance spin-labelling technique (PELDOR) has been used to measure the distances between the spools around which eukaryotic cells wind their DNA. They determined the spacing range from 59 to 70Å[46]. In all, histones make five types of interactions with DNA:

1. Helix-dipoles from alpha-helices in H2B, H3, and H4 cause a net positive charge to accumulate at the point of interaction with negatively charged phosphate groups on DNA;
2. Hydrogen bonds between the DNA backbone and the amide group on the main chain of histone proteins;
3. Nonpolar interactions between the histone and deoxyribose sugars on DNA;
4. Salt bridges and hydrogen bonds between side chains of basic amino acids (especially lysine and arginine) and phosphate oxygen's of the DNA;
5. Non-specific minor groove insertions of the H3 and H2B N-terminal tails into two minor grooves each on the DNA molecule.

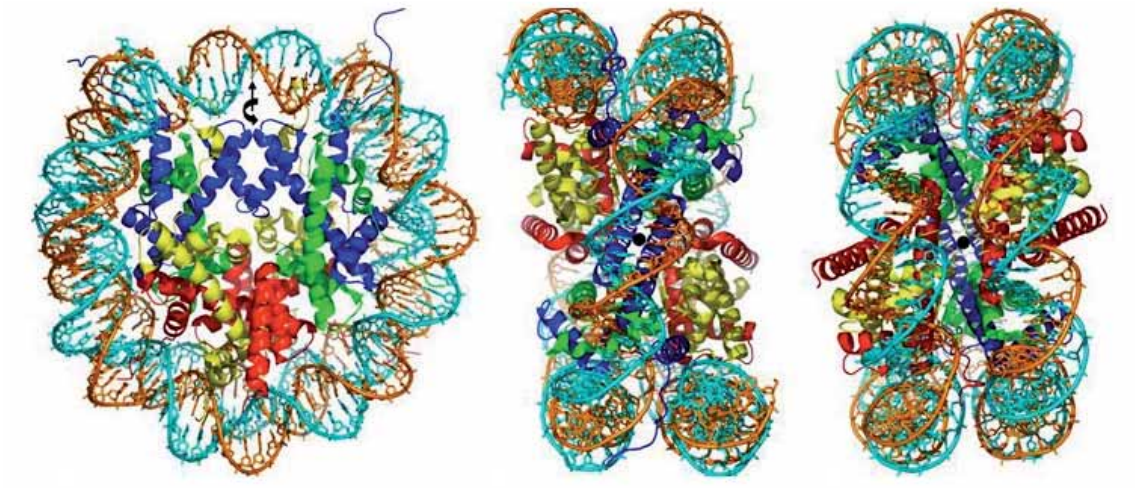


Figure 4. Nucleosome core particle. Side (A), top (B), and bottom (C) views of the 1.9 Å resolution crystal structure of the nucleosome core particle. The two strands of the 147 base pair DNA and the H3 (blue), H4 (green), H2A (yellow), and H2B (red) core. Histone proteins are coloured individually. The view in A is along the DNA super helical axis, with the pseudo two-fold axis of the particle running vertically through the central DNA base pair (straight arrow). In B and C, the view is along the pseudo two-fold axis, shown as a black dot. Taken from Davey *et al.* (2002)[44].

Histone proteins

Histone proteins are easily distinguished from other proteins by four characteristic features: (1) high basicity due to a large amount of basic residues; (2) domain structure in which a part of the molecule

adopts a globular conformation whilst the rest remains unstructured in solution; (3) strong evolutionary conservation; and (4) the presence of non-allelic variants of each histone, except for H4.

Histones are the main effectors and controllers of DNA organisation into chromatin and they either participate in protein–protein interactions among themselves or with other proteins. Five major families of histones exist: H1/H5, H2A, H2B, H3, and H4. Histones H2A, H2B, H3 and H4 are known as core histones, while histones H1 and H5 are known as linker histones[47,48].

Core histones

Core histones have low molecular weight (11-16 kDa) and contain more than a 20% of basic amino acids. H2A and H2B have a larger proportion of lysine while H4 and H3 are richer in arginine[49].

All core histones are highly conserved in length and amino acid sequence through evolution. Histones H3 and H4 are the most preserved: only 8 amino acid differences are found between human and yeast H4 and 19 differences considering *Tetrahymena* and human H4. Most remarkable is the fact that the number of residues in H3 and H4 has been almost exactly conserved in all eukaryotes (~120aa for H4; 135aa for H3). Since H3 and H4 play a central role in the nucleosome and many other chromosomal processes, those requirements will account for their remarkable conservation. In contrast, H2A and H2B are slightly less conserved. H2A histones exhibit their greatest length variability at the C-terminus domain whereas H2B histones diverge more at the N-terminus. This pattern is related to the way they are positioned in the nucleosome and how they bind to DNA[50, 51,] 4.

Histone folding and octamer assembly

The core histone octamer is a tripartite assembly in which a central $(H3\text{-}H4)_2$ tetramer is flanked by a H2A-H2B dimer on either side[52,53,54]. The overall shape is described as a cylindrical wedge with an outer diameter of 6.5nm and a length of 6nm at the wide side of the wedge and ~1nm at the tip. The 3.1Å resolution model[32] accounts for the 70% of the entire mass and comprises the structured parts of the eight core histone chains. The 'missing' mass is attributed to the unstructured parts found in random configuration.

Within the octamer, the four histone dimers are assembled by two distinguishable sets of protein-protein non-covalent interactions based in a hierarchy in the strength of the pairwise histone-histone interactions (H3-H4>H2A-H2B=H2B-H4>H2A-H4)[55.] The first set involves mostly hydrophobic interactions and is responsible for the internal stabilization of the H2A-H2B dimer and the $(H3\text{-}H4)_2$ tetramer. The second set involves the weak association of one $(H3\text{-}H4)_2$ tetramer with two H2A-H2B dimers to form an octamer. The binding of the first dimer to one side of the tetramer enhances four times the intrinsic binding affinity of the second dimer binding site of the tetramer[56]. Thus, binding of H2A-H2B dimers is cooperative4.

The four histone dimers are organized into a left-handed protein supercoil and, in each dimer, each histone chain associates with its partner to form an ellipsoid with one smoothly curving outer face, which becomes the 'footprint' of that dimer on the cylindrical surface of the octamer4.

Although their low sequence homology, the core histones share a common tertiary structural motif known as the 'histone fold'[57] (Figure 5). All four histones contain a globular domain in the C-terminus of the protein which is involved in histone-histone and histone-DNA interactions, and positively charged N-terminus tails containing the bulk of the basic residues. The histone-fold consists of three helices, where the long middle helix is flanked at each end by shorter ones and connected to them by a loop followed by a short β-strand segment[32]. Thus, the histone fold has an internal tandem parallel duplication at

about the middle of the central long helix (helix II), and this unit of the histone fold is called the helix-strand-helix (HSH) motif that leads to several sets of symmetry axes within the octamer[31].

Furthermore, the long central helix acts as an area for dimerization. This domain is preserved through evolution due to the structural restrictions for the nucleosome assembly. Likewise, the sequence of the N-terminal tails is also conserved. These charged tails are the site of several post-translational modifications, which have important consequences for chromatin structure and function, especially during development[58].

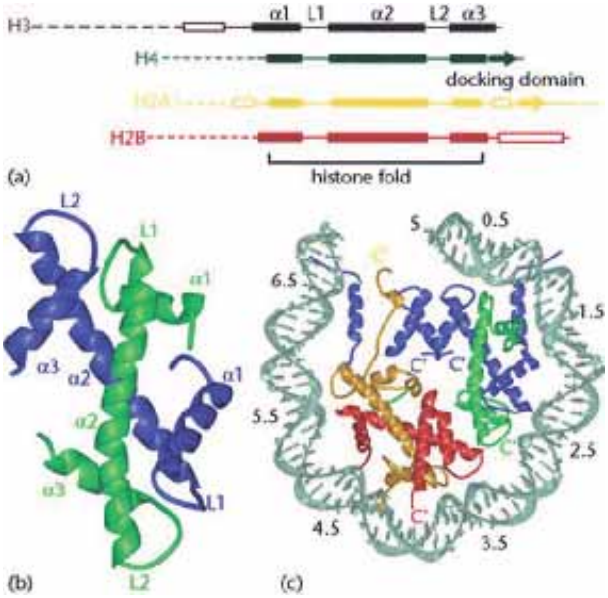


Figure 5. Architecture of the histone fold and of the nucleosome core particle. (a) Schematic overview of the four histone proteins H3, H4, H2A and H2B. α-helices of the histone fold are shown as solid boxes, α-helices and β-strands of the histone fold extensions are shown as open boxes and arrows, respectively. Histone tails are shown as dotted lines. The length of α helices, β strands, loops and tails is shown to scale. The components of the histone fold (α1, L1, α2, L2, α3), and the location of the H2A docking domain are indicated. **(b)** A histone fold dimer (here shown for H3 and H4, shown in blue and green, respectively) is formed by the antiparallel arrangement of two histones. α-helices (α_n) and loops (L_n) of the histone fold are indicated. **(c)** The architecture of the histone octamer is revealed by depicting only half of the DNA and associated proteins. This figure shows the histone fold regions, the histone fold extensions and part of the tails. The C-terminus of each histone protein is labelled as C¢. Note that most of the tails are disordered and are therefore not included in this picture. Regions of protein–DNA interaction are numbered from the centre of the DNA outwards. Taken from Luger, K. (2001)[57].

As a result of the extended structure of the histones and the type of their pairwise association within dimers, each histone chain emerges at multiple places (footprints) on the surface of the octamer. This pattern has important implications for the histone-DNA contacts in the nucleosome: the histone fold portion of each chain, by its occurrence in the four dimers, gives rise to repetitive structural elements: the "parallel beta bridges" and the "paired ends of helix I" motifs, and thus, forming two different surface patches with high electrostatic potential for DNA binding[31].

Post-translational modifications of core histones

Core histones have an important role in genetic regulation. Changes in their stoichiometry, sequence or post-translational modifications (PTM) can lead to remarkable changes in gene expression[59]. Post-translational modifications of the core histones are primarily located on their N-terminal tails (Figure 6)[60,61].

Since nucleosomes can be part of regulatory protein complexes and transcription factors do interact specifically with core histones, the N-terminal tails can act either as activators or suppressors[62] in the transcription of certain genes[63]. Thus, post-translational modifications have an important role in the modulation of chromatin superstructure, delimiting the differences between euchromatin and heterochromatin, and can direct and control the recruitment of the enzymatic complexes that are responsible of the metabolic processes related with DNA, including transcription, replication, DNA repair and condensation of chromatin into chromosomes. The 'histone code', term which will be later expanded on, comprises and describes the role of each modification and their crosstalk in these processes[64].

Post-translational modifications, mainly identified by Mass Spectrometry (MS) and Chromatin Immunoprecipitation on DNA chip (ChIP on CHIP) studies, are cell-cycle dependent and include methylation, acetylation, phosphorylation, SUMOylation, ubiquitination, citrullination, and ADP-ribosylation[64,65,66,67,68].

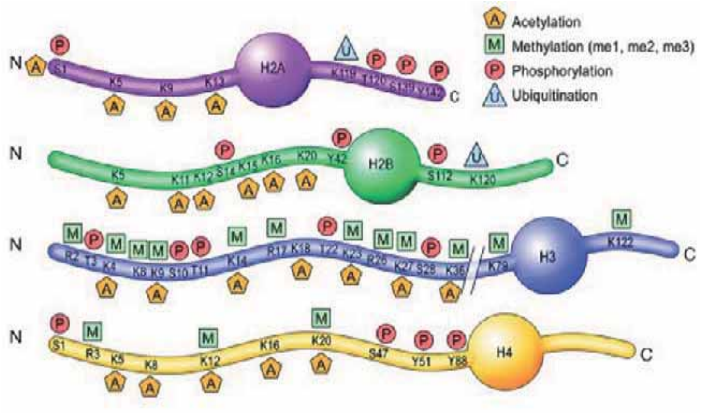


Figure 6. Post-translational modifications on core histones. Schematic representation of the N-termini and C-termini of the core histones with their major PTMs.
C: C-terminal tail; N: N-terminal tail.
Adapted from Franklin TB and Mansuy IM (2010)[68].

Methylation

Histone methylation is the modification of certain amino acids in a histone protein by the addition of one, two, or three methyl groups without altering the charge of the residues. Methylation and demethylation of histones turn the genes in DNA "off" and "on", respectively. In most cases, this PTM acts by either loosening of the binding of their tails and, thereby, allowing transcription factors and other proteins to access the DNA, or by enclosing them around the DNA and, thus, restricting the access to the DNA. This means that this post-translational modification alters the properties of the nucleosome and affects its interaction with other proteins. Broadly speaking, histone methylation is generally associated with transcriptional repression. However, methylation of some lysine and arginine residues of histones results in transcriptional activation. Therefore, the position and level of methylation is mainly related with transcriptional regulation. Recent studies on the pattern of epigenetic modifications and their significance in cellular processes as well as in the development of some diseases, such as cancer, seem to point out how histone methylation could play an important role[69].

Methylation of histone lysines starts just after the nucleosome assembly and reaches its maximal level during mitosis. During the interphase, methylation occurs preferentially in previously acetylated H3 and

H4 histones. However, this could only be an effect of the increase of the accessibility to N-terminal tails due to acetylation[66,67].

While the mechanisms by which arginine methylation regulates transcription are unknown, lysine methylation coordinates the recruitment of chromatin modifying enzymes. Chromodomains (HP1, PRC1), PHD fingers (BPTF, ING2), Tudor domains (53BP1), and WD-40 domains (WDR5) are among a growing list of methyl-lysine binding modules found in histone acetyltransferases, deacetylases, methylases and ATP-dependent chromatin remodelling enzymes. Lysine methylation provides a binding surface for these enzymes, which then regulate chromatin condensation and nucleosome mobility in order to maintain local regions of active or inactive chromatin. In addition, lysine methylation can block binding of proteins that interact with unmethylated histones or directly inhibit catalysis of other regulatory modifications on neighbouring residues. The presence of methyl-lysine binding modules in the DNA repair protein 53BP1 suggests roles for lysine methylation in other cellular processes.

Histone methylation is crucial for proper programming of the genome during development and misregulation of the methylation machinery can lead to diseased states such as cancer. Until recently, methylation was believed to be an irreversible, stable epigenetic mark that is propagated through multiple cell divisions, maintaining a gene in an active or inactive state. While there is no argument that methylation is a stable mark, recent identification of histone demethylases such as LSD1/AOF2, JMJD1, JMJD2 and JHDM1 has shown that methylation is reversible and provides a rational for how genomes might be reprogrammed during differentiation of individual cell lineages[70,71,72,73].

Hence, histone methylation is established through several mechanisms which include the interaction of histone lysine methyltransferase complexes (KMTs) with specific transcription factors, modified chromatin, RNA polymerase II or by directly interacting with specific sequences in the DNA. Lysine demethylases (KDMs) work together in coordination with KMTs for the maintenance of methylation patterns[74].

This modification has been broadly studied in lysine residues K9 and K27 of histone H3 in vertebrates. These residues can incorporate up to three methyl groups and these two methylations are associated with transcriptional repression. Methylation in K9 is important for the recruitment of the heterochromatin protein (HP1). HP1 is released during mitosis, when S10 becomes phosphorylated, and thus, a regulatory mechanism involving two adjacent post-translational modifications, a stable methylation in K9 and a dynamic phosphorylation in S10, is established. Such mechanism is known as 'methyl-phos switch'[75].

Histone H3 is also methylated in lysines K4, K36 and K79. In this case, methylation in these positions is a feature of transcriptionally active chromatin[76,77]. K36 methylation depends on the conformation of proline P38. When P38 is in its *trans* conformation, Set-2 methyltransferase catalyses K36 methylation. If P38 is in *cis* conformation, Set-2 is unable to produce K36 methylation. *Cis/trans* conformation shift of P38 could also have an effect on demethylases activity[78].

Another important function of histone lysine methylation is the demarcation of euchromatin and heterochromatin domains. Heterochromatin borders are delimited by the presence of H3K4 and H3K9 methylation in the adjacent euchromatin domains. Heterochromatin is associated with low levels of acetylation and also with H3K9, H3K27 and H4K20 methylations, related to transcriptional repression. H3K27 methylation is involved in the silencing of HOX gene and the maintenance of the inactive X chromosome. Methylation of H3K9 seems to be important in the maintenance of pericentromeric heterochromatin[64].

Histone H4 is mono- or di-methylated in its lysine K20. That modification is thought to be a signal for DNA repair complexes[79] and it is also related to heterochromatin formation[64].

Arginine methylation is very little characterized. Arginine is known to be mono- or di-methylated, and methylation can be symmetric or asymmetric, potentially with different meanings. It is preferably located in promoter regions and can cause both activation and repression of transcription[64].

Acetylation

Histone acetylation plays an important role in transcriptional regulation and, together with phosphorylation, has been subject of intense interest. Acetylation of the four core histones occurs in all plant and animal species examined[80] and the sites of modification are the lysine residues located in the positively charged N-terminal tails. The incorporation of an acetyl group implies a reduction in histone net positive charge by 1. The number of acetylated lysine residues per histone is determined by equilibrium between histone acetylases and deacetylases. Two populations of acetylated histones appear to exist in a particular cell nucleus: for instance, in the embryonic chicken erythrocyte, 30% of core histones are acetylated in a stable way, while the acetylation status of about 2% changes rapidly. The pattern of specific lysine residues in the histone tails that are acetylated varies among species. However, the non-random usage suggests that sequence specificity for the relevant acetylases and deacetylases may exist[81].

As mentioned, histones are acetylated and deacetylated as part of gene regulation. These reactions are typically catalysed by enzymes with histone acetyltransferase (HAT) or histone deacetylase (HDAC) activity[82,83]. Acetylated histones and nucleosomes represent a type of epigenetic tag within chromatin[60]. Since acetylation removes the positive charge on the histones, thereby decreasing the interaction of the N-terminus of histones with the negatively charged phosphate groups of DNA, the condensed chromatin is transformed into a more relaxed structure that is associated with greater levels of gene transcription[64]. This relaxation can be reversed by HDAC activity. This charge neutralization model has been supported by recent studies, according to which transcriptionally active genes are correlated with hyperacetylation and rapid turnover of histone acetylation[84] (Figure 7). This requires that the HATs and HDACs must act continuously on the affected histone tail.

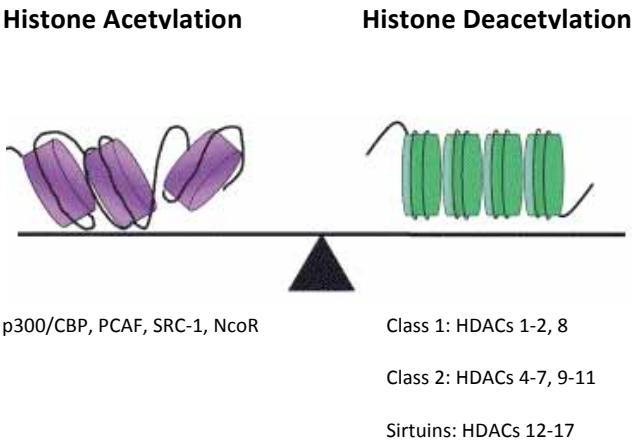


Figure 7. Gene repression and activation are regulated by acetylation of core histones. Histone acetylation is mediated by transcriptional coactivators, which have intrinsic histone acetyltransferase (HAT) activity, whereas repression is induced by histone deacetylases (HDACs), which reverse this acetylation, allowing repackaging of the nucleosomes. From Hayashi R, *et.al.* (2004)[84]

At the same time, acetylation appears to increase the content of α-helix that the N-terminal domains of core histones are able to reach in the presence of TFE[85,86]. In this sense, it has been postulated that the space between the potentially acetylated lysine residues located in H3 and H4 histones tails recalls the super helical turn of an α-helix[87,88].

Hyperacetylation of histone tails leads to small changes in nucleosome conformation[89] and it appears that the more remarkable consequences are be related to protein-protein interactions, either between nucleosomes, with histone H1 or with other non-histone proteins[90].

In vitro studies suggest that histone acetylation has no direct influence in linker histone association with the nucleosome[91]. Nonetheless, acetylated regions seem to be poorer in H1 and acetylation causes chromatin relaxation during the interphase[92].

Moreover, acetylation is also related to DNA repair processes[93,94]. H3K56 acetylation is required for genome stability[95,96]. H3K56 is acetylated by the p300/Rtt109 complex[97,98,99], but is rapidly deacetylated around sites of DNA damage. H3K56 acetylation is also required to stabilise stalled replication forks, preventing dangerous replication fork collapses[100,101]. Although in general mammals make far greater use of histone modifications than microorganisms, a major role of H3K56 acetylation in DNA replication exists only in fungi, and this has become a target for antibiotic development[102].

Thus, histone acetylation appears to be remarkably important in transcriptional regulation as well as in many other processes, indicating that this post-translational modification plays a central biological role[64].

Phosphorylation

Phosphorylation is, after acetylation, the second type of core histone modification that has been more extensively studied. It affects the four core histones and it also has important structural consequences for nucleosome structure and spacing. Addition of a negatively charged phosphate group can lead to major changes in protein structure, leading to the well-characterised role of phosphorylation in controlling protein function. It is not clear what structural implications histone phosphorylation has, but histone phosphorylation has clear functions as a post-translational modification, and binding domains such as BRCT have been characterised[103].

Histone H3 is rapidly phosphorylated on serine residues within its basic N-terminal domain in response to extracellular signals, such as growth factors, that stimulate the proliferation of quiescent cells[104]. H3 phosphorylation is maximal during mitosis, when serine residue S10 is phosphorylated. This modification has been associated with chromosome condensation and it is supposed to modulate the interaction between the H3 tail and the DNA, allowing the access of decondensing factors[105,106].

Cells lacking tumour suppressor genes or expressing oncogenes have significant levels of phosphorylated H1 and H3, which are thought to be responsible for loosening of chromatin structure and aberrant gene expression[107]. Aurora B kinase is responsible for phosphorylation of histone-H3 on serine S10 during mitosis[108,109] and it is overexpressed in many carcinogenic cell lines. This modification is conserved from yeast (where the kinase is named Ipl1) to humans. During mitosis, histone H3 is phosphorylated at Ser10 in all eukaryotes but its function has been extensively debated[108,110,111,112]. Some studies show a strong correlation between Ser10 phosphorylation and chromosome condensation during cell division[112]. However, other genetic data obtained in yeast suggest that histone H3 phosphorylation is not required for chromosome condensation[108]. Despite Aurora B is enriched at centromeres, it localizes diffusely to all chromatin. Hyperphosphorylation is related to the chromosomal instability found in malignant cells[113,114,115]. Furthermore, H3S10 phosphorylation has also been associated with the induction of transcription in response to heat shock[116] and c-Fos and c-jun oncogenes activation[64].

In vivo phosphorylation of H4 and H2A histones occurs in the cytoplasm just after their synthesis[117,118,119] and, together with histone H4 acetylation, these modifications could selectively direct the histones to the molecular chaperones involved in the nucleosome assembly and the replication fork[120,121].

Phosphorylated H2AX (also known as gamma H2AX) is a marker for DNA double strand breaks[122], and forms part of the response to DNA damage[123,124]. H2AX is phosphorylated early after detection of DNA double strand break, and forms a domain extending many kilo bases either side of the damage[122,125,126]. Gamma H2AX acts as a binding site for the protein MDC1, which in turn recruits key DNA repair proteins[127,128] and as such, gamma H2AX forms a vital part of the machinery that ensures genome stability.

Ubiquitination

Ubiquitin is a small regulatory protein (76 amino acids in length and about 8.5 kDa in mass) that has been found in almost all tissues of eukaryotic organisms. Among other functions, the most well-known and important is to label proteins to direct their recycling. The tag can also direct proteins to other locations in the cell, to control other proteins and a wide variety of cellular processes[129,130].

Ubiquitination consists on an enzymatic process in which the carboxylic acid of the terminal glycine from the di-glycine motif in the activated ubiquitin forms an amide bond to the epsilon amine of the lysine in the modified protein. The covalent addition of ubiquitin generates a voluminous protein complex near the C-terminal domain of the protein.

Histones H3, H2B and, specially, H2A, can be modified by ubiquitination[131]**.** However, in the case of H2A and H2B, this post-translational modification is not a signal for protein degradation. Ubiquitinated histone H2A is assembled within the nucleosome and it appears not to experiment important changes in terms of core particle organization[132,133,134]. However**,** H2A ubiquitination could act blocking the interaction between H1 and nucleosomal DNA and also between adjacent nucleosomes. Thus, histone ubiquitination could prevent chromatin condensation and be related with transcriptionally active DNA regions. Despite that, the structural role of core histone ubiquitination and its importance in chromatin dynamics remains still unclear[135].

SUMOylation

Small Ubiquitin-like Modifier (or SUMO) proteins are a family of small proteins (about 12kDa in mass and 100 amino acids in length) that are covalently attached to and detached from other proteins in cells to modify their function. SUMOylation is a post-translational modification involved in various cellular processes, such as nuclear-cytosolic transport, transcriptional regulation, apoptosis, protein stability, response to stress, and progression through the cell cycle[136].

SUMO proteins are similar to ubiquitin, and SUMOylation is directed by an enzymatic cascade analogous to that involved in ubiquitination. Thus, histone conjugation with SUMO leads to a voluminous modification. In contrast to ubiquitin, SUMO is not used to tag proteins for degradation. Mature SUMO is produced when the last four amino acids of the C-terminus have been cleaved off to allow formation of an isopeptide bond between the C-terminal glycine residue of SUMO and an acceptor lysine on the target protein[136].

Core histones can be SUMOylated. The modifications described in yeast are located in K126 in the C-terminal of H2A; K6, K7, K16 and K17 in the N-terminal domain of H2B, and in the five lysine residues present in the N-terminal domain of H4. No data has been reported for H3[137]. In histones, this post-translational modification seems to be related with transcriptional repression processes[64,137].

Citrullination

Citrullination or deimination consists on the post-translational modification of arginine residues within a protein into the amino acid citrulline. This reversible reaction is catalysed by peptidylarginine deiminases (PADs). The conversion of arginine into citrulline can have important consequences for the structure and function of proteins, since arginine is positively charged at a neutral pH, whereas citrulline is uncharged. This fact could increase the hydrophobicity of the protein, leading to changes in protein folding.

Core histones H3 and H4 can both be citrullinated in their arginine residues by PADI4. Since citrulline cannot be methylated[138], this modification can block arginine residues methylation, and lead to indirect effects in the processes modulated by histone methylation.

ADP- Ribosylation

ADP-ribosylation is the reversible addition of one or more ADP-ribose units to a protein[139]. These reactions are involved in cell signaling and the control of many cell processes, including DNA repair and apoptosis[140]. The enzymes that catalyze the reaction are mono-ADP-ribosyl-transferase and poly-ADP-ribosyltransferase activity-transferase (PARPs), according to the incorporation of one or more units of ADP-ribose, respectively.

Some studies on the ADP-ribosylation of histones suggest that this modification could locally increase chromatin decondensation in a local manner. Furthermore, ADP-ribosylation could have an important role in DNA repair[141]. The synthesis of long and negatively charged chains of ADP-ribose could facilitate the interaction of core histones with the competitor polyanion, disrupting the nucleosome integrity.

Histone H1

Histone H1 or linker histone is a protein of about 20kDa very rich in lysine that interacts with the linker region of DNA between adjacent nucleosome cores (internucleosomal DNA). In the absence of histone H1, the nucleosome core particle comprises approximately 147bp of DNA that form 1.75 turns around the outer surface of the histone octamer[142,143]. Histone H1, which is present at a ratio of 0.6-1.3 molecules per nucleosome depending on the cell type, stabilizes about 20 additional nucleotide pairs of DNA. Then, the resulting chromatosome contains two complete turns of 168bp DNA, the core histone octamer and a molecule of H1[144].

Linker histones have an important role in establishing and maintaining chromatin higher-order structure[145,146] and in gene regulation. Histones H1 can act in transcriptional regulation through the modulation of chromatin condensation or as a part of protein complexes responsible of either the activation or repression of certain genes[147,148,149]. H1 also participates in transcriptional regulation through nucleosome positioning[150], stabilization of DNA at the entry/exit of the core particle[149] and preferential binding to SAR regions[151].

Subtypes of histone H1

An important characteristic of H1 is its heterogeneity. There are several subtypes and it is thought they differ in their functions. Functional differentiation is based in their timing of expression[152], moment and extent of phosphorylation[153], turnover rate[153,154] and in their ability to condense chromatin *in vitro*[155,156,157,158]. In vertebrates, the subtypes differ widely in evolutionary stability, suggesting that each subtype may have acquired a unique function[159]. Developmental and gene expression studies also support the concept that the different subtypes play distinct roles in chromatin structure[160,161,162.]

In mammals, nine subtypes (H1a-e, $H1^0$, H1x, H1oo and H1t) have been identified. They share a highly conserved globular domain sequence while exhibiting variation in the N and C-terminal domains[163].

There are up to twelve different nomenclatures for histones H1, which sometimes leads to confusion. The most commonly used are the numerical nomenclature from Doenecke[164], the alphabetical nomenclature from Lennox[165], and the alpha-numerical one, from Parseghian and Hamkalo[134]. In the present study, we used the one from Lennox.

Five of these subtypes (H1a, H1b, H1c, H1d and H1e) are usually referred to as the *somatic subtypes* since they are present in most somatic cells[166,167]. In particular, H1a is restricted to liver, kidney, lung, testis, lymphocytes from thymus and spleen, and immature neuronal cells[168,169]. Its expression is related to cellular division. The levels of H1a, H1b and H1d rapidly decrease in terminal differentiated cells or in G0 phase, whereas H1c and H1e are highly expressed and remain constant during the cell cycle and in almost any cell type. Competition assays carried out to estimate the relative binding affinities of the different H1 subtypes for long chromatin fragments[170], led to a classification of the H1 subtypes into three major binding groups: H1a with low affinity, H1b and H1c with intermediate affinity, and H1d, H1e and $H1^{o}$ with high affinity. The differences in affinity of the H1 subtypes are basically determined by the C-terminal domain of the molecule[171]. This suggests that small differences in the structure of the C-terminus, affecting the spatial distribution of the basic residues, could be fundamental in determining the differential affinity of the subtypes; as a consequence, this characteristic could be functionally relevant and thus, could contribute to the functional differentiation of the subtypes. The subtypes are also non-randomly distributed with respect to active and inactive chromatin. Actively transcribed sequences are presumably located in less condensed regions of chromatin. Gene expression studies localized H1b in regions of active transcription and H1d and H1e in less active or inactive regions[172,173,174], suggesting the association of the subtypes of high binding affinity with more condensed chromatin and of those of intermediate or low binding affinity with less condensed chromatin.

Histone H1 subtypes can be classified in three classes regarding to the type and level of differentiation of the cell where they are being expressed[175.]

Class I. Corresponds to linker histones expressed in germ line cells, where the expression occurs in high levels as a reservoir to be used during primary embryogenesis[176,177]. H1oo is present in mice and human oocytes, whereas CS H1 is present in sea urchin and H1M and B4 are found in frogs. H1t corresponds to male germline, together with two more histone-like proteins: H1T2[178,179] and HILS1[180].

Class II. Corresponds to linker histones which are expressed in somatic cells and whose expression is cell cycle-dependent. All somatic subtypes, together with histone H1t and core histones[181], are coded in the same region of the genome, known as 'histone cluster'. This region is located in chromosome 6 in humans[182] and in chromosome 13 in mice[183].

Class III. Corresponds to histone H1 specific variants involved in cellular differentiation. $H1^{0}$, H5 and H1x[184] belong to this class. H5 is expressed in nucleated chicken erythrocytes. $H1^{0}$ is expressed in terminally differentiated cells and its gene is located in chromosome 22 in humans[182] and in chromosome 15 in mice[185], isolated from the 'histone cluster'. These histone genes have tissue-specific promoter sequences[186] and other regulatory elements that are activated or repressed depending on the tissue and the development state.

Functions of histone H1

The first and earliest role ascribed to H1 is the formation and maintenance of higher order chromatin structure in the form of a 30nm fibre through the organization of nucleosomal spacing and compaction of the 10nm fibre[186].

It is believed that chromatin compaction induced by H1 histone is based in the neutralization of the negatively charged phosphate backbone of DNA. Although the N-terminal domain of the core histones is necessary for the generation of condensed nucleosomal structures regardless of linker histones, H1 is required for the stabilization of these condensed structures[187]. The presence of H1 also modulates the structure of the chromatin fibre[188] and its depletion in mammals causes significant changes to chromatin structure. When chromatin is depleted of H1, there is a global reduction in the nucleosome-repeat length (NRL) and a reduction in the local chromatin condensation[189].

Actually, an interplay between the nucleosome-repeat length (NRL) and the presence of H1 has been an important feature to consider in the description of higher-order chromatin structure models of the 30nm fibre previously extended above[11,15,16,18]. The total level of H1 affects the linker DNA length and thus also the nucleosomal repeat length (NRL) (Figure 8)[190,188,191,192].

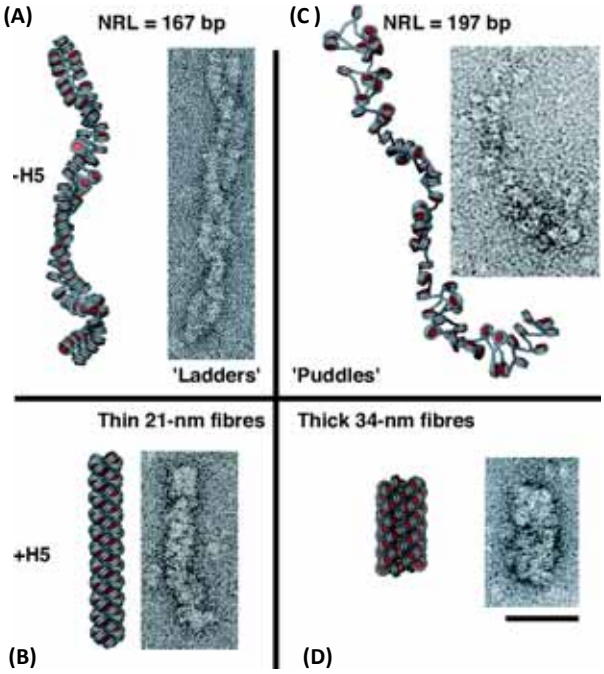


Figure 8. The NRL and the linker histone determine chromatin higher-order structure. (A, C) Without linker histone. **(B, D)** With linker histone. EM micrographs taken from Routh A, *et.al.* (2008)[188]. The 167- and 197-bp NRL arrays have different linker histone-dependent compaction pathways and different structures. **(A)** Unfolded 167-bp NRL fiber showing the two-start helix arrangement typified by the stacking of nucleosome cores in the absence of linker histone. **(B)** Fully folded 167-bp NRL fiber in the presence of saturating linker histone concentrations. **(C)** Unfolded 197-bp NRL array showing the formation of puddles in the absence of linker histone. **(D)** Folded 197-bp NRL fibers in the presence of saturating linker histone concentrations. (Scale bar: 50 nm)

The NRL directly determines the length of internucleosomal DNA. Since native chromatin has an irregular spacing between nucleosomes, the internucleosomal DNA length is heterogeneous and that heterogeneity affects nucleosomes positioning and the interactions among them7. The NRL considerably varies among species and tissues. Cells with a higher H1/nucleosome ratio tend to have a longer NRL, and a reduction in the H1 content leads to a reduction in NRL[193,189,194].

A shorter NRL is often a characteristic of active chromatin domains and rapidly growing cells, such as embryonic stem (ES) cells. In contrast, mature cells with more compact chromatin tend to have longer NRL[195]. Thus, the condensation of the chromatin fibre is determined by the nucleosome/H1 stoichiometry and the NRL. The 30nm fibre formation requires an internucleosomal DNA long enough to allow the attachment of H1 in its appropriate stoichiometry[10].

It appears that there is a linear relationship between H1 stoichiometry and the NRL to maintain a balance in the electrostatic net charge in the nucleus. If the amount of H1 per nucleosome decreases,

the same will do the internucleosomal DNA length in order to compensate the latter[196]. Furthermore, it has been recently seen that H1 subtypes that coexist in most types of cells of higher eukaryotes differ in their effects on the nucleosomal spacing *in vivo*. This suggests that also H1 subtypes play different roles in the organization and functioning of the chromatin fibre[190,197].

It is currently accepted that H1 could have a regulatory role in transcription through the modulation of chromatin folding. H1 has been described as a general repressor of transcription because it contributes to chromatin condensation, which limits the access of the transcriptional machinery to DNA[189,198]. However H1 can be removed by some protein factors to allow transcription[199].

Other studies indicate that H1 might regulate transcription at a more specific level, participating in complexes that either activate or repress specific genes[148,200,92,201,202,203,147,204].

Preferential binding to SAR regions (scaffold-associated regions or scaffold-attachment regions)[151,205] and participation in nucleosome positioning[150] are other mechanisms by which H1 could contribute to transcriptional regulation. SAR regions are very rich A/T regions in homopolymeric tracts (A-tracts) in the DNA where the nuclear matrix attaches to organize the chromatin into structural domains[206]. These regions are thought to be involved in chromosome dynamics[207] and it has been proposed that controlled H1 binding and dissociation to SARs plays a role in localized opening and closure of chromatin loops and, thus, contributing to transcriptional regulation[151].

Modulation of gene expression by H1 can also be a consequence of changes in the accessibility of promoter sequences or transcription factor enhancers, which could affect the binding of transcription factors, limit nucleosome sliding and decrease DNA transient exposure in the nucleosome surface promoting chromatin condensation into higher-order structures

The relationship established between chromatin structure and its transcriptional activity is remarkably important, since DNA compaction within the chromatin fibre represents a barrier for the transcriptional machinery. Transcription factor's binding sites happen to be hidden in the nucleosome structure in several promoter regions and enhancers. Despite that, many of those factors are able to bind the DNA even if it is wrapped around in the nucleosome[208].

In order to permit the initiation of transcription, the nucleosome can be either removed or kept in place, yet with its structure redesigned or altered, allowing the entry of transcription factors[208,209].

H1 has also been implicated in the inhibition of chromatin replication and, more recently, H1 phosphorylation has been linked to DNA repair, apoptosis and ATP-dependent chromatin remodelling[210,211,212].

Structure of histone H1

Histone H1 in vertebrates has a characteristic three-domain structure consisting of a short, flexible N-terminal domain (20-30 amino acids), a central globular domain (80 amino acids) and a long C-terminal domain (~100 amino acids)[213].

The amino- and carboxyl-terminal domains are highly basic and very lysine-rich, specially the C-terminal domain (CTD) which contains up to 40% of lysine residues[213]. Despite they are also highly conserved through evolution, when compared with core histones, linker histones are more variable in both sequence and structure4.

The general tripartite structure of H1 is found only in higher eukaryotes. In fact, in lower eukaryotes, histone H1 does not have this classical structure. In *Tetrahymena*, H1 lacks the globular domain and is

similar to the C-terminal domain of metazoan H1, whereas the counterpart of H1 in *S. cerevisiae*, Hho1p, possesses two globular domains[214,215].

The N-terminal domain

The short N-terminal domain of histone H1 has a length of about 30 amino acids and can be internally subdivided in two different regions[216]. The first region corresponds to the first amino acids of the sequence and thus, to the most distal region of the protein. This region does not have basic residues and it is mainly non-polar due to its high content of alanine and proline. The second region corresponds to the most proximal to the globular domain and it is rich in lysine residues and, because of that, positively charged and highly basic. It is thought that this region could contribute to the stabilization of the globular domain and its interaction with DNA[217].

The N-terminal domain remains unstructured in water solution but it acquires a high proportion of α-helix in the presence of DNA or secondary structure inductors, such as TFE[218,219]. Studies performed by NMR-2D on two isolated peptides corresponding to two different regions of the N-terminal domain of histone $H1^0$ (NH-1 peptide, which contains the residues 1-20 of $H1^0$ and NH-2 peptide, which corresponds to residues 8-29 of $H1^0$) showed that the basic half of the domain become structured into α-helix in the presence of 90% TFE. Subsequent studies by FTIR spectroscopy on DNA-peptide complexes demonstrate that the secondary structure of the N-terminal domain bound to DNA is similar to the one obtained in TFE[218]. Those NMR analysis on NH-1 and NH-2 peptides allowed proposing a structural model for the N-terminal domain of $H1^0$ in which the residues 11-24 are structured in α-helix[216].

Other studies on a peptide corresponding to the basic region of the N-terminal domain of H1e have been performed by NMR and circular dicroism (CD)[219]. In this case, the peptide acquires a 40% of α-helix in the presence of TFE and it is structured into two amphipathic helixes with a Gly-Gly motif in between that works as a hinge. The first helix goes from residues 17 to 27, whereas the second one includes residues 29 to 34, remaining an almost whole α-helix structure along the peptide[219].

The globular domain

The central portion of the protein includes about 80 amino acids. It is highly conserved among different H1 subtypes and species[160] and has a globular conformation. This domain binds to linker DNA where it enters and exits the nucleosome, interacting with about 20bp and protecting them from nuclease micrococcal digestion[142].

The structure of the globular domain is well-known and it has been determined by NMR-2D[220] and X-ray crystallography at 2.5Å resolution[221]. The crystallographic structure of the globular domain of histone H5 (GH5) permitted the classification of the protein as a member of the DNA binding proteins which contain the motif helix-turn-helix (HTH), also known as 'winged-helix' fold[221]. This motif consists of three α-helix followed by an antiparallel β–sheet in its C-terminal end. The HTH motif has been identified in other proteins, such as CAP transcription factor (catabolite gene activator protein) and HNF3γ (hepatocite nuclear factor-3γ) and it is the one motif that permits the binding to DNA. The structure of the globular domain of chicken histone H1.1L resolved by NMR also shows that motif. Considering the homology between chicken and mammals H1/H5 histones, it is reasonable to assume that the same structural pattern might be present in the globular domain of linker histones in mammals.

On the other hand, the striking resemblance with the structures of transcription factors, such as CAP and HNF3-γ, which also bind to DNA, suggest the existence of a primary DNA binding site that would consist of helix III, and its adjacent β-sheet[221,222]. A secondary binding site, which interacts with the minor groove of the internucleosomal DNA adjacent to the nucleosome, has also been described[216,223]. Recent *in vivo* studies show that the globular domain interacts with the major groove of DNA, positioning it in the symmetric axis of the nucleosome via the residues H25, R47, K69, K73, R74 and K85

(Figure 9)[223]. It is thought to involve residues R42, located in between helixes I and II, K52 in helix II and R94 in the β-sheet. These residues are conserved in all mammals' variants, with the exception of H1t[216].

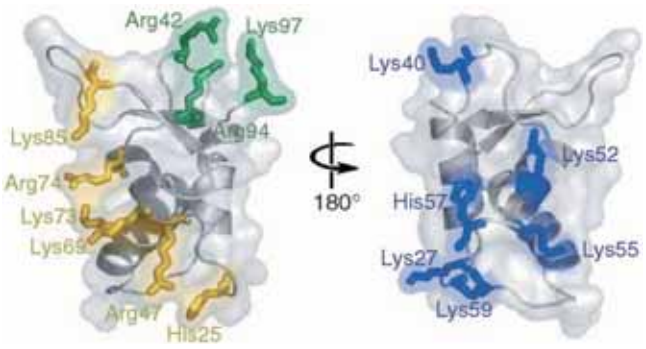


Figure 9. Map of the interaction surface of the globular domain of H1°. In yellow, binding residues in site 1; in green, site 2; in blue, nonbinding residues. From Brown DT, *et.al.* (2006)[223]

The C-terminal domain

The C-terminal domain (CTD) has a length of about 100 amino acids and an approximate content of 40% of basic residues, mainly lysine. It also enriched in alanine, serine and proline. Despite the sequence variability of this domain among different histone subtypes, the charge distribution is quite uniform in all cases[224]. This is due to the high density of positively charged lysine residues within the CTD[224].

Because of this homogeneous positive charge density, it is commonly believed that chromatin condensation is mediated through charge-neutralization of the negatively charged linker DNA. Therefore, this domain probably binds to linker DNA, neutralizing the negative charge of phosphates and facilitating chromatin condensation into the 30 nm fibre and also the intermolecular aggregation[225,226,227]. Studies carried out by circular dicroism (CD) demonstrate that H1 CTD is the main determinant of DNA condensation under physiological salt conditions[228]. Other studies propose the presence of two subdomains within the CTD to be the responsible ones of the linker DNA alteration and the stabilization of more condensed chromatin structures[229]. These two subdomains would be distributed in two discontinuous regions of about 24 residues each, known as subdomains 1 and 3, respectively. In rat H1e, a fragment of about 34 amino acids (residues 145-178) has been identified as the necessary region for chromatin condensation[230].

An important feature of the CTD of linker histones is the SPKK motifs[231] (Figure 10), which are susceptible of phosphorylation by cyclin-dependent kinases (CDKs) and are involved in many structural and functional aspects of H1[232]. SPKK motifs are part of the tetrapeptide repetition (S/T)-P-X-Z in which X can correspond to any amino acid and Z corresponds to a basic residue[233,234]**.** Those motifs are thought to mimic the condensing properties of H1 CTD. Moreover, their phosphorylation in a cell-cycle dependent manner plays an important role in both the condensation status of chromatin and the folding of the protein. This issue will be treated in more detail below.

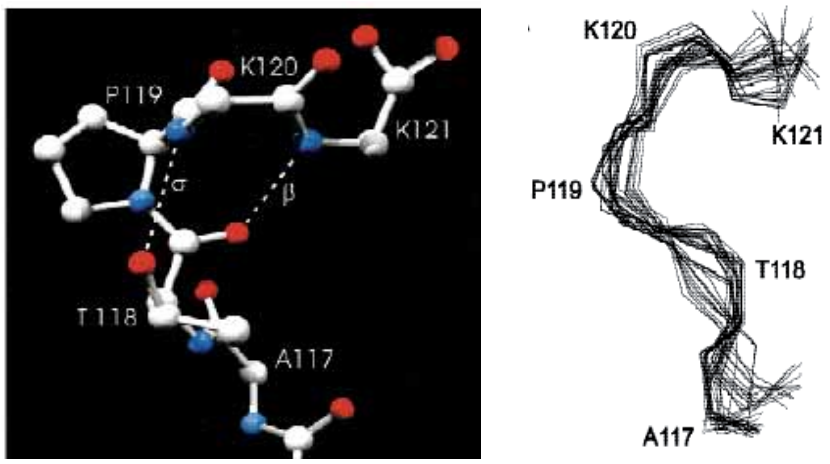


Figure 10. SPKK motif with its characteristic β/σ turn described by NMR in the CTD of H1^0.
From Vila *et.al.*(2000)[231]

The N- and the C-terminal domain of H1 are unstructured in aqueous solution[235,236,237,238]. Nevertheless, in the presence of secondary structure inductors such as TFE, the CTD showed an increase in the content of α-helix[239,240]. This fact led to the hypothesis that, perhaps, the interaction with DNA would also induce an increase of α-helix in the CTD: after all, interaction with DNA would involve charge neutralization, leading to a less polar environment that would facilitate the folding of the domain. For this reason, several studies about the interaction of the CTD of linker histones and DNA were carried out by FTIR spectroscopy[235].

In all cases, the interaction with DNA induced the folding of the CTD. Furthermore, the amount of secondary structure present was dependent on the salt concentration, acquiring a complete folding in 140mM NaCl. Under physiological conditions, no disordered structure was found and the CTD of H1^0 and H1t showed similar percentages of secondary structure elements: 24% α-helix, 25% β-sheet, 18% open loops and 33% turns[235]. The secondary structure induced by the interaction with DNA happened to be extremely stable. This stability could be due to the increase of hydrophobicity caused by the compensation of positive charges in the protein when interacting with the negative phosphate backbone of DNA, as mentioned above[235]**.** The fact that the CTD is unstructured in aqueous solution but becomes completely folded when bound to DNA allowed including this domain in the group of intrinsically disordered proteins with coupled binding and folding.

Post-translational modifications of histone H1

Histone H1, just as the core histones, is the target of several post-translational modifications. Modifications found in H1 primarily include phosphorylation, acetylation, methylation, and ubiquitination. Recently, N-formylation of lysine residues has been detected in linker and core histones[241,242]. However, phosphorylation is considered the most important PTM in histone H1, because it occurs in several sites of the molecule and in a cell cycle-dependent manner. Furthermore, it generates interesting structural and functional effects in the molecule that could be important to understand the role of H1 histone[249,254,243,244].

Phosphorylation

Reversible phosphorylation is the main and most important modification affecting linker histones; also the most extensively studied post-translational modification of chromatin4.

In general, phosphorylation is the most common mechanism of protein function regulation and is often termed 'phosphoregulation'. Reversible phosphorylation of proteins occurs in both prokaryotic and eukaryotic organisms. Many enzymes and receptors are switched *on* or *off* by phosphorylation and dephosphorylation due to the induction of a conformational change in their structure, causing them to become activated or deactivated.

Phosphorylation usually occurs on serine, threonine, tyrosine and histidine residues in eukaryotic proteins[245]. In prokaryotic proteins, phosphorylation can also occur on arginine or lysine residues[245,246]. The addition of a phosphate (PO_4) molecule to a polar R group of an amino acid residue can generate a change in terms of local hydrophobicity and, thus, it can introduce a conformational change in the structure of the protein via interaction with other hydrophobic and hydrophilic residues in the protein.

Most common regulatory roles of phosphorylation include biological thermodynamics of energy-requiring reactions, enzyme inhibition mediation, signal transduction, protein-protein interaction, binding affinity regulation, gene expression and protein degradation. It is estimated that 1/10 to 1/2 of proteins are phosphorylated in some cellular state. Phosphorylation often occurs on multiple sites on a given protein[247,248].

Phosphorylation is also well-known as a mechanism for the control of the cell cycle. In this case, cyclin-dependent kinases (CDKs) are responsible for reversible phosphorylation. CDKs consist of both a catalytic and a regulatory subunit and their enzymatic activity is regulated by post-translational modifications and association to inhibitors[249].

Table 1. Cyclin-dependent kinases related with cell cycle[250]

Cell cycle stage	CDK	Cyclin	Inhibitors
G1	CDK4/6	D	p21, p27,p15,p16
G1/S	CDK2	E	p21, p27
S	CDK2	A	p21
G2/M	cdc2	A	p21
M	cdc2	A, B	

Histone H1 is phosphorylated in a cell-cycle dependent manner by the CDKs on the consensus sequence (S/T)-P-X-Z[233,234]. These conserved sequences are known as SPKK motifs and are localized mostly in the C-terminal domain[251]. They are thought to be sites of DNA binding and function in the compaction of the DNA. The level of phosphorylation depends on a balance of protein phosphatase 1 and cdc2/CDK2 kinase activities in the cell[252].

In addition to these sites, phosphorylation of non-CDK sites has also been reported. Examination of the phosphorylation sites of histone H1b from mitotically enriched CEM cells showed that the pentaphosphorylated H1b was the result of phosphorylation of one of the tetraphosphorylated forms at a nonconsensus motif at T10 in the N-terminal tail. T10 phosphorylation is thought to be the last step in the mitotic-specific cascade of phosphorylation, triggered by events other than those responsible for *TPXZ* motif phosphorylation[243]. Other studies on linker histones PTMs in mouse tissues and human cell lines have found several phosphorylations in non-consensus sites[242,253,]. In general, phosphorylation has been identified in positions S1, T3, T18, S26 and S35 of H1e; S30 of H1.X; T146 and T165 of H1c and T18 and T147 in H1d.

H1 phosphorylation varies progressively through the cell cycle. Studies with mammalian cells in culture showed that phosphorylation of H1 is highest in rapidly dividing cells and decreases in non-proliferating

cells4. Levels of histone H1 phosphorylation are lowest in G1 phase, they rise up during the S phase and become maximal during late G2 and mitosis[254,255,256] During mitosis, phosphorylation of H1 peaks at metaphase, when chromosomes are at their most condensed state: hyperphosphorylation of H1 would permit the binding of chromatin condensing specific factors[257]**.** At the end of telophase, all the phosphate groups have been removed so H1 has no phosphates at the beginning of a new cycle.

Table 2. Maximal Number of Phosphate Groups per Histone Molecule during the G1 and S Phases and Mitosis.

H1 subtype	Late G1	Late S	Mitosis
H1a	1	1	4
H1b	2	3 (3)	5 (5)
H1c	1	1 (1)	4
H1d	1	2 (1)	5
H1e	1	3 (2)	4
$H1^0$	1	1	3

Values are numbers of phosphate groups per molecule of histone variants from mouse fibroblasts and from rat glioma cells (in parentheses). The phosphate groups of $H1^0$ from the G1 and the S phase and of H1a are determined only from mouse fibroblasts. From Sarg, B *et.al.* (2006)[243]

During S phase, H1 exist both as unphosphorylated and low-phosphorylated forms and there is evidence that this partial phosphorylation is important for the process of DNA replication[258,259,260]. A study demonstrated that replicating DNA and phosphorylated H1 colocalize *in vivo*, strongly suggesting that H1 phosphorylation promotes DNA decondensation as precondition for DNA replication[261].

Whereas H1 phosphorylation seems to be involved in chromatin decondensation during the S phase, it is likely linked to condensation during M phase. In fact, it has been shown that H1 hyperphosphorylation is required to maintain metaphase chromosomes in their condensed state[249,257]. Thus, it appears that there is a causal relationship between histone H1 hyperphosphorylation and chromosome compaction. However, this relationship is yet not clear at all, since chromatin condensation can also occur in the absence of H1[262,263,212]. Then, it could be possible that hyperphosphorylation of H1 would decrease the association between H1 and chromatin, allowing the entry of chromatin condensing factors. It is actually known that hyperphosphorylated H1 has less affinity than partial or hypophosphorylated H1 for chromatin[260,264].

Intrinsically disordered proteins are often involved in important biological processes, such as cell cycle control, transcription, translation, transport or signal transduction[265]. Phosphorylation occurs in the intrinsic disordered regions within a protein. This regions have a higher frequency of phosphorylation sites than other regions, thus, phosphorylated sites show a clear preference to be located in intrinsically disordered regions[248].

The effect of phosphorylation in histone H1 secondary structure

Previous studies had demonstrated that phosphorylation of the CTD of H1 has effects on secondary structure and DNA condensation[266,267].

Fourier-transformed infrared spectroscopy (FTIR) of the secondary structure of different phosphorylated forms of H1-CTD bound to DNA[266,267] demonstrated that phosphorylation does not affect the overall conformation of the CTD when free in solution but, when bound to DNA; full phosphorylation (hyperphosphorylation) of the CTD implied an increase of β-structure and a decrease of the proportion of the α-helix. This structural change was dependent on the protein/DNA ratio (r) (Figure 11).

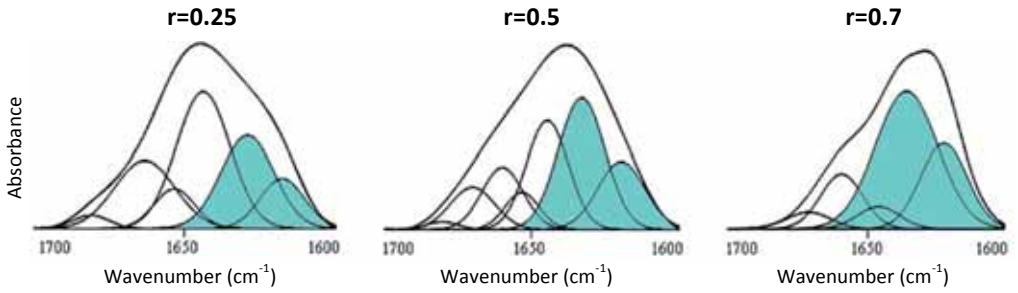


Figure 11. Amide I decomposition of DNA-bound triphosphorylated CH1° at different protein/DNA ratio (r) (w/w). β-structure components are highlighted in blue. Taken from Roque et.al. NAR (2008)[266]

Furthermore, phosphorylation showed to cause a small decrease of the affinity of the CTD for the DNA proportional to the number of phosphate groups present in the protein. When studying the aggregation capacity of DNA fragments by different phosphorylated species by band-shift-gel-electrophoresis, the hyperphosphorylated H1-CTD regained the aggregation capacity of the unphosphorylated C-terminal domain, which aggregated the totality of the DNA fragments[266].

These results, together with the fact that H1-CTD has been defined as an intrinsically disordered protein[235], show that phosphorylation effects on chromatin structure may depend on the level and the position of the phosphate groups incorporated in H1 and that they might be mediated by specific structural changes and are not a simple effect of the net charge. Hence, properties of hyperphosphorylated H1 would not be the result of a continuous variation in molecular properties depending on the number of phosphates, but determined by the specific structures associated with full phosphorylation[266].

Other effects of H1 phosphorylation

H1 phosphorylation is directly related with gene expression. The MMTV promoter (mouse mammary tumour virus) is a suitable system to study the changes in gene expression associated to chromatin remodelling, since nucleosome positioning within this region has been established and it is modulated by hormone addition[268]. Only when H1 is phosphorylated, it can be dissociated from the MMTV promoter by the activity of the glucocorticoid receptor. Prolonged exposure to hormone induces H1 dephosphorylation and the promoter becomes refractory to glucocorticoids[269]. Despite the inducible effect of H1 phosphorylation, the coordinated action of the highly phosphorylated H1 and H3 during mitosis promotes chromatin compaction and thus, transcriptional silencing of MMTV[270].

H1 phosphorylation has also been related to other nuclear processes such as ATP-dependent chromatin remodeling[271], ligation[272], cell differentiation[273] and apoptosis[274].

ATP-dependent chromatin remodeling complexes play key roles in transcriptional regulation, development, DNA repair and cell-cycle control and phosphorylated H1 seems to have an effect in the activity of those complexes. Remodeling activities of the ySWI/SNF, hSWI/SNF, xMi-2 and xACF ATP-dependent complexes are nearly abolished by incorporation of linker histones into polynucleosomal array substrates that consist of eleven repetitions of the 5S rDNA positioning sequence[271]. Moreover,

phosphorylation of H1 by cdc2/Cyclin B kinase can restore the activity of those remodeling complexes. This fact suggests that linker histones play a global role in the control over remodeling activities, implicating the need of a coupling between linker histone kinases and ATP-dependent remodeling enzymes[271].

DNA non-homologous end-joining *in vivo* requires the DNA-dependent protein kinase (DNA-PK) and DNA ligase IV/XRCC4 (LX) complexes. Phosphorylated linker histone takes part in this ligation process. Actually, LX ligation activity *in vitro* is substantially reduced by the incorporation of linker histones but it can be restored by addition of DNA-PK, which efficiently phosphorylates histone H1. Hence, the kinase activity of DNA-PK is essential for the recovery of end-joining and it may act as a linker histone kinase by phosphorylating linker histones in the vicinity of a DNA break and coupling localized histone H1 release from DNA ends, with the recruitment of LX to carry out double-stranded ligation[272].

Terminal differentiation of most of the cell types is related with significant structural changes in chromatin conformation. Dephosphorylation of linker histone has been related with cell differentiation[273]. Murine erythroleukemia (MEL) cells are a model system for erythroid development[275]. As MEL cells differentiate, the binding affinity of all linker histone variants increases due to a progressive loss of CDK activity that leads to the accumulation of dephosphorylated H1 histone and an overall decrease in chromatin flexibility. Either the blockage or compensation of the loss of this CDK activity[276, 277] by adding small amounts of phosphorylated H1 leads to block cell differentiation[273]. This fact implies that the presence of phosphorylated H1 has a dominant effect in contrast with the unphosphorylated linker histones. It seems that phosphorylation reduces H1 affinity for chromatin, but not its immediate release. Then, the effect of H1 phosphorylation could be related with global structural changes of chromatin within the context of assembled chromatin[273].

During S and M phases of the cell cycle, phosphorylated H1 histone has been found to be translocated to the cytoplasm[278]. In neurodegenerative diseases such as Parkinson and Alzheimer diseases, activation of CDK5, a cyclin-dependent kinase similar to cdc2, has been determined. CDK5 is able to carry out H1 phosphorylation in the brain[279]. Thus, the increase of H1 phosphorylation mediated by CDK5 could explain the increase of the cytoplasmic H1 found in neurons and astrocytes which are affected in Alzheimer and Parkinson diseases[280]. Furthermore, cytoplasmic linker histone preferentially interacts with the Aβ-amyloid protein and with α-synuclein in its amyloid conformation, which form amyloid-like fibrils associated with the mentioned neurodegenerative illnesses. Hence, the abnormal phosphorylation of histone H1 and its translocation to the cytoplasm within the cell could explain the cell cycle entry of neurons in Alzheimer and Parkinson diseases[280].

Acetylation

Acetylation of linker histones occurs mainly in cells which proliferate rapidly[242]. Histone H1 is acetylated in the N-terminal residue and also in the ε-amino group of lysine residues.

N^{α}- acetylation of the protein is a common feature in most of eukaryotic proteins: the first methionine is deleted and the amino acid that then becomes the first one is acetylated. This modification occurs during translation and has no sequence specificity. However, in most cases the acetylated residues are serine (S), alanine (A), methionine (M), glycine (G) and threonine (T). N-terminal acetylation has been characterized *in vivo* in most of H1 subtypes[242,281]**.** In H5 and $H1^{0}$, both the acetylated and the unacetylated forms have been found[282,283].

Acetylation in the ε-amino group of lysine residues is quite common in core histones. In histone H1, this modification occurs in the globular domain, specifically in the residues that are thought to be part of the nucleosome binding site with DNA: K52, K64, K85 and K97[284]. In mouse and human, residues K51, K63 and K87, which are also coincident with the primary DNA binding site, are also acetylated[242]. The

acetylation of those residues can interevene in the binding with DNA by giving specific properties to the different subtypes in chromatin.

Methylation

As acetylation, methylation also occurs in cells which proliferate rapidly[242], in this case, in the lateral chain of arginine and lysine residues. Methylation is often found either coupled with acetylation or adjacent to already methylated residues[242].

In some cases methylation is associated with phosphorylation in what is known as 'methyl-phos switch'. This phenomenon has been found in H1e of HeLa cells in which the methylated lysine residue K25 is adjacent to the potentially phosphorylated serine residue S26. When K25 is methylated, it binds HP1 protein in this point. If S26 becomes phosphorylated, then this is interaction is blocked[281]. The methylation of lysine K25 is carried out by the Ezh2 methyltransferase, which has been associated to transcriptional repression[281]. Thus, methylation would be necessary for the interaction whereas phosphorylation would act as a stop signal of the interaction. A similar situation is found in lysine K26 of H1b present in mice fibroblasts[285].

ADP-ribosylation

ADP-ribosylation can include either covalent binding of 8 to 10 molecules of ADP-ribose per molecule of H1 or the non-covalent binding of ramified polymers up to 100-200 units per h1 molecule. The latter one is resistant to the action of strong acids, detergents, chaotropic agents and high ionic strength conditions[286]. The presence of these polymers has a direct effect on the net charge of H1: *in vitro*, poly-ADP-ribosylation lead towards chromatin relaxation but it does not displace H1 from linker DNA[287,288]. *In vivo*, it could play a role in protecting the DNA methylation pattern, because it appears to be necessary for the maintenance of the non-methylated state of CpG islands located in the promoter region of Htf9 gene of fibroblasts[289,290].

Deamidation

Deamidation is a chemical reaction in which an amide functional group is removed from an organic compound. In biochemistry, the reaction is important in the degradation of proteins because it damages the amide-containing side chains of the amino acids asparagine and glutamine, which result into aspartic and glutamic acid, respectively. Generally, it is a non-enzymatic process[291,292,293,294]. Deamidation reactions have been conjectured to be one of the factors that limit the useful lifetime of proteins and it proceeds much more quickly if the susceptible amino acid is followed by a small, flexible residue such as glycine whose low steric hindrance leaves the peptide group open for attack. The reaction also occurs rapidly at high pH (>10) and temperature.

Some studies suggest that deamidation increases with the age in proteins with low or no turnover rate, as it happens with histone $H1^0$, which is deamidated in asparagine N3[283].

Ubiquitination

Two lysine residues have been found ubiquitinated in histone H1: lysine K46 in H1c, H1d and H1e, and lysine K116 in H1a (N156). Taf(II)250 specifically ubiquitinates histone H1 and not the four core histones. Mutations in Taf(II)250 that reduced the enzymatic ubiquitination function reduced the expression of specific genes as well[294]. In the core histones, this modification is associated with an increase of the expression[135], but the effects of this modification in H1 are yet to be determined.

Formylation

This modification has been found in some lysine residues like K64, K85 and K97 (H1c residue numbering) and only in mouse tissues, but not in cell lines[242]. Its potential function and origin are unknown. Formylation can be catalyzed from formyltetrafolate or from formaldehyde released during the di-methylation of lysine residues by a specific enzyme, or by LSD1 dimethylase in the reverse reaction[295].

Histone H1 positioning within the nucleosome

Despite the key role of histone H1 in chromatin structure and dynamics, its location and interactions with nucleosomal DNA are still not elucidated[296]. It is generally accepted that the globular domain of the linker histone is internally located in the 30nm chromatin fibre[297,298] but its exact position within the nucleosome, still remains a subject of debate[299]. Nowadays, three hypothetical models concerning the binding and positioning of H1 histone in the nucleosome have been described(Figure 12)[300,301,222].

The first and more largely accepted model, proposed by Simpson in 1978[300], considers a symmetric location for the globular domain of the linker histone within the nucleosome dyad, in which the protein would be in direct contact with the DNA[142,300]. This model has been questioned, mainly because it would imply a strong interaction between the globular domain of H1 and the minor groove of the DNA, a fact that new experimental data seem to contradict[221,222].

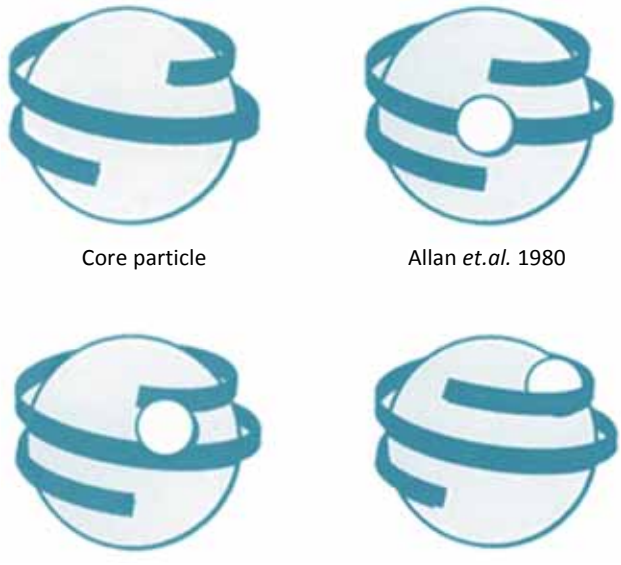


Figure 12. Hypothetical models of H1 histone positioning in the nucleosome. Adapted from Travers A. (1999)[292]

The second model, postulated by Hayes and Wolffe in 1993[301], proposes an asymmetric binding of H1 to the nucleosome and claims that it binds preferentially the DNA associated to the core particle rather than naked DNA. However, the presence of free linker DNA to either sides of H1 is required for this preferential binding. This additional linker DNA would be asymmetrically distributed in respect with the core particle. In this model, the globular domain of H1 would be located 65bp further from the nucleosome dyad, in the inner side of one of the DNA helical turns. This model has been object of criticism because it is based on the results obtained with nucleosomes bound to specific DNA sequences. This objection is questionable since linker histones only have partial sequence specificity for binding to DNA.

A third model, presented in 1998 by Zhou and known as 'bridging model'[222], hypothesizes that the globular domain binds the nucleosome in between of two helical turns of the DNA, forming a bridge between one terminus of chromatosomal DNA and the DNA in the vicinity of the dyad of the core

particle. This model could explain both the symmetric and the asymmetric protection of the linker DNA: the helix III of the globular domain would bind in the major groove of the first helical turn of the chromatosomal DNA, whereas the secondary DNA-binding site on the opposite face of the globular domain would contact with the nucleosomal DNA close to its midpoint.

Protein-DNA crosslinking experiments of core histones within the nucleosome with or without linker histone have revealed important changes in the contacts established between histones H2A and H3 and the DNA[302,303,304], suggesting that the binding of H1 could generate allosteric changes in both the octamer folding and its docking with DNA, resulting in the stabilization of DNA-protein interactions[300,305].

It is important to note that all the models that try to explain the positioning of histone H1 within the nucleosome base their conclusions in the structure of the nucleosome core particle, which could be different to that one of the chromatosome, due to the binding of H1.

A second question not yet resolved concerns the interaction and location of the linker histone C-terminal domain[306]. While the binding to the surface of nucleosomes is clearly known to be directed by the globular domain, the structure and interaction of the lysine-rich CTD, primarily responsible for the chromatin-condensing functions of linker histones, are poorly understood.

Lately, some experimental data obtained with mutant sets of the linker histone[306,307] had started to elucidate which are the interactions established between the C-terminal domain of H1 and the linker DNA.

Recent work provided evidence that two discontinuous regions of the CTD of approximately 24 amino acids played distinct roles in array condensation and the organization of linker DNA[229]. Interestingly, no specific amino acid composition and sequence appears to be important for these functions[229]. As previously mentioned, the CTD has been described as an intrinsically disordered protein[214,308,309] that adopts substantial α-helix structure in the presence of DNA, helix-stabilizing agents and other conditions that mimic the native chromatin environment[310]. Because of that, it has been proposed that the CTD would be similarly folded when bound to DNA and the nucleosome surface[306,307,311].

Analysis of the binding of a set of H1 CTD deletion mutants to nucleosomes containing various lengths of linker DNA indicate that the C-terminal domain of H1 interacts with regions of linker DNA at least 21bp from the edge of the nucleosome core. Furthermore, partial CTD truncations cause an increase in H1 binding affinity and intramolecular fluorescent resonance energy transfer (FRET) studies indicate substantial nucleosome-directed folding of the CTD in a different way from that which occurs when H1 binds to naked DNA[306].

FRET experiments reaffirms that the CTD is disordered when H1 is in solution: the distance between residues 101–173 is about 6.2nm while residues 101–195 are separated by an average distance of 8.0nm. In the nucleosome bound state, H1 CTD residues 101 and 195 are closer in space than residues 101 and 173 (Figure 13). These results support the idea that the CTD is folded into a specific structure upon nucleosome binding and provide important constraints to evaluate models for this structure[311]. It is important to say that the binding of H1 to naked DNA is cooperative whereas H1-nucleosome binding is not[307].

Moreover, analysis of fluorescence quenching showed that H1 residues at either end of the CTD have distinct environments when bound to nucleosomes, and lysine K195 in the CTD is juxtaposed between the two linker DNA helices, suggesting the formation of a stem structure in the H1-bound nucleosome[306].

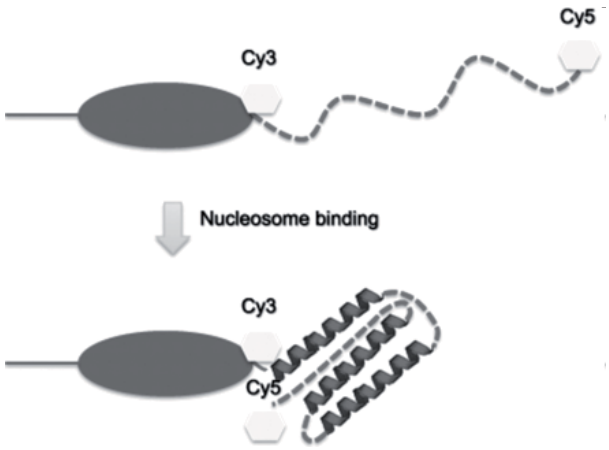


Figure 13. Model for CTD folding upon nucleosome binding in FRET experiments. From Caterino TL, *et.al.* (2011)[306].

The extent of quenching observed in CTD interactions with naked DNA suggests that a single DNA helix lies in the vicinity of residue 101 but when H1 is bound to the nucleosome, more than one helix is juxtaposed to residue K195[311]. This model is consistent with the C-terminal end of the CTD being located within the stem structure formed by both linker DNA segments, arising from the nucleosome along the dyad axis[143,296,312]. On the other hand, recent biophysical analysis and nano-scale modelling of the stem structure indicate a hierarchical structure in which the two linker DNA segments are closely positioned at about 20bp beyond the edge of the nucleosome core region, and intertwine over about 20bp to form a stem structure with a steep super helical pitch of about 120bp[313]. Quenching data suggest that position K195 might be located very near or perhaps even within this dual-helix intertwined stem structure, while position 101 would be located closer to the single super helical turn at the dyad[311].

However, all these results help to understand the interaction of H1 with the nucleosome in chromatin rather than knowing the precise structure of the H1 CTD in that environment. This issue has its origin in difficulties related to the preparation of well-defined nucleosomal samples[296]. Direct binding of linker histone to nucleosomes *in vitro* is inefficient and complicated because of the formation of large aggregates due to the nonspecific association of linker histones with DNA[314,315]. The situation can be considerably improved by using chaperones for linker histone deposition *in vitro*, a mechanism that is likely used *in vivo*[316].

The Histone Code and epigenetics

The histone code is a hypothesis that establishes specific histone-tail modifications, either acting in combination or sequentially in one or more histone, can modulate chromatin affinity of other proteins, DNA accessibility, chromatin condensation level and, eventually, gene expression (Figure 14). Post-translational modifications "crosstalk" with each other; this means that they are interdependent and can either facilitate or repress other modifications[87,317,318].

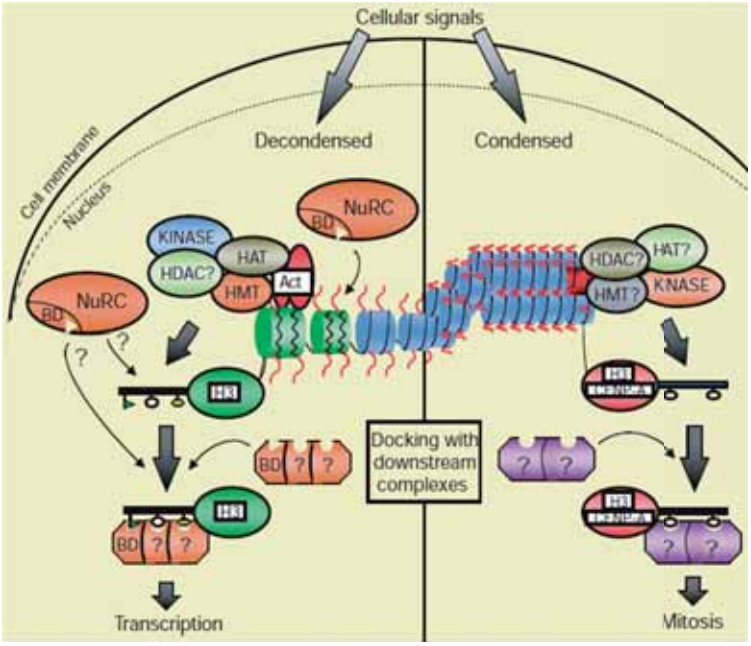


Figure 14. Coordinated recruitment of histone-modifying activities. Distinct histone-modifying activities interact to form multisubunit complexes that work in concert with nucleosome remodeling complexes (NuRCs; for example, Swi/Snf, RSC, NURF) to remodel chromatin. Besides nucleosome remodeling, NuRCs may chemically modify and/or bind histone tails. Binding of a NURC or HAT complex to histone tails may be mediated by the bromodomain (BD). Most of the complexes identified to date are implicated in events leading to transcription (left panel; green nucleosomes), but there are similar complexes that modify and direct chromatin condensation (right panel; red nucleosome). This `code' of modifications may dictate the biological outcome through changes caused in higher-order chromatin structure or may direct the downstream biological effect by recruiting and interacting with docking proteins or complexes that remain to be identified. Taken from Strahl BD&Allis CD.(2000)[87]

Increasing evidence indicates that the post-translationally modified histones serve as extremely selective binding platforms for specific regulatory proteins that drive distinct nuclear processes. Recent findings indicate that certain evolutionarily conserved domains found within specific regulatory proteins possess the ability to selectively bind a certain histone modification with very high affinity. In other words, specific regulatory proteins can 'read' the histone code to initiate DNA-templated programs (Figure 15)[60].

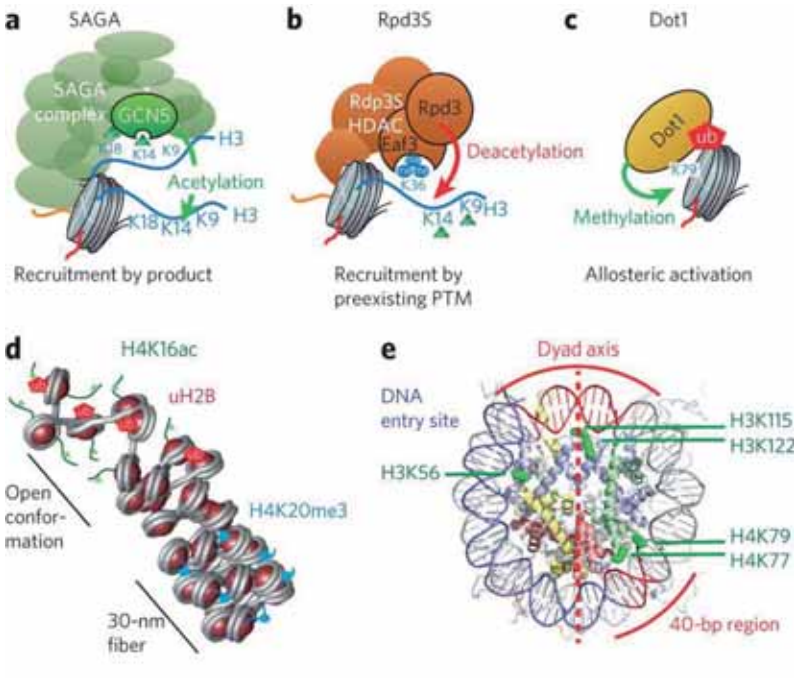


Figure 15. Control of chromatin function by histone marks through *cis*- and *trans*-acting mechanisms. (a–c) Examples of histone mark crosstalk **(a)** when a chromatin effector is recruited to chromatin by a specific mark and then proceeds to install additional copies of the same mark, as in the cooperative acetylation of chromatin by the SAGA complex, **(b)** when an effector is recruited by a mark that is distinct from the one it then acts on, as observed in the Rdp3S HDAC complex, **(c)** Allosteric crosstalk between marks as for the H3K79 methyltransferase Dot1, which is activated by H2B ubiquitylation. (**d**) Certain histone marks intrinsically alter the structural properties of chromatin fibers: H4K20me3 leads to the stabilization of a compact fiber structure, whereas H4K16ac and uH2B impair fiber compaction. (**e**) Histone marks can alter the stability of a single nucleosome: specific acetylation sites in the globular domain of H3 and H4 modulate nucleosomal stability and lead to increased DNA access to downstream effectors. Taken from Beat Fierz & Tom W Muir (2012)[318]

Hypoacetylated loci are typically silent while histone acetylation is important for gene activation. Acetylation of histone tail residues destabilizes the chromatin fibre, likely causing increased mobility of nucleosomes and better access of transcription factor to the DNA. For example, the bromodomain, a conserved motif found within certain transcription factors, binds to acetylated lysine residues on histone H3 and/or H4[319]. This specific binding is believed to result in the stabilization of the transcriptional machinery at the target regions, thus, enhancing transcription of this region. In contrast, regions lacking histone acetylation are unable to effectively bind these transcription factors resulting in transcriptional inactivation of the target region.

It is important to note that inactive regions also contain modified histones which likely help to maintain this inactivated state. For example, the methylation on lysine K9 of histone H3 was shown to occur specifically in inactive chromatin. It was also shown that this modification can recruit and bind heterochromatin protein 1 (HP1) via its evolutionarily conserved chromodomain[320]. This interaction

leads to the structural formation of compacted chromatin that physically inhibits the access of the transcriptional machinery to the underlying DNA[321].

Many other modifications, including phosphorylation, ADP-ribosylation, ubiquitination, and SUMOylation have been characterized and have important roles in setting the order of action and coordinating chromatin and DNA-dependent processes.

Recent studies in budding yeast documented a histone modification pathway associated with RNA polymerase II transcription: ubiquitination of histone H2B leads to methylation of histone H3 on specific lysine residues, such as H3K4 and H3K79[322]. Loss of H2B ubiquitination results in defects in cell growth, septation, and nuclear architecture. Chromatin immunoprecipitation assays demonstrated that loss of H2B ubiquitination alters the distribution of RNApol II and histones in coding regions. Thus, this suggests that ubiquitination of H2B affects transcription elongation and nuclear architecture through its effects on chromatin dynamics[323].

Allis and col. verified that lysine/threonine or lysine/serine pairs act like 'binary molecular switches' that modify histone-effector interactions[324,325.]. They propose that phosphorylation of an adjacent serine or threonine weakens the bond that the effector protein has with the methylated lysines.

It was also observed that post-translational modifications of histones appear frequently in high density, giving rise to 'modification hot spots' which are placed in strategic locations along the histone tail as a mark for different processes, such as gene silencing or gene activation. 'Cross-talk' relationships can occur in the same histone tails (*cis*) or across distinct histone tails (*trans*). It appears that these regulatory pathways determine not only gene expression but also chromatin function during DNA replication and repair, chromosome segregation (in both meiosis and mitosis) and chromatin disruption as cells undergo programmed cell death or apoptosis[326,327,328,329].

Most recently, a theoretical advance suggests that the 'code' may be tied to multiple PTM marks, in one or more histone tails, in one or more nucleosomes. Importantly, this hypothesis, known as the 'multivalency hypothesis', suggests that linked protein modules (PHD fingers, chromodomains, bromodomains, etc.) may work together in reading this histone code[330].

The most important epigenetic modifications are, in fact, histone PTMs and DNA methyaltion. DNA methylation is linked to histone modifications in a mutually-dependent relationship, each modification with its own role in modulating chromatin function; regardless, these molecular phenomena have an important role in regulation of genetic information. Dynamic changes in histone variants and alterations of post-translational modification patterns occur throughout the life of an organism and likely vary in each cell type. The sum of these modifications, variants, and DNA methylation form the molecular basis of epigenetic information. Therefore, epigenetics can be defined as information content that increases the complexity of the genome without changes in the gene sequence[331].

Due to its importance, the epigenetic information has to be duplicated and transmitted in each single cell cycle[332], meaning that the cell has to replicate not only the DNA but also the chromatin structure. The transmission of histone modifications through subsequent generations could be carried out as nucleosomes' inheritance, in which a semiconservative distribution of core histones within new nucleosome units ensures that each nucleosome carries pre-existent modifications that are able to promote the restoration of all original post-translational modifications[333,334,335,336,337,338].

Techniques for the molecular characterization and analysis of histones

Characterization of histone H1 secondary structure by FTIR spectroscopy

Infrared spectroscopy is one of the most used techniques for the determination of secondary structure of proteins, since it allows their analysis under a wide range of conditions[235,339,340,341, 404,407,408]. Infrared spectroscopy has its basis in the absorption of infrared radiation by the molecules. Infrared radiation is located in the frequency range of the electromagnetic spectrum between 14000 and 200 cm-^{1}, but only frequencies from 4000 to 1000 cm^{-1} have biological applicability.

Electromagnetic radiation of an incident light beam on a molecule generates a variation of its energetic state that leads to the vibration of the atomic nucleus with respect to nearby nucleus. Vibration of adjacent atomic nucleus provokes a change the rotational-vibrational movements of the atomic bonds and infrared light is then absorbed. The vibration of a concrete atomic bond depends on the mass of the atoms involved and the atomic force between them. In infrared spectroscopy of proteins, the peptide or amide bond is the one analysed.

Peptide bonds have two vibrational bands: amide-I, which corresponds to the stretching vibration of the C=O bond and generates a band located between 1700 and 1600cm^{-1} frequencies in the infrared spectrum; and amide-II, which corresponds to the bending vibration of the N-H bond and has its characteristic band between 1600 and 1500 cm^{-1} frequencies. There is also a third intense band located around 3300cm^{-1}, referred to as amide-A, also due to the bending vibration of the N-H bond[342,407,408].

Depending on the secondary structure of the protein, the vibrational bands of these bonds will change their location in the frequency range of the infrared spectrum. Secondary structure of proteins by infrared spectroscopy is based on the interpretation of the band peaks of the amide-I, since it is the one with higher signal[342,401] (Table 3).

Table 3. Secondary structure elements assignment in the amide-I region

Structural element	Frequency range (cm^{-1})
Turns	1700-1680
Turns	1680-1658
Turns	1658-1657
α-helix	1656-1647
Random coil	1647-1639
β-sheet	1638-1628
Low frequency β-sheet	1627-1617

Assignment of each structural element to its corresponding frequency range of the amide-I spectrum obtained by infrared spectroscopy (FTIR) of proteins in D_2O. Adapted from Arrondo *et. al.*[401]

In the present study, a Fourier-Transform spectrophotometer has been used for the infrared spectrometry studies (FTIR). These equipments are based in the Michelson interferometer, which consists of two highly polished mirrors, a stationary mirror (M1), and a movable one (M2). In Figure 16, a source S emits light that hits a a half-silvered mirror referred to as beam splitter (surface M in Figure 16), at point C. Since M is partially reflective, one beam is transmitted through to point B while the other is reflected in the direction of A. Both beams recombine at point C' to produce an interference pattern (interferogram) visible to the observer at point E. The intensity as a function of wavenumber can be recovered from the interferogram by an inverse Fourier transform.

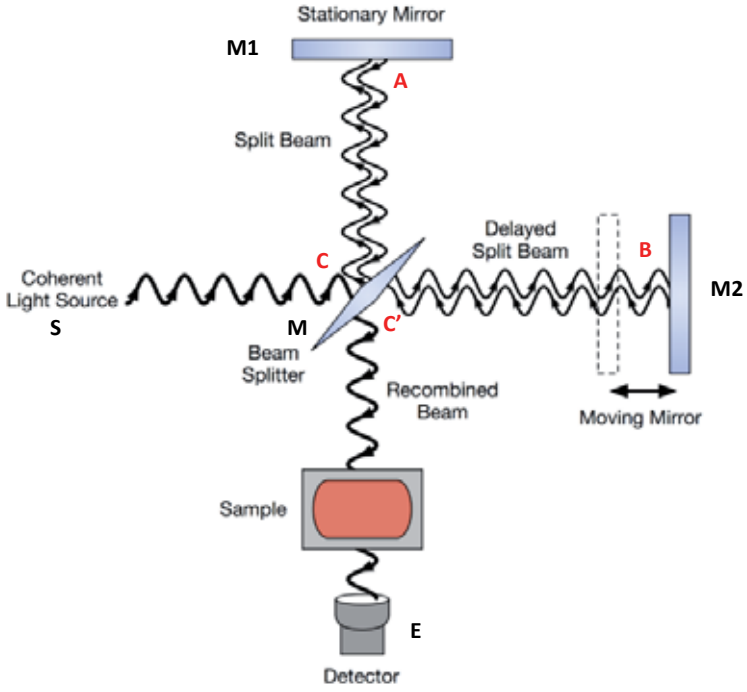


Figure 16. Diagram of a basic Michelson interferometer

Due to the vibration of the O-H bond, water generates two characteristic bands, one at 3400cm^{-1}, and the other one at 1670cm^{-1}. The latter, thus, will mask the amide-I band and counteract the obtention of the information about the secondary structure of the protein. Because of that, water in the sample is usually interchanged with deuterated water (D_2O)[235,340,343,404]. Deuterium atoms are heavier than the hydrogen ones and their bonds vibrate at lower frequencies (2400cm^{-1} and 1200 cm^{-1}, respectively), and thus, they do not interefere in the amide-I band.

FTIR can be applied to the analysis of solids, liquids, and gasses. Because of that, this technique is the appropriate to study the secondary structure of histone H1 in the complexes with DNA, which precipitate upon interaction.

Separation of intact histones, histone variants/subtypes and modified species

High Performance Capillary Electrophoresis (HPCE)

HPCE is a powerful separation technique that combines the advantages of conventional gel electrophoresis and HPLC to separate solutes according to their ability to migrate in an electric field inside a fused-silica capillary[344,345] (Figure 17). HPCE had permitted, over the past 20 years, to overcome the limitations related to common electrophoretic methods, as well as to achieve efficient separations, quantification of the sample and automation of the process. One of the major challenges of histone analysis by HPCE has been the strong interaction of the basic histone tails with the surface of the fused silica, which results in peak tailing, loss of resolution and reduced sensitivity due to adsorption of the histone proteins, which are highly basic, by the negatively charged silanol groups of the capillary surface[345,344]. Coating of the capillary surface with hydrophobic modifiers, such as methylcellulose or hydroxypropylmethylcellullose (HPMC), and the use of low pH phosphate buffers to avoid protein-wall electrostatic interactions[346,347,348,349,350,351] had remarkably improved the method. Other strategies, based on introducing some degree of cation exchange behaviour to the capillary surface had also been developed[352,353].

Eventually, buffer composition has been demonstrated to be a determinant factor for proper separation of histone variants of both core and linker histones. According to their sequence variability and evolutionary divergence, the CE methods developed for histone separation of a definite species can result inappropriate for histone analysis of any other species[349,345]. The influence of buffer concentration and composition on the migration order and resolution has been broadly investigated[347,348,349,350].

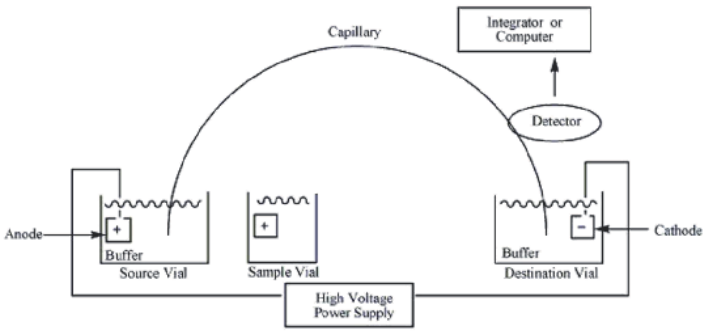


Figure 17. Diagram of a CE intrument

HPCE has been demonstrated to be a well-suited technique for resolving histone variants as well as modified isoforms, especially by acetylation and phosphorylation. Those two modifications greatly affect the protein net charge resulting in the retardation of the migration time along the capillary of the modified species[348,354,350]. Furthermore, histone analysis by mass spectrometry can be coupled with HPCE[351].

Liquid chromatography-based methods

Reverse-phase Liquid Chromatography (RP-HPLC)

RP-HPLC, which separates molecules on the basis of hydrophobicity, is a high-resolution method ideally suited for histone proteins. Acid-extracted or salt-extracted histones can be purified in one run on a standard C_8 or C_{18} column, using an acetonitrile gradient[355]**¡Error! Marcador no definido.** . This fractionation method is generally used for the initial separation of intact histones and histone variants.

This easy first-step is often preferred over other techniques, such as electrophoretic ones, because of several advantages: (1) histones are small hydrophilic proteins that are very soluble and histone separation by RP-HPLC can be baseline resolved; (2) the eluted histone proteins are dissolved in aqueous, salt-free solvents, which are suitable for downstream analysis; (3) The LC can be easily coupled on-line to MS for detection and mass measurement of intact protein mixtures[355]. Reversed-phase separations were extensively studied and improved in the 1980s, primarily by Gurley[356,357,358] and Lindner[359,360,361,362], achieving a high degree of sample recovery and better resolutions. However, the capacity of RP-HPLC to separate histones is limited[363].

RP-HPLC resolution is based in the hydrophobic interactions of the column material with the hydrophobic regions of the protein. In the case of linker histones, the less conserved and highly polar tails are the main responsible for their variability and these regions are also the ones that accumulate the vast majority of modifications. For this reason, RP-HPLC has often not allowed the preparative separation of separating H1 species, making necessary the development of a chromatographic method based on other kind of interactions[355].

Hydrophilic-Interaction Liquid Chromatography (HILIC)

Considering the physicochemical properties of H1 histones, HILIC, a chromatographic technique based on the hydrophilic interactions to separate solutes by utilization of polar stationary phases, was proposed by Alpert in 1990[364]. This chromatographic technique allows protein separation by decreasing the ionic interactions via an increase of the presence of an organic solvent (generally % acetonitrile) in the aqueous phase. Thus, isolation occurs by first separating those variants with lower polarity. Furthermore, the increasing presence of non-polar organic solvents increases even more the hydrophilic interactions of proteins with the column, promoting their retention. However, HILIC methods are based on a non-volatile mobile phase additive, such as chaotropic agents, to improve the resolution or increasing salt gradients for the elution of different histone species. Thus, despite HILIC happens to be the most effective separation of histone proteins, it cannot be directly coupled with other downstream techniques, such as mass spectrometry, which is, as it will be explained later, the most effective for sequence analysis and separation of histones[355].

The fact that HILIC and RPLC have different separation mechanisms suggests a possible use of both for two-dimensional separations (2D-LC) with high orthogonality.

Mass spectrometry

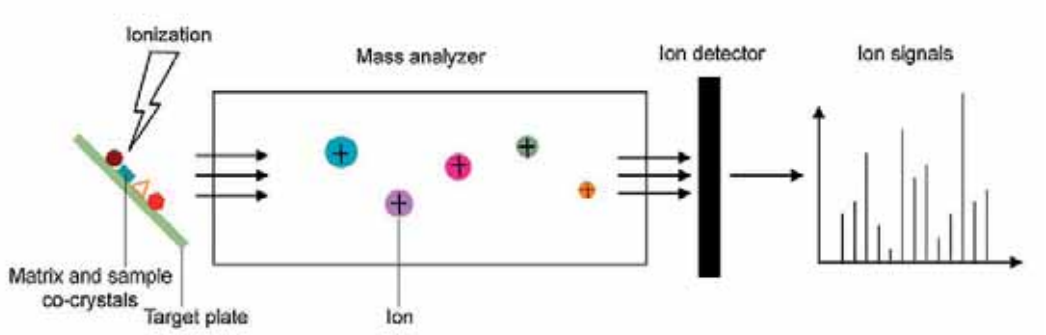
Mass spectrometry (MS) is a physicochemical analysis technique that determines the mass-to-charge ratio (m/z) of gas-phase ions[344]. In the recent years, MS has gained a lot of popularity and attention due to its high sensitivity and accuracy and its ability to sequence proteins and peptides, as well as to perform large-scale analysis. MS is one of the basic tools in proteomics[355,365], and its use comprises mainly four areas[366]: (1) characterization and quality control of recombinant proteins and other macromolecules; (2) protein identification; (3) determination of the molecular weight of a protein and (4) characterization of post-translational modifications and covalent modifications that alter the mass of a protein.

Since most of histone functions are regulated via specific post-translational modifications and their cross-talk, a wide range of MS-based methodologies have been developed for the study and characterization of histones and their PTMs, in order to understand the roles of each histone form in the regulation of chromatin.

Protein characterization by MS-analysis

MS analysis consists on ionization of molecules to generate charged derivatives or ionized molecule fragments and measuring their mass-to-charge ratio (m/z). Classical MS procedure starts with the sample loading onto the mass spectrometer, where it undergoes vaporization by an ion source and components of the sample are ionized resulting in charged particles. Next, the ions entry the mass analyser and are separated by electromagnetic fields according to their mass-to-charge ratio (m/z) and measured in a detector. Finally, the ion signal is processed into a mass spectrum. The ionization process is achieved by two methods (Figure 18): Matrix-Assisted Laser Desorption/Ionization (MALDI) and Electrospray Ionization (ESI).

(A) MALDI-MS



(B) ESI-MS

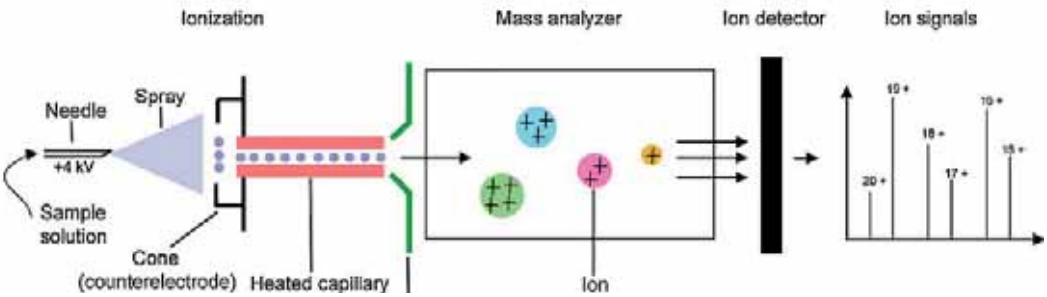


Figure 18. MALDI-MS and ESI-MS procedures. (A) In MALDI-MS, samples are co-crystallized with an organic matrix on a metal target plate. A pulsed laser irradiates the co-crystals, which causes rapid heating and desorption of ions into the gas phase. Ions go through the mass analyser and the detector registers the numbers of ions at each individual mass-to-charge (m/z) value, then the peptide mass fingerprint is generated. MALDI-MS produces relatively simple spectra composed of ions with unit charge. **(B)** In ESI-MS, sample molecules are ionized directly in the analyte solution by passing through a heated capillary device, spraying droplets of solution into a vacuum chamber containing a high-strength electric field. The resulting ions pass through a mass analyser and detector as in *a*. ESI-MS produces complex spectra with multiply charged ions.

MALDI was developed in 1990 by Karas and Hillenkamp[367], and it is based in the co-crystallization of the molecule to be analysed with a volatile organic matrix. Laser excitation is used to desorb/ionize both the matrix and the analyte, generating charged ions. Signal intensities depend on incorporation of the peptides into crystals, their ability for capturing/retaining a proton during the desorption process, and many other factors such as signal suppression effects in peptide mixtures. Because of that, it is difficult to relate peptide peak height (intensity) with the quantity of sample present unless an internal standard is used. Furthermore, the mass range below 500Da is often masked by matrix-related ions[366].

Electrospray mass spectrometry (ESI-MS) was developed in the late 1980s for use in biological mass spectrometry by Fenn *et. al.* [368] .In this ionization process, a fine spray containing the analyte is pumped through a high voltage needle, producing charged droplets. The charged droplets are desolvated before entering the high vacuum of the mass spectrometer to avoid their interfering in the analysis. Formation of ions by electrospray produces multiply charged ions resulting in the observation of a charge envelop distribution of ions in the mass spectrum[365].Furthermore, for a given compound, the signal strength (peak height in the spectrum) increases linearly with the analyte concentration until saturation occurs. Large ions are typically multiply charged, which brings them into the range of mass-to-charge (*m/z*) ratios. Electrospray can be performed in either the infusion mode, the nanoelectrospray format, or in combination with high-performance liquid chromatography[366].

MALDI is mostly applied to the analysis of peptides because proteins generally undergo fragmentation to some extent during the ionization process, resulting in broad peaks and loss in sensitivity. For this reason, MALDI-MS is generally preceded by enzymatic cleavage of the protein to analyse, in order to generate peptides that can be more accurately detected by the mass spectrometer. MALDI has the advantages of speed, sensitivity, and high mass accuracy for peptides, whereas ESI-MS offers higher protein mass accuracy due to greater mass resolving power for protein ions. This means that ESI-MS is the technique of choice if intact proteins had to be analysed. Nevertheless, its use in the analysis of peptides is often performed, since ESI is more easily coupled with previous liquid-based separations (such as HPLC, HILIC or HPCE). Liquid chromatography separations are often used before MS-analysis to reduce suppression effects and/or low ionization efficiency. HPLC eluent can be either directly analysed on-line using ESI-MS or HPLC fractions can be collected and analysed individually by MALDI-MS or ESI-MS (off-line MS)[365,344].

Characterization of proteins by MS-analysis is generally used for the determination of the molecular weight of a protein or different proteins in a complex mixture and thus, identification and characterization of the different proteins or variants and/or determination of the changes of modified isoforms (i.e. Natural variants, isoforms coming from different cell lines, changes due to drug treatment, etc.) or post-translational modifications that alter the mass of a protein. In those cases, MS-analysis is generally coupled with previous liquid-based separations (LC-MS), either on-line or off-line. In the case of histone proteins, many groups had successfully developed and improved those experimental procedures[369,370,371,372,373,374,375,376,377,346,347,378,379,350,353,351].

MS has also been broadly used for peptide mass fingerprint (PMF; also named peptide mass mapping, PMM). This technique is used to identify proteins after digestion by comparing the masses of the observed peptides with the ones of the theoretical protein digest. In PMF, intact proteins are submitted to enzymatic cleavage using appropriate proteolytic enzymes. The experimental peptide masses are determined by MS analysis and searched against theoretical protein digests generated from an appropriate protein database. Matches between the observed and theoretical peptides/proteins are scored. The highest scoring protein is then taken to be the protein undergoing characterization[344].
PMF is also a very efficient method for the identification of PTMs. Each PTM introduces a specific mass shift relative to theoretical mass of an unmodified protein. The location and number of modifications can be determined by the mass shift of the modified peptide[344].

MS analysis of proteins is, in general, not very quantitative. Despite that, relative changes due to PTMs on a specific protein can be determined. The direct comparison of ion abundances between different spectra cannot be made, since ionization of the sample might be different among different runs. However, a comparison of relative abundances of the levels of each isoform within one spectrum is possible. The addition of negatively charged modifications on a protein (i.e., phosphorylation) would reduce its ionization potential in the mass spectrometer. The addition, therefore, of one or two

negatively charged phosphate groups per molecule is not expected to significantly affect the ionization potential of the protein, and this is not expected to significantly affect the relative sensitivity of the differentially phosphorylated protein ions[355].

Liquid chromatography MS (LC-MS) is a sensitive, accurate and fast method that can be used for intact protein analysis. Peptide mass finger-printing (PMF) is an efficient method to map histone PTMs using mass correlation. When it is necessary to locate a PTM or obtain sequence information, experiments by tandem MS (MS/MS) can be performed[344].

Tandem MS

Tandem MS experiments consist of, at least, two stages of mass analysis and it allows the determination of the amino acidic sequence of peptides by interpreting the results obtained from the mass spectrometer (Figure 19). In the first stage of an MS/MS experiment, the peptide species of interest are selected out of a mixture and dissociated by multiple low-energy collisions in the first mass spectrometer. The resulting fragments are separated in the second stage of the experiment, in which the mass-to-charge ratio (m/z) of the product ions after fragmentation is measured. A tandem mass spectrum or MS/MS spectrum is obtained[344,366].

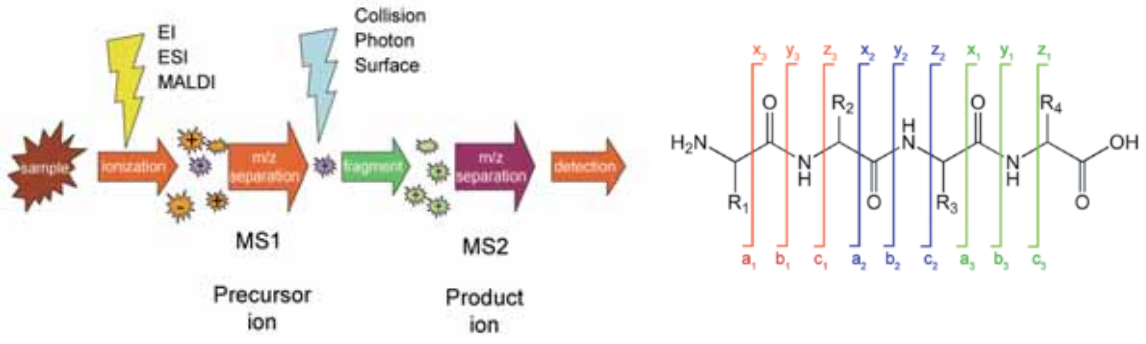


Figure 19. Tandem Mass Spectrometry. (A) Scheme of a Tandem MS experiment. **(B)** Peptide fragmentation notation using the scheme of Roepstorff and Fohlman (1984)[380]

During peptide fragmentation, several bonds along the protein backbone can be broken. Peptide fragments that arise from a tandem mass spectrum are indicated by *a, b*, or *c* if the charge is retained on the N-terminus and by *x, y* or *z* if the charge is maintained on the C-terminus[380]. The most common ion types are the *b* ions and the *y* ions, which denote fragmentation at the amide bond with charge retention on the N-terminus (*b* ions) or the C-terminus (*y* ions), respectively (Figure 19). Neutral losses of ammonia and/or water, in addition to the backbone fragmentation are also useful for sequencing and peptide identification. Also, most proteomics experiments are performed with tryptic peptides, which have arginine (R) or lysine (L) residues as their C-terminal residues. In this case, *y* ions are the predominant ion series observed[366].

Furthermore, by monitoring the relative abundances of precursor or fragment ions generated during the first stage of the MS/MS experiment, relative quantitative changes of corresponding isoforms can be elucidated or deduced in the second stage. Then, in tandem MS, data contain molecular weight information on the peptide species and their amounts. The first mass analyser scans all the masses of the precursor ions. The second mass analyser scans the masses of the product ions selected in MS1. This offset corresponds to a neutral loss that is commonly observed for the class of compounds.

Nano-LC/MS/MS is the ideal approach when the sample is a complex mixture and the purification of individual proteins is challenging[344]. Liquid chromatography (LC) coupled to tandem mass spectrometry (LC-MS/MS or LC-MS), combines efficient separation of biological materials and sensitive identification of the individual components by mass spectrometry. The analyte is first eluted from a reversed-phase column to separate the peptides by hydrophobicity; then, it is ionized and transferred into the mass spectrometer for analysis. As said above, ESI is, generally, the ionization method of choice[366].

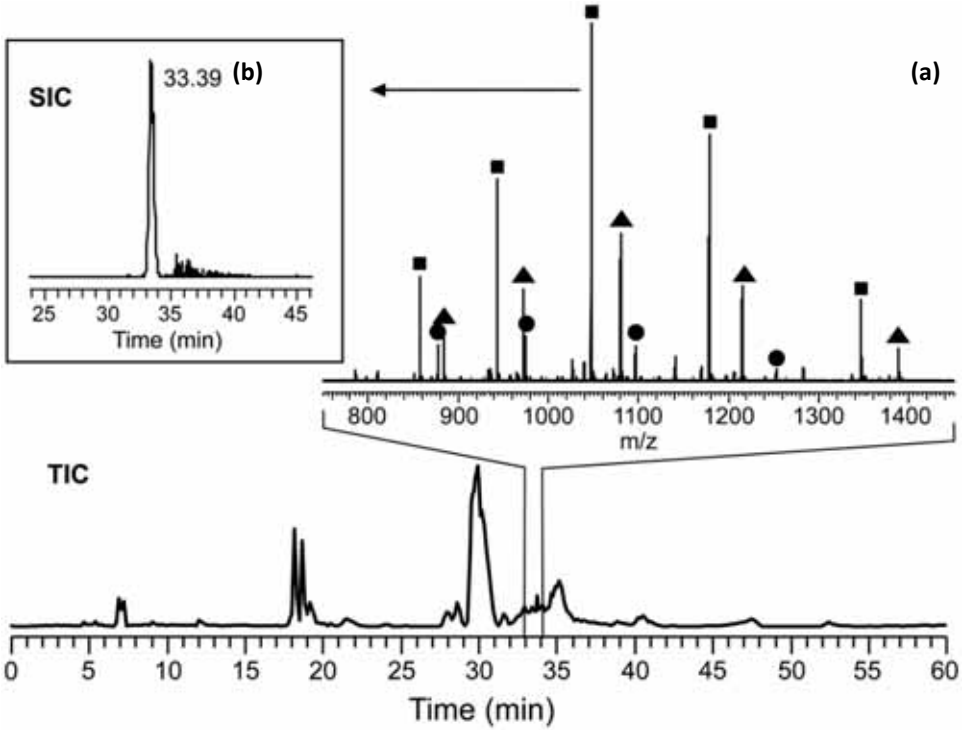


Figure 20. Total ion current (TIC) chromatogram from LC-MS analysis of HDL3. (Inset) Selected ion chromatogram (SIC) for a specific range of time **(a);** and a specific m/z **(b).** Taken as an example from Mazur MT et al. PNAS 2010;107:7728-7733

For each scan, the ion current is plotted as a function of time (TIC: total ion current chromatogram). Afterwards, data can be asked to reveal the ion current at a particular m/z ratio or a single mass spectrum for a selected time (SIC: selected ion current), which will show all the peptides that co-elute at a precise moment in the same ion current (Figure 20). Database search engines such as Mascot, SEQUEST and X! Tandem are valuable software tools for data interpretation.

Analysis of PTMs by Tandem MS

In tandem MS, fragmentation of the peptide produces ions that not only indicate the peptide sequence, but also permit identification of PTMs. Neutral losses and immonium ions are common fragmentation products that have great diagnostic significance. For example, the loss of 59Da is observed for trimetylated lysine residues. For low mass-accuracy instruments, this neutral loss is useful to differentiate between trimetylation and acetylation[381]. Phosphorylation is conveniently detected by mass spectrometry due to a mass increase of 79.966 Da of peptides upon addition of one phosphoryl group, and often also a neutral loss of H_3PO_4 (− 97.982 Da, pSer/pThr) or the presence of an immonium

ion (formed due to peptide fragmentation into amino acids with a structure of $^{+}$NH2=CH-R) in MS/MS spectra[344,355].

The most important consideration in characterizing PTMs is their frequency of occurrence. Since most PTMs are rare events, specific biochemical enrichment strategies need to be used to facilitate their identification and characterization[344,382]. The population of a given peptide sequence is divided in differently modified species during MS, decreasing the abundance and, hence, detectability, of some of the peptide populations of interest. Characterization of some PTMs, such as ubiquitination and phosphorylation, requires more material due to the low stoichiometry of modification and the relatively small amount in comparison to the non-modified peptide[383]. Enrichment of modified peptides can be achieved using immunoaffinity-based strategies or chemical affinity techniques, i.e.,phosphopeptidesare often enriched by immobilized metal-ion affinity chromatography (IMAC) Mass spectrometric analysis can also be targeted to peptides with defined PTM specific fragmentation characteristics, for example by looking for peptides demonstrating neutral loss of labile PTMs (such as phosphoric acid or sulphate), or by performing precursor ion analysis of peptides generating characteristic fragment ions.

Strategies for histone analysis by mass spectrometry

Histones show a high degree of complexity because of the great amount of subtypes, the similarity of their amino acid sequences and the diversity of the PTMs. Therefore, histone characterization demands very specific and sensitive analysis tools. Three main MS-based strategies for histone characterization have been developed: 'bottom-up', 'top-down' and 'middle-down' approaches, respectively[344,355] (Figure 21).

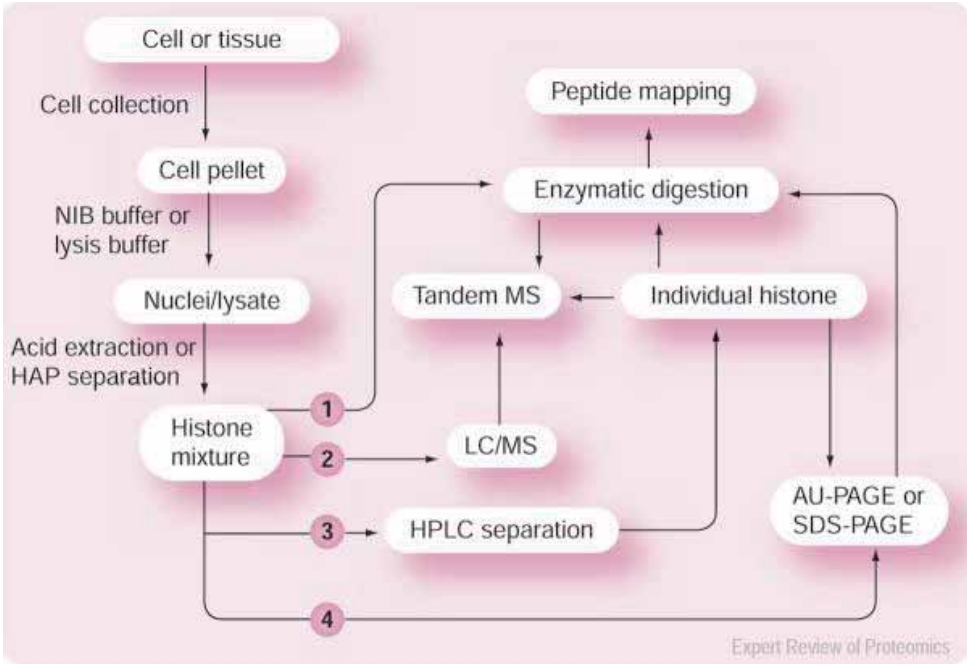


Figure 21. General flowchart of an experimental design for histone characterization 1. Shotgun proteomics; 2. Top-down proteomics; 3. Middle-down proteomics; 4. Bottom-up proteomics. Extracted from Su X, *et.al.* (2007)[344]

The **'bottom-up'** approach (Figure 21, 4) is the most common method to identify proteins and characterize their amino acid sequences and post-translational modifications. It is based in proteolytic

digestion of proteins prior to analysis by mass spectrometry[384]. The proteins may first be purified by gel electrophoresis resulting in one or a few proteins in each proteolytic digest. Alternatively, the whole protein extract can be digested directly, followed by a by liquid-based separation coupled to mass spectrometry. This specific approach of 'bottom-up' proteomics is known as **shotgun proteomics**[385] (Figure 22, 1). The smaller peptides of about 7–25 amino acids in length, are suitable for sequencing by MS/MS and, by comparing the masses of the proteolytic peptides or their tandem mass spectra with those predicted from a sequence database, peptides can be identified and multiple peptide identifications assembled into a protein identification. However, by using this approach the knowledge of the global status of the PTMs along the protein backbone is lost and it is difficult to map multivalent, co-existing PTM patterns[355,365].

'Top-down' proteomics (Figure 22, 2) for protein identification is based on characterization of the intact protein by mass spectrometry[386]. This approach allows for mapping of co-existing PTMs, even those that are located far apart within the histone sequence. The top–down strategy often results in complex MS and MS/MS spectra, due to the higher charge states of intact protein ions generated during the ionization process. Proteins are typically ionized by ESI and trapped in a Fourier transformed ion cyclotron resonance (FT-ICR) or quadrupole ion trap (LTQ) mass spectrometer. Fragmentation for tandem mass spectrometry is accomplished by electron-capture dissociation (ECD) or electron-transfer dissociation (ETD)[355,344]. In order to obtain better results and minimize the excess of ions in the MS/MS spectra of the intact protein, this approach is often coupled to previous LC-separations.

Eventually, the **'middle–down'** approach (Figure 22, 3) is a suitable combination of the bottom–up and top–down methods. Complex protein mixtures require previous separation before MS analysis. This can be crucial in the case of histone proteins, since all histones in the original mixture have similar physico-chemical properties and a high sequence similarity. Isolated histone proteins are then submitted to partial enzymatic cleavage by specific proteases, such as ArgC or Trypsin, generating larger polypeptides that contain several PTMs and that can be purified and/or analysed by MS and MS/MS[355]. 'Middle-down' approach has shown to remarkably reduce technical shortcomings of intact histone analysis while preserving a semi-global overview of coexisting modifications[387].

Mass spectrometry coupled to liquid chromatography (nanoLC-ESI-MS) is the dominant technique for proteomic characterization; and it requires previous enzymatic cleavage of the protein into smaller peptides. Due to the properties of histone proteins, the analysis of histone peptides causes the formation of many short and polar peptides that poorly interact with the RP material and go undetected by conventional LC-ESI-MS[388]

Lately, a new proteomic approach has been successfully developed to overcome this disadvantadge: capillary-electrophoresis electrospray-ionization mass spectrometric (CESI-MS)[388,389] In CESI-MS, the separation of the peptides by m/z is based in their electrophoretic mobilities, not in their hydrophobic nature, which makes it a powerful method in separating individual histones and their modified forms and represents a promising alternative to histone PTMs characterization.

Molecular characterization and analysis of chromatin by optical biophysics techniques

Chromatin structure and dynamics can be broader studied by several techniques. In the present study, chromatin compaction and aggregation mediated by an increase of salt concentration in the medium was studied and analysed by two different optical biophysics techniques light scattering based.

Dynamic Light Scattering (DLS)

Dynamic Light Scattering (sometimes referred to as Photon Correlation Spectroscopy or Quasi-Elastic Light Scattering) is a technique for measuring the size of particles typically in the sub-micron region. This technique measures the diffusion of particles moving under Brownian motion -defined by a property known as the translational diffusion coefficient (D)-, and converts this to size and a size distribution using the Stokes-Einstein relationship;

$$d(H) = \frac{kT}{3\pi\eta D}$$

where *d(H)* is the hydrodynamic diameter; *D* is the translational diffusion coefficient; *k* is the Boltzmann's constant; *T* is the absolute temperature; and η is the viscosity[390,391].

The diameter that is measured in DLS is a value that refers to how a particle diffuses within a fluid so it is referred to as a hydrodynamic diameter (d.nm). The diameter that is obtained by this technique is the diameter of an equivalent sphere. This means, a sphere that has the same translational diffusion coefficient as the particle. If the shape of a particle changes in a way that affects the diffusion speed, then the hydrodynamic size will change. Nevertheless, the translational diffusion coefficient (D) will depend not only on the size of the particle, but also on any surface structure, as well as the concentration and type of ions in the medium.

DLS data analysis: the correlation function

When a monochromatic light source hits small particles, the light scatters in a time-dependent fluctuation in the scattering intensity (Figure 22). This is due to Brownian motion of particles in solution, since the distance between them is constantly changing with time. This scattered light then undergoes either constructive or destructive interference by the surrounding particles, and within this intensity fluctuation, information is contained about the time scale of movement of the scatterers[392,393,394].

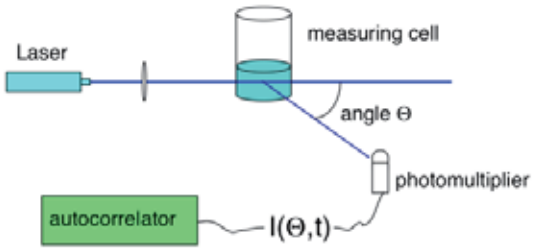


Figure 22. DLS equipment scheme

Scattered light from the samples is back-collected and the intensity correlation function is used to generate a correlation curve. Translational diffusion coefficients (D) are obtained from the correlation function of the scattered intensity, (G(τ)), defined as an exponential decaying function of the correlator time delay (τ);

$$G(\tau) = A\left[1 + B\exp\left(-2\Gamma\tau\right)\right]$$

where *A* is the intercept of the correlation function; *B* is the baseline of the correlation function; and *Γ* is the decay rate, defined as;

$$\Gamma = Dq^2$$

where *D* is the translational diffusion coefficient and *q* is the scattering vector[392,393,394].

At short time delays, the correlation is high because the particles do not have a chance to move to a great extent from the initial state. As the time delays become longer, the correlation decays exponentially, meaning that, after a long time period has elapsed, there is no correlation between the scattered intensity of the initial and final states. This exponential decay is related to the motion of the particles, specifically to the diffusion coefficient (D). With the diffusion coefficient known, the hydrodynamic diameter can be calculated by using a variation of the Stokes-Einstein equation. For a polydisperse sample this curve is a sum of exponential decays. To fit the decay, numerical methods based on calculations of assumed distributions are used[392,393].

The first order result from a DLS experiment is an intensity distribution (%) of particle sizes (nm). The intensity distribution is naturally weighted according to the scattering intensity of each particle population, and this particle scattering intensity is proportional to the square of the molecular weight. Thus, the intensity distribution can be somewhat misleading, in that a small amount of aggregation or presence of larger particle species can dominate the distribution. However this distribution can be used as a sensitive detector for the presence of large material in the sample[395].

Furthermore, the intensity distribution (%) can be converted to a volume distribution (%) using Mie theory. The volume distribution describes the relative proportion of multiple components in the sample based on their mass or volume rather than on their scattering (Intensity). When transforming the intensity distribution to a volume distribution, four assumptions are made: (1) particles are spherical; (2) particles are homogeneous; (3) optical properties of the particles are known –i.e. the refractive index; and (4) there is no error in the intensity distribution[391].

Understanding these assumptions is particularly important since the DLS technique itself produces distributions with inherent peak broadening, so there will always be some error in the intensity distribution. Thus, derived distributions can be used for comparative purposes or for relative proportions estimation but should never be considered as absolute[395]. The same thing has to be considered for the values of the hydrodynamic diameters obtained as an output. Diameter sizes in a DLS experiment are reffer to the diameter of an equivalent sphere. Hence, in the case that the molecules in the sample have not a spherical shape, the values will be an approximation that would consider all the possible orientations of the molecules in solution.

Another important parameter in DLS measurements is the polydispersity index (PdI). It is calculated from a Cumulant analysis of the DLS-measured intensity autocorrelation function and describes the width of the assumed Gaussian distribution. PdI values greater than 0.7 indicate the sample has a very broad size distribution and that probably, it is not suitable for the dynamic light scattering (DLS) technique[395].

DLS applications

DLS is used to characterize size of particles such as polymers, micelles, proteins, carbohydrates and nanoparticles. If the system is monodisperse, the mean effective diameter of the particles can be determined. This measurement depends, as said above, not only on the size of the particle, but also on the surface structures, the concentration, and the ions in the medium[391].

The conformation of proteins and macromolecules is usually dependent on the exact nature of the dispersing medium. As conformational changes will usually affect the diffusion speed, DLS is a very sensitive technique for detecting these changes.

Since DLS essentially measures fluctuations in scattered light intensity due to diffusing particles, the diffusion coefficient of the particles can be determined by the correlation function. DLS software of

commercial instruments typically displays the particle population at different diameters. If the system is monodisperse, there should only be one population, whereas a polydisperse system would show multiple particle populations[392,393,394].

Stability studies can also be done by DLS analyses. Periodical DLS measurements of a sample can show whether the particles aggregate over time by seeing an increase of the hydrodynamic diameter and the emergence of a larger population of particles with a larger radius[390].

Laser diffraction

Laser diffraction analysis or laser diffraction spectroscopy is a technique which utilizes properties of the diffraction patterns of a laser beam passed through a substance or a dispersed particulate sample, to measure the size of particles (µm). This data is then analysed to calculate the size of the particles that created the scattering pattern (Figure 23)[396].

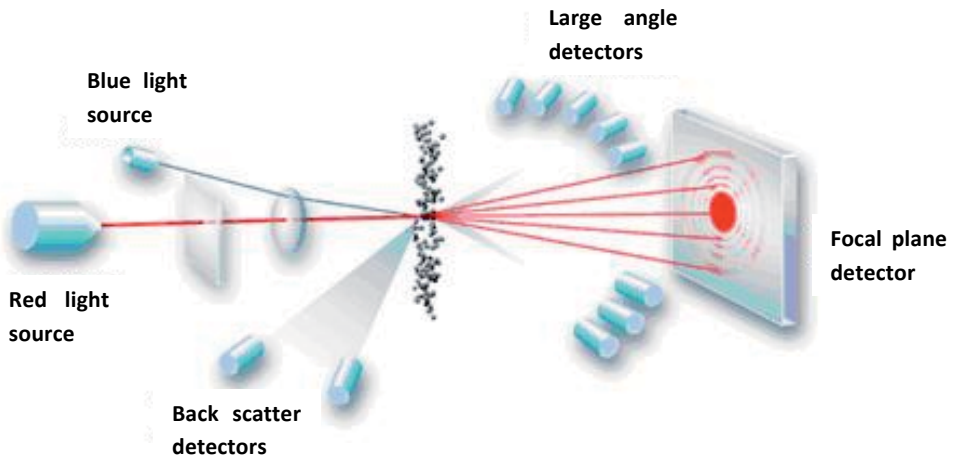


Figure 23. Laser diffraction equipment scheme

Laser Diffraction data analysis

Laser diffraction analysis is based on the Fraunhofer theory of diffraction, which states that the intensity and angle of the light scattered by a particle is directly proportional to the size of the particle[397]. The substance being examined is passed through the laser, and the diffracted light is focused onto a detector which measures the angular distribution of the intensity of the scattered light[398]. As the particle size decreases, the observed scattering angle increases logarithmically.

The Fraunhofer Approximation was used in early laser diffraction instruments, mainly because it is simpler to calculate and does not require knowing the optical properties of the sample. It assumes that: (1) particles being measured are opaque discs; (2) light is scattered only at narrow angles; (3) particles of all sizes scatter light with the same efficiency; (4) the refractive index difference between the particle and surrounding medium is infinite[390,394].

Nowadays, most of the Laser Diffraction equipment uses the Mie theory of light scattering to calculate the particle size distribution, assuming a volume equivalent sphere model. Mie theory, in contrast to The Fraunhofer Approximation, requires knowing the optical properties (refractive index and imaginary component) of both the sample and the dispersant. Mie Theory is based on Maxwell's electromagnetic

field equations and it assumes that: (1) the particles being measured are spherical; (2) the suspension is dilute; (3) the optical properties of the particles and the medium are known; (4) the particles are homogeneous[390,398,396].

Laser Diffraction applications

Laser diffraction is a widely used particle sizing technique for materials ranging from hundreds of nanometers up to several millimeters in size. Thus, it allows particle size determination of larger molecules than the DLS analysis. Nevertheless, since it is based in the Mie Theory, distribution of sizes is given as a Volume distribution (%).

The measurable size ranges from 50nm to 1000µm. The limitations of this method appear for small sizes (< 1µm, Mie theory), for non-spherical particles, and for materials with a low refractive index with respect to the dispersive medium[399].

Objectives

In this PhD thesis, we studied the effects of phosphorylation by CDK2 on histone $H1^0$ secondary structure under different conditions; post-translational modifications of chicken erythrocyte linker histones and the effects of phosphorylation on chromatin aggregation. The thesis was developed according to the following specific objectives:

1. To study the secondary structure and the effect of phosphorylation of the entire $H1^0$ in aqueous solution, bound to DNA and in the presence of the anionic detergent SDS by FTIR spectroscopy.

2. To identify novel post-translational modifications in chicken erythrocyte linker histones after fractionation of native chromatin into soluble and insoluble fractions by Tandem MS.

3. To determine the level and the position of the incorporated phosphate groups by MALDITOF-MS and Tandem MS after *ex vivo* phosphorylation with CDK2 of the SPKK motifs of chicken linker histones.

4. To study the effects of *ex vivo* phosphorylation of chicken linker histones within chromatin in aggregation of the chromatin fragments due to an increase of the ionic strength ($MgCl_2$) by Dynamic Light Scattering.

Materials and Methods

Bacterial strains

E.coli M15 (QIAgen). Strain derived from *E.coli* K12, which contains the plasmid pREP4 used for the expression of genes under control of the inducible Lac operon. M15 cells (*Nal^S, Str^S, FR^S, Thi$^-$, Lac$^-$, Ara$^+$, Gal$^+$, Mtl$^-$, F$^-$, RecA$^+$, Urv$^+$, Lon$^+$*) were used for the expression of recombinant histone H1^0.

Plasmids

pH1^0. Recombinant plasmid derived from pQE-60 vector (QIAgen) which contains Mus musculus histone H1^0 gene cloned Nco I / Bgl II. The commercial vector is ATG-type with a total length of 3431bp. It contains Ampicillin gene as selection marker (Amp^R), the T5 promoter under control of Lac operator, which can be induced by lactose or other analogs as IPTG, and a 6xHis-tag that was fused to the C-terminal of H1^0 for protein purification.

Luria Bertrani (LB) medium

Consists on 10g of Triptone, 10g of NaCl and 5g of Yeast Extract per liter of medium; pH 7.0. The medium was autoclaved for 30 minutes at 121°C. Afterwards, Ampicillin (Roche) was added to a final concentration of 100µg/mL in order to allow the selective growth of *E.coli Amp^R* colonies containing the recombinant plasmid. To prepare solid medium plates, 1% of Agar was added before sterilization.

Agarose gel electrophoresis of DNA

Agarose gels at the appropriate ratio (w/v) were runned in 0.5xTBE buffer (Tris, boric acid 45mM, EDTA 1mM, pH 8.0). DNA bands were visualized by ethidium bromide staining (10µg/mL). Samples were prepared by adding 1:6 (v/v) of loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol, 40% sucrose). The voltage was 1-5 V/cm between the electrodes.

DNA molecular weight markers and DNA ladders were chosen depending of the size of the DNA molecules analysed in each case.

100bp DNA Ladder (Invitrogene): 15 blunt-ended fragments between 100bp and 1500 bp in multiples of 100bp and an additional fragment at 2072 bp.

DNA molecular weight marker VI (Roche): Fragment mixture prepared by cleavage of pBR328 DNA with *Bgl* I, and pBR328 DNA with *Hin*f I. 12 fragments: 154; 220; 234; 298; 394; 453; 517; 653; 1033; 1230; 1766; and 2176 bp

DNA molecular weight marker VII (Roche): Fragment mixture prepared by cleavage of SPP1 DNA with *Eco*R I. 17 fragments: 81; 359; 492; 710; 718; 992; 1164; 1482; 1515; 1882; 1953; 2799; 3639; 4899; 6106; 7427, 8576 bp

Quantification of DNA in solution

DNA concentration in solution was measured in a ThermoScientific NanoDrop 1000 spectrophotometer, directly depositing 2µL of the DNA suspension on the detector and using the Nucleic Acids software installed in the equipment. For each case, the corresponding blank solution was considered. DNA concentration was calculated by measuring the absorbance at 260nm and considering that A_{260}= 1UA is equivalent to a concentration of 50 µg/mL of DNA.

SDS-PAGE gel for protein analysis

Protein electrophoresis was performed in vertical discontinuous SDS gels using a 15% or 18% resolving gel [15% or 18% (w/v) Acrylamide/Bis-acrylamide, 0.375M TrisHCl pH 8.8, 0.1% (w/v) SDS, 0.1% (w/v) N,N,N´,N´-tetramethylethylenediamine (TEMED), 0.1% (w/v) ammonium persulfate (APS)] and a 5% stacking gel [5% Acrylamide/Bis-acrylamide, 0.125M TrisHCl pH 6.8, 0.1% (w/v) SDS, 0.1% (w/v) N,N,N´,N´-tetramethylethylenediamine (TEMED), 0.1% (w/v) ammonium persulfate (APS)]. Protein

samples and a prestained SDS-PAGE low range molecular weight protein standard (BioRad) were mixed 1:6 (v/v) with loading buffer (50mM Tris pH 6.8, 0.1% bromophenol blue, 2% SDS, 10% glycerol). The voltage applied was 100-150V and the running buffer composition was 25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS, pH 8.3. Visualization of proteins was possible due to gel staining with a solution of 0.5% Coomasie Brilliant blue in 20% ethanol at room temperature and low agitation for 2h. Afterwards, distaining was performed in the same solution but without the dye.

SDS-PAGE low range molecular weight protein standard (BioRad): phosphorylase B, 113KDa; bovine serum albumin, 92KDa; ovalbumin, 52.3KDa; carbonic anhydrase, 35KDa; trypsin inhibitor, 28.9KDa; lysozyme, 21KDa)

Competent cells preparation

E.coli M15 strain was used for recombinant protein expression and purification. To obtain a stock of competent M15 cells, 50mL of LB medium were inoculated with 1/100 of an overnight culture. The cell culture was grown at 37°C and 300rpm for 3h30min, until an OD_{600} of 0.6-0.8 was achieved. Afterwards, it was incubated in ice for 15min and transferred into a cooled tube for centrifugation at 630xg (Beckman-Coulter JLA-16.250) and 4°C without brake for 10min. The pellet was recovered and resuspended in 15mL of cold RF1 buffer (100mM RbCl, 50mM $MnCl_2{\cdot}4H_2O$, 30mM KAc, 10mM $CaCl_2{\cdot}2H_2O$, 15% (w/v) glycerol, pH 5.8 adjusted with 0.2N CH_3COOH) filtered with 0,22µm filter (MilliPore). The resuspended pellet was incubated for 10min more in ice and centrifugation was carried out once more under the same conditions. The supernatant was discarded again and the resultant pellet was resuspended in 4mL of cold RF2 buffer (10mM MOPS, 10mM $RbCl_2$, 75mM $CaCl_2{\cdot}2H_2O$, 15% (w/v) glycerol, pH 6.8 adjusted with 0.2N CH_3COOH) filtered with 0,22µm filter (MilliPore). 200 µL aliquots were prepared and stored at -80°C for future use.

Expression and purification of histone H1°

Bacterial culture and protein expression induction

The appropriate volume of LB liquid medium, supplemented with ampicillin (100 µg/mL), was inoculated with 1/20 of a bacterial pre-culture carrying the recombinant plasmid and grown overnight. The culture was grown at 37°C with continuous stirring in an orbital shaker at 250rpm, until an OD_{650} of ~0.8 was reached. At this point, 1mM isopropyl-β-D-thiogalactoside (IPTG) was added to induce the protein expression and, subsequently, the culture was grown for another 4 hours under the same conditions. Finally, the cells were harvested by centrifugation at 5000*xg* and 4°C for 15 min, and the pellet was stored at -20°C for protein purification. 1.5 mL samples of pre- and post-induction culture were taken to monitor the process.

Preparation of the bacterial lysate

The pellet of each litre of bacterial culture was resuspended vortexing it in 20mL of Lysis Buffer (10mM K_3PO_4, pH 7.0) plus 4M guanidine hydrochloride to avoid degradation and aggregation of the expressed protein. The suspension was kept at room temperature for 15min until the lysis was completed and then, centrifuged at 20000×*g* for 25 min (Beckman-Coulter JA-25.50 rotor). The soluble fraction (protein extract) was collected in a clean tube for purification.

Hydroxyapatite column chromatography

For protein purification, a CHT-II cartridge filled with ceramic hydroxyapatite type II (Bio-Rad) was used. Hydroxyapatite is a crystalline form of calcium phosphate [$Ca_{10}(PO_4)_6(OH)_2$] that has a high binding capacity for basic proteins like histone H1.

The cellular lysate was added to the column, which was previously equilibrated with lysis buffer (10mM K_3PO_4, 4M guanidine hydrochloride, pH 7.0). To remove the unbound protein, the column was washed

with lysis buffer until the baseline was restored. Elution of the recombinant protein was performed with 200mM K_3PO_4, 4M guanidine hydrochloride, pH 7.0 and samples of 5mL of the eluted fraction were collected. The process was carried out with a flow rate of 2 ml/min and monitored by a detector with a filter of 210nm coupled to a recorder (Pharmacia).

Gel filtration chromatography

Desalting of the purified protein was performed by gel filtration through a 30mL Sephadex G-25 column (Amersham Biosciences). The latter was equilibrated with H_20 miliQ and 5mL of the eluted fraction were loaded on the column with a flow rate of 0.5 mL/min. The whole process was monitored as described above.

Quantification of purified protein

Protein concentration was calculated by measuring the UV absorption of a 1/10 dilution of the sample at 210nm and considering that A_{210}= 205 is equivalent to a concentration of 10 mg/mL of histone H1. The absorbance measurements were performed in a 1 cm cuvette using a Cary100 Bio UV/visible spectrophotometer (Varian) and its Scan software.

Protein lyophilisation

The purified protein was frozen at -80°C for a minimum of 2 hours and placed in the lyophilizer (Virtis) at -100 ° C and 12 millitorr until all the water was removed.

In vitro phosphorylation of histone $H1^0$

Histone $H1^0$ was phosphorylated with CDK2-cyclinA kinase (Sigma-Aldrich), a specific enzyme that phosphorylates S/T residues within the sequence (S/T-P-X-K/R). The reaction was carried out in 50mM Tris-HCl, 10mM $MgCl_2$, 0.1mM EGTA, 2mM dithiotreitol, pH 7.5, plus 200µM ATP and 1µL of CDK2-cyclin A per 20 µg of protein. The mixture was incubated at 30°C for 1 h and the reaction buffer was eliminated by gel filtration on a 5 mL Sephadex G-25 desalting column (Amersham Biosciences) with a flow rate of 2 ml/min. The process was monitored as described above. Phosphorylation sites were S118, S140 and S152 relative to the entire recombinant $H1^0$.

The extent of phosphorylation was evaluated by MALDITOF mass spectrometry, carried out in the Proteomics facility from UAB, a member of ProteoRed-ISCIII network. The incorporation of each phosphate to the protein results in an increase of ~ 80 Da in its molecular weight.

FTIR spectroscopy

In order to determine the secondary structure of histone $H1^0$ under different conditions, FTIR spectra of the unphosphorylated $H1^0$ ($H1^00P$) and fully phosphorylated $H1^0$ ($H1^03P$), both in 10mM HEPES 140mM NaCl pH 7.0 and in the presence of DNA in the mentioned buffer, were recorded. It was also analysed the structure of $H1^00P$ and $H1^03P$ in the presence of anionic detergent SDS in 10mM HEPES, pH 7.0, plus 10mM or 140mM NaCl. Protein concentration was 5 mg/mL in all cases. Protein/DNA complexes were prepared at two protein/DNA (w/w) ratios (r): r=0.5 and r=0.7 with the appropriate amount of sonicated salmon sperm DNA (~500bp). The protein and the DNA were reconstituted in D_2O separately before the analysis and the complex formation was performed on the measurement cell. The SDS/H1° molar ratio was 14:1, which corresponds to an SDS/CTD molar ratio of 7:1, assuming a uniform distribution of SDS along the polypeptide chain. To avoid the interference of the vibration of the O-H bond of H_2O at around 1652cm^{-1}, which masks the amide I region[400,401,402], aqueous solution was evaporated in a Speed-Vac (Savant) evaporator and then reconstituted in D_2O buffer.

Measurements were performed on a FT600 Bio-Rad spectrometer equipped with a MCT detector, using a dismountable liquid cell with calcium fluoride windows and 50µm spacers. Typically, 1000 scans for each background and sample were collected and the spectra were obtained with a nominal resolution of

2 cm^{-1} [403], at 22 °C[235]. Data treatment and band decomposition of the original amide I' band have been described previously[404]. The baseline contributed by the solvent was removed before deconvolution[405]. The DNA contribution to the spectra of the complexes with $H1^0$ was subtracted using a DNA sample of the same concentration; the DNA spectrum was weighted so as to cancel the symmetric component of the phosphate vibration at 1087 cm^{-1} in the difference spectra, as described in Vila *et al.*[406].

The spectra were deconvolved in order to find the number and position of the individual Lorentzian components of the amide I' band -which corresponds to the amide I band in deuterated solutions- with Rampon and GplotC. Curve fitting of the component bands to the original spectrum was performed using CURVEFIT running under SpectraCalc (Galactic Inc.; Salem, NH). For each component, four parameters were considered: band position, band height, band width and band shape. In the curve fitting, Gaussian components were used. Initial heights were set at 90% of those of the original spectrum for the bands in the wings and for the most intense component, and at 70% of the original intensity for the other bands[404]. The curve-fitting procedure was accomplished in two steps: (1) The band position was fixed, allowing widths and heights to approach final values, and (2) band positions were left to change. The quality of the fitting was evaluated visually by overlapping the reconstituted overall curve on the original spectrum and by examining the residual obtained by subtracting the fitting from the original curve. The mathematical solution of the decomposition may not be unique, but if restrictions are imposed, such as maintenance of the initial band positions in an interval of ±1 cm^{-1}, preservation of the bandwidth within the expected limits, or in agreement with theoretical boundaries or predictions, the result becomes, in practice, unique.

The different secondary structure elements in a protein were assigned depending on the position of the vibration band in the amide I' region (1700-1600cm^{-1}), where the stretching vibration of the amide bond (C=O) results to a characteristic band[407,408,409]. Band assignment of secondary structure elements was performed as shown in Table 3 of the Introduction.

Optical microscopy of histone $H1^0$

Histone $H1^0$, both unphosphorylated ($H1^0$0P) and fully phosphorylated in its CTD ($H1^0$3P), was mixed with anionic SDS detergent at a detergent/protein molar ratio of 14:1 in 10mM HEPES, pH 7.0, plus 140mM NaCl. The concentration of protein was 5 mg/mL. Samples were observed by phase-contrast microscopy (Leica DMRB). Complexes were incubated for 5 min with 50 µM thioflavin in 10mM phosphate buffer, 140mM NaCl, pH 7.0. The excitation wavelength interval was 450–490nm, and the emission was recorded using a longpass 520nm optical filter. Complexes were also incubated for 30 min with 10mM Congo red in 20 mM HEPES, 0.1 mM EDTA, pH 7.4. The resulting birefringence was observed with a polarizing microscope (Leitz Axiotron) using crossed polarizers in the excitation and emission paths[410].

Chicken erythrocyte isolation

Chicken blood was obtained from the Farm and Experimental Fields Facility (Servei de Granges i Camps Experimentals de la UAB) located at the University Campus. After heart puncture of the chickens, the extracted blood was placed into tubes containing heparine and preserved on ice. For purification, 1/10 volumes of NKM buffer (130mM NaCl, 5.2mM KCl, 7.5mM $MgCl_2$, 0.4mM PMSF) were added to the collected blood and the mixture was divided and centrifuged for 15 minutes at 2500*xg* and 4°C in a Heraeus 1.0R centrifuge. The supernatant, containing the plasma, was discarded and the superior whitish layer of leucocytes was removed using a Pasteur pipette connected to a vacuum pump. This step was repeated using a volume of NMK buffer until the complete cleaning of the pellet, where the erythrocytes remained. Clean erythrocyte pellets were stored at -20°C.

Purification of chicken erythrocyte nuclei

The erythrocytes were resuspended in 60mL of Buffer E1 (0.25M Sucrose, 1N SSC buffer (140mM NaCl, 10mM sodium citrate), 0.1mM PMSF) and the mixture was slightly stirred to its complete resuspension. An equal amount of Buffer E1 supplemented with 0.6% (w/v) saponine was added. The final suspension was 120mL Buffer E1 plus 0.3% (w/v) saponine, which was centrifuged for 20 min at 2500*xg* and 4°C. The supernatant containing the detergent was discarded, and the nuclei were resuspended again in Buffer E1. This washing step was repeated twice to achieve the complete removal of saponine. Next, the nuclei were recovered and resuspended in 30mL of Buffer E1 and 90mL of Buffer E2 (2.3M Sucrose, 1N SSC, 0.1mM PMSF) were added to increase the volume up to 120mL.

20 mL of this suspension were carefully placed above 10 mL of a sucrose cushion (2.3M Sucrose, 1N SSC) Differential ultracentrifugation through the sucrose cushion was carried out for 2h at 10000*xg* and 4°C using a SW 32Ti rotor (Optima L-100 XP Ultracentrifuge, Beckman Coulter). After ultracentrifugation, erythrocyte nuclei were pelleted at the bottom of the tube and resuspended in 2mL each of 2x Buffer C (60mM KCl, 3mM $MgCl_2$, 10mM Tris-HCl, 1% thiodiglycol pH 7.4) adding 0.1mM PMSF. DNA concentration of the nuclei was quantified by measuring the absorbance at 260nm of a sample diluted in 1% SDS with the ThermoScientific NanoDrop 1000. Immediately afterwards, 50% glycerol was added to the nuclei and 3-5mL aliquots were stored at -80°C.

Preparation of erythrocyte chromatin

Stored erythrocyte nuclei were centrifuged in a Heraeus 1.0R centrifuge at 2880*xg* and 4°C for 20min and then, washed twice in 0.25M Sucrose, 10mM Tris, 140mM NaCl, pH 8.0 and resuspended in the same buffer to a final DNA concentration of 2mg/mL. The nuclei were pre-incubated for 5min at 37°C and the nuclease micrococcal digestion was carried out in 0.3mM $CaCl_2$ at 37°C adding 1U of enzyme per 20µg of DNA. The time for nuclease digestion was adjusted for every new batch of erythrocyte nuclei, according to the results obtained in the kinetics assay (see below). The reaction was stopped with 10mM EDTA.

Nuclei were centrifuged again for 20min in a Heraeus 1.0R centrifuge at 2880*xg* and 4°C. Eventually, nuclei were resuspended in half the original volume with TE 1x and incubated on ice for 30min to allow cellular lysis. After incubation, lysated nuclei were centrifuged for 15min at 16100*xg* and 4°C (Eppendorf 5415R). The supernatant, corresponding to the soluble fraction of chromatin, was collected and the DNA concentration of the fraction was calculated after measuring the OD_{260nm} in a ThermoScientific NanoDrop 1000 spectrophotometer. The size of the chromatin fragments was analysed by DNA electrophoresis after proteinase K digestion.

To assess the time of digestion of each batch of erythrocyte nuclei in order to obtain chromatin fragments of the desired size, a kinetic assay with micrococcal nuclease was performed. 1mL sample of the purified erythrocyte cell nuclei was divided in 5 aliquots and digested with micrococcal nuclease for 0'', 30'', 60'', 90'' and 120''. The size of the fragments was analysed by agarose gel electrophoresis after digestion with proteinase K and a time was selected for the preparation of erythrocyte chromatin.

Proteinase K digestion for DNA isolation

Samples of soluble chromatin were digested with proteinase K in order to analyse the size of the DNA fragments obtained after the nuclease micrococcal digestion. The reactions were performed by adding 1µL of proteinase K (10 mg/mL stock (Roche)) per 10µL of chromatin in the presence of 1/10 (v/v) proteinase K buffer (100mM Tris-HCl pH 8.0, 50mM EDTA, 5% SDS). The reactions were incubated overnight at 37°C. DNA was further analysed by agarose gel electrophoresis.

Ex vivo chromatin phosphorylation

Phosphorylation of linker histones bound to the extracted chicken erythrocyte chromatin was performed considering that 1µg of chromatin contains 9.75pmol of H1 histones and establishing an average molecular weight of 22000 g/mol. Phosphorylation with CDK2-cyclinA kinase (Sigma-Aldrich) was carried out in 50mM TrisHCl, 10mM $MgCl_2$, 1mM EGTA, 20mM dithiotreitol, pH 7.5, plus 200µM ATP and 1µL of CDK2-cyclin A per 20 µg of H1. 1µL of a cocktail of protease inhibitors (cOmplete Mini Protease Inhibitor Cocktail Tablets; Roche Applied Biosciences) per 100µL of reaction was added. A master mix reaction was prepared and a sample for "time zero" was collected before adding the kinase. Afterwards, the enzyme was added and the mixture was incubated at 30°C at different times. After the phosphorylation time of the corresponding sample was accomplished, they were stored at 4°C in order to develop posterior analyses of all of them in parallel.

In the radioactive CDK2 phosphorylation performed in order to corroborate the specific phosphorylation of linker histones within chromatin, 10 µCi [ɣ-32P] ATP (3000-1500 Ci/mmol) were added per 100µL of the standard phosphorylation reaction (see above). After 30 minutes of incubation at 30°C, the phosphorylation reactions at different times were mixed 1:6 (v/v) with SDS-PAGE loading buffer (50mM Tris pH 6.8, 0.1% bromophenol blue, 2% SDS, 10% glycerol), heated for 5 min at 95°C, and analysed by SDS-PAGE. After SDS-PAGE, the gel was stained with Coomasie blue and then exposed in Imagining Screen Biorad plates (35x43) to obtain an autoradiography of the gel in a Personal Molecular Imager (BioRad).

Proteomic study of chromatin samples

Isolation of linker histones from erythrocyte chromatin by perchloric acid extraction

After chicken erythrocyte nuclei disruption following digestion with micrococcal nuclease, two fractions were separated by centrifugation: soluble and insoluble chromatin. Soluble chromatin was phosphorylated *ex vivo* at different times, to study the effects of phosphorylation. Part of all the soluble chromatin samples and the insoluble fraction were extracted with perchloric acid to analyse the native post-translational modifications and the extent of phosphorylation by the CDK2.

Perchloric acid extraction of the chromatin samples was performed with 5% (w/v) $HClO_4$ and 0.05M NaH_2SO_3. The suspension was incubated for 1h at room temperature (RT) by mixing with an orbital shaker and centrifuged for 15min at 16100*xg* and 4°C (Eppendorf 5415R). The supernatant, containing the soluble linker histones, was preserved. The chromatin pellet was re-extracted for 30min more using the same protocol and re-centrifuged under the same conditions. The supernatants corresponding to the same fraction were put together and linker histones were precipitated overnight at 4°C with 20% (w/v) TCA (trichloroacetic acid). Next, the suspension was centrifuged (Eppendorf 5415R) for 15min at 16100*xg* and 4°C and the protein pellets were washed three times with acetone and dried at RT. Extracted linker histones were resuspended in MilliQ distilled water, quantified and analysed by SDS-PAGE. Samples containing linker histones and no protein degradation were dried in a Sped Vac (Savant) and stored until proteomic analyses.

High performance capillary electrophoresis (HPCE)

HPCE was carried out in order to resolve and identify the different linker histone subtypes present in the samples and for qualitative identification of the phosphorylated variants generated by *in vitro* phosphorylation of chromatin. It was also used for primary determination of the quality and purity of the samples and as a control of enzymatic cleavage reactions (see below).

HPCE was performed on an electrophoresis system (P/ACE™ 2100 and System Gold™ software or P/ACE™ MDQ Capillary Electrophoresis System and 32 Karat™ 5.0 Software; Beckman Instruments, Palo Alto, CA, USA). These softwares were also used for determination of peak heights. The capillary cartridge was fitted with a 75µm ID fused silica with a total length of 50cm to the detector. Protein samples (1 µg/µL) were injected under pressure for 5s, and detection was performed by measuring UV absorption at 200 nm. An untreated capillary was used in all the experiments, and after every 5-10 injections, it was rinsed with water, 0.1N NaOH, water, 1N H_2SO_4, water and finally with the running buffer, washing for 2 min with each of the solvents mentioned. Running buffer consisted on 500mM or 100mM H_3PO_4 containing 0.02% HPMC, adjusting pH with TEA to pH 2.0 and filtering it through a 0.45 µm filter before every use. All runs were performed at constant voltage (12 kV) and at a capillary temperature of 25°C.

Reversed phase high performance liquid chromatography (RP-HPLC)

RP-HPLC was used for initial fractionation of histone H5 and H1 variants after acid extraction of the whole linker histone fraction (H1+H5). This method separates molecules on the basis of hydrophobicity using an acetonitrile gradient.

Fractionation of histone variants H5 and H1s was performed on a Nucleosil 300-5 C_4 column (10 × 4 mm inner diameter; 5-µm particle pore size; 30-nm pore size; end-capped; Macherey-Nagel). The lyophilized proteins were dissolved in water (1 µg/µL), and the appropriate amount of sample (20-100 µg) was injected onto the column. The equipment used consisted of a 127 Solvent Module and a model 166 UV-visible region detector (Beckman Instruments). The effluent was monitored at 210 nm, and the peaks were recorded using Beckman System Gold software.

Histone H1/H5 fractions from chicken erythrocytes samples were separated by chromatography within 25 minutes at a constant flow of 0.5 ml/min with a linear acetonitrile gradient starting (solvent A: solvent B 62:38; being solvent A water containing 0.1% trifluoroacetic acid (TFA), and solvent B, 70% acetonitrile and 0.1% TFA). The concentration of solvent B was increased from 38% to 50% during a period of 25 minutes. Fractions were collected adding 5µl of EtSH 200mM, lyophilized and stored at -20 °C. H5 fraction came out of the column before than the H1 variants fraction.

Enzymatic cleavage

Chymotrypsin digestion. 5-15 µg of each of the different whole chicken histone H1 or H5 fractions obtained by RP-HPLC were digested with α-chymotrypsin [EC 3.4.21.1] (Sigma type I-S, 1/150 (w/w)) in the presence of 100 mM sodium acetate buffer (pH 5.0). The reaction was performed for 30min at room temperature in a final volume of 20 µL. The cleavage product was verified by HPCE analysis and, afterwards, subjected to MALDI analysis for determination of the peptide masses and thus, the phosphorylation level.

Arg-C digestion. 5 µg of each of the different chicken H5 fractions obtained by RP-HPLC were digested with Arg-C (EC 3.4.21.1) (sequencing grade, 1:40, w/w; Roche Applied Science) in the presence of 5mM NH_4HCO_3 buffer (pH 8.0) for 5min at 37°C in a final volume of 20 µL. The cleavage product was verified by HPCE analysis and then submitted to LC-ESI-MS for MS/MS sequencing.

Trypsin digestion. 5 µg of each of the different chicken H1 fractions obtained by RP-HPLC were digested with trypsin (Roche Applied Science; sequencing grade, 1:100 w/w) in the presence of 5 mM NH_4HCO_3 buffer (pH 8.0) for 15min at room temperature in a final volume of 20µL. The cleavage product was verified by HPCE analysis and, afterwards, LC-ESI-MS was performed.

Mass spectrometric analysis

MALDI-TOF-MS

Matrix-assisted laser-desorption ionization coupled with time-of-flight mass-spectrometry (MALDITOF-MS) was used for determination of the molecular masses of the peptide fractions obtained after α-chymotrypsin cleavage of chicken H1 and H5 histones and identification of post-translational modifications. Mass spectrometric measurements were performed on a 4800 Plus MALDI TOF/TOF Analyser instrument (Applied Biosystems, Foster City, CA) using the 4000 Series Explorer Software V3.6 for data acquisition and processing, and samples were measured in the mid mass range (4.0 – 25 kDa).

Sample preparations were purified with PerfectPure C-18 Tips (Eppendorf) before MALDI-MS analysis. The detailed protocol was as follows: (i) Aspirate and expel 10 µL Pre-Wet solution (50% Acetonitrile in 0.1% TFA) two times to wet the PerfectPure C-18 Tip. (ii) Equilibrate two times with 10 µL equilibration / wash solution (0.1% TFA). (iii) Aspirate and expel the acidified sample (approx. 0.2% TFA final concentration) 3 – 10 times to bind the analyte to the C-18 matrix. (iv) Wash with 5 µL equilibration / wash solution two times. (v) Elute the sample with approx. 0.8 µL matrix solution (4 mg/mL *alpha*-cyano-4-hydroxycinnamic acid (CHCA) solved in 60% Acetonitrile and 0.1% TFA) and spot it directly onto the MS targe. The matrix-analyte droplet is then slowly dried in air.

The experimental data was compared and contrasted with the theoretical values of peptide mass obtained when submitting the protein to *in silico* cleavage with α-chymotrypsin digestion. Expasy Tools Proteomics Software (PeptideCutter and PeptideMass) and GPMAW 7.02 Lighthouse Data Software were used with this purpose.

MS/MS sequencing: LC-ESI-MS

Arg-C cleavage of H5 histone and Trypsin cleavage of H1 histones at different phosphorylation times were analysed using nano-HPLC consisting of an UltiMate 3000 system (Dionex Corp.) connected online to a linear ion trap mass spectrometer (LTQ Velos ThermoScientific) equipped with a nanospray ionization source for MS/MS sequencing and phospho-sites mapping. A homemade fritless fused silica microcapillary column (75 µm i.d. x 280 µm o.d.) packed with 10 cm of 5 µm reverse-phase C_{18} material (MagicC18) was used. The gradient (solvent A: 0.1% formic acid; solvent B: 0.1% formic acid in 85% acetonitrile) started at 4% B. The concentration of solvent B was increased linearly from 4% to 40% during 35 min and from 40% to 100% during 5 min. A flowrate of 300nL/min was applied. The nanospray voltage was set at 1.6 kV, and the heated capillary was held at 200 °C. Data-dependent neutral loss MS^3 was used for precise localization of phosphorylation. MS/MS (MS^2) and MS^3 spectra were searched against a histone data base using SEQUEST in Proteome Discoverer software (Version 1.2, ThermoScientific) and validated manually. The identified peptides were further evaluated using charge state *versus* cross-correlation number (Xcorr). The criteria for positive identification of peptides were Xcorr > 1.5 for singly charged ions, X^2corr > 2.0 for doubly charge ions, and X^3corr > 2.5 for triply charge ions. Only best matches were considered. MS/MS tolerances of +0.8 Da were allowed.

Transmission Electron Microscopy (TEM)

Samples of native chromatin stored at 4°C and *ex vivo* phosphorylated chromatin for 1h, 5h and overnight were diluted, respectively, to a final concentration of 2µg/mL in 5mM TEA 35mM NaCl plus 1mM $MgCl_2$ and fixed with 0.1% GTA for 30 minutes at 4°C. Copper grids covered with a carbon film and activated by glow-discharge for 1 minute were used as a support for the sample. 10µL of sample were deposited over the grid for 5 minutes. After this period, excess of sample was removed and a drop of 0.2% uranyl acetate for sample negative staining was added and immediately removed. Sample grids

were air dried and visualized in a JEM-1400 TEM microscope at the Servei de Microscòpia Electrònica (Campus UAB).

Dynamic Light Scattering (DLS)

Dynamic Light Scattering (DLS) measurements were carried out in order to study potential changes in chicken erythrocyte chromatin due to partial phosphorylation.

DLS measurements were carried out in a ZetaSizer Nano ZS particle analyser (Malvern Instruments, UK) at the Institute of Materials Science of Barcelona (ICMAB-CSIC) and in the Nanotechnology Laboratory at Matgas 2000 A.I.E. (Campus UAB). This equipment permits to measure the particle size distribution in the range comprised between 0.6nm and 6microns. ZetaSizer Nano ZS is a high performance two angle particle and molecular size analyser for the enhanced detection of aggregates and measurement of small or dilute samples, and samples at very low/high concentration using dynamic light scattering with non-invasive backscatter optics (NIBS).

The light source was a He-Ne laser (633nm) that uses 5mW power at the same wavelength. Scattered light from the samples was collected at an angle of 173° and the intensity correlation function was used to generate a correlation curve.

For each measurement, three records were performed. Each record consisted in 14 to 18 cumulant scans, depending on the measurement position within the cuvette of the Nano ZS. This parameter can be changed by moving the focusing lens and is determined automatically by the Nano software Nano DTS. The measurement position is determined automatically through a combination of the intercept of the correlation function and the intensity of the light scattered to achieve the optimal signal depending on the sample.

Samples of *ex vivo* phosphorylated chromatin at 0h (native chromatin), 1h, 5h and overnight were diluted to achieve a final concentration of 0.2mg/mL in a final volume of 1mL with Tris 10mM pH 7.0 plus 35mM NaCl, 1mM $MgCl_2$. After the initial measurement at 1mM of $MgCl_2$ the concentration of this salt was increased up to 1.6mM, incubated for 5 min at RT and measured again. For each phosphorylation time, a control was made with a second sample incubated at 30ªC without kinase. The spectra obtained were processed with the high resolution option of the software and the number and size in nm of the peaks for each sample was analysed.

Laser Diffraction

Laser diffraction measurements were carried out, as DLS measurements, to study the changes in chicken erythrocyte chromatin due to partial phosphorylation. Laser diffraction is a widely used particle sizing technique for particle size determination of larger molecules than the DLS analysis.

Laser light scattering measurements were carried out in a Mastersizer 2000 particle analyser (Malvern Instruments, UK) in the Nanotechnology Laboratory of Matgas 2000 A.I.E. (Campus UAB). This equipment permits to measure the particle size distribution in the range comprised between 0.02µm and 2000µm. The light sources were a red light He-Ne laser (633nm) that uses 4mW power at the same wavelength; and a blue light LED source (470nm) that uses 0.3mW power at the same wavelength. Lens arrangement was Reversed Fourier (converged beam). A Hydro 2000µP wet sample dispersion unit was used for measuring the samples. This accessory needs 18-20mL sample volume and it is provided with a controlled sample pump/stirrer and a controlled sonication system.

To perform each Mastersizer measurement, the Hydro 2000µP unit had to be first filled manually with the dispersant medium to measure the background (blank) and, afterwards, the sample could be added. When the obscuration of the medium was in range (>1%), the scattered light by the molecules in the

sample could be measured. Three measurements of 10 scans each were recorded in all cases. After that, an average measurement of these three was obtained automatically.

Optical parameters (refractive index) were set in 1.33 for the dispersant (Tris 10mM pH 7.0, NaCl 35mM, 1mM $MgCl_2$) and 1.45 for the sample (chromatin). The lowest obscuration range allowed to record data was 1% and the analysis model of data was set to the multiple narrow modes of the software.

Samples of *ex vivo* phosphorylated chromatin at 1h, 5h and overnight coming from the DLS measurements (0.2mg/mL) were measured in a final volume of 18-20mL of Tris 10mM pH 7.0, 35mM NaCl, 1.6mM $MgCl_2$. For each phosphorylation time, a measurement was made with a second sample incubated at 30^0C without kinase.

Results and Discussion

Chapter One

Folding and fibrillation of histone H1

Previous studies on isolated histone H1 domains determined the significance of the interaction with DNA in the folding of these two domains[235,218] and, furthermore, the importance of phosphorylation, together with DNA interaction, in the folding of the CTD[266].

The C-terminal domain, which represents the 52% of the entire protein, is the primary determinant of histone H1 binding to chromatin *in vivo*. In aqueous solution, the CTD is unfolded and it mainly contains random coil and turn-like conformations. It folds cooperatively upon interaction with DNA in a highly-stable structure[235,306]. The CTD is, thus, considered an intrinsically disordered protein, with coupled binding and folding[411,412,413].

The aim of the present chapter was to study the secondary structure of the entire $H1^0$ in aqueous solution, bound to DNA and in the presence of anionic detergent SDS and to compare these results with the ones obtained with the isolated domains. The preferred biophysical technique to study the secondary structure of histone H1 was infrared spectroscopy (FTIR). Histone H1 remains unstructured and soluble in aqueous solution but it precipitates upon interaction with DNA. FTIR can be applied to the analysis of solids, liquids, and gasses. Hence, this technique resulted more useful than others, such as CD or NMR, which do not allow the analysis of the protein in the complexes with DNA.

Histone $H1^0$ secondary structure in aqueous solution

Histone H1 contains three distinct domains: the short N-terminal domain, the central and well-structured globular domain, and the C-terminal domain. In aqueous solution at physiological salt concentration, only the globular domain is structured in a three-helix bundle and a β-hairpin[414] whereas the N- and C-terminal domains remain basically unfolded.

Infrared spectroscopic analyses in D_2O of the entire histone $H1^0$ were performed in aqueous solution in order to determine the secondary structure elements present in the whole protein (Table 1.1.). These results were contrasted with the ones reported for the isolated domains and were used as a reference spectrum of the whole protein to compare the secondary structure of $H1^0$ under different conditions by FTIR.

Table 1.1. Secondary structure motifs (%) of $H1^0$ in aqueous solution.

Secondary structure assignment (%)	$H1^0$ domains			$H1^0$ estimated values ʰ	$H1^0$ experimental values *
	$NH1^0$§	$GH1^0$γ	$CH1^0$¤		
Turns	5	0	18	23	21
α-helix	1	19	0	20	23
Random coil/ flexible regions	6	11	27	44	35
β- structure	0	6	7	13	21
∑ % Ŧ	**12**	**36**	**52**	**100**	**100**

* Experimental data obtained by FTIR spectroscopy in D_2O of histone $H1^0$. Secondary structure elements (%) were obtained by curvefitting of the amide I region of the protein at 5 mg/ml in aqueous solution. The buffer was HEPES 10mM NaCl 140mM, pH 7.0

§ Secondary structure elements of the N-terminal domain of histone $H1^0$. Data taken from Vila *et.al.*(2001)[218]. The N-terminal domain corresponds to the 12% of histone $H1^0$.

γ Secondary structure elements of the globular domain of histone $H1^0$ resolved by NMR. Data taken from Cerf *et.al.*(1994)[220]. The globular domain corresponds to the 36% of histone $H1^0$.

¤ Secondary structure elements of the C-terminal domain of histone $H1^0$. Data taken from Roque *et.al.*(2008)[266]. The C-terminal domain corresponds to the 52% of histone $H1^0$.

ʰ Estimation of the percentages of secondary structure elements of $H1^0$ in solution obtained from the reported data for individual domains (§,γ,¤). Secondary structure elements of the isolated domains(§,γ,¤) are referred as a percentage (%) of the full-length protein and were all obtained in HEPES 10mM NaCl 140mM pH 7.0.

Ŧ Sum of percentages (%) of secondary structure motifs. Percentages of each domain to the full length protein (200aa).

In buffer solution, the entire protein contained 23% α-helix, which can be attributed to the globular domain, since this domain is already structured in physiological salt[220]. The amount of β-structure was 21%. This percentage is significantly higher than the estimated one (Table 1.1.), which derives from the sum of the β-structure content from the isolated C-terminal and globular domains (~13%). On one hand, the secondary structure of the globular domain is stable and well-defined[220]; on the other hand, the N-terminal domain itself is too short to explain this difference. Because of that, we suppose that this increase in β-structure in the entire protein may be mainly contributed by the C-terminal domain, which is basically unstructured in aqueous solution and, hence, capable of gaining β-structure. Note that, when comparing the expected values with the experimental ones, the percentage of β-structure obtained in the FTIR experiments was higher whilst the amount of random coil and open loops decreased with respect to which was estimated. Thus, the presence of the globular domain seems to stabilize the β-structure of the CTD in the entire protein, which suggests the presence of some degree of structural coupling between the CTD and the globular domain.

Effect of phosphorylation

Previous studies by FTIR spectroscopy of the isolated C-terminal domain of the protein showed that the CTD has a remarkable conformational flexibility. In particular, phosphorylation at the consensus sequence of cyclin-dependent kinases, S/T-P-X-K/R (SPKK motifs), greatly affects its DNA bound structure and DNA condensing capacity[266]. Histone $H1^0$ has three SPKK motifs located in its CTD, in the positions S118, S140 and S152.

In the FTIR studies of the isolated CTD[266], only a slight increase in β-structure was observed due to phosphorylation in buffer solution. In order to determine if there were any changes in the full-length protein, FTIR spectroscopy analysis of the unphosphorylated ($H1^0$0P) and the phosphorylated ($H1^0$3P)

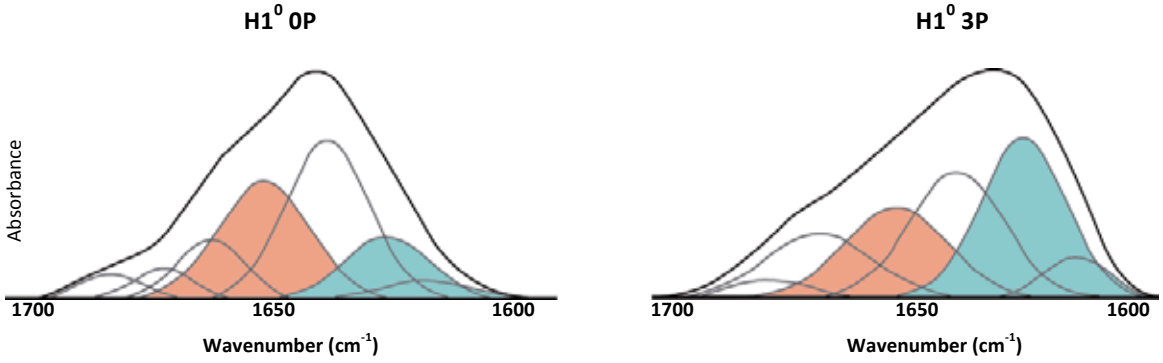


Figure 1.1. Amide I decomposition of unphosphorylated ($H1^0$0P) and phosphorylated ($H1^0$3P) histone $H1^0$. FTIR spectra in D_2O. The buffer was 10 mM HEPES 140 mM NaCl, pH 7.0. The protein concentration was 5 mg/ml. The α-helix component is highlighted in orange and the β-structure components are highlighted in blue. The positions of phosphorylation were 118/140/152. The positions are defined relative to the entire protein.

histone $H1^0$ were performed in D_2O (Figure 1.1.)

Phosphorylation of the three SPKK motifs induced an increase of the percentage of β-structure of histone $H1^0$ in aqueous solution, indicating that phosphorylation leads to a structural change in the protein (Figure 1.1.). The values of β-structure increased from 21% in the unphosphorylated protein to a considerable 35% in the phosphorylated protein. The proportion of α-helix (~20%) was attributed to the globular domain, since the values remained similar to the ones in the unphosphorylated protein and,

thus, to the estimated ones (Table 1.2.). These results were also observed in the FTIR analysis of rat histone H1e[267].

Table 1.2. Secondary structure motifs (%) of unphosphorylated ($H1^0$0P) and triphosphorylated ($H1^0$3P) histone $H1^0$ in aqueous solution.

Secondary structure assignment (%)	$H1^0$ 0p		$H1^0$ 3p	
	Band cm^{-1}	%	Band cm^{-1}	%
Turns	1683	4		
Turns	1672	3	1678	3
Turns	1663	14	1667	15
α-helix	1652	23	1652	20
Random coil/flexible regions	1640	35	1640	27
β-sheet	1628	17	1627	30
Low frequency β-sheet	1619	4	1616	5

Experimental data obtained by FTIR spectroscopy in D_2O of histone $H1^0$. Secondary structure elements (%) were obtained by curvefitting of the amide I region of the protein at 5 mg/ml in aqueous solution. The buffer was HEPES 10mM NaCl 140mM, pH 7.0. Values have been rounded off to the nearest integer.

These results indicate that phosphorylation favours the induction of β-structure in histone $H1^0$. This fact is mainly determined by the conformational change of the C-terminal domain following phosphorylation, since all three phosphorylation sites (SPKK motifs) are located in this domain.

Secondary structure of $H1^0$-DNA complexes

FTIR spectroscopy studies of the isolated C-terminal domains of histones $H1^0$, H1t and H1e bound to DNA showed that the CTD has little structure in aqueous solution but becomes extensively folded upon interaction with DNA[235,267].

Histone-DNA binding is electrostatic and it involves reciprocal charge compensation of the DNA phosphates and the Lys ε-amino groups in the protein, which are more abundant in the CTD. Neutralization of the Lys-charge in the isolated CTD induced the folding of the CTD with proportions of secondary structure motifs similar to those in the DNA complexes. Thus, increased hydrophobicity of the amphipathic Lys side chains accompanying charge neutralization is responsible for the folding of the CTD upon DNA interaction[310].

In order to examine if those structural changes were also present in the full-length protein, FTIR spectroscopic analysis in D_2O of the entire histone $H1^0$ bound to DNA were carried out. Infrared spectra were performed at physiological salt concentration (10mM HEPES 140mM NaCl, pH7.0) for both, the unphosphorylated ($H1^0$0P) and the phosphorylated ($H1^0$3P) histone $H1^0$, in the complexes with DNA at two different protein/DNA ratios, r(w/w), r=0.5 and r=0.7 (Table 1.3.).

The binding of unphosphorylated histone $H1^0$ to DNA induced an increase in the proportion of α-helix from a 23% in the protein free in solution to a 34% when bound to DNA. This change implied a decrease in the proportion of turns and random coil, whereas the amount of β-structure remained the same (Table 1.2., Table 1.3., Figure 1.2.). The increase in α-helix is supposedly a consequence of the folding of the terminal domains upon DNA interaction[218,235]. The percentages (%) were nearly the same for the two protein/DNA ratios tested (±5%) in the DNA complexes with $H1^0$0p. Thus, only the results obtained in the complexes at r=0.7 are shown (Table 1.3, Figure 1.2.).

Table 1.3. Secondary structure motifs (%) of unphosphorylated ($H1^0$0P) and triphosphorylated ($H1^0$3P) histone $H1^0$ in aqueous solution and in the complexes with DNA.

Secondary structure assignment (%)	$H1^0$ 0p/DNA r=0.7		$H1^0$ 3p/DNA r=0.5		$H1^0$ 3p/DNA r=0.7	
	Band cm^{-1}	%	Band cm^{-1}	%	Band cm^{-1}	%
Turns			1685	4	1685	2
Turns	1673	8	1673	15	1672	13
Turns	1662	12	1662	9	1663	4
α-helix	1650	34	1653	13	1653	15
Random coil/flexible regions	1640	22	1643	23	1644	12
β-sheet	1631	14	1629	25	1630	44
Low frequency β-sheet	1619	10	1617	11	1618	10

Experimental data obtained by FTIR spectroscopy in D_2O of histone $H1^0$ in the complexes with DNA. Secondary structure elements (%) were obtained by curvefitting of the amide I region of the protein at 5mg/ml at physiological salt concentration. The buffer was HEPES 10mM NaCl 140mM, pH 7.0. Two protein/DNA ratios, r(w/w), were considered, r=0.5 and r=0.7. Values have been rounded off to the nearest integer.

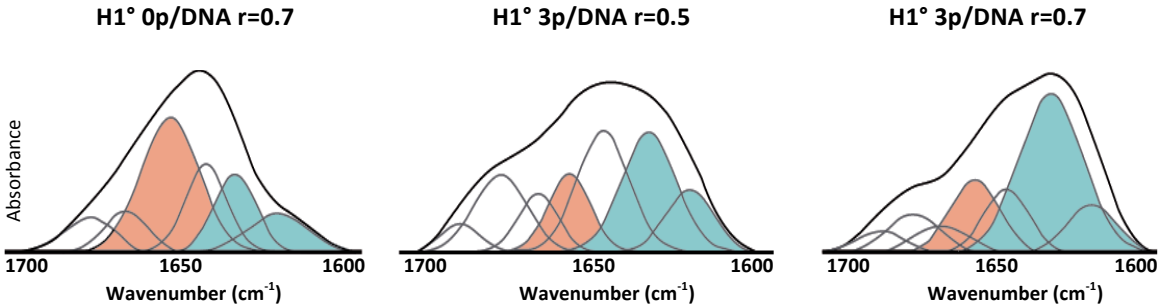


Figure 1.2. Amide I decomposition of FTIR spectra in D_2O of unphosphorylated ($H1^0$0P) and phosphorylated ($H1^0$3P) histone $H1^0$ bound to DNA. Protein/DNA ratios (w/w) correspond to the spectra at r=0.7 for the complexes with the unphosphorylated $H1^0$ and r=0.5 and r=0.7, respectively, for the complexes with the phosphorylated protein. The buffer was 10 mM HEPES 140 mM NaCl, pH 7.0. The protein concentration was 5 mg/ml. The α-helix component is highlighted in orange and the β-structure components are highlighted in blue. The positions of phosphorylation were 118/140/152. The positions are defined relative to the entire protein.

These results suggest that DNA binding to the unphosphorylated $H1^0$ induces its folding in α-helix conformation, as it happens with the isolated CTD. Thus, histone $H1^0$ can be also classified as an intrinsically disordered protein, in which DNA binding leads to protein folding.

Effect of phosphorylation

Phosphorylation of the SPKK motifs present in the CTD induced a significant change in the structure of the isolated domain when bound to DNA, remarkably increasing its proportion of β-structure[266].

In the entire protein, the most remarkable change also occurred in the amount of secondary structure elements of the phosphorylated histone $H1^0$ ($H1^0$3P) bound to DNA at physiological salt concentration. Under these conditions, the amount of β-structure increased considerably in contrast with the results obtained for $H1^0$0P bound to DNA. The amount of β-structure increased from 24% in the unphosphorylated $H1^0$ to 54% in the complexes $H1^0$3P-DNA at r(w/w)=0.7. This increase was accompanied by a significant decrease in the amount of α-helix, from 34% to 15% (Table 1.3.). In this case, there is an effect of protein/DNA ratio, the proportion of β-structure increased from a 36% at r(w/w)=0.5 to a 54% when r(w/w)=0.7, with a concomitant decrease of random coil percentage. The amount of α-helix remained unchanged (Figure 1.2.).

DNA binding increases the folding of the phosphorylated protein (Table 1.2., Table 1.3.), which implies a remarkable gain of β-structure, as it happens in the isolated CTD. Note that, at a protein/DNA ratio of 0.7 (w/w), 54% of β-structure was reached, which is even a little bit higher than the contribution of the CTD to the entire protein (52%). Therefore, these results suggest that phosphorylation, coupled with DNA binding, induces the folding of the CTD of histone $H1^0$ in an all-β conformation in the entire protein. Furthermore, there is also a remarkable decrease of α-helix in contrast with $H1^0$0P bound to DNA, which could be due to the perturbation of the globular domain associated to the propagation of the β-structure from the CTD towards the rest of the protein.

$H1^0$ secondary structure in the presence of SDS

In addition to binding to nucleosomes and linker DNA and processes related to genome packing and regulation that are carried out in the cell nuclei, histone H1 has been found outside the nucleus and even outside the cell, where it could perform several functions that would involve the interaction with membrane lipids[415,416,417,418,419,420,421,422].

Since the folding of histone H1 may be determined by an increase of the hydrophobicity followed by charge compensation of the Lys positive charge in the CTD by the DNA phosphates, these results suggest that the interaction with hydrophobic ligands could also trigger the folding of the intrinsically disordered CTD and, thus, the entire protein, as efficiently as DNA does. To test this hypothesis, the induction of secondary structure in the presence of the anionic detergent SDS, which would act as a hydrophobic ligand, was studied by FTIR spectroscopy.

Anionic SDS has been shown to induce and stabilize secondary structure in polypeptides and it is used to model hydrophobic environments such as membranes. Due to its anionic character, SDS is capable of establishing electrostatic and hydrophobic interactions, mimicking the environment found in the complexes with DNA.

FTIR spectroscopic analyses in D_2O of histone $H1^0$ in the presence of anionic detergent SDS were carried out. Infrared spectra were performed at physiological salt concentration (10mM HEPES 140mM NaCl, pH7.0) for both the unphosphorylated ($H1^0$0p) and the phosphorylated ($H1^0$3p) histone $H1^0$ at a SDS/$H1^0$ molar ratio of 14:1. Previous analyses with the isolated CTD showed that, in the interaction with SDS, both the level of phosphorylation and the detergent/CTD molar ratio affected the proportions of secondary structure motifs. The greatest effects were observed at a molar ratio of 7:1 SDS/CTD: the unphosphorylated CTD showed proportions of secondary structure motifs similar to those found in the DNA complexes; and the phosphorylated CTD became an all-β protein with 83% β-structure and no α-helix[423]. Since we assumed a uniform distribution of SDS along the polypeptide chain, a 14:1 SDS/H1 molar ratio was tested when studying the effects of SDS in the secondary structure of the entire protein.

At a 14:1 SDS/H1^0 molar ratio, the amount of α-helix increased up to 30% while the proportion of β-structure (21%) remained unaltered. The effect of SDS in the unphosphorylated protein is, thus, the same that occurs due to the binding of DNA (Table 1.4., Figure 1.3.).

Table 1.4. Secondary structure motifs (%) of the unphosphorylated (H1^{0}0P) and the fully-phosphorylated (H1^{0}3P) histone H1^0 in the presence of SDS at a molar ratio 14:1.

Secondary structure assignment (%)	SDS/H1^0 0p 14:1		SDS/H1^0 3p 14:01	
	Band cm^{-1}	%	Band cm^{-1}	%
Turns	1683	1	1681	1
Turns	1672	11	1671	6
Turns	1660	16	1661	11
α-helix	1651	30	1651	13
Random coil/flexible regions	1640	21	1642	14
β-sheet	1630	14	1632	34
Low frequency β-sheet	1617	7	1620	21

Experimental data obtained by FTIR spectroscopy in D_2O of histone H1^0 in aqueous solution, in the complexes with DNA and in the presence of SDS. Secondary structure elements (%) were obtained by curvefitting of the amide I region of the protein at 5mg/ml at physiological salt concentration. The buffer was HEPES 10mM NaCl 140mM, pH 7.0. SDS/protein molar ratio was 14:1. Values have been rounded off to the nearest integer.

The triphosphorylated protein in the presence of SDS at a molar ratio of 14:1, also showed changes in secondary structure similar to those found in the complexes with DNA at r(w/w)=0.7. Interaction with anionic SDS led to an increase of β-structure up to 55%, accompanied by a decrease in α-helix up to a 13% (Table 1.4., Figure 1.3.); as it occurred in the H1^{0}3p/DNA complexes, where the amounts of β-structure and α-helix were 54% and 15%, respectively.

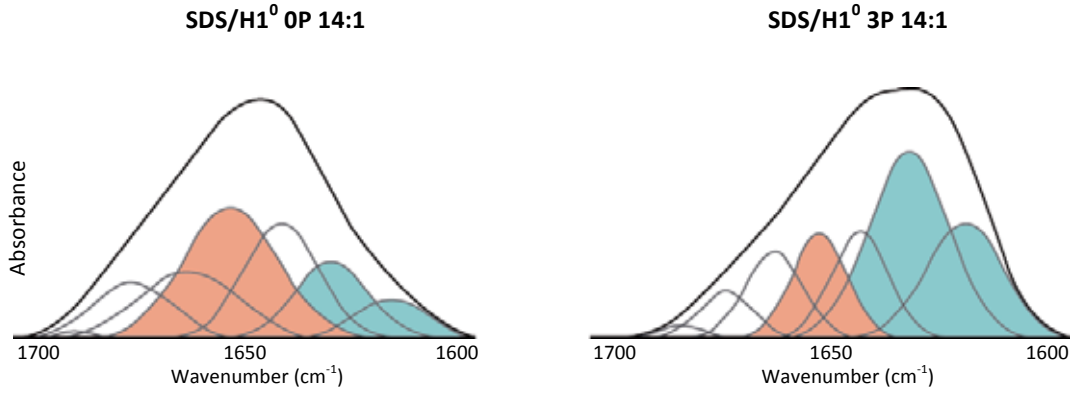


Figure 1.3. **Amide I decomposition of FTIR spectra in D_2O of unphosphorylated (H1^{0}0P) and phosphorylated (H1^{0}3P) histone H1^0 in the presence of SDS.** Amide I decomposition of the unphosphorylated H1^0 (H1^0 0P) and triphosphorylated H1^0 (H1^{0}3P). IR-spectra in D_2O. The buffer was 10 mM HEPES plus 140 mM NaCl, pH 7.0. The protein concentration was 5 mg/ml. The α-helix component is highlighted in red and the β-structure components are highlighted in blue. The SDS/H1 molar ratio was 14:1. The positions of phosphorylation were 118/140/152, defined relative to the entire protein.

As previously mentioned, the globular domain corresponds to a 36% of the entire protein and it is structured in a three-helix bundle and a β-hairpin[220], whereas the CTD corresponds to the 52% of the entire H1 and remains unstructured in buffer solution[205]. Considering that, the results for secondary structure elements of the phosphorylated protein in the complexes with DNA and in the presence of anionic SDS seem also to indicate that the CTD of histone H1 was in an all-β conformation and that the β-structure had propagated, in part, into the globular domain.

Fibrillation of histone H1 in the presence of anionic SDS

Phosphorylated CTD in the presence of SDS (7:1, detergent/CTD molar ratio) became an all-β protein, with 83% β-structure and no α-helix[423]. The CTD in all-β conformation formed ribbon-like fibres of the amyloid type, as indicated by the strong birefringence in the presence of Congo red and ThT fluorescence enhancement[423]. In the entire H1, interaction with anionic SDS led to an increase of β-structure up to 55% indicating that the CTD within histone H1 was also in an all-β conformation.

In order to observe if the entire phosphorylated histone $H1^0$ also formed amyloid fibres in the presence of SDS as the CTD did, thioflavin phase-contrast microscopy and polarized microscopy with Congo red staining were carried out.

The fully phosphorylated histone H1 formed ribbon-like fibres in the presence of SDS at a molar ratio of 14:1 (SDS/histone) in 10mM HEPES pH 7.0 plus 140mM NaCl. Fibres were of the amyloid type, as judged by strong birefringence in the presence of Congo red. Fibres were also stained with thioflavine T (ThT), a fluorescent dye that shows some degree of specificity for amyloid fibres. ThT fluorescence enhancement of the amyloid fibres was also observed.

The conditions for the observation of amyloid fibres were those in which histone $H1^0$ is fully phosphorylated and has its CTD structured in an all-β conformation (Figure 1.4.). In the unphosphorylated $H1^0$, only amorphous aggregates of varied morphology were observed. Occasionally, non-fibrillar aggregates showed ThT fluorescence enhancement. (Figure 1.5.).

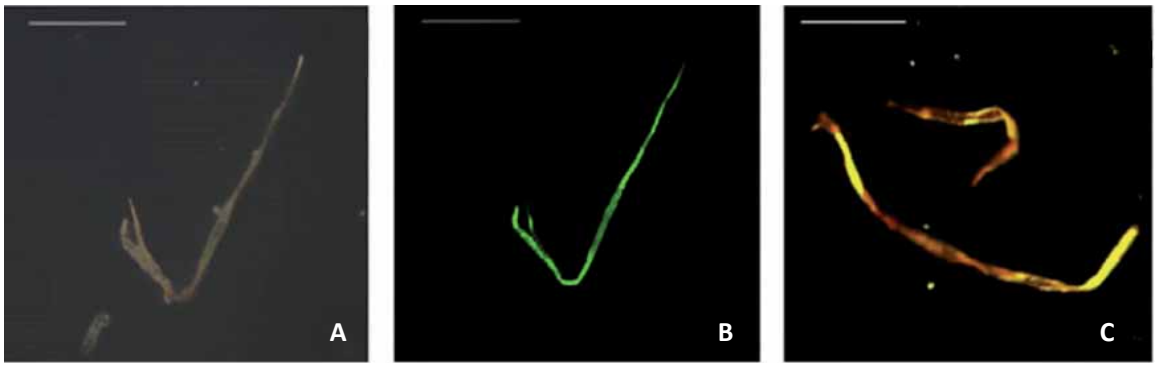


Figure 1.4. Fibrillation of histone H1 fibrils formed by the triphosphorylated H1 in the presence of SDS. The SDS/H1 molar ratio was 14:1. The concentration of protein was 5 mg/ml in 10mM HEPES plus 140mM NaCl, pH 7.0. All experiments were carried out at room temperature. (A) Phase-contrast microscopy image. (B) Complexes stained with thioflavin. (C) Birefringence of fibrils after staining with Congo red. The scale bars are 200 μm.

The entire H1 formed amyloid fibres in the presence of SDS, when the CTD was fully phosphorylated and adopted an all-β conformation. These results suggest that amyloid fiber formation requires the joint effects of hydrophobic interactions and some degree of charge neutralization, together with the all-β potential associated with full phosphorylation. More generally, our results reinforce the idea that amyloid formation is favoured by hydrophobic environments[424,425,426].

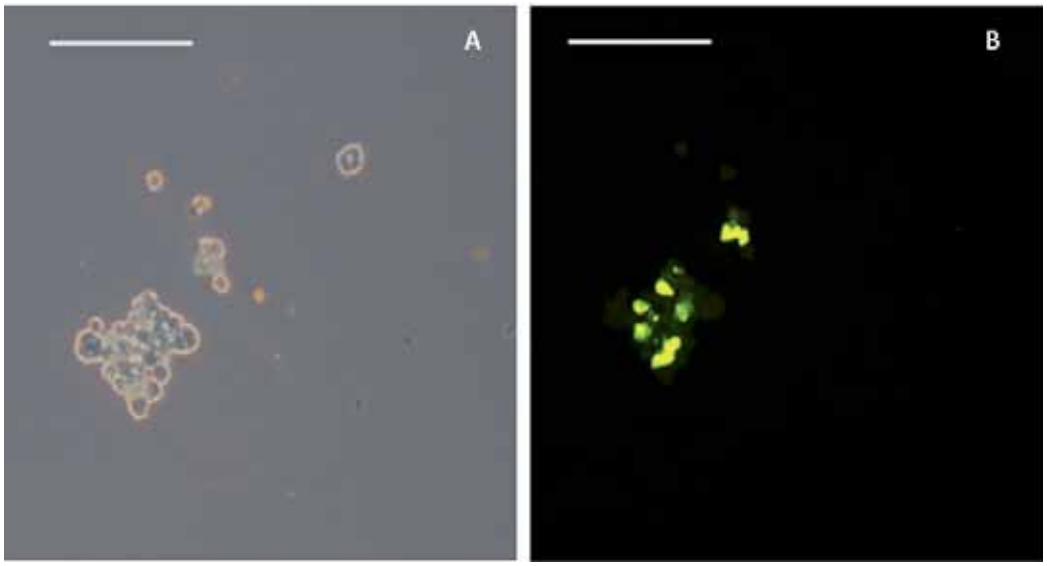


Figure 1.5. Aggregates formed by $H1^0$ in the presence of SDS. The concentration of protein was 5 mg/ml in 10 mM HEPES plus 140 mM NaCl, pH 7.0. All experiments were carried out at room temperature. The detergent/$H1^0$ molar ratio was 14:1. (A) Phase-contrast microscopy images of unphosphorylated $H1^0$ in the presence of SDS. (B) Unphosphorylated $H1^0$ in the presence of SDS stained with thioflavin. The scale bars are 200 μm.

In summary, increased hydrophobicity via interaction with hydrophobic ligands or via Lys positive charge compensation led to the folding of histone $H1^0$ with similar proportions of secondary structure motifs as in the DNA complexes. In general, phosphorylation increased the proportion of the β-structure and decreased that of α-helix. In the triphosphorylated H1, SDS, which combines hydrophobic and electrostatic effects, induced an all-β conformation (55% of β-structure in the entire protein) in the CTD, which corresponds to a 52% of the entire protein. The greatest effects were observed with the fully phosphorylated $H1^0$ in the presence of SDS (14:1 detergent/H1 molar ratio). In these conditions, the protein formed amyloid fibres. The folding of histone H1 upon interaction with hydrophobic ligands and the modulation of the secondary structure by phosphorylation could be relevant to the roles of histone H1 outside cell nuclei.

Chapter Two

Proteomic characterization and identification of novel post-translational modifications in chicken erythrocyte linker histones

Mature chicken erythrocyte nuclei contain highly condensed and inert chromatin, mainly consisting of DNA and histone proteins, which is characteristic of this type of differentiated cells[354].

Chicken erythrocyte nuclei contain six different subtypes of histone H1 (H1.01, H1.02, H1.03, H1.10, H1.1L and H1.1R) plus the avian specific subtype, histone H5[354,441]. Histone H5 replaces histone H1 in mature erythrocytes. Histone H5 has a natural variant in which an arginine residue is replaced by a glutamine (R16→Q16). The amount of histone H5 is higher than that of H1 subtypes. Histone stoichiometries in nuclei from different species and organs were determined by Bates and Thomas[427]. In their study, they found that, in chicken erythrocytes, core histones were present in equimolar amounts within the nucleosome, whereas H1 and H5 together constituted 1.3 moles per mole of the octamer. From those, 0.9 moles were contributed by H5 and 0.4 moles were contributed by H1 subtypes[427]. Thus, the 30% of nucleosomes contain two linker histones.

In recent experimental studies, H5 has been shown to represent approximately 60% of the total amount of linker histones[354,428]. This exchange has shown to reduce nucleosome mobility[150] resulting in increased chromatin stability, so that the histone H5 produces extensively repressed regions in the erythroid genome[428]. Studies performed in rat sarcoma cells showed that avian histone H5 has a higher binding affinity, compared to rat H1 subtypes[170] and this, together to its high content of arginine residues, locate H5 in repressed regions in the rat genome and allow H5 to generate more compacted chromatin fibres[429].

As mentioned above, in spite of the presence of nucleus, transcriptional activity and mitosis are almost totally repressed in avian erythrocytes[430]. Because of that, avian linker histones are thought to be poorly phosphorylated and uniform in terms of post-translational modifications (PTMs).

In the present chapter, chicken erythrocyte nuclei were treated with micrococcal nuclease and native chromatin was fractionated into soluble and insoluble fractions. Linker histones were analysed in order to observe their relative proportions in each chromatin fraction and to determine their post-translational modifications.

Characterization of the differences in linker histones between soluble and insoluble chromatin fractions and determination of their PTMs by proteomic methods were carried out during a PhD stay at the Division of Clinical Biochemistry, headed by Dr. Prof. Herbert Lindner in the Biocenter - Medizinische Universität Innsbruck (Austria). Figure 2.1. shows a scheme of the experimental procedure carried out.

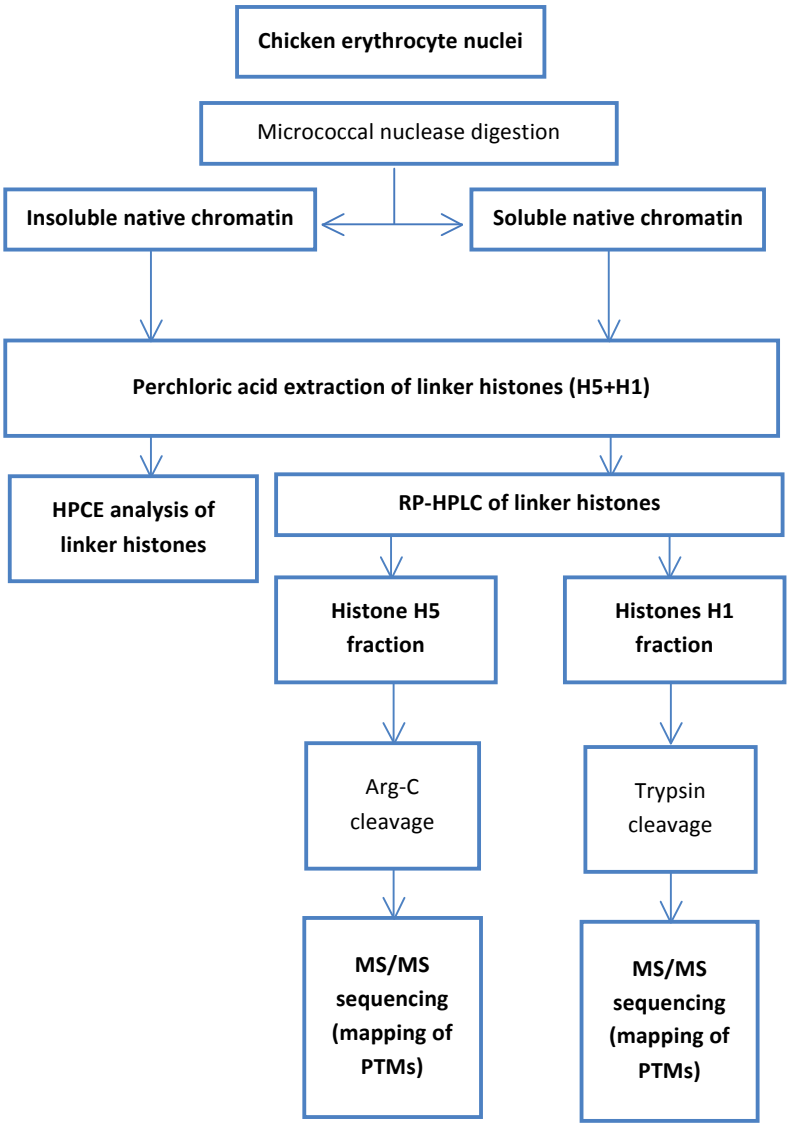


Figure 2.1. Scheme of the experimental procedure for the analysis of chicken linker histones from native chromatin by several proteomic techniques.

Analysis of linker histones by HPCE

Separation of chicken linker histones proved difficult for many years due to their chemical and physical properties and also because of the similarity and complexity of the protein members of the histone family[378,345]. Shannon and Wells[441] first separated three linker histone subfractions from chicken erythrocytes, but the presence of two variants of H5, six different subtypes of H1 and the potential post-translational modifications of all of them placed very high demands on the methods applied to overcome the resolution problems for histones H1/H5 isolation and identification. Application of different phosphate buffers, together with other changes in the analytical methods, such as the use of coated capillaries with HPMC, finally allowed the resolution of H1 and H5 subfractions by HPCE[349,350].

HPCE has been demonstrated to be a well-suited technique for resolving histone variants as well as modified isoforms, especially by acetylation and phosphorylation. Those two modifications greatly affect

the protein net charge resulting in the retardation of the migration time along the capillary of the modified species[348,350,354]. Because of that, this was the method chosen to assess the quality and purity of the samples and to identify the two linker histones (H1+H5) in each chromatin fraction.

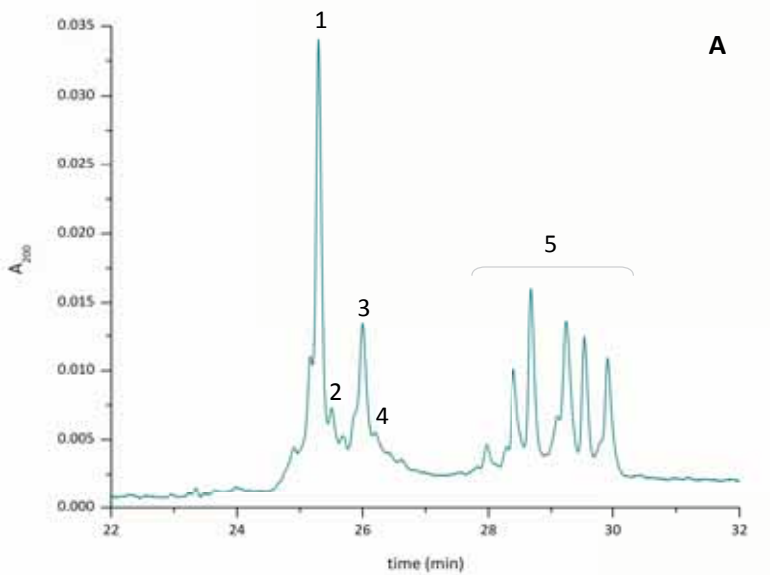
As already mentioned, our method was based in chromatin fractionation with micrococcal nuclease into two fractions: the soluble chromatin and the insoluble one. We analysed both pools by HPCE in order to determine if there were differences in the amount of H1 and H5 in the two fractions and/or in the basal phosphorylation level between them.

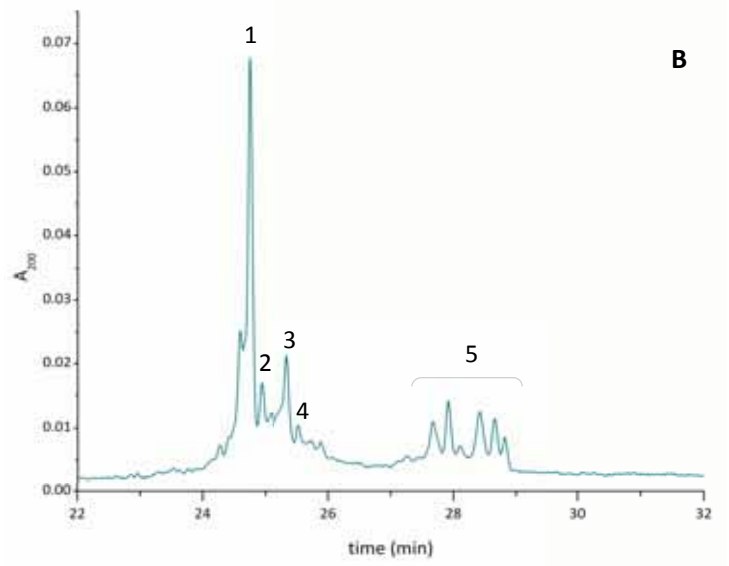
Two different regions, associated with the migration time of each linker histone in the capillary, were identified in the HPCE profiles (Figure 2.2. A/B). The first region, with migration times between 24 and 27 minutes, corresponded to histone H5 and its post-translationally modified species. The second region of the HPCE profile comprised migration times between 28 and 31 minutes and corresponded to histone H1 subtypes. Capillary electrophoresis (CE) allows separating species based on their size to charge ratio in the interior of a small capillary. Linker histones were positively charged in our buffer and pH conditions. Histone H5 is smaller than histones H1 and, because of that; its migration/retention time in the capillary was shorter. H5 natural variant was not identified in the HPCE profile, since a single amino acidic change does not introduce a change in mass enough to be detected by this technique. Hence, histone H5 and its natural variant had the same retention time. However, the addition of a chemical group due to a PTM leads to an increase of the retention time of the modified protein in the capillary with respect to the unmodified one, which allows the identification of modified species in the sample.

In the HPCE separations of purified linker histones from soluble and insoluble native chromatin extracted from chicken erythrocyte nuclei, the fractions that correspond to avian histone H5 (H5ac0; Figure 2.2., A/B, peak 1) and its acetylated species, H5ac1 (Figure 2.2., A/B, peak 3), as well as all six histone H1 subtypes (Figure 2.2., A/B, peaks 5) were identified. Secondary peaks, due to basal phosphorylation of the these species were also identified (Figure 2.2., A/B, peaks 2 and 4), as previously reported[243,282,344-350]. Regarding to the basal level of phosphorylation, there were no significant differences between those two HPCE profiles.

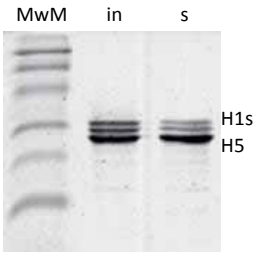
Note that the HPCE profile is represented as a function of the absorbance (A_{200}), which is determined by the amount of sample injected in the capillary. The intensity of each peak can be associated with the relative amount of each protein within the sample. In both samples (insoluble and soluble) histones H1 had similar intensities -with peak heights around 0.015UA- whereas the peak heights of histone H5 were considerably different between the insoluble (Figure 2.2.; A) and the soluble (Figure 2.2.; B) fractions. However, if we consider the maximum absorbance of the sample and the relative absorbance between the peaks corresponding to each linker histone (H1 vs. H5), we will see that, in the insoluble fraction, H5 is present in lower proportion with respect to H1s peaks. This fact suggests that the relative amount of histones H1 with respect to H5 histone is higher in the insoluble chromatin.

In order to confirm these differences, total linker histone fractions from insoluble and soluble chromatin fractions were separated and quantified in an SDS-PAGE (Figure 2.2, C/D).





C



D

		Band intensity	Total Intenstity	H1/H5 ratio
insoluble fraction	**H1s**	1,12 1,50	2,62	0,79
	H5	3,30	3,30	
soluble fraction	**H1s**	0,42 0,66	1,08	0,31
	H5	3,54	3,54	

Figure 2.2. HPCE separations of purified H5/H1 histone proteins extracted from erythrocyte nuclei chromatin, both insoluble (A) and soluble (B) chromatin fractions. SDS-PAGE of total linker histones fraction from chicken erythrocyte chromatin (C) and quantification of H1/H5 relative amounts in each fraction (D). In A and B, the run was performed on a Beckman system P/ACE 2100. Running conditions were as follows: 500 mM phosphate buffer, pH 2.0, containing 0.02% HPMC; untreated capillary (50cmx75µm); injection time, 5s; constant voltage (12 kV); temperature, 25°C; detection at 200nm. The marked peaks in the HPCE spectra correspond, in both cases, to: **1.** H5ac0p0: unacetylated H5 variant, unphosphorylated; **2**. H5ac0p1: unacetylated H5 variant, mono-phosphorylated; **3.** H5ac1p0: acetylated H5 variant, unphosphorylated; **4.** H5ac1p1: acetylated H5 variant, mono-phosphorylated; **5.** H1 histones fraction. In C, MwM: SDS-PAGE low range molecular weight protein standard (BioRad); in: linker histones from insoluble chromatin fraction; s: linker histones from soluble chromatin fraction. In D, quantification of the bands' intensity in the gel (OD*mm^2) and calculation of the H1/H5 ratio.

The intensity of the bands corresponding to histone H5 was approximately the same in both chromatin fraction, whereas the intensity of the bands corresponding to histones H1 in the insoluble fraction (2.62UA*mm^2) was twice the intensity of the soluble one (1.08UA*mm^2) (Figure 2.2, C and D). Calculation of the ratio between linker histones (H1/H5) confirmed the results observed in the HPCE analysis: H1/H5 ratio in the insoluble chromatin fraction was around 0.8 whereas it was 0.31 in the soluble fraction (Figure 2.2, D).

This data indicate that the amount of H5 is the same in both fractions but that there are more histone H1 molecules in the insoluble chromatin. This fact suggests that the 30% of nucleosomes that contain two histone molecules are located in the insoluble chromatin fraction.

Novel post-translational modifications in chicken erythrocyte linker histones

After chicken native chromatin fractionation with micrococcal nuclease, linker histones were isolated by perchloric acid extraction. Next, histone H5 and histones H1 were further separated by RP-HPLC and enzymatically cleaved with Arg-C and Trypsin, respectively. Afterwards, tandem mass spectrometry (LC-ESI-MS/MS) was carried out in order to map the post-translational modifications of all linker histones subtypes in each chromatin fraction (Figure 2.1).

Previous and similar studies on linker histone PTMs characterization[253,431,243], were considered in order to adjust the conditions of the experiment. In all the references consulted, enzymatic cleavage had been carried out under similar buffer conditions and temperature of reaction. However, in most cases, cleavage was performed for long periods (1-6h, even 24h in some case) and with high amounts of enzyme (w/w).

Different enzymes, protein concentrations, enzyme/protein ratios and times of enzymatic cleavage were tested in order to optimize the results in the LC-ESI-MS/MS experiment and achieve high sequence coverages that will permit to map as many sites as possible. Times of reaction and enzyme/protein ratios similar to those found in bibliography resulted in low sequence coverage due to extensive digestion of the proteins and, thus, sample loss. These parameters were then scaled down until maximum sequence coverage could be determined. Shorter times of reaction than the ones finally chosen were also tested, but they resulted in very little cleavage of the proteins.

The enzymatic cleavage conditions that ensured the best sequence coverage in the Tandem MS experiments were as follows: H1 fraction was digested with Trypsin (1:100, w/w) for 15 minutes whereas H5 fraction was digested for 5minutes with Arg-C (1:40, w/w). All six isoforms of the chicken histone H1 and the avian isoform H5 were detected in both samples. Table 2.1. shows the best sequence coverages (%) obtained in the MS/MS experiments per histone subtype and chromatin fraction.

Coverage values oscillated between the 52% and the 88% of coverage, with only a few proteins covered above an 80%. Peptides corresponding to the N-terminal and the globular domain of the proteins were found in most cases. C-terminal domains of linker histones, which is very lysine-rich and, thus, more susceptible to enzymatic cleavage, were the region of the proteins with less coverage.

Table 2.1. Sequence coverage (%) of chicken histones from native chromatin in MS/MS experiments.

Swiss-Prot Acc. Num.	Histone variant	Sequence coverage of each variant per fraction (%)	
		Insoluble fraction	Soluble fraction
P02259	H5	81	82
P08284	H1.01	60	59
P09987	H1.02	57	52
P08285	H1.03	66	67
P08286	H1.10	59	59
P08287	H1.1L	59	88
P08288	H1.1R	54	81

A nanoLC-ESI-MS/MS strategy was carried out to determine the PTMs present in chicken erythrocyte linker histones. Arg-C cleavage products of histone H5 and trypsin cleavage products of H1 histones at every chromatin phosphorylation time were analysed using nano-HPLC connected online to a linear ion trap mass spectrometer (LTQ Velos ThermoScientific) equipped with a nanospray ionization source for MS/MS sequencing and PTMs mapping. Data-dependent neutral loss MS^3 was used for precise localization of phosphorylation. MS^2 and MS^3 spectra were searched against a histone data base using SEQUEST and validated manually. All possible modifications were considered as variable modifications in all searches. Peptides scoring X^2corr values <2.5 in their MS^2 spectra were rejected whereas all those modified peptides with X^2corr ≥2.5 were subjected to manual verification (Tables 2.3. and 2.5.).

Snijders *et.al.* already used mass spectrometry to characterize post-translational modifications of the linker histones H1 and H5 from chicken erythrocytes[432]. In their study, they purified linker histones directly from cell nuclei under non denaturing conditions. Thus, they determined a wide range of post-translational modifications in chicken linker histones extracted from bulk chromatin, considering linker histones PTMs to be uniform in the whole chromatin. They determined post-translational modifications including N^{α}-acetylation, phosphorylation, acetylation, methylation and deamidation.

In their study, they determined the presence of N^{α}-peptides from H5 and all six H1 subtypes in both the acetylated and unacetylated forms. Both peptides showed to be widespread and prevalent in the analysis, indicating that the abundance on the unacetylated N^{α}-peptides may reflect the mature nature of the cells used[432], in contrast with other studies on human H1[242]. Phosphorylation on the chicken linker histones was also identified by Snijders *et.al.* The majority of it occurred in histone H5 in contrast to H1 subtypes, and predominantly on non-CDK sites of the N-terminal region of the protein[432]. Different sites of acetylation and methylation (including lysine mono- and di-methylation) were also identified in their study. Some of these modifications were predominant in the N- and C-terminal domains of chicken H1 subtypes, and several sites showed to have different modifications in the same lysine residue[432].

In our study, we fractionated native chromatin from chicken erythrocytes into soluble and insoluble fractions. Novel post-translational modifications in linker histones extracted from those two fractions were identified by ESI/LC-MS/MS, including N^{α}-acetylation, lysine acetylation and methylation and phosphorylation. This analysis also allowed to find differences in specific PTMs between the histones extracted from insoluble and soluble fractions of native chromatin, indicating the existence of a different pattern of post-translational modifications between those two fractions.

Novel post-translational modifications of histone H5

Avian histone H5 (P02259) is a histone protein of 190 amino acids and a molecular weight of ~20734Da. Histone H5 can be N^{α}-acetylated and it could be potentially phosphorylated at four serine (S) residues (UniProt). Table 2.2. shows the theoretical data entered in the UniProt database for histone H5.

Table 2.2. Theoretical data of chicken histone H5 (UniProt)

Swiss-Prot accession number	P02259
Length	190 aa
Mw (average mass)	20733.69 Da
Mw (monoisotopic mass)	20721.19 Da
Initiator methionine	M1
Chain	2-190
Natural variant	16 R→Q
Probable phosphorylation sites	S23, S30, S146, S167
Acetylation	T2

Numbering does not consider N^{α}-terminal methionine cleavage

Two novel post-translational modifications, together with others already described, were identified for histone H5 in this study. Table 2.3. and Figure 2.2 summarize the post-translational modifications found for this linker histone.

A. PTMs in histone H5 from insoluble chromatin fraction

```
         10         20         30         40         50         60
TESLVLSPAP AKPKRVKASR RSASHPTYSE MIAAAIRAEK SRGGSSRQSI QKYIKSHYKV

         70         80         90        100        110        120
GHNADLQIKL SIRRLLAAGV LKQTKGVGAS GSFRLAKSDK AKRSPGKKKK AVRRSTSPKK

        130        140        150        160        170        180
AARPRKARSP AKKPKATARK ARKKSRASPK KAKKPKTVKA KSRKASKAKK VKRSKPRAKS

       189
GARKSPKKK
```

B. PTMs in histone H5 from soluble chromatin fraction

```
         10         20         30         40         50         60
TESLVLSPAP AKPKRVKASR RSASHPTYSE MIAAAIRAEK SRGGSSRQSI QKYIKSHYKV

         70         80         90        100        110        120
GHNADLQIKL SIRRLLAAGV LKQTKGVGAS GSFRLAKSDK AKRSPGKKKK AVRRSTSPKK

        130        140        150        160        170        180
AARPRKARSP AKKPKATARK ARKKSRASPK KAKKPKTVKA KSRKASKAKK VKRSKPRAKS

       189
GARKSPKKK
```

Figure 2.2. Post-translational modifications (PTMs) identified in histone H5 extracted from insoluble (A) and soluble (B) chromatin fractions. Initiator methionine has been removed from the sequence. The globular domain appears underlined in each case. The DNA binding sites are highlighted in bold italics typing. The amino acid modifications are highlighted as follows: red=acetylation; green=phosphorylation; light blue=oxidation.

The first novel PTM was K14 acetylation. This modification was identified in both chromatin fractions. No modifications had been previously reported in this position for chicken erythrocyte H5. MS^2 spectrum for this modification is shown in Figure 2.3 (panel A).

The second novel PTM was T84 phosphorylation, which was found in the soluble fraction but not in the insoluble one. It is interesting to note that T84 is right next to K85, which is actually part of the primary binding site of DNA (Figure 2.2). Phosphorylation reduces the binding affinity of linker histones for DNA, relaxing the chromatin. Thus, the fact that this modification was found in H5 from the solubilised chromatin fraction suggests that phosphorylation in this position might play a role in the differences in chromatin compaction between this two fractions. Figure 2.3 (panel B) shows the mass spectrum for this modification.

In the present study, M31 oxidation was also identified. Nevertheless, we do not consider it a truly PTM since oxidation of the sulphur group of methionine residues is a common and well-known post-translational modification that increases with age[433]. Taking into account that chicken erythrocytes are terminally differentiated cells; this modification could be somehow expected. However, it has never been reported before. Snijders *et.al.* treated it as a variable modification in their searches, together with N^{α}-acetylation[432] and, although they found the latter in high amounts; they did not report the oxidation of M31.

Table 2.3. Modified peptides identified by ESI/LC- MS/MS in chicken erythrocytes histone H5

Sample	Peptide sequence	Position	PTM	MS^2 Xcorr
H5 insoluble native chromatin	TESLVLSPAPAKPK	1-14	T1-Acetyl	3.75
			S3-Phospho	3.50
			S7-Phospho	3.46
			T1-Acetyl, S3-Phospho	3.40
			T1-Acetyl, S7-Phospho	3.01
			T1-Acetyl, K14-acetyl	4.09
			K12-Acetyl	3.68
	SASHPTYSEMIAAAIR	22-38	M10-Oxidation	3.89
H5 soluble native chromatin	TESLVLSPAPAKPK	1-14	T1-Acetyl	3.41
			S3-Phospho	3.61
			S7-Phospho	3.82
			T1-Acetyl, S3-Phospho	2.58
			T1-Acetyl, S7-Phospho	3.58
			T1-Acetyl, K14-Acetyl	3.15
	QTKGVGASGSFR	83-95	T2-Phospho	3.22
	SASHPTYSEMIAAAIR	22-38	M10-Oxidation	4.21

N^{α}-acetylation of proteins is the most common post-translational modification in eukaryotic proteins, since it takes place together with N^{α}-terminal methionine cleavage[434]. Sjniders *et.al.* reported T1 as a phosphorylated and N^{α}-acetylated residue[242,432]. In the present study, only N^{α}-acetylation in T1 was identified; in both, soluble and insoluble, chromatin fractions. However, both the acetylated and unacetylated forms of this peptide were found, suggesting that the unacetylated peptide is also present in mature erythrocytes. S3 and S7 were also found phosphorylated in the N^{α}-peptides of H5 from both chromatin fractions, as already reported by Sjniders *et.al.* These modifications were identified in combination with T1-acetylation and as a single modification of the peptide. According to PHOSIDA (phosphorylation site database), S3 is the residue that can be phosphorylated in the canonical motif of GSK3 kinase (**S**-X-X-X-S), which is involved in the regulation of metabolism, the cytoskeleton, and gene expression[435] (GSK3B in *G. gallus*). S7 is the residue that is phosphorylated in CK1 kinase motif (S/T-X-X-X-**S**).

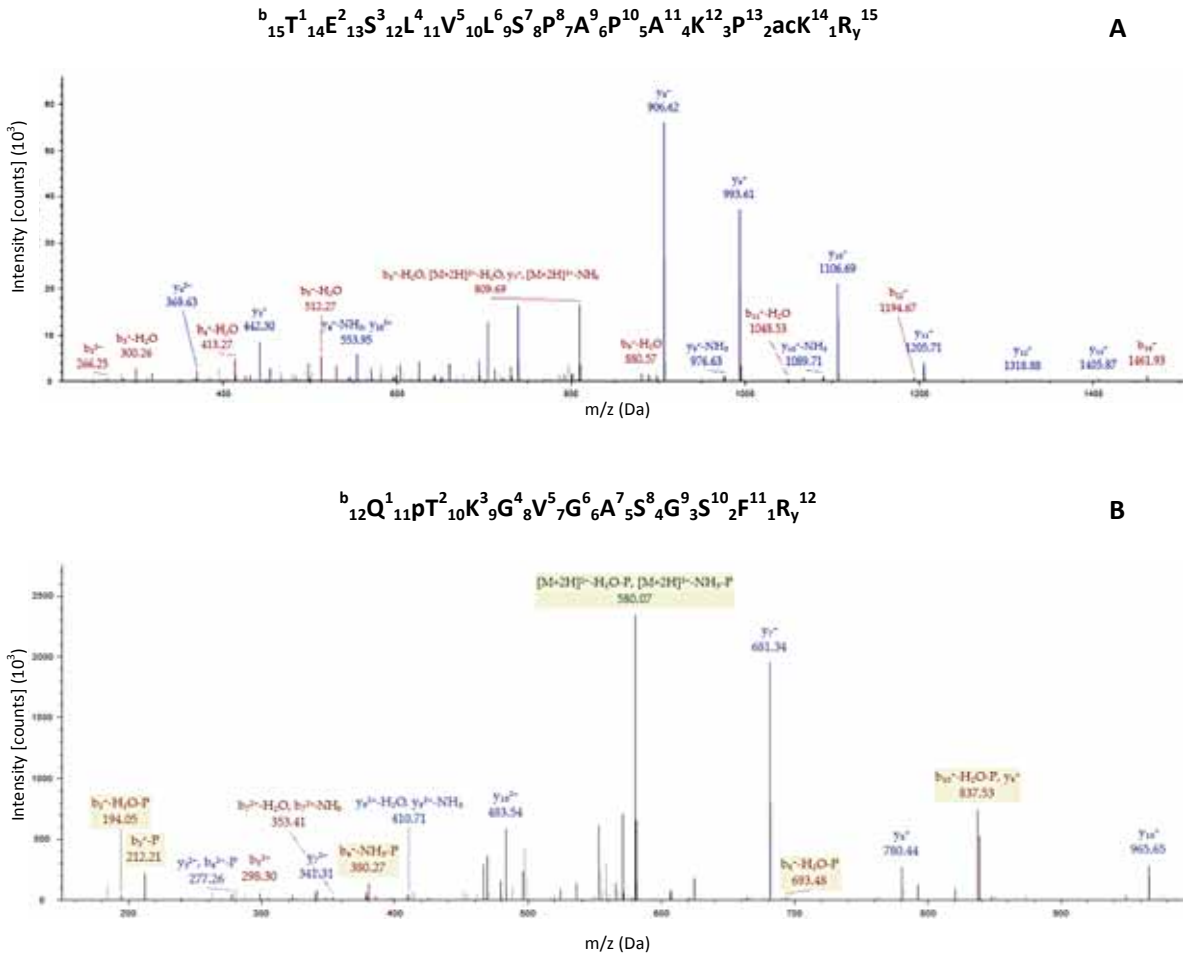


Figure 2.3. Mass spectra of identified peptides of chicken histone H5 containing novel post-translational modifications. Peaks are labelled in the mass spectra and the fragment ions are indicated in the peptide sequence. Peptides identification with SEQUEST (v1.13). Fragment match tolerance used for search: 0.8 Da **A.** MS^2 spectra of novel lysine acetylation (ac) on residue K14. XCorr: 3.45; Probability: 71.80. **B.** MS^3 spectra of novel threonine phosphorylation (p) on residue T84. Xcorr: 2.22; Probability: 32.36. In green, neutral loss of phosphate group.

Phosphorylation of non-CDK1 sites within the N-terminal tail has already been reported. The combination of N^{α}-acetylation and phosphorylation in (S/T)1 and phosphorylation in (S/T)3 leads to the SET motif, which has been found in histone H1 from mouse spleen[242] and human[242,436], This motif has also been identified in chicken histone H5 [432].

In the present study, K12-acetylation was determined. This PTM was only found in H5 histone extracted from the insoluble fraction of native chromatin. In their study, Sjniders *et.al.* identified two different modifications for K12 residue, monomethylation and acetylation[432].

In contrast to Snjiders *et.al.,* S22 and S24 phosphorylation were not identified in the present study. Neither S129 (SPKK-3), which was reported as phosphorylated, could be identified because the coverage of the experiment did not included the sequence in which this motif is present. None of the probable phosphoserines reported for chicken H5 in the Uniprot database (Table 2.2) have never been identified.

Sequence coverages for histone H5 were around 80% but the overall content of identified PTMs was low in both chromatin fractions. Most of the PTMs were found in the N-terminal and the globular domains.

Novel post-translational modifications of histone H1 subtypes

Histone H1 fraction in mature chicken erythrocytes is composed of six different subtypes[437], named H1.01, H1.02, H1.03, H1.10, H1.1L and H1.1R, according to the six histone H1 genes identified by Coles *et. al.* in 1987[438]. Histone H1.02 (P09987) is referred to as histone H1 in the Swiss-Prot database and it corresponds to the H1.02 gene; so it will be named in the present study (Table 2.1. and Table 2.4). Most of the N^{α}-terminal peptides are reported to be acetylated but no other PTMs are entered for none of the histones H1 in the UniProt database (Table 2.4.)

Table 2.4. Theoretical data of chicken histone H1 subtypes (UniProt)

Accession number (Swiss-Prot)	Histone H1 subtype/gene	Length (aa)	Mw_{av} (Da)	Amino acid modifications (Uniprot)
P08284	H1.01	219	21913,31	acetyl-S1
P09987	H1.02* (H1)	218	21750,21	acetyl-S1
P08285	H1.03	224	22395,06	-
P08286	H1.10	220	21872,29	acetyl-S1
P08287	H1.1L	225	22396,99	acetyl-S1 (by similarity)
P08288	H1.1R	219	21672,14	-

*Swiss-Prot accession number P09987 is referred to as Histone H1 in the Swiss-Prot database and it corresponds to the H1.02 gene sequence. It is referred to as H1.02 in this study.

Figure 2.4. shows the Percent Identity Matrix obtained from the sequence alignment of the six chicken histone H1 subtypes in Clustal O (1.1.1). As it can be seen, the six subtypes have a high sequence similarity and share a percentage of identity that varies from 88.48% to 94.04%.

```
H1.01_chick   100.00
H1.02_chick    91.71  100.00
H1.03_chick    88.48   86.57  100.00
H1.10_chick    93.55   91.67   88.99  100.00
H1.1L_chick    91.74   89.40   88.79   92.24  100.00
H1.1R_chick    91.51   88.63   90.78   91.55   94.04  100.00
```

Figure 2.4. Percent Identity Matrix of chicken histone H1 subtypes multiple sequence alignment in Clustal O (1.1.1.)

In the present study, six novel post-translational modifications were identified in histone H1 subtypes from chicken erythrocytes chromatin. Table 2.5. and Figure 2.5. summarize all the modified peptides found. Considering the highest sequence identity between H1 subtypes, some peptides matched with more than one H1 subtype (in these cases, a list of the histone subtypes that share the same peptide is shown in the table). When a PTM is found in a 'shared-sequence' peptide, one cannot assure whether the modification is present in one or more H1 subtypes from the mixture. Figure 2.6. shows the mass spectra of modified peptides with novel PTMs identified in the present study.

A. PTMs in histones H1 from insoluble chromatin fraction. Sequence alignment in Clustal O (1.1.1.)

```
sp|P08284|H101      SETAPA----AAPDAPAPGAKAAAKKPKKAAGGAKARKPAGPSVTELITKAVSASKERKG   56
sp|P09987|H102      SETAPV----AAPAVSAPGAKAAAKKPKKAAGGAKPRKPAGPSVTELITKAVSASKERKG   56
sp|P08285|H103      AETAPVAAPDVAA-APTPAKAAPAKKPKKAAGGAKARKPAGPSVTELITKAVSASKERKG   59
sp|P08286|H110      SETAPA----AAPAVAAPAAKAAAKKPKKAAGGAKARKPAGPSVTELITKAVSASKERKG   56
sp|P08287|H11L      SETAPAPAAEAAPAAAPAPAKAAAKKPKKAAGGAKARKPAGPSVTELITKAVSASKERKG   60
sp|P08288|H11R      AETAPAAAPA---AAPAPAAKAAAKKPKKAAGGAKARKPAGPSVTELITKAVSASKERKG   57
                    :****.        .       * ************ ************************

sp|P08284|H101      LSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGSFRLNKKPGEVKE  116
sp|P09987|H102      LSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGSFKLNKKPGETKA  116
sp|P08285|H103      LSLAALKKALAAGGYDVEKSNSRIKLGLKSLVSKGTLVQTKGTGASGSFRLSKKSGDVKE  119
sp|P08286|H110      LSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGSFRLSKKPGEVKE  116
sp|P08287|H11L      LSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGSFRLSKKPGEVKE  120
sp|P08288|H11R      LSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGSFRLSKKPGEGLE  117
                    ******************.*****************************:*.** *:

sp|P08284|H101      KAPRKRATAAKPKKPAAKKPAAAAKKPKKAAAVKKSPKKAKKPAAAATKKAAKSPKKAAK  176
sp|P09987|H102      KATKK-KPAAKPKKPAAKKPAAAAKKPKKAAAVKKSPKKAKKPAAAATKKAAKSPKKATK  175
sp|P08285|H103      KAPKKKTPAAKPKKPAAKKPAAAAKKPKKAVAVKKSPKKAKKPAAAATKKAAKSPKKVTK  179
sp|P08286|H110      KAPRKRTPAAKPKKPAAKKPASAAKKPKKAAAAKKSPKKAKKPAAAATKKAAKSPKKATK  176
sp|P08287|H11L      KAPKKKASAAKPKKPAAKKPAAAAKKPKKAVAVKKSPKKAKKPAASATKKSAKSPKKVTK  180
sp|P08288|H11R      KAPKKKASAAKPKKAAAKKPAAAAKKPKKAVAVKKSPKKAKKPAASATKKSVKSPKK---  174
                    ** :*   ****** ******:********.*.************:****:.*****

sp|P08284|H101      AGRPKKA--AKSPAKAKAVKPKAAKPKATKPKAAKAKKTAAKKK  218
sp|P09987|H102      AGRPKKT--AKSPAKAKAVKPKAAKSKAAKPKAAKAKKAATKKK  217
sp|P08285|H103      AAKPKKAVAVKSPAKAKAVKPKAAKPKATKPKAAKAKKAAPKKK  223
sp|P08286|H110      AAKPKKAATAKSPAKAKAVKPKAAKPKAAKPKAAKAKKAAAKK-  219
sp|P08287|H11L      AVKPKKAVAAKSPAKAKAVKPKAAKPKAAKPKAAKAKKAAAKKK  224
sp|P08288|H11R      AAKPKKAVAAKSPAKAKAVKPKAAKPKAAKPKAAKAKKAAAKKK  218
                    * :***:  .*************** **:**********:* **
```

B. PTMs in histones H1 from soluble chromatin fraction. Sequence alignment in Clustal O (1.1.1.)

```
sp|P08284|H101      SETAPA----AAPDAPAPGAKAAAKKPKKAAGGAKARKPAGPSVTELITKAVSASKERKG   56
sp|P09987|H102      SETAPV----AAPAVSAPGAKAAAKKPKKAAGGAKPRKPAGPSVTELITKAVSASKERKG   56
sp|P08285|H103      AETAPVAAPDVAA-APTPAKAAPAKKPKKAAGGAKARKPAGPSVTELITKAVSASKERKG   59
sp|P08286|H110      SETAPA----AAPAVAAPAAKAAAKKPKKAAGGAKARKPAGPSVTELITKAVSASKERKG   56
sp|P08287|H11L      SETAPAPAAEAAPAAAPAPAKAAAKKPKKAAGGAKARKPAGPSVTELITKAVSASKERKG   60
sp|P08288|H11R      AETAPAAAPA---AAPAPAAKAAAKKPKKAAGGAKARKPAGPSVTELITKAVSASKERKG   57
                    :****.        .       * ************ ************************

sp|P08284|H101      LSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGSFRLNKKPGEVKE  116
sp|P09987|H102      LSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGSFKLNKKPGETKA  116
sp|P08285|H103      LSLAALKKALAAGGYDVEKSNSRIKLGLKSLVSKGTLVQTKGTGASGSFRLSKKSGDVKE  119
sp|P08286|H110      LSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGSFRLSKKPGEVKE  116
sp|P08287|H11L      LSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGSFRLSKKPGEVKE  120
sp|P08288|H11R      LSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGTLVQTKGTGASGSFRLSKKPGEGLE  117
                    ******************.*****************************:*.** *:

sp|P08284|H101      KAPRKRATAAKPKKPAAKKPAAAAKKPKKAAAVKKSPKKAKKPAAAATKKAAKSPKKAAK  176
sp|P09987|H102      KATKK-KPAAKPKKPAAKKPAAAAKKPKKAAAVKKSPKKAKKPAAAATKKAAKSPKKATK  175
sp|P08285|H103      KAPKKKTPAAKPKKPAAKKPAAAAKKPKKAVAVKKSPKKAKKPAAAATKKAAKSPKKVTK  179
sp|P08286|H110      KAPRKRTPAAKPKKPAAKKPASAAKKPKKAAAAKKSPKKAKKPAAAATKKAAKSPKKATK  176
sp|P08287|H11L      KAPKKKASAAKPKKPAAKKPAAAAKKPKKAVAVKKSPKKAKKPAASATKKSAKSPKKVTK  180
sp|P08288|H11R      KAPKKKASAAKPKKAAAKKPAAAAKKPKKAVAVKKSPKKAKKPAASATKKSVKSPKK---  174
                    ** :*   ****** ******:********.*.************:****:.*****

sp|P08284|H101      AGRPKKA--AKSPAKAKAVKPKAAKPKATKPKAAKAKKTAAKKK  218
sp|P09987|H102      AGRPKKT--AKSPAKAKAVKPKAAKSKAAKPKAAKAKKAATKKK  217
sp|P08285|H103      AAKPKKAVAVKSPAKAKAVKPKAAKPKATKPKAAKAKKAAPKKK  223
sp|P08286|H110      AAKPKKAATAKSPAKAKAVKPKAAKPKAAKPKAAKAKKAAAKK-  219
sp|P08287|H11L      AVKPKKAVAAKSPAKAKAVKPKAAKPKAAKPKAAKAKKAAAKKK  224
sp|P08288|H11R      AAKPKKAVAAKSPAKAKAVKPKAAKPKAAKPKAAKAKKAAAKKK  218
                    * :***:  .*************** **:**********:* **
```

Figure 2.5. Post-translational modifications (PTMs) identified in histones H1 extracted from insoluble (A) and soluble (B) chromatin fractions. Sequence alignment in Clustal O (1.1.1). Initiator methionine has been removed from the sequence. The globular domain appears underlined. The amino acid modifications are highlighted as follows: red=acetylation; green=phosphorylation; magenta=di-methylation; light blue=oxidation.

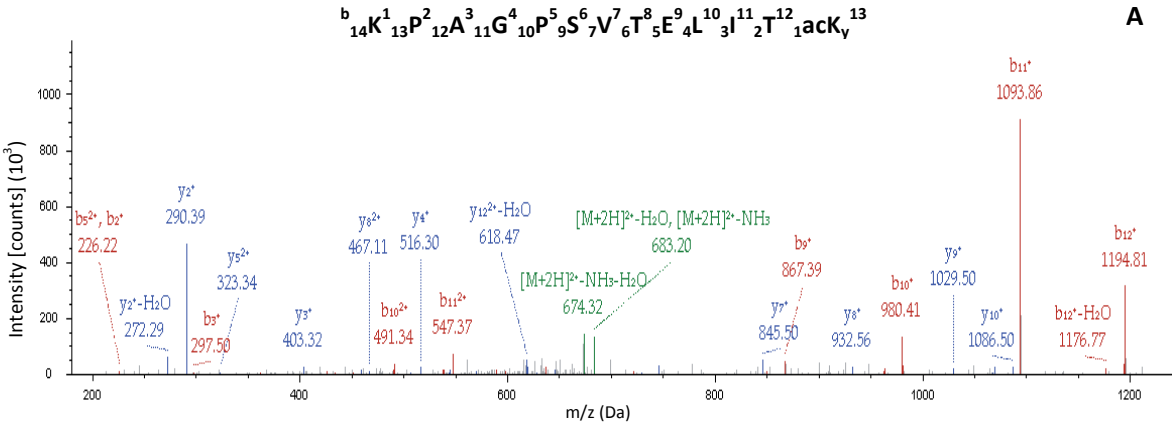
Three novel PTMs were identified at the beginning of the globular domain of histone H1 subtypes, in the shared peptide KPAGPSVTELITK: K34-acetylation, S39-phosphorylation and K46-acetylation (positions in H1.01, for equivalent numbering in the other H1 subtypes, Figure 2.5.). K34-acetylation and S39-phosphorylation were found in a shared-peptide of all histone H1 subtypes extracted from both fractions of chromatin (Figure 2.6., E and F), with the exception of H1.02, in which S39 phosphorylation was only found in the insoluble fraction. K46-acetylation was also found in this shared-peptide, but only in the insoluble fraction (Figure 2.6., A). Equivalent positions of K46 had been reported to be also acetylated in several mouse and human histones H1[242].

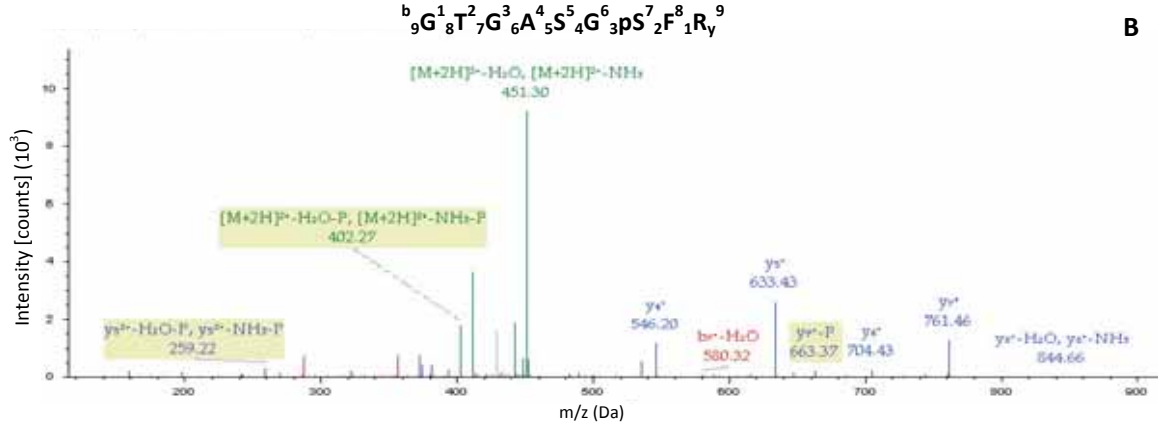
Snijders *et. al.*[432] reported T3 phosphorylation in H1 subtypes H1.1L, H1.03 and H1.1R. This modification, as already explained for H5, resembles the SET motif found in several histones subtypes. In the present study, this modification was not identified in H1.03 neither in H1.1R; but it was present in H1.1L, which had already been described, and in H1.02, which is a novel PTM. T3 was found in these two subtypes extracted from both fractions of chromatin (Figure 2.6., F). This fact indicates again that, when fractionating bulk native chromatin into two pools, post-translational modifications of linker histones are not uniform between fractions.

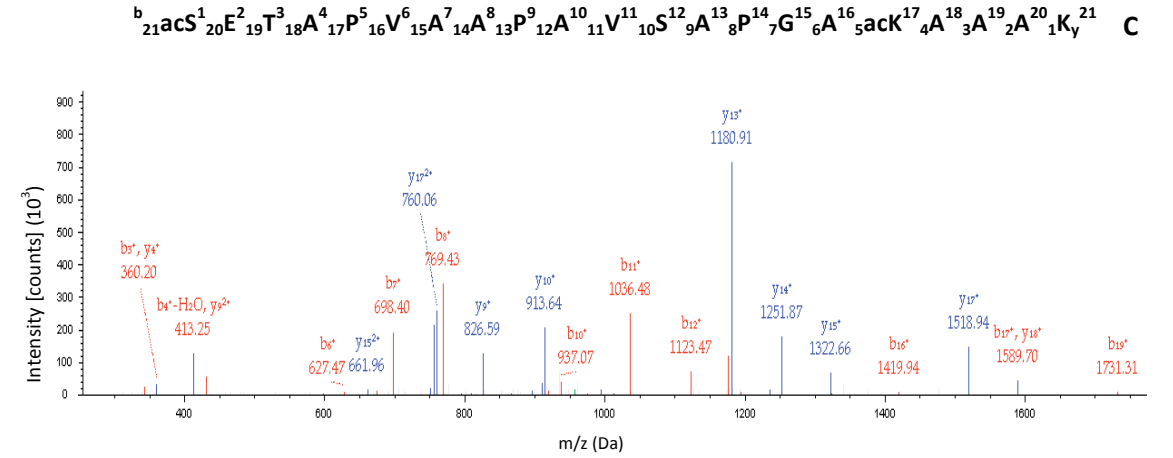
Table 2.5. Modified peptides identified by ESI/LC-MS/MS in chicken erythrocyte histones H1

Sample	Peptide sequence	Positions	H1 histone variant	PTM	MS^2 Xcorr
H1 insoluble chromatin	SETAPAAAPDAPAPGAK	1-17	H1.01	S1-Acetyl	3.70
	SETAPVAAPAVSAPGAK(AAAK)	1-21	H1.02	S1-Acetyl	3.81
				K17-Acetyl	2.60
				T3-Phospho	2.79
	AETAPVAAPDVAAAPTPAK	1-19	H1.03	A1-Acetyl	2.18
	SETAPAAAPAVAAPAAK	1-17	H1.10	S1-Acetyl	3.10
	SETAPAPAAEAAPAAAPAPAK	1-21	H1.1L	S1-Acetyl	3.81
				T3-Phospho	3.91
	AETAPAAAPAAAPAPAAK	1-18	H1.1R	A1-Acetyl	2.99
	KAVAAKSPAKAK	186-197 (H1.1L)	H1.1L, H1.1R	K6-Acetyl	2.66
	SLVSKGTLVQTK	86-97 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K5-Acetyl	3.07
	GTGASGSFR	98-106 (H1.01)	H1.01, H1.03, H1.10, H1.1L, H1.1R	S7-Phospho	3.06
	ARKPAGPSVTELITK	32-46 (H1.01)	H1.01, H1.03, H1.10, H1.1L, H1.1R	S8-Phospho	3.19
	KPAGPSVTELITK	34-46 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K1-Acetyl	3.78
				S6-Phospho	3.22
				K12-Acetyl	2.65
	KALAAGGYDVEK	64-75 (H1.01)	H1.01, H1.02, H1.10, H1.1L, H1.1R	K12-Dimethyl	3.22
	ALAAGGYDVEK	65-75 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K11-Dimethyl	2.49
H1 soluble chromatin	SETAPAAAPDAPAPGAK	1-17	H1.01	S1-Acetyl	3.17
	SETAPVAAPAVSAPGAK	1-17	H1.02	S1-Acetyl	3.80
				T3-Phospho	3.08
	AETAPVAAPDVAAAPTPAK	1-19	H1.03	A1-Acetyl	3.59
	SETAPAAAPAVAAPAAK	1-17	H1.10	S1-Acetyl	3.63
	SETAPAPAAEAAPAAAPAPAK	1-21	H1 1L	S1-Acetyl	3.45
				T3-Phospho	3.07
	AETAPAAAPAAAPAPAAK	1-18	H1 1R	A1-Acetyl	2.28
	ARKPAGPSVTELITK	32-46 (H1.01)	H1.01, H1.03, H1.10, H1.1L, H1.1R	S8-Phospho	3.60
	KPAGPSVTELITK	34-46 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K1-Acetyl	3.81
	ALAAGGYDVEK	65-75 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K11-Dimethyl	2.44
	SLVSKGTLVQTK	86-97 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K5-Acetyl	2.52
	KAVAAKSPAKAK	186-197 (H1.1L)	H1.1L, H1.1R	K6-Acetyl	2.61

Modified residues positions are referred considering their location within the peptide found in the analysis.
Position of each peptide within the protein sequence is given in the second column.
For 'shared peptides' in different histone subtypes, the positions given correspond to the histone subtype in brackets.
For equivalent numbering in the other H1 subtypes, see amino acid sequences displayed in Figure 2.5.


$^{b}{}_{14}K^{1}{}_{13}P^{2}{}_{12}A^{3}{}_{11}G^{4}{}_{10}P^{5}{}_{9}S^{6}{}_{7}V^{7}{}_{6}T^{8}{}_{5}E^{9}{}_{4}L^{10}{}_{3}I^{11}{}_{2}T^{12}{}_{1}acK_{y}{}^{13}$
A
Intensity [counts] (10^3)
m/z (Da)
b11+ 1093.86
y2+ 290.39
b5 2+, b2+ 226.22
y8 2+ 467.11
y4+ 516.30
y12 2+-H2O 618.47
[M+2H]2+-H2O, [M+2H]2+-NH3 683.20
y5 2+ 323.34
[M+2H]2+-NH3-H2O 674.32
b9+ 867.39
y9+ 1029.50
b12+ 1194.81
y2+-H2O 272.29
b3+ 297.50
y3+ 403.32
b10 2+ 491.34
b11 2+ 547.37
y7+ 845.50
y8+ 932.56
b10+ 980.41
y10+ 1086.50
b12+-H2O 1176.77


$^{b}{}_{9}G^{1}{}_{8}T^{2}{}_{7}G^{3}{}_{6}A^{4}{}_{5}S^{5}{}_{4}G^{6}{}_{3}pS^{7}{}_{2}F^{8}{}_{1}R_{y}{}^{9}$
B
Intensity [counts] (10^3)
m/z (Da)
[M+2H]2+-H2O, [M+2H]2+-NH3 451.30
[M+2H]2+-H2O-P, [M+2H]2+-NH3-P 402.27
y5+ 633.43
y4+ 546.20
y6+ 761.46
b6+-H2O 580.32
y6+-P 663.37
y4+ 704.43
y6+-H2O, y6+-NH3 844.66
y5 2+-H2O-P, y5 2+-NH3-P 259.22


$^{b}{}_{21}acS^{1}{}_{20}E^{2}{}_{19}T^{3}{}_{18}A^{4}{}_{17}P^{5}{}_{16}V^{6}{}_{15}A^{7}{}_{14}A^{8}{}_{13}P^{9}{}_{12}A^{10}{}_{11}V^{11}{}_{10}S^{12}{}_{9}A^{13}{}_{8}P^{14}{}_{7}G^{15}{}_{6}A^{16}{}_{5}acK^{17}{}_{4}A^{18}{}_{3}A^{19}{}_{2}A^{20}{}_{1}K_{y}{}^{21}$
C
Intensity [counts] (10^3)
m/z (Da)
y13+ 1180.91
y17 2+ 760.06
b8+ 769.43
b3+, y4+ 360.20
b4+-H2O, y9 2+ 413.25
b7+ 698.40
y10+ 913.64
b11+ 1036.48
y9+ 826.59
y14+ 1251.87
b12+ 1123.47
y15+ 1322.66
y17+ 1518.94
b6+ 627.47
y15 2+ 661.96
b10+ 937.07
b16+ 1419.94
b17+, y18+ 1589.70
b19+ 1731.31

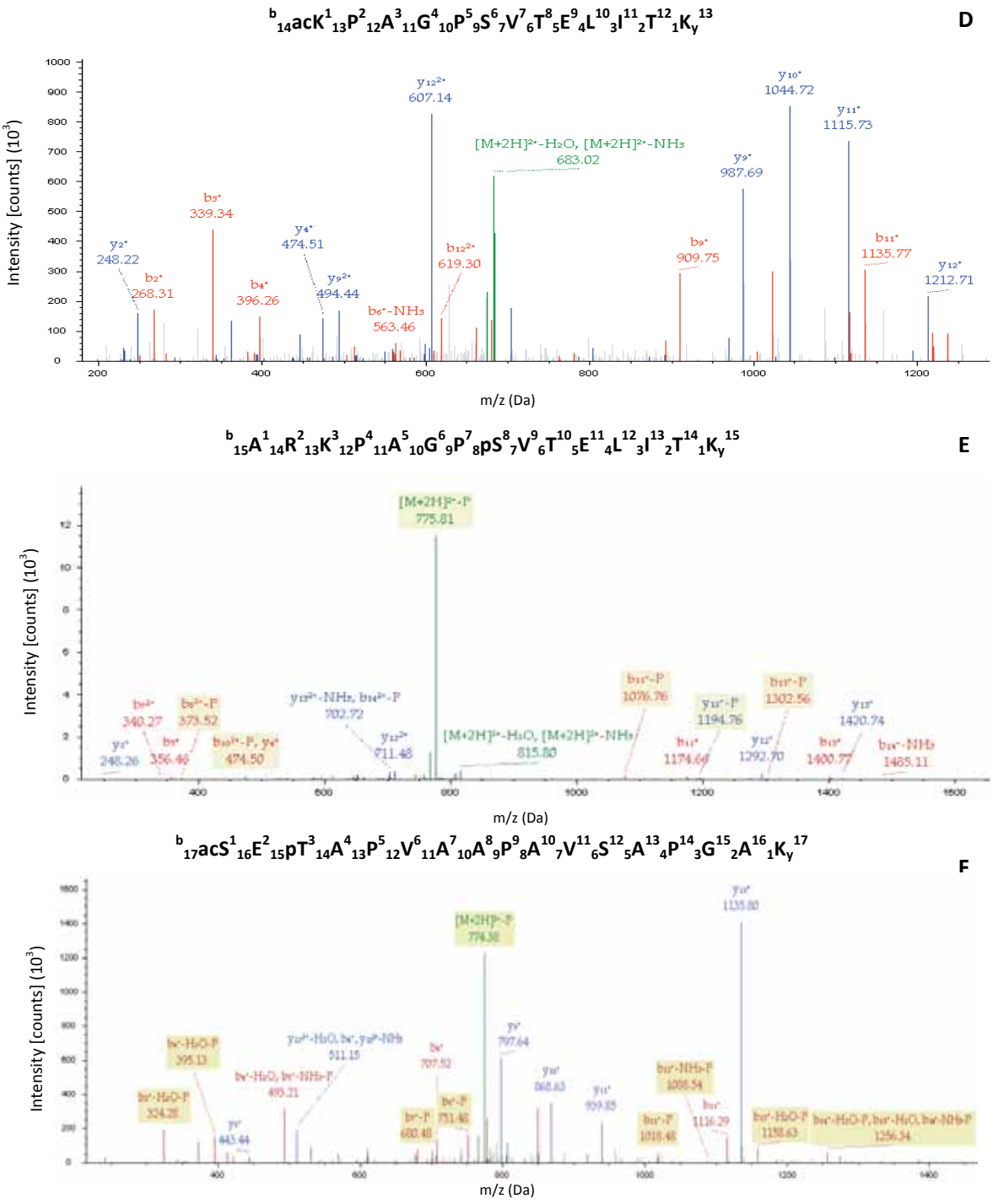


Figure 2.6. Mass spectra of identified peptides of chicken histone H1 containing novel post-translational modifications. Panels A, B, C: PTMs only present in the insoluble chromatin fraction. Panels D, E, F, G: PTMs present in both chromatin fractions. Peaks are labelled in the mass spectra and the fragment ions are indicated in the peptide sequence. Peptide identification with SEQUEST (v1.13). Fragment match tolerance used for search: 0.8 Da **A.** MS^2 spectrum of novel lysine acetylation (ac) on residue K46. XCorr: 2.11; Probability: 18.70. **B.** MS^3 spectrum of novel serine phosphorylation (p) on residue S104. XCorr: 2.23; Probability: 33.40. **C.** MS^2 spectrum of novel lysine acetylation (ac) on residue K17. XCorr: 2.80; Probability: 1.00. **D.** MS^2 spectrum of novel lysine acetylation (ac) on residue K34. XCorr: 3.75; Probability: 7.24. **E.** MS^3 spectrum of novel serine phosphorylation (p) on residue S39. XCorr: 3.14; Probability: 26.46. **F.** $MS2^2$ spectrum of novel threonine phosphorylation (p) on residue T3 and serine S1 N^{α}-acetylation (H1.02). XCorr: 2.81; Probability: 1.00.

Novel acetylation in K17 of histone H1.02 was also identified. This modification was only found in H1.02 of the insoluble fraction (Figure 2.6., C).

Novel S104 phosphorylation was identified in a peptide that matched all histone H1 variants purified from the insoluble fraction of chromatin (Figure 2.6., B), except for histone H1.02. The peptide sequence in which this modification was found is GTGASGSFR which is followed by the residues LS in histones H1.10, H2.03, H1.1L and H1.1R. This combination leads to the canonical phosphorylation site of GSK3 kinase (**S**-X-X-X-S). H1.01 and H1.02 sequences contain the peptide GTGASGSFR(L**N**) and GTGASGSFK(L**N**), respectively, and thus, do not match this motif. The peptide GTGASGSFR containing this PTM was found to be shared by histones H1.01, H1.03, H1.10, H1.1L and H1.1R but not in H1.02. H1.01 and H1.02 contain the peptide sequence but not the canonical phosphorylation site and, thus, H1.01 might not ne phosphorylated. Snijders *et.al.* did not identify this modification[432], what could mean that this phosphorylation is present in very small amounts.

All H1 subtypes were found N^{α}-acetylated in all cases, as it has been previously reported. N^{α}-acetylation occurs as a result of the N^{α}-terminal methionine cleavage[432]. In the Uniprot Data Base, only H1.01 (P08284), H1.02 (P09987), H1.10 (P08286) and, by similarity, H1.1L (P08287), are reported to be N^{α}-acetylated in S1 residue. In the present study, all histone H1 subtypes showed this modification. It has been determined that both the acetylated and unacetylated forms exist in all human, mouse and avian linker histone variants[242,282,283].

Di-methylation in K75 was also identified in a peptide shared by all histone H1 subtypes and in both fractions. Snijders *et.al.* already found this residue dimethylated/formylated and they described this PTM only for H1.03 subtype[432], since they identified these PTMs in a longer peptide than ours that was specific of H1.03. In our study, K75 dimethylation was found in a shorter petide that matched all H1 subtypes, which did not allow us to know wether this PTM was present in all the subtypes or only in some of them.

Acetylation in K91 was identified in the shared-peptide SLVSKGTLVQTK in histone subtypes from both chromatin fractions. This PTM was already identified[432].

Finally, K191 in H1.1L and its equivalent residue, K185, in H1.1R were found to be acetylated in both the soluble and the insoluble fraction of chromatin. Snijders *et.al.* [432] reported the presence of two different PTMs for this position, acetylation and methylation. K191 is followed by an SPKK motif (SPAK) in both subtypes. In our study, we identified this residue to be acetylated, but the fact that Snijders *et.al.* found methylated K191 next to a phosphorylation site suggests that the binding of DNA in this position could be regulated by a methyl-phos switch. According to this model, phosphorylation of a site adjacent to (or nearby) a methyl mark that engages an effector module could lead to consecutive loss of binding to that factor.

In contrast with their study, in which they found acetylated K142 and phosphorylated S154 in all histone subtypes[432], we did not found these two modifications.

In summary, we identified fifteen PTMs in chicken linker histones. From those, eight were novel post-translational modifications: two novel PTMs were identified in histone H5 and six in histone H1 subtypes. Most of the PTMs were found in peptides corresponding to the N- and globular domains. Some of the PTMs found have also been identified in other species, such as humans or mouse, which indicates that they are conserved through evolution. Fractionation of chromatin into soluble and insoluble fractions and separated analysis of linker histones from each fraction demonstrated that some of these PTMs were not uniform along chromatin.

Chapter Three

Proteomic characterization of linker histones after *ex vivo* chromatin phosphorylation

Previous studies showed that phosphorylation of histone H1 (Chapter One) and its CTD[235,266] induced a structural change in the protein that became even more remarkable when it was bound to DNA. Furthermore, full phosphorylation of the C-terminal domain (CTD) of histone H1 affects not only its affinity for DNA fragments, but also its DNA aggregation capacity[266].

As previously explained, phosphorylation of SPKK motifs of linker histones, most of them located in the C-terminal domain, is the main PTM affecting histone H1, and it is known that the maximal number of phosphate groups corresponds to the number of SPKK sites of the molecule[266]. Histone H1 is phosphorylated in a cell cycle-dependent manner[254,249,266]: hyperphosphorylation is related to condensed chromatin in metaphase[255] while unphosphorylation is usually associated with condensed interphasic chromatin and partial phosphorylation is linked to relaxed chromatin[147].

Taking into account all those results, it can be thought that the effects of phosphorylation on chromatin structure may depend on the number and the position of the phosphate groups incorporated in H1. These effects might be mediated by specific structural changes and not be a simple effect of the reduction of the net charge associated to phosphorylation[266]. Since histone H1 is part of native chromatin and it is there where it develops its main function in chromatin condensation and regulation of gene expression, it is of great interest to know the effect of histone H1 phosphorylation on chromatin aggregation.

Chicken erythrocyte is a popular system to study histone modifications. Chicken erythrocytes have a low transcriptional activity with very few genes being expressed. Because of that, some authors have used this system, with such a little background, to study the structure and mechanisms of chicken erythrocyte transcriptionally active and potentially active gene chromatin[439].

Its poor content in non-histone proteins together with the fact that chicken erythrocytes are terminally differentiated cells and, thus, they are supposed to have a low basal level of phosphorylation, were the reasons for choosing this system for the study of *ex vivo* chromatin phosphorylation. We wanted to specifically phosphorylate linker histones within chromatin and, because of that; it was of a great importance to avoid other proteins background in the original material. As mentioned in Chapter Two, chicken erythrocyte nuclei contain six different subtypes of histone H1 plus the specific avian subtype, histone H5[354,441].

Native chromatin from chicken erythrocyte nuclei was purified and phosphorylated *ex vivo* with CDK2 kinase, which specifically phosphorylates the SPKK motifs present in linker histones. The aim was to study the changes in chromatin aggregation due to *ex vivo* phosphorylation. For this purpose, we first extracted the linker histones from the *ex vivo* phosphorylated chromatin, in order to determine if phosphorylation had occurred in H1/H5 within chromatin.

Ex vivo phosphorylation of chicken erythrocytes chromatin in the presence of ɣ-ATP-P^{32}

The first challenge was to find out if *ex vivo* phosphorylation of chromatin led to linker histones phosphorylation and thus, changes in chromatin could be associated to the phosphorylation of H1 within chromatin. In order to corroborate the specific phosphorylation of linker histones within chromatin; we performed a radioactive CDK2-phosphorylation reaction in the presence of ɣ-ATP-P^{32}. Afterwards, total protein fraction of chromatin was analysed by SDS-PAGE (Figure 3.1).

The radioactive assay consisted in two parallel chromatin phosphorylation reactions: one with CDK2 kinase and the other one, without; both in the presence of ɣ-ATP-P^{32}. Reactions were carried out under the same conditions and chromatin samples were taken after 30 minutes, 3 hours and overnight in

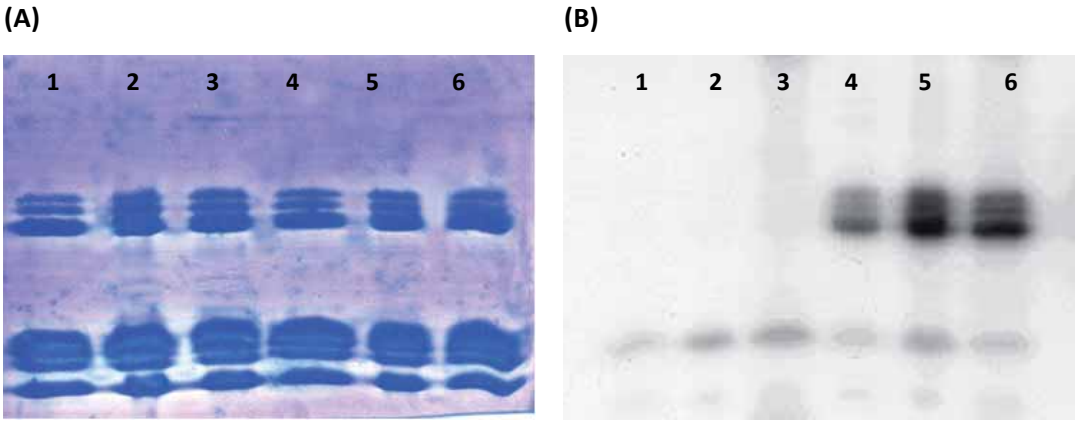


Figure 3.1. Radioactive phosphorylation assay. SDS-PAGE of total protein of *ex vivo* phosphorylated chicken erythrocytes chromatin. Comassie blue staining (A) and autoradiography (B). Phosphorylation reaction of chicken erythrocyte chromatin in the absence (lanes 1, 2, 3) or the presence (lanes 4, 5, 6) of CDK2 kinase. Lanes correspond to different times of reaction: 30 minutes (lanes 1 and 4); 3 hours (lanes 2 and 5) and overnight (lanes 3 and 6).

order to observe the incorporation of P^{32}.

In the SDS-PAGE stained with Comassie blue from Figure 3.1., it is important to note that there are not significant amounts of other proteins, apart from histones, in chicken erythrocytes chromatin. Samples mainly contained core histones (lower bands, 11-16kDa) and linker histones (upper bands, ~20kDa). The autoradiography showed that phosphorylation only occurred in the samples containing CDK2 kinase (lanes 4, 5 and 6 in Figure 3.1., B) and P^{32} was only incorporated in the linker histones fraction (Figure 3.1., left; lanes 4, 5 and 6; upper bands). An increase of the phosphorylation time led to a progressive incorporation of P^{32} in linker histones, as it can be seen when comparing lanes 4 (30 minutes), 5 (3 hours) and 6 (overnight) in the autoradiography. A weak radioactive signal appeared in all the lanes in the area of the gel corresponding to the core histones. This signal may be due to unspecific interactions between the ɣ-ATP-P^{32} and the core histones, and not a consequence of phosphorylation, since it appears in all the samples, even in those without CDK2 kinase (lanes 1, 2, 3). This is in agreement with the absence of CDK2 motifs in core histones.

This result showed that *ex vivo* phosphorylation of linker histones within chromatin was achieved and that it specifically occurred in linker histones, not in core histones. Furthermore, phosphorylation increased with time.

Proteomic characterization of linker histones after *ex vivo* phosphorylation of chromatin

Characterization of the post-translational modifications (PTMs) present in the linker histones extracted from the *ex vivo* phosphorylated chromatin by proteomic methods were carried out during a PhD stay at

the Division of Clinical Biochemistry, headed by Dr. Prof. Herbert Lindner in the Biocenter - Medizinische Universität Innsbruck (Austria).

Once *ex vivo* phosphorylation of linker histones within chromatin was corroborated, our purpose was to identify the different histone subtypes and to determine the level of phosphorylation and the position of the phosphate groups after *ex vivo* phosphorylation.

Our samples consisted on linker histones extracted from native soluble chromatin from chicken erythrocyte nuclei after *ex vivo* phosphorylation with CDK2 for 15 minutes, 30 minutes, 1 hour, 3 hours, 5 hours and overnight.

In order to characterise linker histones after *ex vivo* chromatin phosphorylation, several analyses by different proteomic techniques were carried out. Figure 3.2. shows a scheme of the experimental procedure.

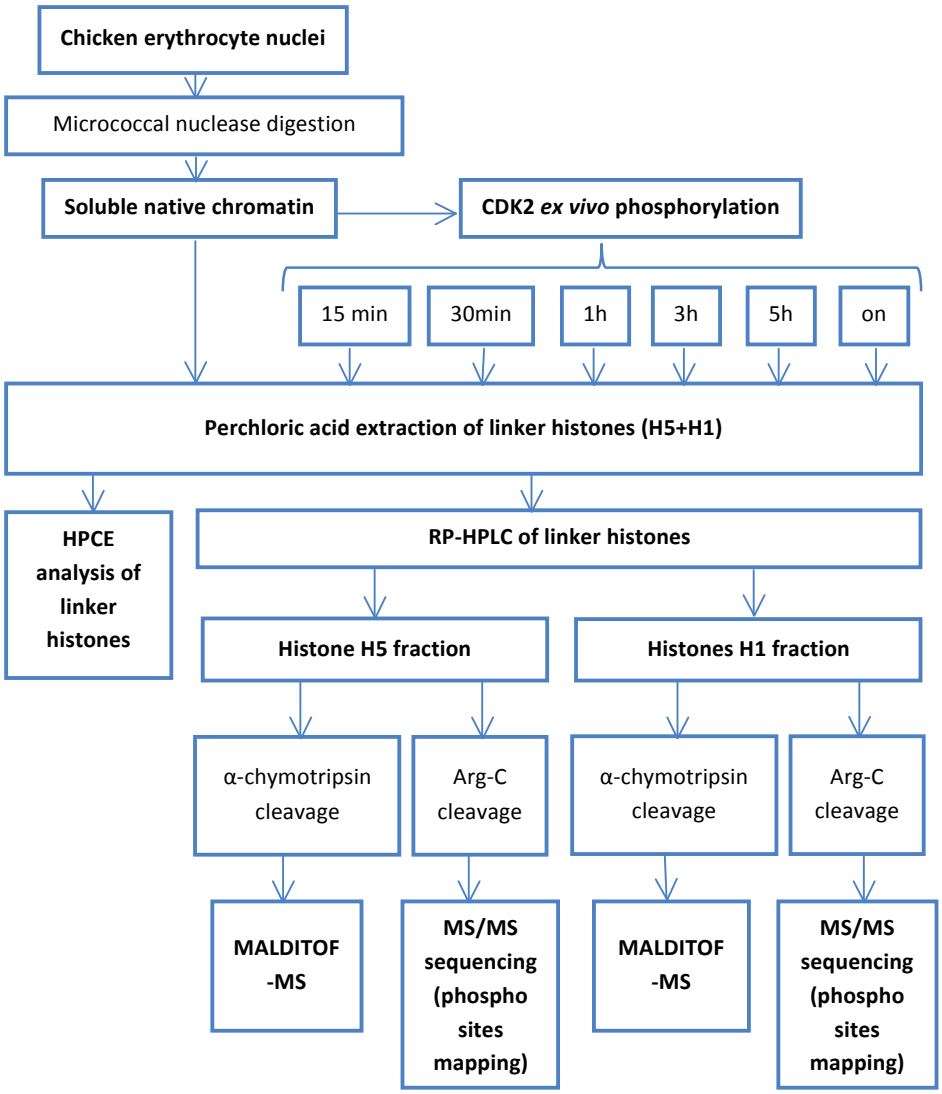


Figure 3.2. Scheme of the experimental procedure for the analysis of chicken linker histones from *ex vivo* phosphorylated chromatin by several proteomic techniques.

Analysis of linker histones by HPCE

As already mentioned in Chapter Two, HPCE allows to successfully separate distinct histone subtypes and their post-translationally modified species with high sensitivity. The addition of a chemical group due to a PTM leads to an increase of the retention time of the modified protein in the capillary with respect to the unmodified one. Thus, HPCE allows to distinguish phosphorylated derivatives of individual histone H1 variants[348,350,354]. Because of that, this method was suitable for qualitative identification of the phosphorylated species that appeared after *ex vivo* phosphorylation of chromatin.

Phosphorylation was performed *ex vivo* with the solubilized chromatin fraction. For that reason, the HPCE profile of native chromatin (Chapter Two) was also a control of basal phosphorylation to compare the HPCE pattern of phosphorylated samples, with respect to the unphosphorylated chromatin.

An HPCE separation profile, in which both H5 variants and six H1 subtypes were easily identified, was successfully obtained in all protein samples with a signal of high intensity and an optimal and reproducible separation profile, which indicated a high quality and purity of the original samples.

Figure 3.3 shows the HPCE separations of purified H5/H1 histone proteins extracted from erythrocyte nuclei soluble chromatin after ex vivo phosphorylation at different times.

As it can be seen in Figure 3.3., there is a progressive increase in the number of secondary peaks present in the HPCE separation profiles. This change could be attributed to an increase in the number of phosphate groups after *ex vivo* chromatin phosphorylation.

The changes in the elution profile were progressive and became more apparent with increasing phosphorylation time: the longer the phosphorylation reaction, the greater the number and height of the secondary peaks found in the HPCE profiles. This fact shows that linker histones within soluble chromatin were phosphorylated *ex vivo* by CDK2 kinase and, furthermore, that the phosphorylation levels increased with reaction time, in agreement with the results of the radioactive assay.

Another interesting result was that the changes in the HPCE separation profile occurred differently for histones H5 and H1. At shorter phosphorylation times, most of the HPCE profile changes occurred in histone H5 variants, whereas changes in the six H1 subtypes did not appeared before 1 hour of phosphorylation.

After 15 minutes of phosphorylation (Figure 3.3, B), the mono-phosphorylated unacetylated H5 peak (H5ac0p1; Figure 3.3., peak 2) remained unaltered, but a slight increase in the di-phosphorylated unacetylated H5 variant peak (H5ac0p2; Figure 3.3., peak 3) could already be identified. This change was also seen in the acetylated H5 variant (H5ac1; Figure 3.3., peak 4). Despite the peak for the mono-phosphorylated H5ac1 (H5ac1p1; Figure 3.3., peak 5) remained also unaltered, two small peaks appeared at higher retention times, which could be attributed to the di- and tri-phosphorylated H5ac1 (H5ac1p2 and H5ac1p3; Figure 3.3., peaks 6 and 7, respectively). At this phosphorylation time, no changes were observed in the HPCE separation profile of H1 subtypes.

The HPCE separation profile described was almost the same at 30 minutes and 1 hour of phosphorylation (Figure 3.3, C and D). Among those samples, only a low and progressive increase in the proportion of H5ac0p2 (Figure 3.3., peak 3) was observed.

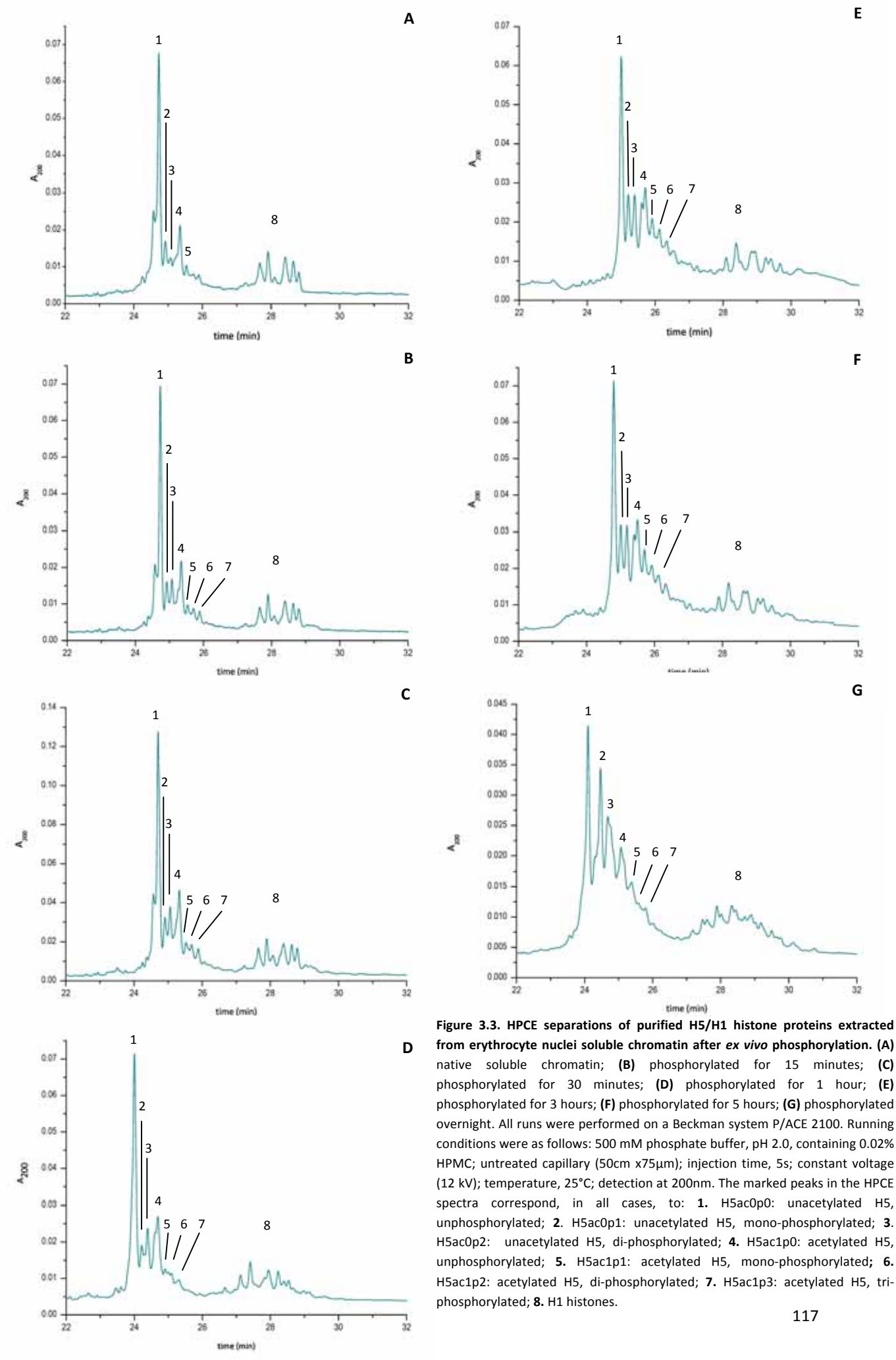


Figure 3.3. HPCE separations of purified H5/H1 histone proteins extracted from erythrocyte nuclei soluble chromatin after *ex vivo* phosphorylation. (A) native soluble chromatin; **(B)** phosphorylated for 15 minutes; **(C)** phosphorylated for 30 minutes; **(D)** phosphorylated for 1 hour; **(E)** phosphorylated for 3 hours; **(F)** phosphorylated for 5 hours; **(G)** phosphorylated overnight. All runs were performed on a Beckman system P/ACE 2100. Running conditions were as follows: 500 mM phosphate buffer, pH 2.0, containing 0.02% HPMC; untreated capillary (50cm x75µm); injection time, 5s; constant voltage (12 kV); temperature, 25°C; detection at 200nm. The marked peaks in the HPCE spectra correspond, in all cases, to: **1.** H5ac0p0: unacetylated H5, unphosphorylated; **2**. H5ac0p1: unacetylated H5, mono-phosphorylated; **3**. H5ac0p2: unacetylated H5, di-phosphorylated; **4.** H5ac1p0: acetylated H5, unphosphorylated; **5.** H5ac1p1: acetylated H5, mono-phosphorylated**; 6.** H5ac1p2: acetylated H5, di-phosphorylated; **7.** H5ac1p3: acetylated H5, tri-phosphorylated; **8.** H1 histones.

After 3 hours of phosphorylation (Figure 3.3, E), the number and height of secondary peaks increased for all the phosphorylated species. At this point, secondary peaks in the H1 fraction could also be observed (Figure 3.3 peaks 8). This HPCE separation profile remained quite similar for 5 hours of phosphorylation (Figure 3.3, F). When the phosphorylation reaction was carried out overnight, the amount of secondary peaks was even higher and, to an extent, difficult to determine due to the overlapping of peaks (Figure 3.3, G).

These results showed that the changes in phosphorylation levels primarily occurred in histones H5 at low phosphorylation times, whereas histone H1 subtypes needed longer times of phosphorylation to incorporate that PTM. Furthermore, it seems that phosphorylation lead to an increase, first, of the mono-phosphorylated species and, later on, to the appearance of di- and tri-phosphorylated species.

Analysis of chicken linker histones phosphorylation by MALDITOF-MS and mapping of PTMs by Tandem MS after *ex vivo* phosphorylation of chromatin

The total linker histone fraction (H5+H1) was separated by Reverse Phase High Performance Liquid Chromatography (RP-HPLC) and, for each phosphorylation time, two fractions of protein were obtained (H5 and H1) (Figure 3.2.).

In order to determine the level of phosphorylation, the samples were analysed by Matrix-Assisted Laser-Desorption Ionization coupled with Time-Of-Flight Mass-Spectrometry (MALDITOF-MS). A middle-down proteomic strategy was carried out. The two fractions were partially digested with α-chymotrypsin in order to obtain smaller peptides derived from the original proteins with an optimal ionization level, and thus, a signal with higher quality in MALDITOF-MS. After each α-chymotrypsin digestion, an HPCE separation was run to ensure a proper enzymatic cleavage (Figure 3.2.).

MALDITOF-MS was used to obtain the peptide mass through the time of flight (TOF) of the ionized peptides and their charge. The addition of a phosphate group due to protein phosphorylation implies an increase of the theoretical original mass of the peptide without PTMs. This increase should be equal to the mass of the incorporated modification, so an increase of approximately 80Da would be expected for the incorporation of every phosphate group. The phosphorylated species were identified by comparing the experimental peptide masses in the MALDITOF spectra with the expected theoretical masses of the unphosphorylated peptides. Moreover, data obtained in the MALDITOF-MS spectra allowed to calculate the percentages of the phosphorylated species in each sample and the proportion of phosphorylated CDK2 sites.

In parallel and with the purpose of determining the positions of the incorporated phosphate groups after *ex vivo* phosphorylation, RP-HPLC purified fractions of histones H5 and H1 were also analysed using enzymatic digestions followed by nanoflow reverse-phase high-performance liquid chromatography tandem mass spectrometry (LC-ESI-MS/MS) (Figure 3.2.). Enzymatic cleavage generates peptides that are more readily characterised. All six subtypes of the chicken histone H1 (H1.01, H1.02, H1.03, H1.10, H1.1R and H1.1L) and the specific avian isoform H5 were mapped.

As already explained in Chapter Two, enzymatic cleavage conditions were optimized to ensure a good result in the LC-ESI-MS/MS experiments and achieve high sequence coverages: H1 fraction was digested with Trypsin (1:100, w/w) for 15 minutes whereas H5 fraction was digested for 5minutes with Arg-C (1:40, w/w). Table 3.1. shows the best sequence coverages (%) obtained in the MS/MS experiments.

Table 3.1. Sequence coverage (%) of chicken histones in MS/MS experiments.

Swiss-Prot Acc. Num.	Histone variant	Sequence coverage of each variant per experiment (%)						
		Native soluble	15min	30min	1h	3h	5h	on
P02259	H5	82	82	76	73	64	66	63
P08284	H1.01	59	62	62	55	55	55	57
P09987	H1.02	52	56	48	52	52	52	38
P08285	H1.03	67	68	71	61	67	61	62
P08286	H1.10	59	49	62	46	46	46	48
P08287	H1.1L	88	66	66	65	65	59	71
P08288	H1.1R	81	65	64	59	56	47	63

In most of the cases, the values oscillated between the 45% and the 88% of coverage, with only a few proteins covered above an 80%. Peptides corresponding to the N-terminal and the globular domain of the proteins were found in most cases. C-terminal domain of linker histones, which is very lysine-rich and, thus, more susceptible to enzymatic cleavage, was the region of the proteins with less coverage.

Mass spectrometry analyses enabled the identification of a wide range of post-translational modifications including acetylation, methylation and phosphorylation. In the present chapter, we focused our analysis in the identification of phosphorylated CDK2 sites. Only modified peptides scoring X^2corr ≥2.5 were subjected to manual verification (Tables 3.4 and 3.9). The concrete modifications will be broadly discussed for each linker histone.

Proteomic characterization of chicken histone H5 after ex vivo phosphorylation of chromatin

Analysis of chicken histone H5 by MALDITOF-MS

Histone H5 (P02259; 190 aa; 20734Da) has a natural variant[361], in which arginine R16 is replaced by a glutamine (Q) residue, is present in lower proportion and has a molecular weight of 20705.06Da. As already mentioned in Chapter Two, UniProt database theoretical data reports histone H5 acetylation at threonine T1 (Mw_{ac}=42Da) and four probable phosphorylation sites at serine residues S23, S30, S146, S167 (Mw_{phos}=80Da). H5 has five SPKK motifs that could be phosphorylated by CDK2 (Figure 3.4.).

>sp|P02259|H5_CHICK Histone H5 OS=Gallus gallus PE=1 SV=2

```
        10         20         30         40         50         60         70
TESLVLSPAP AKPKRVKASR RSASHPTYSE MIAAAIRAEK SRGGSSRQSI QKYIKSHYKV GHNADLQIKL
        80         90        100        110        120        130        140
SIRRLLAAGV LKQTKGVGAS GSFRLAKSDK AKRSPGKKKK AVRRSTSPKK AARPRKARSP AKKPKATARK
       150        160        170        180        189
ARKKSRASPK KAKKPKTVKA KSRKASKAKK VKRSKPRAKS GARKSPKKK
```

Figure 3.4. Sequence of chicken histone H5 (P02259). The SPKK motifs are highlighted in yellow. α-chymotrypsin cleavage sites are marked in red, α-chymotrypsin cleavage occurs in the C-term of F/Y/W, not before P.

Once the histone H5 fraction was purified by RP-HPLC, it was partially cleaved with α-chymotrypsin, which divided the protein into two peptides (Table 3.2.): the first one, with a molecular weight of ~9989Da, corresponded to residues 2 to 94 and it included the N-terminal domain and most part of the globular domain (residues 25-98); and the second peptide, from residues 95 to 190 and a molecular weight of ~10634Da, mainly consisting of the C-terminal domain of H5. This mixture was analysed by MALDITOF-MS analysis and MALDITOF spectra were obtained for the identification of peaks.

Table 3.2. Cleavage of chicken histone H5 with α-chymotrypsin

$Mass_{av}$ (Da)	Positions	# Missed cleavages	Sequence
9988.56	2-94	3	TESLVLSPAPAKPKRVKASRRSASHPTYSEMIAAAIRAEKSRGGSSRQSIQKYIKS HYKVGHNADLQIKLSIRRLLAAGVLKQTKGVGASGSF
10633.96	95-190	0	RLAKSDKAKRSPGKKKKAVRRSTSPKKAARPRKARSPAKKPKATARKARKKSRA SPKKAKKPKTVKAKSRKASKAKKVKRSKPRAKSGARKSPKKK

α-chymotrypsin cleavage occurs in the C-side of F/Y/W, not before P. Missed cleavages are shown in red.
Data generated with the PeptideMass application on ExpasyTools
Mass corresponds to the average value (Da)

MALDITOF-MS is a high sensitivity technique for proteomic analysis and it allows the determination of the masses of molecules with only a deviation of ±5Da (m/z) in relation to the theoretical mass. This fact allows the identification of almost every single modification in the protein.

Two regions were identified in the MALDITOF-MS spectra of histone H5, corresponding to the two peptides obtained by α-chymotrypsin cleavage. Figure 3.5. shows the MALDI-TOF spectra of histone H5 from native soluble chromatin and phosphorylated chromatin for 15 minutes, 3 hours and overnight. Spectra not shown for other phosphorylation times are collected in the Appendix I.

In the first region of the spectra, peptides of positions 2-94, which corresponds to the N-terminal domain plus the globular domain of H5, were detected. In this region, the unacetylated species (NH5R16ac0, Mw_{av}=9988,56Da; theoretical mass; Figure 3.5, peak 2) and, in a lower proportion, the acetylated one (NH5R16ac1, Mw_{av}=10030,56Da; theoretical mass; Figure 3.5, peak 3) could be identified, since the acetyl group gives a difference of 42Da between those two peaks.

When HPCE separations were carried out, both the acetylated and unacetylated variants of histone H5 could be identified (Figure 3.2). However, the sensitivity of the technique was not enough to identify the natural variant 16(R→Q). In the MALDITOF-MS analysis, the unacetylated natural variant (NH5Q16ac0, Mw_{av}=9960,56Da; theoretical mass; Figure 3.5, peak 1) was also found in the spectra of most of the samples. Most likely, the acetylated natural variant was also present in the mixture but it could not be detected. This fact could be explained because the peak corresponding to NH5Q16ac1 (Mw_{av}=10002,56Da; theoretical mass) has only a difference of 14Da with the main and widest peak in the profile (NH5R16ac0), and thus, these two peaks may overlap, remaining hidden the one corresponding to NH5R16ac1.

Phosphorylation was only found in the peptide 2-94 of the unacetylated H5R16 histone extracted from the overnight phosphorylated chromatin (NH5R16ac01P, Mw_{av}=10068,56Da; Figure 3.5, peak 8), despite there are no consensus motifs for CDK2 phosphorylation in this region.

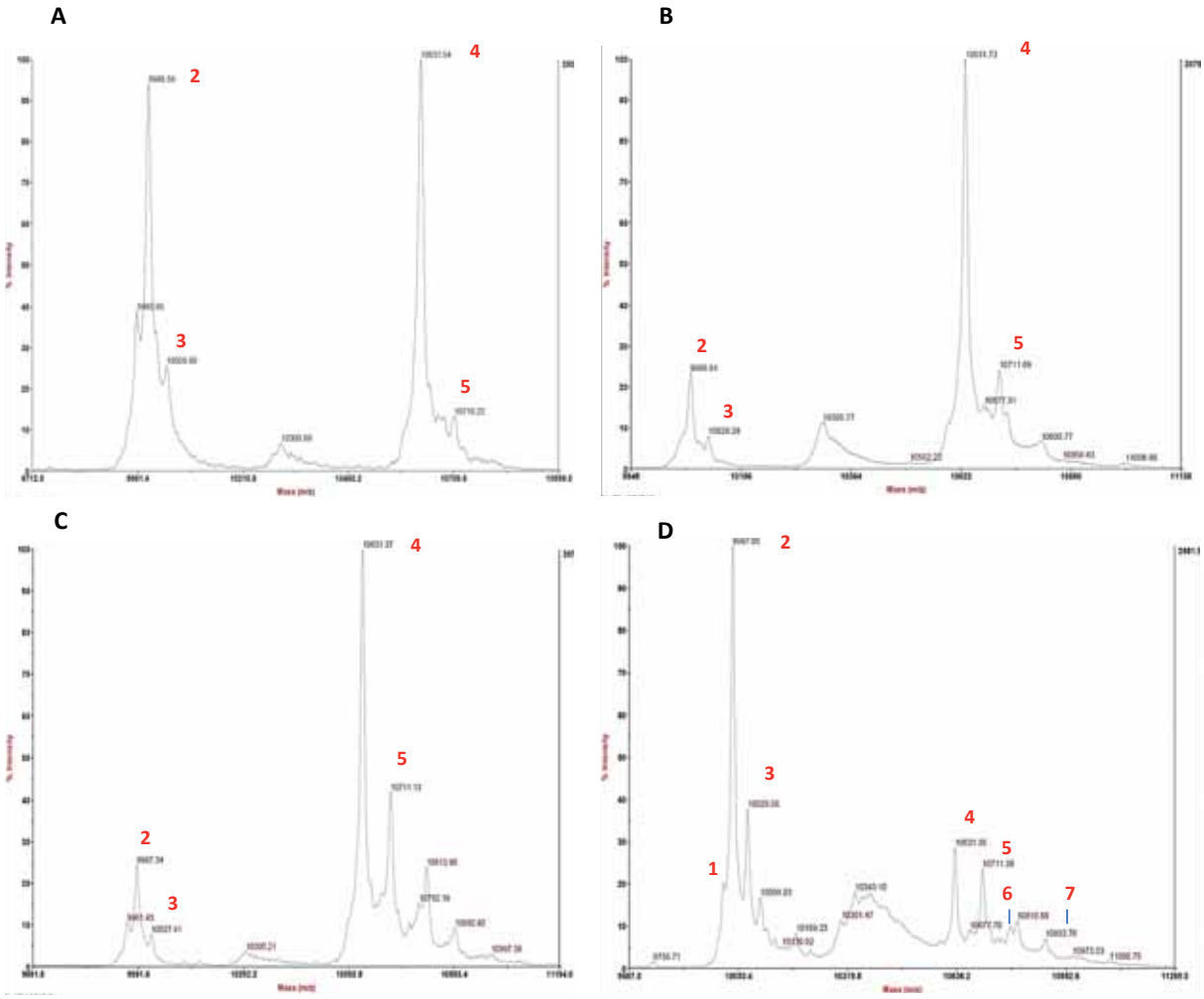


Figure 3.5. MALDITOF-MS spectrum of chicken histone H5 from native soluble chromatin (A) and *ex vivo* phosphorylated chromatin for 15 minutes (B); 3 hours (C); and overnight (D). Peak numbering is defined as follows for all the spectra: **(1)** NH5Q16ac0; **(2)** NH5R16ac0; **(3)** NH5R16ac1; **(4)** CH50P; **(5)** CH51P; **(6)** CH52P; **(7)** CH53P; **(8)** NH5R16ac01P. NH5Q16ac0=N-terminal and globular domain (2-94) of unacetylated H5Q16 natural variant; NH5R16ac0=N-terminal and globular domain (2-94) of unacetylated H5R16 natural variant; NH5R16ac1=N-terminal and globular domain (2-94) of monoacetylated H5R16 natural variant; CH50P=C-terminal domain (95-190) of unphosphorylated H5; CH51P=C-terminal domain (95-190) of monophosphorylated H5; CH52P=C-terminal domain (95-190) of diphosphorylated H5; CH53P=C-terminal domain (95-190) of triphosphorylated H5; NH5R16ac01P= N-terminal and globular domain (2-94) of unacetylated and monophosphorylated H5R16 natural variant.

In the case of the second peptide (residues 95-190), which corresponded to the C-terminal domain of histone H5, phosphorylated species were also identified (Figure 3.5.). The main peak was, for all samples, that corresponding to the unphosphorylated C-terminal domain of H5 (CH50P, Figure 3.5, peak 4), which has a theoretical molecular weight of 10633,96Da. Following that one, at a distance of ~80Da, a second peak for the mono-phosphorylated H5 (CH51P; Figure 3.5, peak 5), was found in all the samples, including those coming from native chromatin (Figure 3.5.; A). The height of this peak increased with the time of phosphorylation. This increase was first seen after 15 minutes of phosphorylation (Figure 3.5.; B), in agreement with the results obtained in the HPCE separation profiles

(Figure 3.3.), which already showed an increase in height of H5's secondary peaks. The samples of 30 minutes and 1 hour of phosphorylation showed similar profiles in the MALDITOF spectrum (see Appendix I).

In the MALDITOF-MS spectra of H5 after 3 hours of chromatin phosphorylation (Figure 3.5.; C), the height of the peak of the mono-phosphorylated CH5 (Figure 3.5, peak 5) considerably increased with respect to the previous samples and in contrast to the CH50P peak (Figure 3.5, peak 4). Another peak, located at a distance of ~160Da from the unphosphorylated CH5 appeared, indicating, thus, the presence of di-phosphorylated species of H5 (CH52P; Figure 3.5, peak 6). These results confirmed, as well, the increase of secondary peaks that had already been observed in the HPCE profile of this sample (Figure 3.3).

For longer phosphorylation times -this means for the CH5 samples of 5 hours and overnight-phosphorylated chromatin -, the peaks of CH51P and CH52P continued to increase in relation to the previous samples and the difference in height with the CH50P peak was lower, up to a point that, in the overnight sample (Figure 3.5; D), the CH50P and the CH51P peaks gave a similar intensity in the MALDITOF-MS spectrum and even a tri-phosphorylated species (CH53P) appeared.

Thus, these results confirmed the ones obtained in the HPCE separations about the presence of phosphorylated species and their increase with the extent of phosphorylation time.

From the data obtained by MALDITOF-MS, it is possible to identify the presence and the intensity (peak height) of the different phosphorylated species in the sample, but the absolute amount of each one of them cannot be determined. This is because the intensity of the peaks directly depends on the ionization of the peptide, which is, in the last term somehow dependent on the MALDITOF-MS run. For a better comprehension, the difference in the intensities in the MALDITOF-MS spectra between the two peptides obtained by α-chymotrypsin cleavage of H5 can be taken as an example. Depending on the sample, the intensities of those two peptides are significantly different (i.e. Figure 3.5. C vs. D). Since both peptides are generated from the cleavage of the same H5 sample, both peptides are present at equimolar concentration. Then, the difference in intensity is not related with the amount of each peptide but with its ionization capacity in each run. Nevertheless, for each run and peptide, the ratio of the peak's heights can be used to determine the relative amount (%) of the phosphorylated species identified.

Moreover, if we consider the number of CDK2 sites per H5 molecule (5 CDK2 sites) susceptible to be phosphorylated and the relative amounts (peak height) of the phosphorylated species found, we can also calculate de percentage (%) of phosphorylated CDK2 sites (SPKK motifs) per H5 molecule (Table 3.3):

$$\%\ phosphorylated\ SPKK\ sites = \frac{\sum_{n=0}^{k} nh_n}{k\sum_{n=0}^{k} h_n} * 100$$

where *k* is the total number of SPKK motifs in the molecule; *n* corresponds to each phosphorylated species identified in the experiment and h_n is the height of the peak of each phosphorylated species identified in the experiment.

Table 3.3. MALDI analysis of C-terminal domain of H5 histone after cleavage with α-chymotrypsin

Sample	Species	Mw_{av} (Da)	ΔMw (Da)	Peak height	% 1P	% 2P	% (1P+2P)	% phos-SPKK
H5 native soluble chromatin	0P	10631.54		234	12,36			2,47
	1P	10710.22	78.68	33				
H5 15min phosphorylated chromatin	0P	10631.73		2805	19,99			4,00
	1P	10711.69	79.96	701				
H5 30min phosphorylated chromatin	0P	10632.16		952	20,67			4,13
	1P	10712.22	80.05	248				
H5 1h phosphorylated chromatin	0P	10631.72		277	21,31			4,26
	1P	10711.38	79.65	75				
H5 3h phosphorylated chromatin	0P	10631.37		367	27,05	10,10	37,16	9,45
	1P	10711.13	79.76	158				
	2P	10792.18	81.04	59				
H5 5h phosphorylated chromatin	0P	10632.01		123	26,26	11,62	37,88	9,90
	1P	10711.02	79.01	52				
	2P	10793.16	82.14	23				
H5 overnight phosphorylated chromatin	0P	10631.35		708	37,64	16,80	54,44	14,25
	1P	10711.38	80.03	585				
	2P	10792.98	81.60	261				

%1,%2P: percentages (%) of the mono- and diphosphorylated species of H5 CTD calculated from peak heights.
(%1+%2P): percentages (%) of all the phosphorylated species of H5 CTD in the sample calculated from peak heights.
%phos-SPKK: percentages (%) of phosphorylated SPPK motifs in the sample considering the potentially phosphorylated CDK2 sites. Calculated with the formula above.

The calculated percentages between different phosphorylated species of histone H5 confirm the progressive increase of phosphorylation with time. As mentioned before, di-phosphorylation first appeared after 3 hours of phosphorylation (~10%) and, at this time of reaction, the mono-phosphorylated species had increased up to a 27% in the sample. The total percentage of phosphorylated species present in this sample was around a 37%. At 5 hours of phosphorylation, the mono-phosphorylated species remained the same, whereas there was a slight increase of the intensity of the CH52P. However, the total amount of phosphorylation (~38%) remained stable and similar to that of H5 after 3 hours of phosphorylation. The overnight phosphorylation introduced a remarkable change in the percentages of the mono- and the di-phosphorylated species with respect to the unphosphorylated one. The percentage of CH51P increased up to a ~38% and CH52P species corresponded to a ~17% of this sample. After overnight phosphorylation, the percentage of phosphorylated species in the sample (1P+2P) corresponded to a ~54% of the total species identified (0P+1P+2P). These results indicated that phosphorylation first increased the amount of the monophosphorylated species and, only then, di- and triphosphorylated species appeared.

It is important to note that the latter percentages of phosphorylated species were calculated with respect to the total species identified in each sample. Histone H5 contains five CDK2 sites (SPKK motifs) that can be phosphorylated, but only 0P, 1P, 2P and 3P species were identified. The triphosphorylated species were not considered in these calculations, since they were present in little amounts and, thus, it was difficult to precisely determine the peak height. Despite the phosphorylated species consisted in a ~55% of the total species found after overnight phosphorylation, these only corresponded to a 14.25% of the potential phosphorylation of histone H5. This result indicates that there were many

phosphorylated species (~54%) although they were partially phosphorylated, with only few modified SPPK sites per molecule (2.5%-14.25%) (Table 3.3.). However, the amount (%) of phosphorylated SPKK motifs increased with phosphorylation time.

Note that other peaks were also detected in the MALDITOF-MS spectra, in some cases, even with higher intensities than the ones of the H5's phosphorylated species. Those ones were not considered in any case, since the mass differences could not be attributed to any modification of the protein or to a different α-chymotrypsin miscleavage. Most likely, some of those peaks might be a due to protein degradation or to the combination of several PTMs. Phosphorylation was the only PTM identified in the C-terminal domain peptides of H5, probably because other PTMs, despite they might be present, they might be in lower proportion and, thus, could not be detected.

Mapping of histone H5 post-translational modifications

MALDITOF-MS allows the identification of phosphorylated species, but not the determination of the position in which the PTM has been incorporated. Histone H5 has five SPKK motifs in its C-terminal domain that can be potentially phosphorylated by CDK2, together with two possible phosphorylation sites in this region entered in the UniProt database. Mono-, di- and triphosphorylated species have been detected in this part of the study. In order to determine which of these positions had been phosphorylated, a tandem MS/MS to map the PTMs had to be performed.

In order to determine the positions of the incorporated phosphate groups after *ex vivo* phosphorylation, RP-HPLC purified fraction of histone H5 was enzymatically cleaved with Arg-C for 5 minutes and sequenced by Tandem MS. Table 3.4. summarizes all the post-translational modifications found in histone H5. As mentioned, all modified peptides with X^2corr values above 2.5 were validated manually for precise verification and localization of the post-translational modification.

In all the samples, both the acetylated and unacetylated N^{α}-terminal peptides were identified, together with other PTMs that were already present in the native chromatin (Chapter Two).

Table 3.4. Modified peptides identified by ESI/LC- MS/MS in chicken erythrocytes H5 histone

Sample	Peptide sequence	Position	PTM	MS^2 Xcorr
H5 soluble native chromatin	TESLVLSPAPAKPK	1-14	T1-Acetyl	3.41
			S3-Phospho	3.61
			S7-Phospho	3.82
			T1-Acetyl, S3-Phospho	2.58
			T1-Acetyl, S7-Phospho	3.58
			T1-Acetyl, K14-Acetyl	3.15
	QTKGVGASGSFR	83-95	T2-Phospho	3.22
	SASHPTYSEMIAAAIR	22-38	M10-Oxidation	4.21
H5 15min phosphorylated chromatin	TESLVLSPAPAKPK	1-14	T1-Acetyl	3.91
			S3-Phospho	2.91
			S7-Phospho	3.51
			T1-Acetyl, S3-Phospho	3.00
			T1-Acetyl, S7-Phospho	3.01
			T1-Acetyl, K14-Acetyl	3.96
			K12-Acetyl	2.95
	SASHPTYSEMIAAAIR	22-38	M10-Oxidation	3.97

H5 30min phosphorylated chromatin	TESLVLSPAPAKPK	1-14	T1-Acetyl	3.48
			S3-Phospho	3.96
			S7-Phospho	2.63
			T1-Acetyl, S3-Phospho	3.01
			T1-Acetyl, S7-Phospho	3.77
	RLLAAGVLKQTKGVGASGSFR	74-94	K9-Acetyl	3.25
	SASHPTYSEMIAAAIR	22-38	M10-Oxidation	3.63
			S3-Phospho	3.11
H5 1hour phosphorylated chromatin	TESLVLSPAPAKPK	1-14	T1-Acetyl	3.87
			S3-Phospho	2.83
			S7-Phospho	2.52
			T1-Acetyl, S3-Phospho	3.21
			T1-Acetyl, S7-Phospho	3.65
			K14-Acetyl	3.28
	KKSRASPK	143-150	S6-Phospho	2.09
	SASHPTYSEMIAAAIR	22-38	M10-Oxidation	3.55
			S3-Phospho	3.57
H5 3hour phosphorylated chromatin	TESLVLSPAPAKPK	1-14	T1-Acetyl	3.48
			S3-Phospho	2.70
			S7-Phospho	2.43
			T1-Acetyl, S3-Phospho	2.97
			T1-Acetyl, S7-Phospho	3.22
	SASHPTYSEMIAAAIR	22-38	M10-Oxidation	3.69
H5 5hour phosphorylated chromatin	TESLVLSPAPAKPK	1-14	T1-Acetyl	3.83
			S3-Phospho	4.43
			S7-Phospho	2.48
			T1-Acetyl, S3-Phospho	2.77
			T1-Acetyl, S7-Phospho	3.48
			T1-Acetyl, K14-Acetyl	3.82
	KKSRASPK	143-150	S6-Phospho	2.28
	QTKGVGASGSFR	83-95	T2-Phospho	2.30
	SASHPTYSEMIAAAIR	22-38	M10-Oxidation	3.74
			S3-Phospho	3.23
H5 overnight phosphorylated chromatin	TESLVLSPAPAKPK	1-14	T1-Acetyl	3.99
			S3-Phospho	4.43
			S7-Phospho	2.81
			T1-Acetyl, S3-Phospho	3.40
			T1-Acetyl, S7-Phospho	3.01
	SASHPTYSEMIAAAIR	22-38	M10-Oxidation	3.55
			S3-Phospho	3.23
	KKSRASPK	143-150	S6-Phospho	2.09

As introduced above, one of the aims of the present study was to determine the position of the phosphate groups incorporated in SPKK motifs after phosphorylation with CDK2 for different times.

Sequence coverages of histone H5 were the highest obtained in the Tandem MS analysis but only a few peptides containing SPKK motifs were detected, mainly because the regions that were poorly covered were the ones containing the SPKK motifs. Table 3.5. summarizes the results for the SPKK motifs that could be found in ESI/LC-MS/MS experiments for histone H5. If not detected, they appear marked by a hyphen (-) in the corresponding experiment in the table. If the SPKK motif was found but it was not phosphorylated, a cross mark (X) appears in the corresponding table cell; if found and phosphorylated, the peptides are marked by a check symbol (√). Sequence coverage of each sample is given in (%).

Two important considerations have to be done in relation to the identification of modified peptides and sequence coverage in a Tandem MS analysis. First, if sequence coverage includes a certain peptide containing an SPKK motif, it will be detected by the equipment, modified or not, and in both cases its detection can be validated. Second, when the certain peptide is not covered in the analysis, it is not found, and then, whether it is modified or not, cannot be determined. The non-detection of a peptide does not mean that it is not present in the protein or in the cleaved sample; but that the product of the enzymatic cleavage was not sufficiently ionized to be detected. Despite the initial amount of peptides from the protein is the same, a complex series of factors result in very different signal intensities for each of these peptides[440]. Thus, some of the peptides will be clearly visible and some will not. This usually happens for complex mixtures and for modified peptides that often do not ionize as well as non-modified peptides. Because of that, strategies like direct "online" electrospray mass spectrometry (LC/MS) are carried out after enzymatic cleavage. Nevertheless, even the combination of those does not necessarily result in measurement of peptides covering the complete protein. Very short peptides of only a few amino acids or very long peptides of >30 amino acids are often not detected[440].

Table 3.5. SPKK motifs identification and their phosphorylation status by ESI/LC-MS/MS analyses in H5

SPKK motif	S residue	peptide coverage and identification						
		H5 from soluble fraction of native chromatin (82%)	H5 from phosphorylated chromatin					
			15min (82%)	30min (76%)	1h (73%)	3h (64%)	5h (65%)	o/n (63%)
SPGK	S104	X	-	X	X	-	-	-
SPKK	S117	-	-	-	-	-	-	-
SPAK	S129	X	X	X	-	-	-	-
SPKK	S148	X	X	X	√	X	√	√
SPKK	S185	-	-	X	-	-	-	-

SPKK motifs are (S/T)-X-Z-(K/R)
hyphen (-); not detected peptide
check symbol (√); detected peptide/ phosphorylation
cross (X); detected peptide/ no phosphorylation
o/n; abbreviation for 'overnight'

As it can be seen in Table 3.5, only SPKK motif in position S148 was identified in all the samples but it was found modified only in the samples phosphorylated for 1 hour, 5 hours and overnight. These results suggest that phosphorylation in this position might be due to CDK2 kinase reaction and that this phosphate group was incorporated between 30 minutes and 1 hour of reaction. Note that this modification was not detected in the sample phosphorylated for 3 hours, whereas it was in the ones from 5 hours and overnight phosphorylation. Since all the phosphorylated samples were taken at the corresponding time from the same phosphorylation reaction, the detection of phosphorylation at 1 hour but its non-detection at 3 hours suggests that this modification might be present in very little amounts at these phosphorylation times, whereas it might had increased later on with time of phosphorylation, since it was easily detected and highly present after 5 hours and overnight phosphorylation. Figure 3.6. shows the MS^2 spectrum of the peptide containing this modification in the sample of 5 hours phosphorylated chromatin.

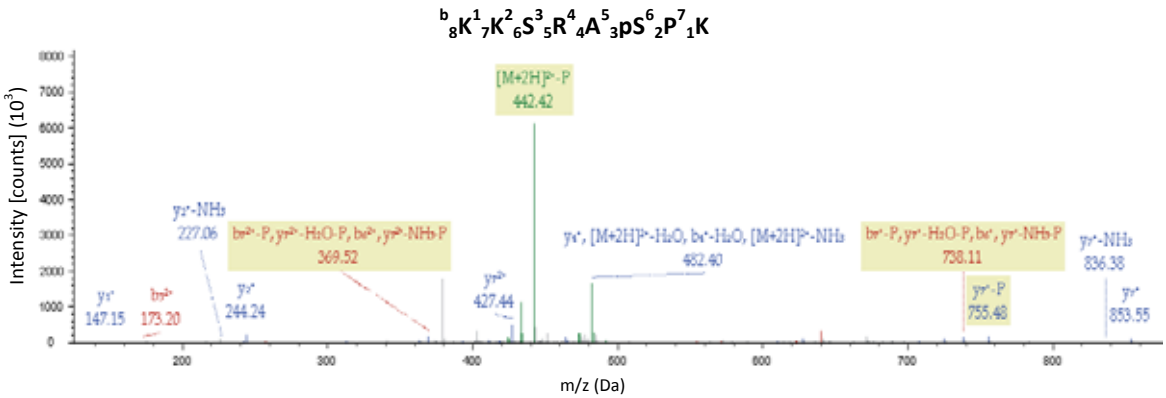


Figure 3.6. Mass spectra of the phosphorylated peptide of chicken histone H5 corresponding to the SPKK motif in position S148. Peaks are labelled in the mass spectra and the fragment ions are indicated in the peptide sequence. Peptide identification with SEQUEST (v1.13). Fragment match tolerance used for search: 0.8 Da. MS^2 spectra of the phosphorylated peptide. XCorr: 2.45; Probability: 11.29. The MS^2 corresponds to the H5 sample after 5 hours of chromatin *ex vivo* phosphorylation.

None of the other SPKK motifs found (S104, S129 and S185) were identified as phosphorylated peptides after short phosphorylation times (<1 hour). For longer phosphorylation times, the peptides were not identified in the analysis. S185 was only identified in the sample of H5 after 30 minutes of phosphorylation. The SPKK in position S117 was never detected. All the SPKK motifs identified in the sample of native chromatin were always not phosphorylated. This result indicates that the identified phosphorylations were incorporated due to CDK2 phosphorylation.

Proteomic characterization of chicken histones H1 after ex vivo phosphorylation of chromatin

Analysis of chicken histones H1 by MALDITOF-MS

As already mentioned in Chapter Two, histone H1 fraction is composed of six different subtypes[441] (H1.01, H1.02, H1.03, H1.10, H1.1L and H1.1R). Histone H1 subtypes are larger than H5 (about 220 amino acids vs 189 amino acids), but they are considered to have similar functions *in vivo*[442,443,444].

In terms of proteomic analysis, the methods applied for the identification of histone H1 variants in some species are not useful for the analysis of H1s of other species[345], since linker histones usually exhibit high sequence divergence.

The six isoforms of the chicken histone H1 have a sequence identity of 76%, sharing 172 identical positions and 21 similar ones (data for global identity in multiple sequence alignment by Clustal O; Figure 3.7. The percent matrix Identity showed in Chapter Two shows higher sequence identity between each pair of H1 subtypes). Apart from N^{α}-terminal acetylation, no other PTMs are entered for none of the histones H1 in the UniProt database, despite they have, at least, three SPKK motifs.

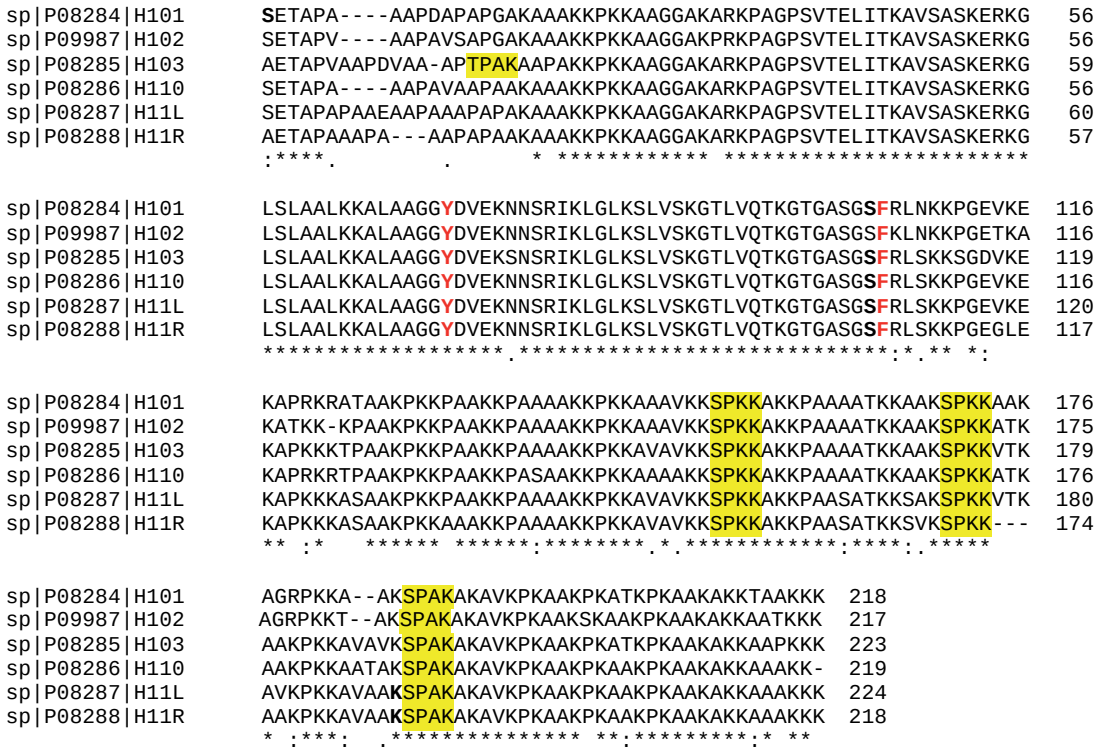


Figure 3.7. Clustal O (1.1.0) multiple sequence alignment of chicken histone H1 subtypes. Histone sequences extracted from the UniProt database. The SPKK motifs are highlighted in yellow. SPKK motif: (S/T)-P-X-Z, where *X* is any amino acid and *Z* is a basic amino acid. α-chymotrypsin cleavage sites are marked in red. α-chymotrypsin cleavage occurs in the C-term of F/Y/W, not before P.

When HPCE separations were carried out, six peaks that corresponded to the six chicken histone H1 subtypes could be identified, in addition to the two H5 variants (Figure 3.2). After HPCE identification, the whole histone H1 fraction, composed of the six mentioned variants, was separated from the H5 fraction by RP-HPLC. However, isolation of every single H1 variant from the mixture was not possible. After RP-HPLC separation, the whole six chicken histone H1 fraction was partially cleaved with α-chymotrypsin for posterior analysis by MALDITOF-MS.

As in the case of H5, partial cleavage with α-chymotrypsin divided each of the six H1 subtypes into two peptides (Table 3.6): the first one, with a molecular weight between 10268.98 Da and 10635.36 Da, included the N-terminal domain and most of the globular domain; and the second peptide, mainly consisting on the C-terminal domain of histones H1, with a molecular weight between 11370.11 Da and 11821.72 Da. This mixture was analysed by MALDITOF-MS.

Table 3.6. Cleavage of chicken histones H1 with α-chymotrypsin

H1 subtype	$Mass_{av}$ (Da)	Positions	# Missed cleavages	Sequence
H1.01	10296.95	2-106	1	SETAPAAAPDAPAPGAKAAAKKPKKAAGGAKARKPAGPSVTELIT KAVSASKERKGLSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGT LVQTKGTGASGSF
	11636.38	107-219	0	RLNKKPGEVKEKAPRKRATAAKPKKPAAKKPAAAAKKPKKAAAVK KSPKKAKKPAAAATKKAAKSPKKAAKAGRPKKAAKSPAKAKAVKP KAAKPKATKPKAAKAKKTAAKKK
H1.02	10325.05	2-106	1	SETAPVAAPAVSAPGAKAAAKKPKKAAGGAKPRKPAGPSVTELITK AVSASKERKGLSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGTL VQTKGTGASGSF
	11445.18	107-218	0	KLNKKPGETKAKATKKKPAAKPKKPAAKKPAAAAKKPKKAAAVKK SPKKAKKPAAAATKKAAKSPKKATKAGRPKKTAKSPAKAKAVKPK AAKSKAAKPKAAKAKKAATKKK
H1.03	10593.36	2-109	1	AETAPVAAPDVAAAPTPAKAAPAKKPKKAAGGAKARKPAGPSVT ELITKAVSASKERKGLSLAALKKALAAGGYDVEKSNSRIKLGLKSLVS KGTLVQTKGTGASGSF
	11821.72	110-224	0	RLSKKSGDVKEKAPKKKTPAAKPKKPAAKKPAAAAKKPKKAVAVK KSPKKAKKPAAAATKKAAKSPKKVTKAAKPKKAVAVKSPAKAKAV KPKAAKPKATKPKAAKAKKAAPKKK
H1.10	10268.98	2-106	1	SETAPAAAPAVAAPAAKAAAKKPKKAAGGAKARKPAGPSVTELIT KAVSASKERKGLSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGT LVQTKGTG ASGSF
	11623.33	107-220	0	RLSKKPGEVKEKAPRKRTPAAKPKKPAAKKPASAAKKPKKAAAAKK SPKKAKKPAAAATKKAAKSPKKATKAAKPKKAATAKSPAKAKAVKP KAAKPKAAKPKAAKAKKAAAKK
H1.1L	10635.36	2-110	1	SETAPAAAPAVAAPAAKAAAKKPKKAAGGAKARKPAGPSVTELIT KAVSASKERKGLSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKGT LVQTKGTGASGSF
	11781.65	111-225	0	RLSKKPGEVKEKAPKKKASAAKPKKPAAKKPAAAAKKPKKAVAVKK SPKKAKKPAASATKKSAKSPKKVTKAVKPKKAVAAKSPAKAKAVKP KAAKPKAAKPKAAKAKKAAAKKK
H1.1R	10322.05	2-107	1	AETAPAAAPAAAPAPAAKAAAKKPKKAAGGAKARKPAGPSVTELI TKAVSASKERKGLSLAALKKALAAGGYDVEKNNSRIKLGLKSLVSKG TLVQTKGTGASGSF
	11370.11	108-219	0	RLSKKPGEGLEKAPKKKASAAKPKKAAAKKPAAAAKKPKKAVAVK KSPKKAKKPAASATKKSVKSPKKAAKPKKAVAAKSPAKAKAVKPKA AKPKAAKPKAAKAKKAAAKKK

α-chymotrypsin cleavage occurs in the C-side of F/Y/W, not before P. Missed cleavages are shown in red.
Data generated with the PeptideMass application on ExpasyTools
Mass corresponds to the average value (Da)

A MALDITOF-MS spectrum of the histone H1 fraction of each chromatin phosphorylation time was obtained. Identification of peaks in these MALDITOF-MS spectra was difficult due to the impossibility of previous isolation of the six histone H1 variants by RP-HPLC and the similarity of the masses of some of the peptides obtained by enzymatic cleavage. As already mentioned, enzymatic cleavage generated two peptides of each variant, which eventually meant at least twelve different peptides to be identified, aside from the potential presence of post-translational modifications.

As mentioned above, MALDITOF-MS allows the identification of peaks with an experimental error of only ±5Da. Molecular masses of the peptides, when not modified, would have a difference in Daltons

sufficient to be distinguished with this technique. However, the great heterogeneity of the sample, the presence of PTMs in the peptides, the differences in the ionization of the peptides in each run and, to some extent, the presence of protein degradation, did not allow, in some cases, the identification of all the six histone H1 variants.

The data obtained was analysed by comparing the m/z value of each peak with the theoretical average molecular masses (Da) of the peptides of all chicken histone H1 subtypes. Post-translational modifications were also considered in the identification of peaks: acetylation (ac1) in peptide 1, which will correspond with the reported acetylation in serine S1 (Table 2.4.); and phosphorylation (mono-, di- and tri-) for both peptides. Acetylation in serine S1 is only described for subtypes H1.01, H1.02, H1.10 and H1.1L in the UniProt database. Even so, we considered this PTM to be present in all the subtypes. All chicken histone H1 subtypes have three SPKK motifs, located in their C-terminal domain that can be potentially phosphorylated by CDK2 kinase. Histone H1.03 has another SPKK motif in its N-terminal domain (Figure 3.7.).

In Table 3.7., a summary of the detected chicken H1 variants and their modified forms is presented for each sample.

We found phosphorylated species for peptides 1 and 2 and also acetylated species for the N-terminal peptides (peptide 1). Figure 3.8. shows the MALDITOF spectra of the H1 fraction of native soluble chromatin and of the phosphorylated chromatin after 5 hours to illustrate the results obtained. The MALDITOF-MS spectra corresponding to the rest of the samples are compiled in the Appendix I. In histones H1, the difference in mass between peptide 1 and peptide 2 is about 1000Da and, because of that, these two peptides appear in different regions of the MALDITOF spectra. The identified H1 subtypes and their post-translational modifications are indicated in the spectrum. H1 subtypes will be named with a preceding N when the spectra corresponds to peptide 1 (N-terminal plus globular domain) and with a preceding C when the spectra corresponds to peptide 2 (C-terminal domain).

The first observation that can be extracted from the results shown in Table 3.7. is that, in most cases, the peptides found corresponded to the same histone H1 subtypes. The peptide 1 and the peptide 2 of histone H1 variants H1.1L and H1.1R were always present. In some cases, while the presence of one of the peptides of a certain subtype could be identified, the other was not; i.e. for the H1 fraction derived from 1 hour phosphorylated chromatin, the N-terminal peptide of H1.03 was identified whereas its C-terminal peptide was not. In contrast, for H1.1R, both peptides could be identified, plus a mono-phosphorylated variant of its N-terminal peptide.

The similarity of the masses between the different peptides led to the overlapping of m/z peaks in three cases. This fact did not allow to assign the peak to a concrete histone H1 subtype, since the differences between the overlapped peaks were lower than the experimental error value of MALDITOF-MS assays (±5Da). In Table 3.7., overlapping of peptide pairs are marked with a super index symbol (†,* and °).

The first overlap (marked with † in Table 3.7.) occurred between unacetylated peptides 1 of subtypes H1.1R and H1.02, which differ only in 3Da (10321.05Da vs. 10324.05Da, respectively). MALDITOF-MS spectra showed a single peak at about an m/z value of 10322.6Da, which is, in fact, the average of these two masses. Thus, this m/z value could not be assigned for sure to any of the two mentioned H1 subtypes. Most likely, both peaks were present but, since their molecular mass values are so close, they might had been detected as a single peak with the average molecular mass by the MALDITOF-MS equipment.

The second overlap was detected between the m/z peaks corresponding to the acetylated peptides of these two subtypes (marked with * in Table 3.7.). The addition of the same post-translational modification to peptides 1 of subtypes H1.1R and H1.02 leads to the same difference in the molecular

masses of the peptides. In this case, a single peak with an experimental m/z value of 10363.05Da was detected in the MALDITOF-MS spectra. This m/z value exactly corresponded to the theoretical value of H1.1R's acetylated peptide 1. The difference between this experimental value (10363.05 Da) and the theoretical one for the acetylated H1.02's peptide 1 (10366,05 Da) is within the experimental error value (±5Da) of MALDITOF-MS analyses. Taking all this into account, it could be that: (a) the detected peak corresponded to the acetylated peptide 1 of H1.02; (b) the peak corresponded to the acetylated peptide 1 of H1.1R, since the experimental m/z value matches the exact mass; and (c) the experimental peak was an average of the latter ones and the acetylated peptides 1 of the two subtypes were present in the sample, as it has been discussed for the unacetylated peptide. In fact, both subtypes have been shown to be acetylated in Chapter Two of the present study.

However, a third explanation is possible, if we consider together the first and the second overlaps. In this case, the experimental peak around 10322.6Da, discussed as the first overlap, could correspond to the unacetylated peptide 1 of H1.1R, for which no acetylation has been reported, whereas the experimental peak discussed as the second overlap (~10363.05Da) will correspond to the acetylated peptide 1 of the H1.02, for which the acetylation at S1 is reported.

The third overlap (marked with ° in Table 3.7.) occurred between the potentially acetylated peptide 1 of H1.03 (Mw_{th}=10634,36Da) and the unacetylated peptide 1 of H1.1L (Mw_{th}=10634,36Da). These two peptides have the same theoretical mass value. However, H1.03 has no reported acetylation in the UniProt database.

At this point, none of these explanations can be confirmed in the MALDITOF analysis and, because of that, mapping of the peptides by MS/MS sequencing was performed.

As already introduced, several phosphorylated species were also detected in the MALDITOF-MS spectra phosphorylation was identified in peptides 1 and 2 of different histone H1 subtypes. Chicken histone H1 subtypes have three SPKK motifs located in their CTD (peptide 2), and H1.03, also has a fourth SPKK site in the N-terminal domain (peptide 1). As mentioned, SPKK motifs are phospho-sites of CDK2 kinase that can be phosphorylated after treatment with this enzyme. However, it is important to consider that basal phosphorylation could be already present in native chromatin.

In the MALDITOF-MS analysis of native soluble chromatin, phosphorylation was only found in H1.1R subtype (Table 3.7, Figure 3.8.), in which a tri-phosphorylated species of peptide 2 was identified. This species correspond to basal phosphorylation of chromatin, since this sample was used as a control and it was not *ex vivo* phosphorylated. This species was not identified in the phosphorylated samples, which indicates that it might be present in very little amounts.

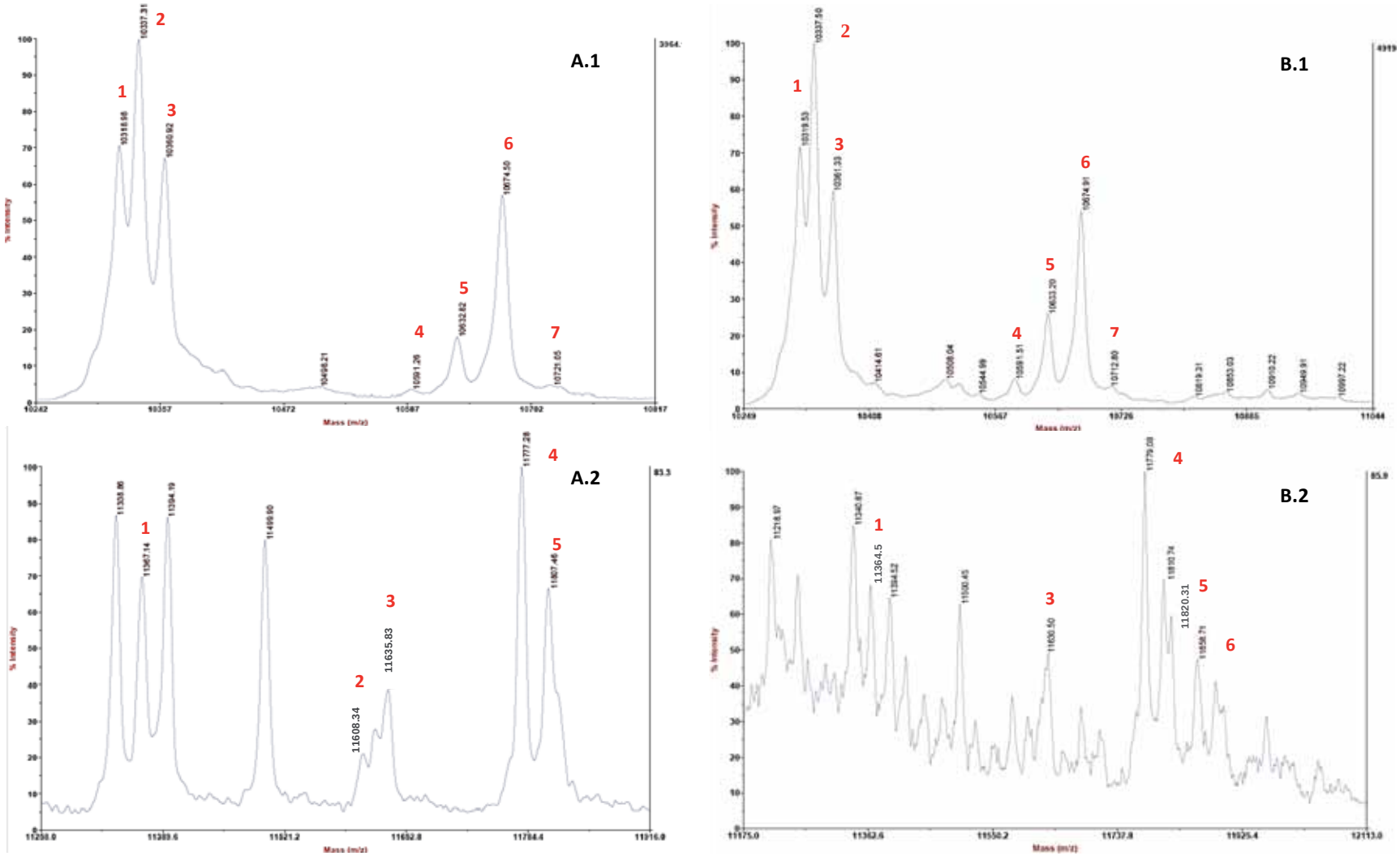


Figure 3.8. MALDITOF-MS spectra of α-chymotrypsin cleaved H1 fraction from chicken erythrocyte native chromatin (A.1 and A.2) and from *ex vivo* phosphorylated chromatin for 5 hours (B.1 and B.2). A.1 and B.1 correspond to the region of the MALDITOF spectra of peptide 1. Peak numbering (in red) in these two spectra is as follows: 1. NH1.1Rac0 or NH1.02ac0; 2. NH1.01ac1; 3. NH1.1Rac1 or NH1.02ac1; 4. NH1.03ac0; 5. NH1.1Lac0 or NH1.03ac1; 6. H1.1Lac1; 7. H1.1Lac0-1P or H1.03ac1-1P. NH1.1R and NH1.02 have similar masses and could be overlapped. The same happens for NH1.1Lac0 and NH1.03ac1. A.2 and B.2 correspond to the region of the MALDITOF spectra of peptide 2. Peak numbering (in red) in these two spectra is as follows: 1. CH1.1R-0P; 2.CH1.1R-3P; 3. CH1.01-0P; 4. CH1.1L-0P; 5. CH1.03-0P; 6. CH1.1L-1P. ac0=unacetylated; ac1=monoacetylated; 0P=unphosphorylated; 1P=monophosphorylated.

Table 3.7. Summary of experimental data obtained by MALDITOF-MS analysis of chicken histone H1 variants.

Histone H1 subtype	Theoretical Average Molecular Mass (Da)							Experimental H1 variants identified in MALDITOF-MS spectra						
	Thr. Av. (Da) of subtype	Thr. Av. (Da) of peptides						chromatin phosphorylation time						
				0P	+1P	+2P	+3P	soluble native	15min	30min	1h	3h	5h	overnight
H1.01	21913,30	**peptide 1 (Nter+GD)**	**ac0**	10295,96	10375,96	10455,96	10535,96	-	-	-	-	-	-	-
			ac1	10337,96	10417,96	10497,96	10577,96	0P	0P / 2P	0P	0P	0P	0P	0P
		peptide 2 (Cter)		11635,37	11715,37	11795,37	11875,37	0P	0P	-	-	-	0P	0P / 1P
H1.02	21750,21	**peptide 1 (Nter+GD)**	**ac0**	10324,05	10404,05	10484,05	10564,05	0P†	0P†	0P†	0P†	0P†	0P†	0P†
			ac1	10366,05	10446,05	10526,05	10606,05	0P*	0P*	0P*	0P*	0P*	0P*	0P*
		peptide 2 (Cter)		11444,18	11524,18	11604,18	11684,18	-	0P /1P	-	-	-	-	0P /1P
H1.03	22395,06	**peptide 1 (Nter+GD)**	**ac0**	10592,36	10672,36	10752,36	10832,36	0P	0P	-	0P	0P	0P	0P
			ac1	10634,36	10714,36	10794,36	10874,36	0P°	0P°	0P°	0P° / 1P°	0P° /3P°	0P° / 1P°	0P° / 1P°
		peptide 2 (Cter)		11820,72	11900,72	11980,72	12060,72	0P	0P / 1P	-	-	-	0P	-
H1.10	21872,29	**peptide 1 (Nter+GD)**	**ac0**	10267,98	10347,98	10427,98	10507,98	-	-	-	-	-	-	-
			ac1	10309,98	10389,98	10469,98	10549,98	-	-	-	-	-	-	-
		peptide 2 (Cter)		11622,33	11702,33	11782,33	11862,33	-	-	-	0P	-	-	-
H1.1L	22396,99	**peptide 1 (Nter+GD)**	**ac0**	10634,36	10714,36	10794,36	10874,36	0P°	0P°	0P°	0P° / 1P°	0P° /3P°	0P° / 1P°	0P° / 1P°
			ac1	10676,36	10756,36	10836,36	10916,36	0P	0P	0P	0P	0P	0P	0P
		peptide 2 (Cter)		11780,65	11860,65	11940,65	12020,65	0P	0P /1P	-	0P	-	0P/1P	0P / 1P /2P
H1.1R	21672,14	**peptide 1 (Nter+GD)**	**ac0**	10321,05	10401,05	10481,05	10561,05	0P†	0P†	0P†	0P†	0P†	0P†	0P†
			ac1	10363,05	10443,05	10523,05	10603,05	0P*	0P*	0P*	0P*	0P*	0P*	0P*
		peptide 2 (Cter)		11369,11	11449,11	11529,11	11609,11	0P /3P	0P /1P /2P	-	-	0P / 1P	0P	0P /1P

Super index symbols (†,*,°) mark possible overlapping of peaks in the MALDI spectra due to similarity of molecular masses between peptide pairs.

Hypens (-) mark peptides not found in the MALDI spectra.

In the MALDI-TOF analyses of the *ex vivo* phosphorylated samples, mono-, di- and tri-phosphorylated peptides were found for the N-terminal peptide (peptide 1), together with the unphosphorylated species (Table 3.7.). In H1.01, a di-phosphorylated species was detected after 15 minutes of phosphorylation. However, this PTM was not found for longer phosphorylation times, and only the unphosphorylated species of peptide 1 was identified. Unacetylated peptide 1 of H1.1L and acetylated peptide 1 of H1.03 also presented phosphorylated species of their N-terminal peptides, which appeared after 1hour of phosphorylation. Mono-phosphorylation was found in the samples of 1 hour, 5 hours and overnight phosphorylation. Instead of this species, a tri-phosphorylated peptide 1 was found in the sample of 3 hours of phosphorylation. Note that unacetylated peptide 1 of H1.1L and acetylated peptide 1 of H1.03 is one the already mentioned possible overlaps (marked with ° in Table 3.7.), and, because of that, the phosphorylated species cannot be attributed to either subtype. On the other hand, these phosphorylated species appear, as mentioned, after 1 hour of phosphorylation, but only H1.03 has an SPKK motif in this region of the protein. It could be assumed then, that the PTM detected corresponds to CDK2 phosphorylation of peptide 1 of H1.03. The presence of a tri-phosphorylated species in the sample of 3 hours, which was not found for other phosphorylation times, suggests that there might be other phosphorylated no-CDK2 sites in the protein and that this phosphorylations must be already present in the protein but they were not always detected indicating that they might be present in little amounts.For the other histone H1 subtypes, no phosphorylation peaks were detected in the MALDI-TOF spectra region of peptide 1.

In the MALDITOF analysis of the C-terminal peptides (peptide 2) of histones H1, mono-phosphorylated species were identified after 15 minutes of phosphorylation for H1.02, H1.03 and H1.1R. For the latter, also a di-phosphorylated species was identified (Table 3.7.). These phosphorylated species were not detected in the native sample, and, because of that, it can be assumed that they are due to CDK2 phosphorylation of SPKK motifs. Even though, in the samples of 30 minutes, 1 hour, 3 hours and 5 hours phosphorylated chromatin, very little peaks corresponding to the C-terminal domain of histones H1 (peptide 2) were found or could be assigned (Figure 3.8.). In the H1 fraction from overnight phosphorylated chromatin, peaks corresponding to unphosphorylated and mono-phosphorylated species of peptide 2 were detected again.

The lack of peak identification for the C-terminal peptides in the samples of 30 minutes, 1hour, 3hours and 5hours phosphorylated chromatin was due to the limitations of the technique and the complexity of the samples, in which the amount of peaks per spectrum generated a lot of background. It has to be considered that phosphorylated species are expected to be in a lower proportion than the unphosphorylated ones and, thus, presumably, the identification of the corresponding peaks was not possible because their intensities were lower or became overlapped with other peaks.

This fact did not allow determining an increase of phosphorylation with time of reaction, neither the proportions of phosphorylated species in relation with the unphosphorylated peptides (relative ratios), as we were able to do in the analysis of histone H5.

Mapping of histone H1 post-translational modifications

After separation from H5 fraction by RP-HPLC, each H1 fraction was enzymatically cleaved with trypsin for 15minutes and then, sequenced by Tandem MS for PTMs mapping. As already explained, after RP-HPLC, the histone H1 fraction consisted in a mixture of the six H1 subtypes. Isolation of every single H1 subtype from the total mixture was not possible. The cleavage of H1 subtypes with trypsin generated peptides of ~20 residues that were analysed by LC-ESI-MS/MS and searched against a histone data base that allowed assigning an H1 subtype to a certain peptide sequence. As mentioned, all modified

peptides with X^2corr values above 2.5 were validated manually (Table 3.8.) for precise verification and localization of the post-translational modification.

Table 3.8. summarizes all the post-translational modifications found in histones H1. Considering the highest sequence similarity between H1 subtypes, some peptides matched with more than one variant (in **Table 3.8.,** a list of the histone subtypes that share the same peptide is shown in these cases). Thus, when a post-translational modification is found in a 'shared-sequence' peptide, one cannot assure if that modification is present in one or more H1 subtypes from the mixture. Despite that, conserved regions usually mean conserved functions, so it is possible to think that the modification is present in all variants. Furthermore, depending on the sequence of each variant, its amount in the sample and the peptides found in MS/MS sequencing experiments, the sequence coverage was different for each variant, as it has been shown in Table 3.1.

In all the samples, both the acetylated and unacetylated N^{α}-terminal peptides were identified, together with other PTMs that were already present in the native chromatin (Chapter Two).

Table 3.8. Modified peptides identified by ESI/LC- MS/MS in chicken erythrocytes H1 histone variants

Sample	Peptide sequence	positions	H1 histone variant	PTM	MS^2 Xcorr
H1 soluble native chromatin	SETAPAAAPDAPAPGAK	1-17	H1.01	S1-Acetyl	3.17
	SETAPVAAPAVSAPGAK	1-21	H1.02	S1-Acetyl	3.80
				T3-Phospho	3.08
	AETAPVAAPDVAAAPTPAK	1-19	H1.03	A1-Acetyl	3.59
	SETAPAAAPAVAAPAAK	1-17	H1.10	S1-Acetyl	3.63
	SETAPAPAAEAAPAAAPAPAK	1-21	H1 1L	S1-Acetyl	3.45
				T3-Phospho	3.07
	AETAPAAAPAAAPAPAAK	1-18	H1 1R	A1-Acetyl	2.28
	ARKPAGPSVTELITK	32-46 (H1.01)	H1.01, H1.03, H1.10, H1.1L, H1.1R	S8-Phospho	3.60
	KPAGPSVTELITK	34-46 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K1-Acetyl	3.81
	SLVSKGTLVQTK	86-97 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K5-Acetyl	2.52
	KAVAAKSPAKAK	186-197 (H1.1L)	H1.1L, H1.1R	K6-Acetyl	2.61
H1 15min phosphorylated chromatin	SETAPAAAPDAPAPGAK	1-17	H1.01	S1-Acetyl	3.10
	SETAPVAAPAVSAPGAK	1-17	H1.02	S1-Acetyl	2.68
				S1-Acetyl, T3-Phospho	3.94
	AETAPVAAPDVAAAPTPAK	1-19	H1.03	A1-Acetyl	3.04
				T16-Phospho	3.22
				A1-Acetyl, T16-Phospho	2.86
	SETAPAAAPAVAAPAAK	1-17	H1.10	S1-Acetyl	3.77
	SETAPAPAAEAAPAAAPAPAK	1-21	H1.1L	S1-Acetyl	2.68
	AETAPAAAPAAAPAPAAK	1-18	H1.1R	A1-Acetyl	3.94
	KPAGPSVTELITK	34-46 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K1-Acetyl	4.05
	SLVSKGTLVQTK	86-97 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K5-Acetyl	3.22
	KAVAAKSPAKAK	186-197 (H1.1L)	H1.1L, H1.1R	K6-Acetyl	3.34
	ALAAGGYDVEK	65-75 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K11-Dimethyl	2.53
	ARKPAGPSVTELITK	32-46 (H1.01)	H1.1L	K15-Dimethyl	3.12

H1 30min phosphorylated chromatin	SETAPAAAPDAPAPGAK	1-17	H1.01	S1-Acetyl	2.99
	SETAPVAAPAVSAPGAK	1-21	H1.02	S1-Acetyl	3.97
	AETAPVAAPDVAAAPTPAK	1-19	H1.03	A1-Acetyl	2.27
				T16-Phospho	3.22
				A1-Acetyl, T16-Phospho	3.17
	SETAPAAAPAVAAPAAK	1-17	H1.10	S1-Acetyl	3.14
	SETAPAPAAEAAPAAAPAPAK	1-21	H1.1L	S1-Acetyl	3.67
				S1-Acetyl, T3-Phospho	2.77
	AETAPAAAPAAAPAPAAK	1-18	H1.1R	A1-Acetyl	3.50
	KPAGPSVTELITK	34-46 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K1-Acetyl	3.80
	SLVSKGTLVQTK	86-97 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K5-Acetyl	3.09
	KAVAAKSPAKAK	186-197 (H1.1L)	H1.1L, H1.1R	K6-Acetyl	2.78
				K6-Acetyl, S7-Phospho	2.85
	ALAAGGYDVEK	65-75 (H1.01)	H1.1L, H1.03	K11-Dimethyl	3.98
H1 1hour phosphorylated chromatin	SETAPAAAPDAPAPGAK	1-17	H1.01	S1-Acetyl	3.74
	SETAPVAAPAVSAPGAK	1-21	H1.02	S1-Acetyl	3.10
				S1-Acetyl, T3-Phospho	2.89
	AETAPVAAPDVAAAPTPAK	1-19	H1.03	A1-Acetyl	3.18
				T16-Phospho	3.26
				A1-Acetyl, T16-Phospho	2.99
	SETAPAAAPAVAAPAAK	1-17	H1.10	S1-Acetyl	3.23
	SETAPAPAAEAAPAAAPAPAK	1-21	H1.1L	S1-Acetyl	3.58
				S1-Acetyl, T3-Phospho	2.90
	AETAPAAAPAAAPAPAAK	1-18	H1.1R	A1-Acetyl	3.67
	KPAGPSVTELITK	34-46 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	S6-Phospho	2.27
				K1-Acetyl	3.58
	ARKPAGPSVTELITK	32-46 (H1.01)	H1.01, H1.03, H1.10, H1.1L, H1.1R	S8-Phospho	3.10
	KAVAAKSPAKAK	186-197 (H1.1L)	H1.1L, H1.1R	K6-Acetyl, S7-Phospho	2.92
	SLVSKGTLVQTK	86-97 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K5-Acetyl	2.06
	KALAAGGYDVEK	64-75 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L	K12-Dimethyl	3.25
	ALAAGGYDVEK	65-75 (H1.01)	H1.01, H1.02, H1.10, H1.1L	K11-Dimethyl	2.22
	ALAAGGYDVEKSNSR	68-82	H1.03	K11-Acetyl	2.02

H1 3hours phosphorylated chromatin	SETAPAAAPDAPAPGAK	1-17	H1.01	S1-Acetyl	3.68
	SETAPVAAPAVSAPGAK	1-21	H1.02	S1-Acetyl	3.77
				S1-Acetyl, T3-Phospho	3.08
	AETAPVAAPDVAAAPTPAK	1-19	H1.03	A1-Acetyl	3.16
				T16-Phospho	3.43
				A1-Acetyl, T16-Phospho	2.55
	SETAPAAAPAVAAPAAK	1-17	H1.10	S1-Acetyl	3.40
	SETAPAPAAEAAPAAAPAPAK	1-21	H1.1L	S1-Acetyl	3.43
				S1-Acetyl, T3-Phospho	3.66
	AETAPAAAPAAAPAPAAK	1-18	H1.1R	A1-Acetyl	3.57
	SLVSKGTLVQTK	86-97 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K5-Acetyl	2.14
	KPAAAAKKPKKAVAVKKSPK	139-158 (H1.1L)	H1.03, H1.1L, H1.1R	S18-Phospho	4.24
	AVAVKSPAK	186-194	H1.03	S6-Phospho	2.61
	KAVAAKSPAKAK	186-197 (H1.1L)	H1.1L, H1.1R	K6-Acetyl	3.89
				K6-Acetyl, S7-Phospho	2.68
	KPAGPSVTELITK	34-46 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K1-Acetyl	4.00
				S6-Phospho	3.87
	ARKPAGPSVTELITK	32-46 (H1.01)	H1.01, H1.03, H1.10, H1.1L, H1.1R	S8-Phospho	2.94
	GTGASGSFR	101-109 (H1.03)	H1.10, H1.03, H1.1R	S5-Phospho	2.41
	KALAAGGYDVEK	67-78	H1.03	K12-Dimethyl	3.79
H1 5hours phosphorylated chromatin	SETAPAAAPDAPAPGAK	1-17	H1.01	S1-Acetyl	3.50
	SETAPVAAPAVSAPGAK	1-21	H1.02	S1-Acetyl	3.87
				S1-Acetyl, T3-Phospho	3.65
	AETAPVAAPDVAAAPTPAK	1-19	H1.03	A1-Acetyl	3.33
				T16-Phospho	3.48
				A1-Acetyl, T16-Phospho	3.26
	SETAPAAAPAVAAPAAK	1-17	H1.10	S1-Acetyl	3.56
	SETAPAPAAEAAPAAAPAPAK	1-21	H1.1L	S1-Acetyl	3.51
				S1-Acetyl, T3-Phospho	3.81
	AETAPAAAPAAAPAPAAK	1-18	H1.1R	A1-Acetyl	3.08
	KAVAAKSPAK	186-196 (H1.1L)	H1.1L, H1.1R	K6-Acetyl	2.99
				K6-Acetyl, S7-Phospho	2.67
	AVAVKSPAK	186-194	H1.03	S6-Phospho	2.23
	KPAGPSVTELITK	34-46 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K1-Acetyl	3.98
				S6-Phospho	2.22
	ARKPAGPSVTELITK	32-46 (H1.01)	H1.01, H1.03, H1.10, H1.1L, H1.1R	S8-Phospho	3.24
	SLVSKGTLVQTK	86-97 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K5-Acetyl	2.25
	KALAAGGYDVEK	67-78	H1.03	K12-Dimethyl	3.57

	SETAPAAAPDAPAPGAK	1-17	H1.01	S1-Acetyl	3.17
	SETAPVAAPAVSAPGAK	1-21	H1.02	S1-Acetyl	3.67
				S1-Acetyl, T3-Phospho	2.99
	AETAPVAAPDVAAAPTPAK	1-19	H1.03	A1-Acetyl	3.25
				T16-Phospho	3.30
				A1-Acetyl, T16-Phospho	2.94
	SETAPAAAPAVAAPAAK	1-17	H1.10	S1-Acetyl	3.09
	SETAPAPAAEAAPAAAPAPAK	1-21	H1.1L	S1-Acetyl	3.57
				S1-Acetyl, T3-Phospho	3.25
	AETAPAAAPAAAPAPAAK	1-18	H1.1R	A1-Acetyl	2.64
	KAVAVKKSPK	148-159 (H1.03)	H1.03, H1.1L, H1.1R	S8-Phospho	3.26
				K6-Acetyl, S8-Phospho	3.97
H1 overnight phosphorylated chromatin	KPAGPSVTELITK	34-46 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K1-Acetyl	4.23
				S6-Phospho	2.29
	KAVAAKSPAKAK	186-197 (H1.1L)	H1.1L, H1.1R	K6-Acetyl	3.69
				S7-Phospho	3.11
				K6-Acetyl, S7-Phospho	3.91
	KPAGPSVTELITKAVSASKER	34-54	H1.01	K1-Acetyl	2.95
				S6-Phos	2.75
	KPAGPSVTELITKAVSASKER	34-54 (H1.02)	H1.02, H1.10	S8-Phospho	4.42
				K6-Acetyl, S8-Phospho	3.97
	KAVAAKSPAKAKAVKPK	186-202	H1.1L, H1.1R	S7-Phospho, K12-Acetyl	3.22
	LSKKSGDVKEKAPK	110-123	H1.03	K9-Acetyl	3.03
	AVAVKSPAK	186-194	H1.03	S6-Phospho	4.10
	SLVSKGTLVQTK	86-97 (H1.01)	H1.01, H1.02, H1.03, H1.10, H1.1L, H1.1R	K5-Acetyl	3.14

As said before, sequence coverage in these experiments was not the optimal in the CTD region of proteins; especially when adding the condition of an initial complex mixture of H1 variants. Identifying differential phosphorylation in SPKK sites in relation with the time of phosphorylation would have allowed not only to determine which positions had been phosphorylated, but also in which order.

Trypsin cleaves the C-terminal of H1s to arginine and lysine, generating fragment ions containing basic residues that produce more intense peaks, easier to sequence by MS. Despite that, the high content of these residues in the proteins and the high sensitivity to cleavage of this region generated too short precursor ions to be detected[440].

Table 3.9. shows the SPKK motifs found in the H1 fraction, which were only identified for H1.03, H1.1L and H1.1R. These subtypes were also the ones with higher sequence coverage. No peptides containing an SPKK motif were found for H1.01, H1.10 and H1.02, and because of that, they are not shown in the table.

Table 3.9. SPKK motifs identification and their phosphorylation status by ESI/LC- MS/MS analyses in H1 variants

H1 variant	SPKK motif	native soluble chromatin		15min phosphorylated chromatin		30 min phosphorylated chromatin		1h phosphorylated chromatin	
		status	coverage	status	coverage	status	coverage	status	coverage
H1.03	T16	X	66,82%	√	66,37%	√	71,30%	√	60,99%
	S155	-		-		-		-	
	S173	-		-		-		-	
	S191	X		X		X		X	
H1.1L	S155	-	87,50%	-	58,48%	-	65,63%	-	64,73%
	S174	-		-		-		-	
	S192	X		X		√		√	
H1.1R	S153	-	81,19%	-	53,67%	-	64,22%	-	59,17%
	S171	-		-		-		-	
	S186	X		X		√		√	

H1 variant	SPKK motif	3h phosphorylated chromatin		5h phosphorylated chromatin		on phosphorylated chromatin	
		status	coverage	status	coverage	status	coverage
H1.03	T16	√	68,82%	√	60,99%	√	61,88%
	S155	√		-		√	
	S173	-		-		-	
	S191	√		√		√	
H1.1L	S155	√	64,73%	-	58,48%	√	70,98%
	S174	-		-		-	
	S192	√		√		√	
H1.1R	S153	√	56,42%	-	56,42%	√	63,30%
	S171	-		-		-	
	S186	√		√		√	

hyphen (-); not detected peptide

check symbol (√); detected peptide/phosphorylation

cross (X); detected peptide/ no phosphorylation

SPKK motifs are (S/T)-X-Z-(K/R)

The progression of phosphorylation of SPKK motifs with time could be seen in the first CDK2 site of H1.03 (T16) and in the third CDK2 site of H1.03, H1.1L and H1.1R (positions S191, S192 and S186, respectively). In the other cases, peptides containing an SPKK motif were not found in all the samples, which did not allow to see progression of phosphorylation. Nevertheless, if a certain peptide was found in the soluble native chromatin –which is actually the one *ex vivo* phosphorylated, and then, used as a control- and in one or more phosphorylated samples, some changes, likely due to phosphorylation by CDK2 kinase can be discussed.

H1.03 contains four SPKK sites: one in the N-terminal domain and the other three, in its C-terminal domain (Figure 3.7.). The first SPKK site (TPAK) was found in all the samples and, after 15 minutes of phosphorylation, it appeared to be modified in its T16 residue. The phosphorylated peptide was found in all the samples in high amounts, with the exception of native chromatin, where the modified peptide was not identified. This suggests that this CDK2 site is phosphorylated from the very beginning. Figure 3.9. shows the MS^2 spectrum of one of the peptides in which this phosphorylation was identified.

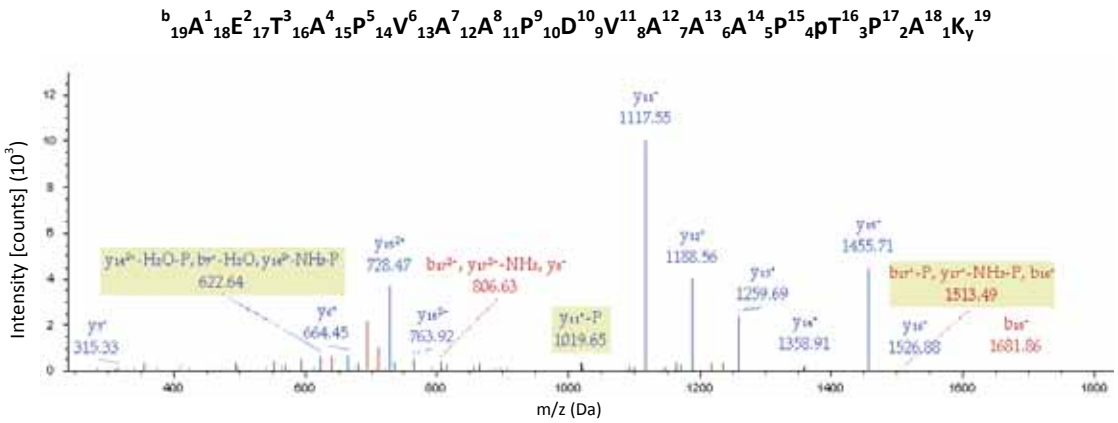


Figure 3.9. Mass spectra of the phosphorylated peptide of chicken histone H1.03 corresponding to the SPKK motif in position T16. Peaks are labelled in the mass spectra and the fragment ions are indicated in the peptide sequence. Peptide identification with SEQUEST (v1.13). Fragment match tolerance used for search: 0.8 Da. MS^2 spectra of the phosphorylated peptide. XCorr: 3.49; Probability: 102.63. The MS^2 corresponds to the H1 sample after 15 minutes of chromatin *ex vivo* phosphorylation, in which this modification was first identified.

The peptide containing the SPKK motif located in S191 of H1.03 variant was also identified in all experiments and, therefore, a progression due to CDK2 *ex vivo* phosphorylation can also be seen in this case. This SPKK site was not phosphorylated in native chromatin either after 15 minutes, 30 minutes or 1hour of phosphorylation reaction, but it already appeared phosphorylated after 3 hours of reaction, and this species was also identified in the 5 hours and overnight phosphorylated samples. Figure 3.10. (A) shows the MS^2 spectrum of one of this modified peptide.

In the case of H1.1L and H1.1R, the equivalent SPKK motif is in positions S192 and S186, respectively. This SPKK was also found in all the samples and it appeared phosphorylated after 30 minutes of reaction. This two subtypes share the same peptide sequence in which this SPKK was found, whereas the same peptide differs in a residue in the case of H1.03. As said, in H1.03, this SPKK motif appeared phosphorylated after 1 hour of reaction. This result indicates that the phosphorylation of this SPKK motif is due to CDK2 treatment and that it occurred before in H1.1L/H1.1R than in H1.03. Since both H1.1L and H1.1R share the peptide, it is not possible to determine if the modification occurs in either or both histone H1 subtypes. Figure 3.10. (B) shows the MS^2 spectrum of one of this modified and shared peptide of H1.1L and H1.1R.

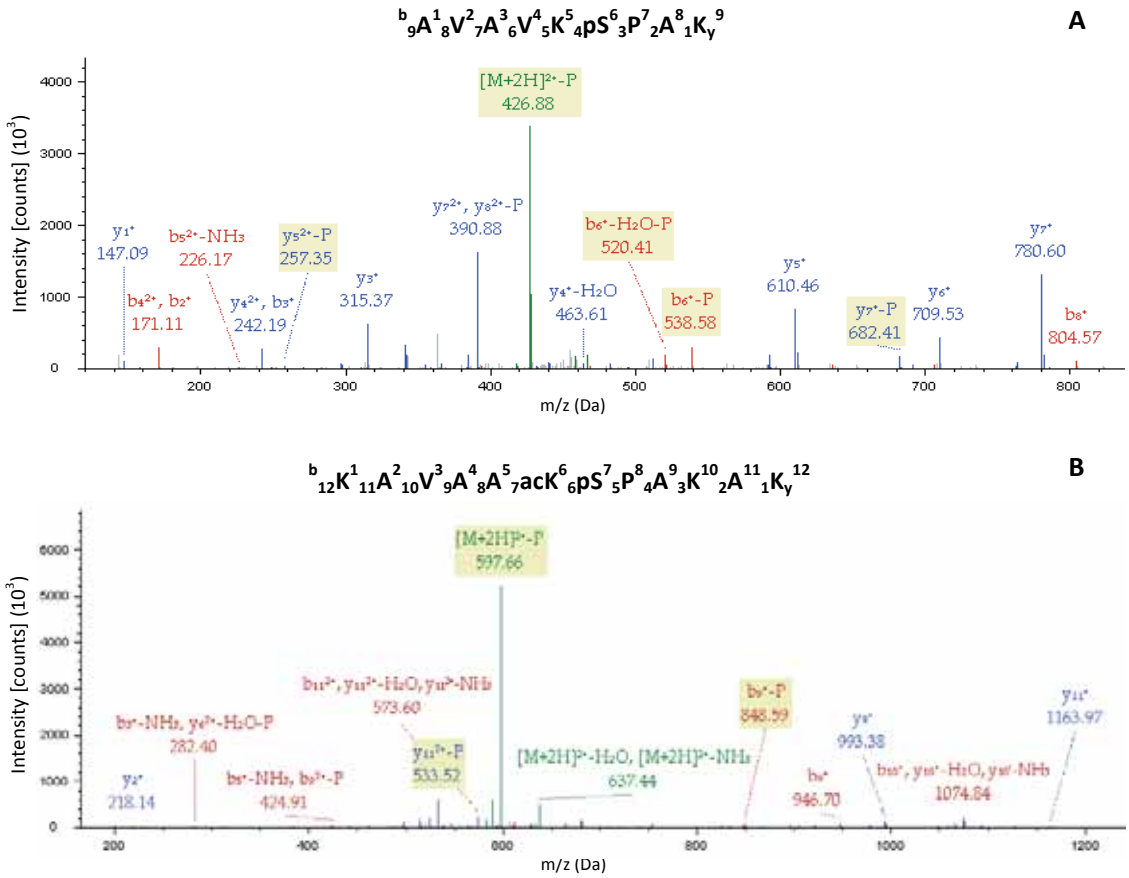


Figure 3.10. Mass spectra of the phosphorylated peptides of chicken histones H1.1L, H1.1R and H1.03 corresponding to the SPKK motifs in positions S192, S186 and S191, respectively. Peaks are labelled in the mass spectra and the fragment ions are indicated in the peptide sequence. Peptide identification with SEQUEST (v1.13). Fragment match tolerance used for search: 0.8 Da. **A.** MS^2 spectra of the phosphorylated peptide in S191 of H1.03. XCorr: 2.61; Probability: 15.99. The MS^2 corresponds to the H1 sample after 3 hours of chromatin *ex vivo* phosphorylation, in which this modification was first identified. **B.** MS^2 spectra of the phosphorylated peptide in S192 of H1.1L and equivalent position S186 in H1.1R. XCorr: 3.65; Probability: 24.39. The MS^2 corresponds to the H1 sample after 30 minutes of chromatin *ex vivo* phosphorylation, in which this modification was first identified.

A third phosphorylated SPKK motif (S155) was identified for H1.03 in the samples of 3 hours and overnight phosphorylated chromatin. Despite that, the peptide containing this SPKK was not found in the native chromatin sample and, thus, whether the phosphorylation was already present in the native sample or it was due to CDK2 phosphorylation could not be determined. Figure 3.11. shows the MS^2 spectrum of one of the peptides in which this phosphorylation was identified.

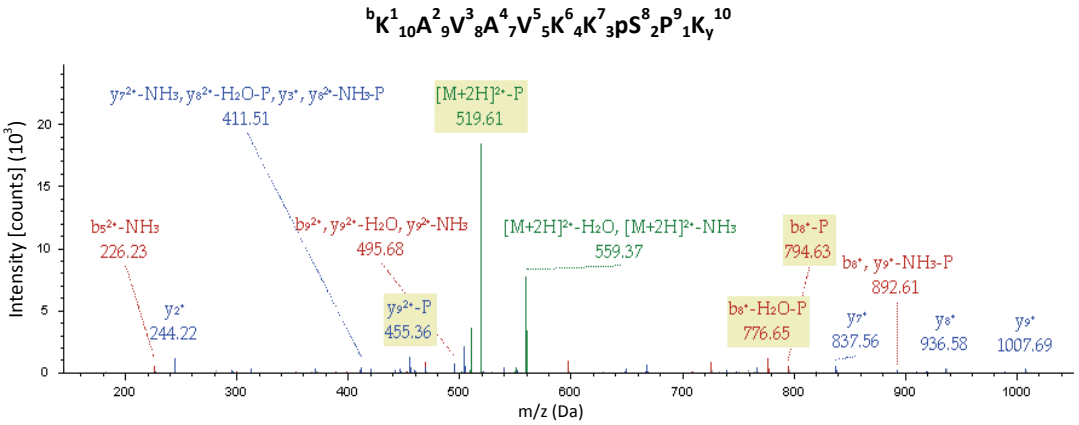


Figure 3.11. Mass spectra of the phosphorylated peptide of chicken histone H1.03 corresponding to the SPKK motif in position S155. Peaks are labelled in the mass spectra and the fragment ions are indicated in the peptide sequence. Peptide identification with SEQUEST (v1.13). Fragment match tolerance used for search: 0.8 Da**.** MS^2 spectra of the phosphorylated peptide. XCorr: 2.51; Probability: 20.06. The MS^2 corresponds to the H1 sample after overnight *ex vivo* phosphorylation of chromatin.

In none of the samples and for none of the subtypes, a peptide containing the SPKK motif in position S173 (H1.03 numbering) was identified.

Note that all the SPKK motifs found in native chromatin happened to be unphosphorylated. The number of phosphorylated SPKK motifs identified increased with the phosphorylation time. After 3 hours of phosphorylation, all the SPKK mapped had added a phosphate group (Table 3.9).

In summary, phosphorylation of SPKK motifs increased with time due to *ex vivo* chromatin phosphorylation with CDK2. This increase was observed in both linker histones and by all the techniques performed (radioactive phosphorylation assay, HPCE analysis, MALDITOF-MS and Tandem MS). The HPCE and the MALDITOF-MS analyses showed that this increase first occurred by increasing the amount of monophosphorylated species, followed by the appearance of di- and triphosphorylated species. Calculation of the percentages of the phosphorylated species of H5 CTD from the MALDITOF-MS analyses showed that, after overnight phosphorylation, the amount of phosphorylated species (%1P+%2P) had increased up to 54%. However, this phosphorylation was partial, since only around 14% of potentially phosphorylated CDK2 sites of histone H5, which has five SPKK motifs, were phosphorylated. In the case of histone H1 subtypes, which contain at least three SPKK sites, these calculations could not be performed. However, in the MALDITOF-MS analyses of histone H1 subtypes, mono- and diphosphorylated species were detected and, in the Tandem MS experiments, all the SPKK motifs were found to be modified after overnight phosphorylation. Hence, the level and quality of phosphorylation was more significant for H1 subtypes although it was also partial, because it was extended to all the CDK2 sites.

Chapter Four

Effects of histone H1 phosphorylation on chromatin aggregation

Phosphorylation is the most important modification affecting linker histones and the most extensively studied post-translational modification of chromatin4. Reversible phosphorylation of linker histones by cyclin-dependent kinases (CDKs) occurs in SPKK motifs[233,234]. H1 phosphorylation is involved in chromatin decondensation during the S phase, whereas hyperphosphorylation is related to its condensation during M phase and has been shown to be required for the maintenance of condensed metaphase chromosomes[249,257]. It might be, then, a direct relationship between chromatin compaction and linker histones phosphorylation in which the latter acts as a global regulator of chromatin structure[260,261,264].

The soluble chromatin fraction was phosphorylated *ex vivo* for 1 hour, 5 hours and overnight with CDK2 kinase. The possibility of achieving H1 phosphorylation within chromatin and to determine the level and position of phosphorylation in the protein has been broadly explained and demonstrated in Chapter Three of the present study. Once phosphorylation was confirmed, we studied the effect of phosphorylation on chromatin aggregation.

It is well known that addition of 0.1-1.5mM $MgCl_2$ leads to hierarchical condensation transitions of chromatin arrays[445]. Chromatin, in which the DNA negative charge is not fully neutralized by histones, is self-repulsive in low ionic strength buffers, but that charge repulsion is reduced or eliminated in the presence of divalent and polyvalent cations[446]. At 0.5mM of $MgCl_2$, chromatin arrays achieve their maximal folding into the 30nm fibre compaction and the cooperative oligomerization into higher order polymeric species begins at 0.6-0.7mM $MgCl_2$. At 1.5mM of $MgCl_2$, half-maximal intermolecular self-association of chromatin arrays is achieved[445]. Thus, the presence of Mg^{2+} affects the intramolecular folding of chromatin (condensation) and, increasing the amount of $MgCl_2$ above this concentration leads to intermolecular chromatin self-association (aggregation).

Considering the dramatic effects in changes of conformation that isolated chromatin undergoes in response to the addition of $MgCl_2$, we speculated about the possibility that phosphorylation of histone H1 within chromatin would have an observable effect in this chromatin. Histone H1 is associated with chromatin compaction and aggregation. H1 phosphorylation itself is known to reduce the affinity of H1 for chromatin, relaxing the fibre. This effect of histone H1 phosphorylation on chromatin compaction (intramolecular condensation of chromatin) suggests the possibility of an effect in chromatin aggregation (intermolecular self-association of chromatin). Thus, salt-dependent aggregation of phosphorylated chromatin might occur in a different way: unphosphorylated and more condensed chromatin would be more aggregated than the partially phosphorylated one in the presence of divalent cations.

In order to test this hypothesis, we analysed the induction of chromatin aggregation mediated by an increase in $MgCl_2$ concentration (from 1mM up to 1.6mM $MgCl_2$) of *ex vivo* phosphorylated chromatin by Dynamic Light Scattering (DLS). This technique allows the measurement of particle size distributions within a sample and the detection of the appearance of larger aggregates. Laser Diffraction was also used to determine the presence or absence of larger aggregates above the DLS range.

Initial analysis of *ex vivo* phosphorylated chromatin

Ex vivo phosphorylation of chromatin was carried out at 30°C in a single reaction and samples were taken at 1 hour, 5 hours and overnight and stored at 4°C after phosphorylation. In parallel, an incubation control reaction was made with a second sample of the same chromatin, at 30°C without kinase.

In Chapter Three, we also studied *ex vivo* phosphorylation of chromatin after 15 minutes, 30 minutes and 3 hours of reaction. These samples were not studied in the present Chapter, since the amount of

chromatin per experimental point needed to perform a series of Dynamic Light Scattering analyses with a good signal was high. Because of that, we reduced the number of experimental samples by choosing an intermediate phosphorylated sample (1 hour) and two extensively phosphorylated samples (5 hours and overnight), in which we expected to observe the most important effects of phosphorylation on chromatin aggregation.

Transmission Electron Microscopy (TEM)

Samples of native and *ex vivo* phosphorylated chromatin, together with their corresponding controls were analysed by Transmission Electron Microscopy (TEM) in order to ensure chromatin integrity. Samples were in 5mM TEA, 35mM NaCl plus 1mM $MgCl_2$, which was the initial ionic strenght for all the posterior analysis. Figure 4.1. shows the images obtained of chicken native chromatin and phosphorylated chromatin fragments for 1 hour, 5 hours and overnight.

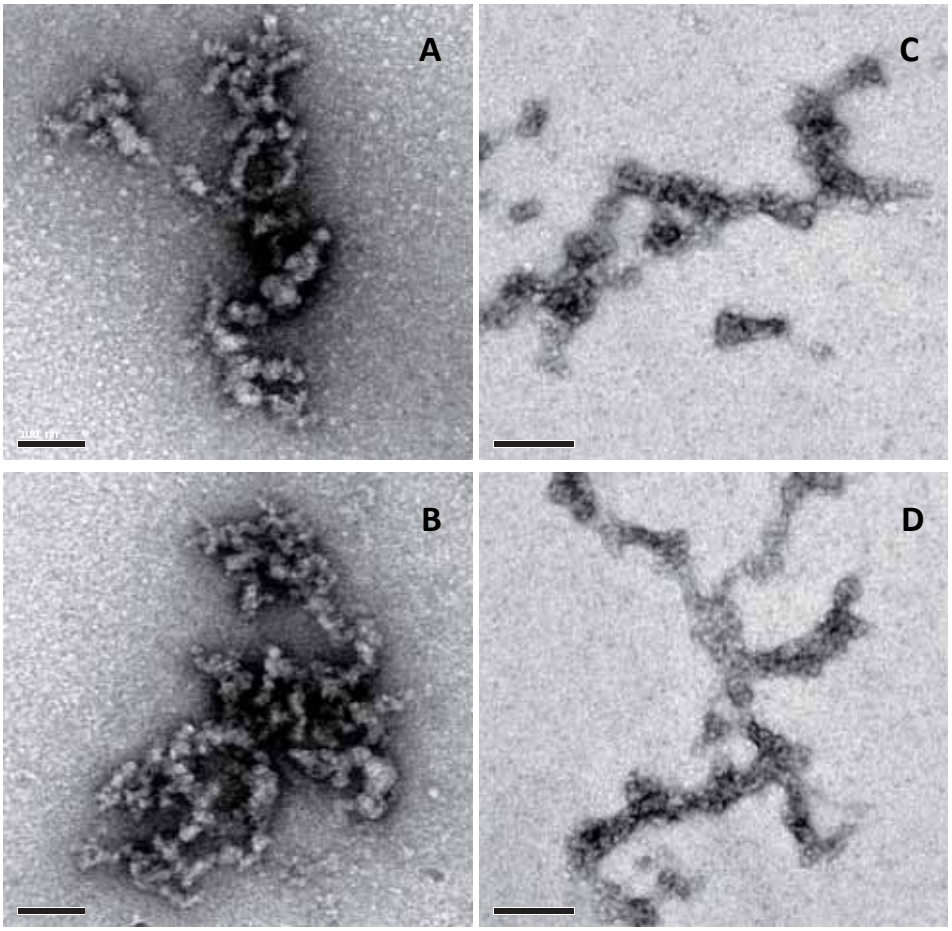


Figure 4.1. Transmission Electron Microscopy (TEM) of *ex vivo* phosphorylated chicken erythrocyte chromatin. (A) Chicken native chromatin; **(B)** 1 hour phosphorylated chromatin; **(C)** 5 hours phosphorylated chromatin; **(D)** overnight phosphorylated chromatin. Chromatin concentration was 2µg/mL in TEA 5mM NaCl 35mM plus 1mM $MgCl_2$. Samples were fixed with 0.1% GTA and laid individually on copper grids covered with a carbon film that was activated by glow-discharge. Negative staining was performed by adding 0.2% uranyl acetate. Visualization of samples was performed in a JEM-1400 microscope. Scale bar: 100nm

At 1mM $MgCl_2$, chromatin samples were in a 30nm-fibre conformation and appeared not to be aggregated. No differences were detected between each phosphorylation sample and the corresponding control (images not shown). No changes were either visualized due to phosphorylation (Figure 4.1.). Chromatin structure was not affected after phosphorylation reaction and, thus, it was considered to be in good conditions for posterior analyses.

Dynamic Light Scattering (DLS)

Chromatin is a polyelectrolyte[447,448] which folds and self-associates in a salt-dependent manner, mainly by the addition of milimolar concentrations of Mg^{2+} and/or higher concentrations of monovalent salt (~140mM NaCl)[449]. Chromatin condensation by Mg^{2+} produces larger aggregates of 30 nm fibres, affects its mobility and increases its sedimentation coefficient[447]. Polynucleosomes and chromatin fibres become less flexible when increasing the ionic strength.

As mentioned above, Dynamic Light Scattering (DLS) was the optical technique used to study the changes in size distributions following chromatin aggregation induced by an increase in the ionic strength ($MgCl_2$) of the medium.

DLS allows measuring the diffusion of particles moving under Brownian motion -defined by the translational diffusion coefficient (D)-, and converts this to size and a size distribution using the Stokes-Einstein relationship. The diameter measured in DLS is the diameter of an equivalent sphere of the particle and is related to how this particle diffuses within a fluid. It is referred to as hydrodynamic diameter (d.nm). The first order result from a DLS experiment is an intensity distribution (%) of particle sizes (d.nm). It allows the detection of particles with sizes comprised between 0.3nm and 10µm. This data can be coverted to a volume distribution (%) according with the Mie theory.

The conformation of macromolecules is usually dependent on the exact nature of the dispersing medium (ie. concentration, ions). If the shape of a particle changes in a way that affects the diffusion speed, then the hydrodynamic size will change. Hence, chromatin aggregation due to progressive addition of divalent Mg^{2+} would lead to a change in the hydrodynamic diameter of individual chromatin molecules. Furthermore, DLS measurements of a sample can show whether the particles aggregate over progressive $MgCl_2$ addition by seeing an increase of the hydrodynamic diameter and the appearance of a population of particles with a larger diameter. Because of that, Dynamic Light Scattering (DLS) measurements were carried out to study potential changes in chicken erythrocyte chromatin aggregation due to phosphorylation.

An important parameter in DLS measurements is the polydispersity index (PdI), which represents the width of the size distribution, which is the measurement of the homogeneity of a dispersion, ranging from 0.0 (monodisperse) to 1.0 (very heterogeneous).

In our case, increasing the $MgCl_2$ concentration of the sample leads to progressive aggregation of chromatin molecules. This results in an increase of turbidity and heterogeneity of the sample due to the appearance of another population of aggregated molecules. PdI values greater than 0.7 indicate the sample has a very broad size distribution and that it is not suitable for the dynamic light scattering (DLS) technique.

The advantage of DLS analyses is that they allow not only to observe a decrease in the amount of soluble chromatin but also to follow the transition towards aggregation and to determine the particle size of both, the soluble chromatin and the aggregates.

In all the DLS analyses performed in this study, chromatin samples were at a final concentration of 0.2mg/mL in Tris 10mM pH 7.0, NaCl 35mM plus 1mM $MgCl_2$. An initial DLS measurement was made in order to determine if chromatin was aggregated under these initial conditions. In all cases, a single peak appeared in the DLS analyses, indicating that chromatin was not aggregated at this $MgCl_2$ concentration.

Table 4.1. shows the DLS data obtained in the analysis of native chromatin (stored at 4°C, not incubated, not phosphorylated), the controls and the phosphorylated chromatin samples. Figure 4.2. shows the DLS scans for native chromatin and the incubation controls at 30°C. No significant differences were observed in the DLS scans of the phosphorylated samples at 1mM $MgCl_2$ (data not shown).

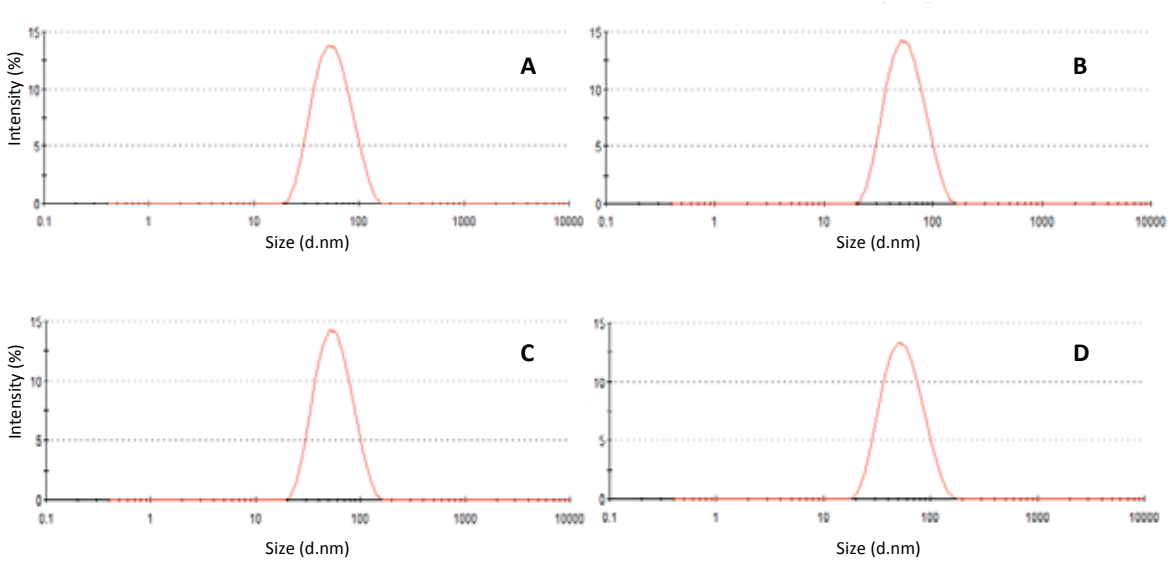


Figure 4.2. Dynamic Light Scattering (DLS) of native chromatin controls at 1mM $MgCl_2$. Hydrodynamic diameter (d.nm) size distribution by intensity (%). Samples were stored at 4°C (A) and incubated for 1 hour (B); 5 hours (C); and overnight (D) under phosphorylation reaction conditions (30°C) in the absence of kinase, as a control of chromatin phosphorylation and chromatin incubation at phosphorylation reaction temperature. Measurements were performed in Tris 10mM NaCl 35mM 1mM $MgCl_2$ at room temperature (RT). Chromatin concentration was 0.2mg/mL. The intensities (%) plotted correspond to one of the three measurements performed. Each measurement report is obtained from 12-15 cumulant scans of each sample.

Intensity distributions of particle sizes were of 100%, consisting of a single peak with hydrodynamic diameter sizes varying between 45 and 49 nm. The polydispersity indexes (PdI) went from 0.147 to 0.267 (Table 4.1.). These values indicated that the samples were monodisperse and that the scattering intensities were due to a single particle size population. The volume distributions (%) also indicated the presence of a single population of molecules in the sample.

Table 4.1. Data obtained from DLS measurements of chicken erythrocyte chromatin at 1mM $MgCl_2$. Native chromatin after isolation (4°C), incubation controls of the phosphorylation reaction (at 30°C, without kinase) after the corresponding time; and *ex vivo* phosphorylated chromatin samples after the corresponding time.

Sample	Polydispersity Index (PdI)	Diameter (d.nm)	% Intensity	% Volume
Native chromatin (4°c)	0.229	46.99±0.4	100	100
Control 30°C – 1h	0.193	45.15±1.0	100	100
Control 30°C – 5h	0.242	48.35±2.1	100	100
Control 30°C – on	0.173	44.72±1.3	100	100
Phosphorylated chromatin – 1h	0.207	48.98±1.5	100	100
Phosphorylated chromatin – 5h	0.178	49.47±1.5	100	100
Phosphorylated chromatin – on	0.148	47.34±0.9	100	100

Data shown corresponds to the average of the three measurements performed. Each measurement report is obtained from 12-15 cumulant scans of each sample. Standard deviation is indicated for the hydrodynamic diameters. Measurements were performed at RT at a final concentration of chromatin of 0.2mg/ml in Tris 10mM pH 7.0, NaCl 35mM and 1mM $MgCl_2$.

Several $MgCl_2$ titration assays were performed between 1mM and 2mM $MgCl_2$ in order to determine the exact $MgCl_2$ concentration in which aggregation of the samples occurred and a second peak, corresponding to the population of aggregated molecules, could be observed. The titration assays determined that increasing the salt concentration up to 1.6mM $MgCl_2$ led to the appearance of aggregates in all the samples. Increasing the salt concentration above 1.6mM $MgCl_2$ led to an increase of turbidity that gave polydispersity index (PdI) values greater than 0.7, which indicated that the samples were not suitable for DLS measurements. For this reason, 1.6mM $MgCl_2$ was the chosen concentration to induce chromatin aggregation and perform the DLS assays. It will be, sometimes here on, referred as critical $MgCl_2$ concentration.

It is important to consider that the phosphorylation reaction occurred at 30°C, whereas chromatin is usually manipulated at low temperatures to preserve its integrity. In general, temperature is known to promote macromolecular aggregation and, since the aim was to observe differential aggregation of chromatin in the presence of Mg^{2+} ions (1.6mM $MgCl_2$) due to phosphorylation, it was really important to ensure that the changes observed in the hydrodynamic diameter of the molecules by increasing the amount of divalent cations were a consequence of phosphorylation and not due to incubation at 30°C.

Because of that, we performed a DLS analysis of native chromatin stored at 4°C and native chromatin after incubation at 30°C for 1 hour. After the initial measurement at 1mM $MgCl_2$ (DLS spectra not shown because of equivalence with data in Table 4.1. and Figure 4.2.; ionic strength was rised up to 1.6mM $MgCl_2$ to induce chromatin aggregation and samples were measured again in the DLS equipment. Figure 4.3. and Table 4.2 show the DLS scans and the DLS data obtained for these two samples.

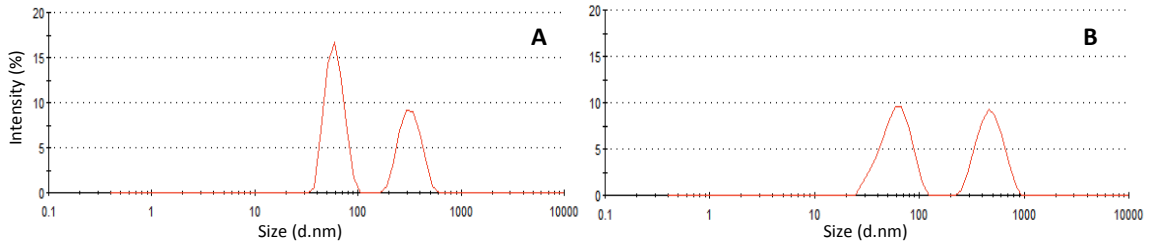


Figure 4.3. Dynamic Light Scattering (DLS) scans in 1.6mM $MgCl_2$ **of chicken native chromatin stored at 4°C (A) and after incubation at 30°C for 1 hour (B). Hydrodynamic diameter (d.nm) size distribution by intensity (%).** Measurements were performed in Tris 10mM NaCl 35mM 1.6mM $MgCl_2$ at room temperature (RT). Chromatin concentration was 0.2mg/mL. The intensities (%) plotted correspond to one of the three measurements performed. Each measurement report is obtained from 12-15 cumulant scans of each sample.

At 1.6mM $MgCl_2$, two peaks were observed in the DLS scans in both chromatin samples. The first peak (Peak 1 in Table 4.2.) was equivalent to the one observed in the measurements at 1mM $MgCl_2$, in which chromatin was not aggregated. The second peak (Peak 2 in Table 4.2.) corresponded to the appearance of a second population of molecules, much larger, due to chromatin aggregation induced by an increase of $MgCl_2$ concentration.

In both samples, the mean value of the hydrodynamic diameter of peak 1 was around 60nm. The most remarkable change between those two samples was the difference in size of the second peak, which corresponds to the aggregated population of molecules. As it can be observed in Table 4.2., the particle size of the second peak in the sample that had been incubated at 30°C was 486nm, whereas in native chromatin stored at 4°C, the second peak had a hydrodynamic diameter of 324nm. Thus, the aggregated molecules were 162nm larger after incubation at 30°C for 1 hour.

Table 4.2. Data obtained from DLS measurements at 1.6mM $MgCl_2$ of chicken erythrocyte native chromatin stored at 4°C and after incubation at 30°C for 1 hour.

Native chromatin sample	PdI	Peak	Diameter (d.nm)	% Intensity	% Volume
Stored at 4°C	0.365	Peak 1	60.09±1.4	60.2	90.0
		Peak 2	324.3±5.2	39.8	10.0
Incubation control (1h at 30°C)	0.496	Peak 1	60.82±1.2	53.9	85.3
		Peak 2	486.3±17.3	46.1	14.7

Data shown corresponds to the average of three measurements per sample, obtained from the cumulant data of 12-15 scans each. Standard deviation is indicated for the hydrodynamic diameters. Measurements were performed at RT at a final concentration of chromatin of 0.2mg/ml in Tris 10mM pH 7.0, NaCl 35mM and 1.6mM $MgCl_2$.

The intensity distributions were also different between those samples (Figure 4.3). In the DLS scan of native chromatin stored at 4°C (Figure 4.3, A), the maximum intensity of the first peak was higher than that of the second peak, whereas in the DLS scan of the incubation control at 30°C for 1 hour (Figure 4.3, B), both peaks had similar maximum intensities. Intensity distribution of the population of aggregated molecules (peak 2) increased a 6% after 1 hour of incubation at 30°C (Table 4.3).

Polydispersity values (PdI) were <0.7, indicating that the measurements were in range despite the presence of aggregates. Posterior laser diffraction measurements (Mastersizer 2000, Malvern Instruments) confirmed the absence of larger aggregates at this ionic strength in the samples (data not shown).

These results indicate that incubation at 30°C for 1 hour promoted by chromatin aggregation at a given $MgCl_2$ concentration. Incubation of chromatin at 30°C led to the formation of larger aggregates (diameter), together with an increase of their amount (intensity distribution). Hence, chromatin aggregation depends of both the presence of divalent cations (Mg^{2+}) in the medium and of the temperature of incubation. Consequently, for each phosphorylation time, a control of unphosphorylated chromatin was made with a second sample incubated at 30°C without kinase.

After each DLS measurement, samples were analysed by Laser Difraccion (Mastersizer 2000, Malvern Instruments) in order to ensure that there were not larger aggregates out of the DLS range.

Effects of histone H1 phosphorylation on chromatin aggregation

Preparation of chromatin fragments from chicken erythrocyte nuclei by micrococcal nuclease digestion led to chromatin fragments with heterogeneous sizes within a range in each experiment. All the chromatin samples were obtained from the same chicken erythrocytes' nuclei pool.

We considered that the size of the chromatin fragments could have an important effect in their aggregation capacity and, because of that; we analysed chromatin samples of different sizes by using the same experimental procedure.

Chromatin was *ex vivo* phosphorylated for 1h, 5h and overnight. A control reaction was made with a second sample of the same chromatin. All the samples were prepared as mentioned before and an initial DLS measurement at 1mM of $MgCl_2$ was performed to ensure that chromatin was not aggregated under this conditions. After this first measurement, $MgCl_2$ concentration was increased up to 1.6mM $MgCl_2$ to induce chromatin aggregation and to determine if there were any differences that could be associated to linker histones phosphorylation.

After the DLS measurement at 1.6mM $MgCl_2$, all the samples were also analysed by laser diffraction (Mastersizer 2000, Malvern Instruments) to determine if there were any larger aggregates out of the range of the DLS equipment. This verification was performed even when the PdI values in the DLS analyses were in range (PdI<0.7, d(nm)> 10µm).

Effects of histone H1 phosphorylation in the aggregation of short chromatin fragments

Short chromatin fragments were obtained by micrococcal nuclease digestion and a DNA gel electrophoresis was run after treatment with proteinase K in order to determine their range of sizes. As it can be seen in Figure 4.4., micrococcal digestion led to chromatin fragments ≥600bp and most of them were around 1500bp, which corresponds to ~7.5 nucleosomes (200bp).

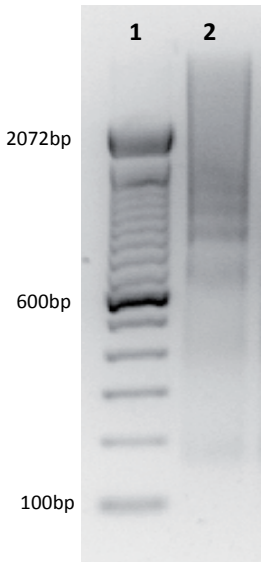


Figure 4.4. Agarose gel electrophoresis of DNA extracted from chicken erythrocyte native chromatin. (1) 100bp DNA ladder (Invitrogene); **(2)** DNA from chicken erythrocyte chromatin fragments after proteinase K digestion. 1% agarose gel stained with ethidium bromide.

100bp DNA Ladder consists of 15 blunt-ended fragments between 100bp and 1500 bp in multiples of 100bp and an additional fragment at 2072 bp.

An initial DLS measurement at 1mM of $MgCl_2$ was made for all the samples in order to determine that chromatin was not aggregated. The results obtained were equivalent to those shown in Figure 4.2**.** To induce chromatin aggregation, $MgCl_2$ concentration was increased up to 1.6mM $MgCl_2$ in both, the incubation controls (unphosphorylated chromatin) and the phosphorylated samples, and they were measured again by DLS.

Figure 4.5. shows the DLS scans of the unphosphorylated (controls) and the phosphorylated samples obtained in 1.6mM $MgCl_2$. Two peaks could be observed in these scans, the first one had a size distribution in which the hydrodynamic diameter size varied from 48.7(±3.09)nm to 53.7(±1.1)nm, depending on the sample (Table 4.3., Peak 1). As already explained, this peak corresponds to the same one already observed in the DLS measurements at 1mM $MgCl_2$, in which chromatin was not aggregated (Figure 4.2.). The second peak, which corresponds to chromatin aggregates, had a diameter size between 171.9(±2.1) and 304.5(±34.6) (Table 4.3., Peak 2), much larger than the first one.

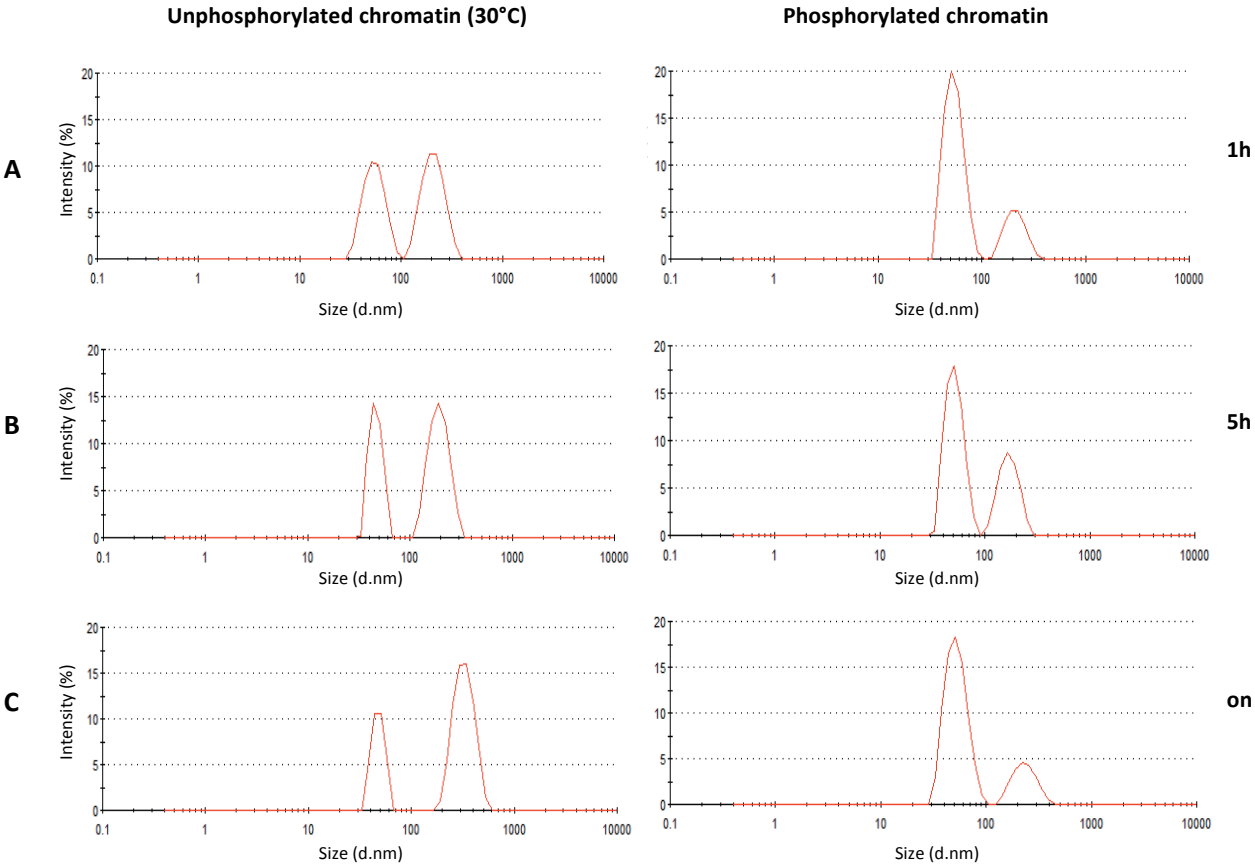


Figure 4.5. Dynamic Light Scattering (DLS) of native chromatin controls (left) and phosphorylated chromatin (right) of short chromatin fragments. Hydrodynamic diameter size distribution (d.nm) by intensity (%). Samples correspond to native chromatin incubated for 1 hour (A); 5 hours (B); and overnight (on) (C) under phosphorylation reaction conditions (30°C); in the absence (left panels) or presence (right panels) of CDK2 kinase. Measurements were performed in Tris 10mM NaCl 35mM 1.6mM $MgCl_2$ at room temperature (RT). Chromatin concentration was 0.2mg/mL. The intensities (%) plotted correspond to one of the three measurements performed. Each measurement report is obtained from 12-15 cumulant scans of each sample.

The intensity distributions (%) were significantly different when comparing the incubation control samples with the phosphorylated chromatin ones (Figure 4.5.; Table 4.3.). The controls, in which chromatin had been incubated under phosphorylation conditions but in the absence of kinase - and were, consequently, not phosphorylated- the intensity distribution of both peak 1 and peak 2 was quite similar in all the incubation times (1h, 5h and overnight). Intensities for the first peak in the controls varied from 46.8% to 33.2%, whereas the intensities of the second peak varied from 53.2% to 66.8%. On the contrary, the intensity distribution values obtained for phosphorylated chromatin varied from 79.8% to 65.2% for the first peak and from 21% to 34.4% for the second one. Moreover, the values of volume distribution (%), which describe the relative proportion of the different components or populations in the sample, confirmed the observed by intensity.

Table 4.3. Data obtained from DLS measurements of chicken erythrocyte short chromatin fragments. Uphosphorylated samples (incubation controls at 30°C) and *ex vivo* phosphorylated for 1h, 5h and overnight (on).

Sample		PdI	Peak	Diameter (d.nm)	% Intensity	% Volume
1h	**unphosphorylated chromatin (control at 30°C)**	0.404	Peak 1	53.7±1.1	46.8	90.7
			Peak 2	200.1±18.1*	53.2	9.3
	phosphorylated chromatin	0.241	Peak 1	53.2±2.5	78.0	97.4
			Peak 2	201.6±12.5*	22.0	2.6
	Δ peaks		**Peak 1**	**0.5**	**-31.2**	**-6.7**
			Peak 2	**-1.7**	**31.2**	**6.7**
5h	**unphosphorylated chromatin (control at 30°C)**	0.397	Peak 1	48.7±3.09	40.2	91.1
			Peak 2	194.6±0.07*	59.8	8.9
	phosphorylated chromatin	0.254	Peak 1	50.1±2.4	65.2	95.6
			Peak 2	171.9±2.1*	34.8	4.4
	Δ peaks		**Peak 1**	**-1.4**	**-25.0**	**-4.5**
			Peak 2	**22.7**	**25.0**	**4.5**
on	**unphosphorylated chromatin (control at 30°C)**	0.488	Peak 1	49.1±1.8	33.2	83.2
			Peak 2	304.5±34.6*	66.8	16.8
	phosphorylated chromatin	0.210	Peak 1	52.8±2.2	78.9	97.8
			Peak 2	235.9±16.7*	21.1	2.2
	Δ peaks		**Peak 1**	**-3.7**	**-45.5**	**-14.6**
			Peak 2	**68.6**	**45.7**	**14.6**

Data shown corresponds to the average of three measurements per sample, obtained from the cumulant data of 12-15 scans each. Standard deviation is indicated for the hydrodynamic diameters. Measurements were performed at RT at a final concentration of chromatin of 0.2mg/ml in Tris 10mM pH 7.0, NaCl 35mM and 1.6mM $MgCl_2$. The differences between the experimental values obtained in peak size, intensity and volume of the incubation control and the phosphorylated chromatin were calculated for each time.

Δ peaks correspond to the difference of the experimental values obtained in the control samples minus the ones obtained for the phosphorylated samples in each case.
The asterisks (*) indicate the values represented in Figure 4.6.

For every incubation/phosphorylation pair, the differences between the incubation control and the phosphorylated sample were calculated for peak 1 and peak 2. As it can be observed in Table 4.3., the intensity distributions (%) of the second peak were always higher in the incubation controls, indicating that those were more aggregated than the phosphorylated samples. Moreover, this difference became more remarkable with the increase of the incubation/phosphorylation time. After 1 hour of reaction, the difference in intensity of the second peak between the control and the phosphorylated sample was around 31% but it increased up to 46% after overnight incubation. This change led to a simultaneous decrease in the amount of non-aggregated molecules (peak 1). This could also be observed in terms of volume distribution.

The greater differences were observed in the hydrodynamic diameters (d.nm) between the unphosphorylated chromatin (control) and the corresponding phosphorylated sample. The sizes of the particles of the first population of molecules (peak 1) did not vary significantly between these two samples for none of the reaction periods (Table 4.3.). In contrast, the diameter sizes of the second peak were higher in the controls than in the phosphorylated chromatin, indicating that the aggregates were larger. Figure 4.6. shows the average size of the aggregates (peak 2) in each sample. Note that the

difference in size of the aggregates between the controls and the phosphorylated samples became more remarkable with the increase of the incubation/phosphorylation time. After overnight incubation, the particles of the aggregated population were ~69nm larger in the incubation control than in the overnight phosphorylated chromatin (Table 4.3.).

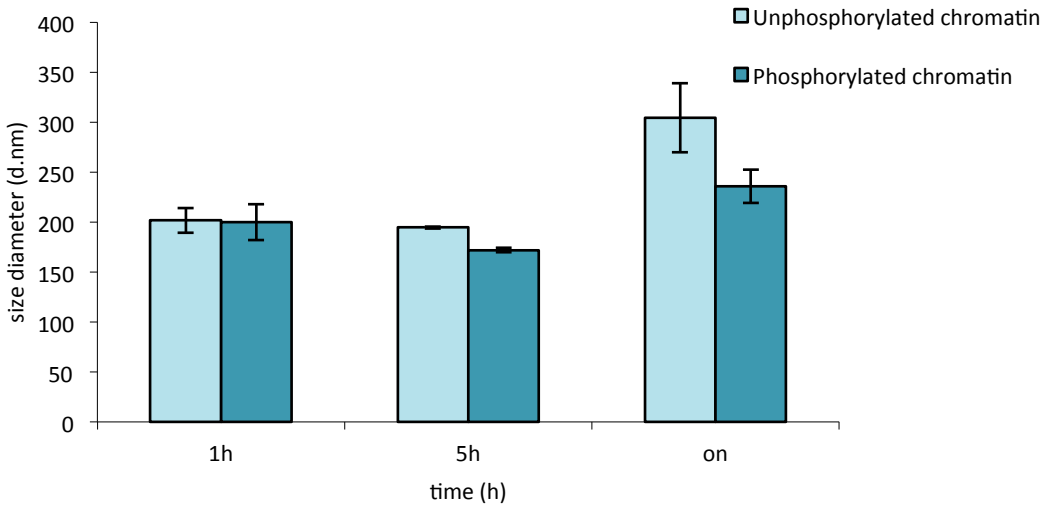


Figure 4.6. Average of the hydrodinamic diameter (nm) of the aggregated population (peak 2/peak 3) of each incubation/phosphorylation time of short chromatin fragments. For each time, data for the unphosphorylated chromatin samples (incubation controls) is shown in light blue and data for the phosphorylated chromatin samples appears in dark blue. Data shown corresponds to the average of three measurements per sample, obtained from the cumulant data of 12-15 scans each. Standard deviation of sizes is shown with a statistics bar in each case.

The PdI values were in range in all the measurements performed (<0.7), but note that these were higher in the DLS reports obtained in the control samples (~0.425) than in the phosphorylated chromatin (~0.235), indicating that the turbidity of the sample was higher than in the phosphorylated chromatin. This feature was also observable during the performance of the experiments: when increasing the ionic strength in the samples before the DLS measurement, the phosphorylated ones remained clear, whereas the controls became slightly whitish. After the DLS measurement, samples were analysed in the Mastersizer equipment (data not shown). No signal was detected in none of the samples, indicating that there were not larger aggregates than the ones detected in the DLS analyser.

Several considerations could be extracted from all these results. First, that the appearance of the second peak was a consequence of the addition of $MgCl_2$ to the samples, which induced the aggregation of soluble chromatin. Self-association of soluble chromatin molecules with hydrodynamic diameters around 50nm (Table 3.1.) defined by peak 1 caused, thus, the appearance of a second population of larger molecules that resolved in the second peak. Nevertheless, the results indicated that the tendency towards aggregation was lower in the phosphorylated chromatin with respect to that observed in the unphosphorylated ones (incubation controls).

The results showed that phosphorylated samples were not as aggregated as the unphosphorylated ones and, moreover, that the aggregates were smaller after phosphorylation. Hence, these results indicated that phosphorylation impaired chromatin aggregation at 1.6mM $MgCl_2$.

Effects of histone H1 phosphorylation in the aggregation of large chromatin fragments

Large chromatin fragments from chicken erythrocyte nuclei were also obtained by micrococcal nuclease digestion. A DNA gel electrophoresis was run after their purification and treatment with proteinase K to determine the sizes of the fragments. In this case, micrococcal digestion led to chromatin fragments >1000bp (Figure 4.7.). Most of the fragments had sizes above 8576bp, which corresponds to, at least, 43 nucleosomes.

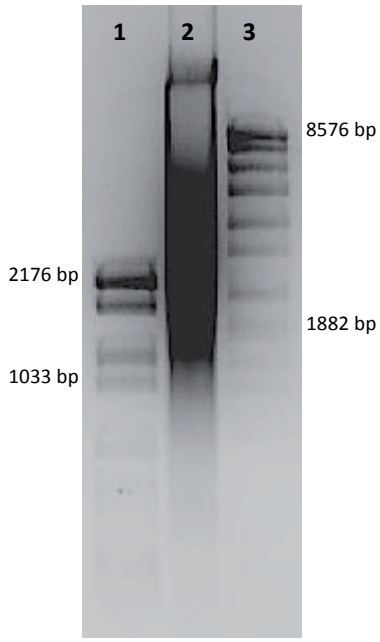


Figure 4.7. Agarose gel electrophoresis of DNA extracted from chicken erythrocyte native chromatin. (1) DNA molecular weight marker VI (Roche); **(2)** DNA of chicken erythrocyte native chromatin fragments after proteinase K digestion; **(3)** DNA molecular weight marker VII (Roche). 1% agarose gel stained with ethidium bromide.

DNA molecular weight marker VI (Roche): Fragment mixture prepared by cleavage of pBR328 DNA with *Bgl* I, and pBR328 DNA with *Hin*f I. 12 fragments: 154; 220; 234; 298; 394; 453; 517; 653; 1033; 1230; 1766; and 2176 bp
DNA molecular weight marker VII (Roche): Fragment mixture prepared by cleavage of SPP1 DNA with *Eco*R I. 17 fragments: 81; 359; 492; 710; 718; 992; 1164; 1482; 1515; 1882; 1953; 2799; 3639; 4899; 6106; 7427, 8576 bp

After *ex vivo* phosphorylation and sample preparation as previously described, the initial DLS measurement at 1mM of $MgCl_2$ was performed. Large chromatin fragments were neither aggregated under these conditions. Afterwards, salt concentration was increased up to 1.6mM $MgCl_2$ to induce chromatin aggregation and the samples were analysed by DLS.

Figure 4.8. shows the DLS scans of the controls and the phosphorylated samples obtained at 1.6mM $MgCl_2$.

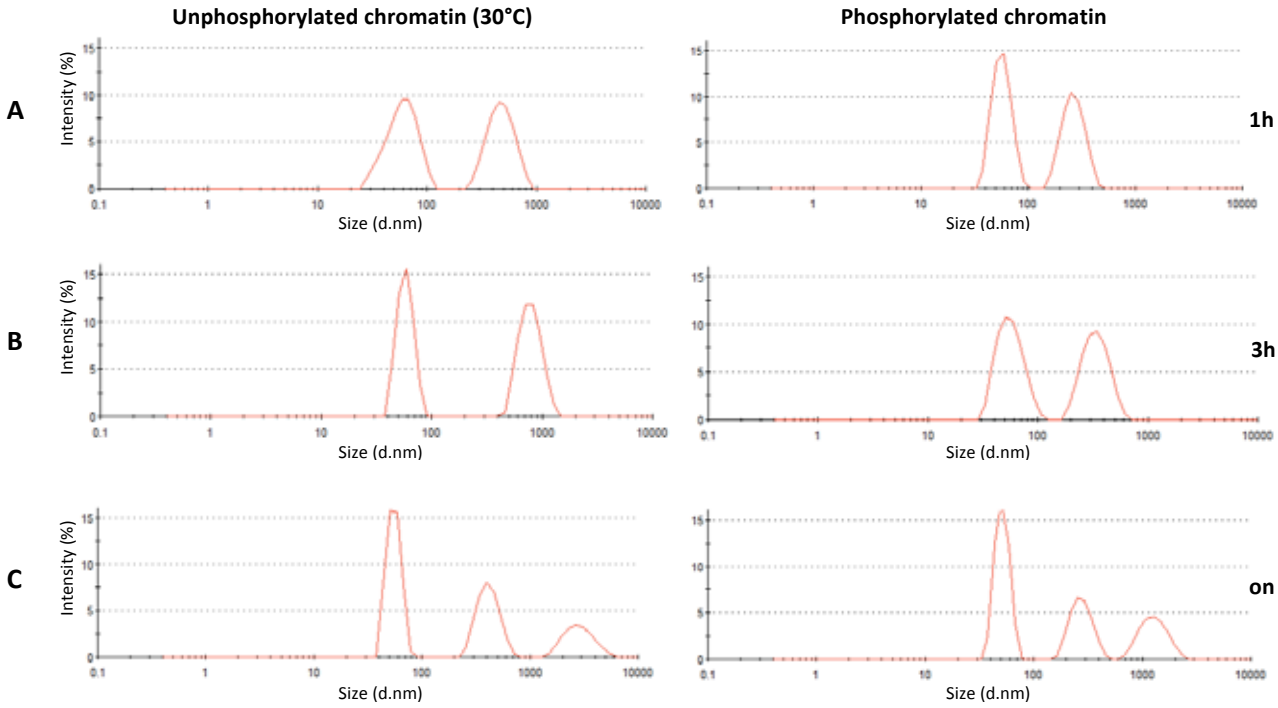


Figure 4.8. Dynamic Light Scattering (DLS) of native chromatin controls (left) and phosphorylated chromatin (right) of large chromatin fragments. Hydrodynamic diameter size distribution (d.nm) by intensity (%). Samples were incubated for 1 hour (A); 5 hours (B); and overnight (C) under phosphorylation reaction conditions (30°C); in the absence (left panels) or presence (right panels) of CDK2 kinase. Measurements were performed in Tris 10mM NaCl 35mM 1.6mM $MgCl_2$ at room temperature (RT). Chromatin concentration was 0.2mg/mL. The intensities (%) plotted correspond to one of the three measurements performed. Each measurement report is obtained from 12-15 cumulant scans of each sample.

In large chromatin fragments, the increase in the ionic strength also led to the appearance of a second peak in the DLS spectra due to the aggregation of chromatin fragments in all the samples. A third peak of even larger aggregates appeared in the samples that had been incubated/phosphorylated overnight.

Table 4.4. shows the DLS data obtained in this experiment. Despite the profiles of the intensity distributions of DLS scans appeared to be quite different between the samples (Figure 4.8.), note that the intensity percentages of each peak did not vary significantly between each pair of samples (Table 4.4). This was also observed in terms of volume distribution.

The most significant changes were observed in the diameter of the peaks, as it already happened with short chromatin fragments. In all the samples, the hydrodynamic diameter of the first peak, which corresponded to the population of molecules that were not aggregated (Table 4.4., peak 1), varied between 50.2(±0.5)nm and 58.1(±1.4)nm. The values obtained were equivalent to those in the DLS measurements in 1mM $MgCl_2$ (data not shown), where a single population is present. Note that the size values obtained in this case were slightly higher than the ones in the experiment performed with shorter chromatin fragments (Table 4.3.).

Table 4.4. Data obtained from DLS measurements of chicken erythrocyte large chromatin fragments. Unphosphorylated samples (incubation controls at 30°C) and *ex vivo* phosphorylated for 1h, 3h, 5h and overnight (on).

Sample		PdI	Peak	Diameter (d.nm)	% Intensity	% Volume
1h	unphosphorylated chromatin (control at 30°C)	0.496	Peak 1	56.9±5.5	53.9	85.3
			Peak 2	482.6±5.16*	46.1	14.7
	phosphorylated chromatin	0.380	Peak 1	56.7±1.5	55.0	90.6
			Peak 2	255.9±19.16*	45.0	9.4
	Δ peaks		**Peak 1**	**0.2**	**-1.1**	**-5.3**
			Peak 2	**226.7**	**1.1**	**5.3**
5h	unphosphorylated chromatin (control at 30°C)	0.667	Peak 1	51.5±1.0	51.5	36.3
			Peak 2	795.5±51.9*	48.5	63.3
	phosphorylated chromatin	0.476	Peak 1	58.1±1.4	53.4	89.6
			Peak 2	344.6±1.1*	46.6	10.4
	Δ peaks		**Peak 1**	**-6.6**	**-1.9**	**-53.3**
			Peak 2	**450.9**	**1.9**	**53.3**
on	unphosphorylated chromatin (control at 30°C)	0.690	Peak 1	53.5±2.4	47.7	64.4
			Peak 2	285.9±10.4	33.3	8.7
			Peak 3	1375±93.34*	18.9	26.9
	phosphorylated chromatin	0.683	Peak 1	50.2±0.5	47.9	69.6
			Peak 2	127.3±18.74	27.8	4.2
			Peak 3	721±12.1*	24.2	26.2
	Δ peaks		**Peak 1**	**3.3**	**-0.2**	**-5.2**
			Peak 2	**158.6**	**5.5**	**4.5**
			Peak 3	**654**	**-5.3**	**0.7**

Data shown corresponds to the average of three measurements per sample, obtained from the cumulant data of 12-15 scans each. Standard deviation of hydrodynamic diameters is indicated in each case. Measurements were performed at RT at a final concentration of chromatin of 0.2mg/ml in Tris 10mM pH 7.0, NaCl 35mM and 1.6mM $MgCl_2$. The differences between the experimental values obtained in peak size, intensity and volume of the incubation control and the phosphorylated chromatin were calculated for each time.

Δ peaks correspond to the difference of the experimental values obtained in the control samples minus the ones obtained for the phosphorylated samples in each case.
The asterisks (*) indicate the values represented in Figure 4.9.

The hydrodynamic diameters obtained for the second peak were more variable between the different samples. The mean values were, in general, much higher than the ones obtained with shorter chromatin fragments. In all pairs of samples, the aggregated population (peak 2) had higher sizes in the control than in the phosphorylated sample (Table 4.4.), as it already happened with shorter fragments (Table 4.3.). At every incubation/phosphorylation time, the differences between the control and the phosphorylated sample were calculated for peak 1 and peak 2. As it can be observed in Table 4.4., the differences in size (nm) of the aggregated molecules became greater with the increase of the incubation/phosphorylation time (Table 4.4.). After 1 hour, the aggregated population of the phosphorylated chromatin had a mean size for the hydrodynamic diameter of 255.9(±19.16)nm whereas in its incubation control, the mean size was 482.6(±5.16)nm. Thus, the aggregated population in the control was around 230nm larger than in the phosphorylated sample (Δpeaks). After 5 hours, the second peak of the unphoshorylated control increased up to 795.5(±51.9)nm of size, whereas the phosphorylated sample only increased up to 344.6(±1.1)nm. At this point, the difference (Δpeaks)

between those peaks was of ~451nm. After overnight phosphorylation, a third peak appeared in both the unphoshorylated and the phosphorylated sample. The second peak measured 285.9(±10.4)nm in the control and 127.3(±18.74)nm in the phosphorylated sample and the third peak was 1375(±93.34)nm and 721.0(±12.1)nm, respectively. The difference between the peaks (Δpeaks) of the most aggregated population (peak 3) for this pair of samples was around 654nm.

Figure 4.9. represents the variation in the average size of the most aggregated population of each pair of samples. Data used in this graphic is marked with an asterisk (*) in Table 4.4. As it already happened for short chromatin fragments, the difference in size of the aggregates between the controls and the phosphorylated samples became more remarkable with the increase of the incubation/phosphorylation time.

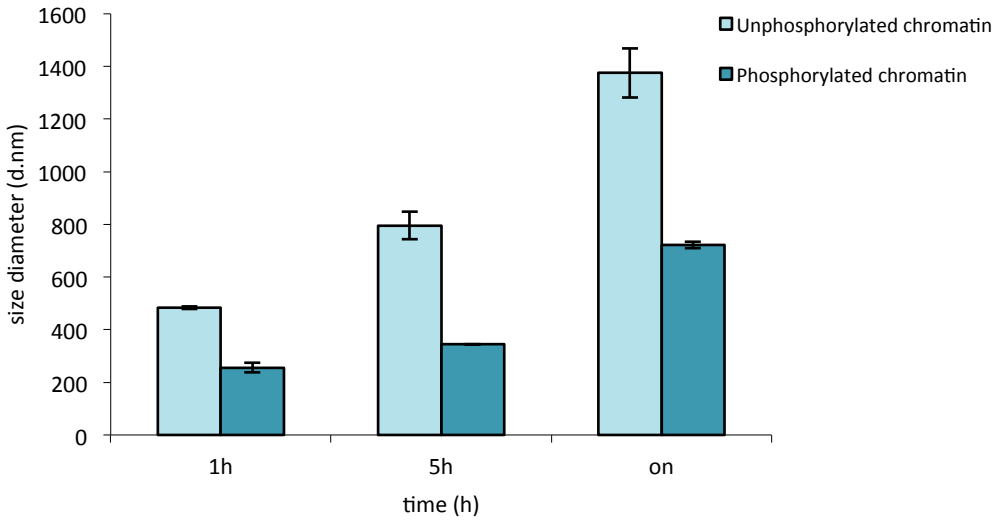


Figure 4.9. Average of the hydrodinamic diameter (nm) of the aggregated population (peak 2/ peak 3) for each incubation/phosphorylation time of large choramotin fragments. For each time, data for the unphosphorylated chromatin samples (incubation controls) is shown in light blue and data for the phosphorylated chromatin samples appears in dark blue. Data shown corresponds to the average of three measurements per sample, obtained from the cumulant data of 12-15 scans each. Standard deviation of sizes is shown with a statistics bar in each case.

The appearance of a third peak in the overnight sample seemed to indicate that the aggregates of the population described by peak 2 in the samples after 5 hours of incubation/phosphorylation progressively aggregated between them, leading to the appearance of a new population of larger aggregates (peak 3). Size values of peak 2 and 3 of the overnight sample were below and above of the values obtained for peak 2 after 5 hours (Table 4.4.). However, the aggregated molecules continued to be larger in the incubation control than in the phosphorylated chromatin (Figure 4.9.).

The PdI values were in range in all the measurements performed (<0.7), but they were close to the limit value (Table 4.4.), indicating that the turbidity of the sample was high due to the presence of higher aggregates and more polydisperse samples than the ones obtained for shorter fragments. Because of that, posterior measurement of the samples by laser diffraction (Mastersizer 2000) was of great importance in order to determine if there were larger particles out of the DLS range. Mastersizer measurement showed the presence of a little proportion of larger aggregates above the measurable DLS range, which justified the high PdI values obtained in these analyses.

The results led, once more, to the conclusion that phosphorylation of chromatin impaired the aggregation of chromatin fragments, even when those were larger: phosphorylated chromatin not only formed smaller aggregates than the incubation controls, but they also showed up more progressively with the increase of the time of reaction.

Moreover, the data clearly suggest that there is an effect of the size of the initial chromatin fragments used in the experiment. Under the same experimental conditions (1.6mM $MgCl_2$), larger chromatin fragments led to larger aggregates than those observed with shorter chromatin. For short chromatin fragments, the maximum difference between the unphosphorylated and the phosphorylated sample of the most aggregated population was of ~69nm (overnight sample), whereas in larger chromatin fragments this difference was of ~654nm. Thus, the effect observed due to chromatin phosphorylation was intensified by the length of the initial chromatin fragments.

In summary, DLS measurements at 1.6mM $MgCl_2$ showed that increasing the ionic strength of the medium led to the differential aggregation of phosphorylated and unphosphorylated chromatin fragments. The most remarkable difference associated to *ex vivo* phosphorylation of linker histones within chromatin was a decrease in the hydrodynamic diameter of the aggregated. These differences became greater with the increase of phosphorylation time and with the size of the chromatin fragments. For short chromatin fragments, the intensity distributions (%) of the aggregates were also higher in the unphosphorylated chromatin than in the phosphorylated ones, indicating that those were more aggregated. For large chromatin fragments, the DLS measurements showed the appearance of a third peak of even larger aggregates in the samples that had been incubated/phosphorylated overnight.

General Discussion

Folding and fibrillation of histone $H1^0$

Histone H1 consists of three structural domains: a central and globular domain flanked by the N- and C-terminal domains, which are known to be unfolded in aqueous solution at physiological salt concentration. The C-terminal domain represents the 52% of the entire protein and it is the primary determinant of histone H1 binding to chromatin *in vivo*. Despite being the most variable region among different linker histones in terms of sequence, the net charge is quite uniform along this lysine-rich domain. The C-terminal domain is considered an intrinsically disordered protein[235], since it folds cooperatively upon interaction with DNA into a highly stable structure[235,306]. This stability is supposed to be due to the compensation of positive charges in the protein when interacting with the negative phosphate backbone of DNA, which would cause an increase on hydrophobicity. Phosphorylation of the CTD in its SPKK motifs is the most common PTM affecting this domain and it plays an important role in many structural and functional aspects of H1, having a direct effect in the structure of the CTD and its interaction with DNA[266].

In the present study, we have studied the effects of phosphorylation and DNA binding in the secondary structure of the entire protein ($H1^0$) by infrared spectroscopy in order to asses the folding of the CTD within the entire protein.

In the FTIR analyses of histone $H1^0$ in aqueous solution, the full-length protein showed a percentage in α-helix (23%) that could be attributed to the globular domain, which is known to be structured in a three-helix bundle and a β-hairpin[414], with an estimated content of 19% of α-helix. The amount of β-structure significantly increased (21%) with respect to that estimated from the sum of the percentages of the isolated domains (13%). The increase in the content of β-structure of the entire protein could be contributed mainly by the C-terminal domain, since the N-terminal domain is too short to explain this difference and the amount of this structure element is well-defined in the globular domain. The isolated CTD has little structure in aqueous solution (7% β-structure) and, thus, is the only of the three domains capable of folding in β-structure. The increase of β-structure in the entire H1 was followed by a decrease in the amount of flexible regions, in comparison with the estimated values. It is important to consider that at physiological salt, the isolated CTD is unfolded and, hence, the differences observed in the percentage of the mentioned structure elements may be due to the structural coupling between the CTD and the globular domain which would stabilise the β-structure within the CTD.

Phosphorylation of histone $H1^0$ by CDK2 kinase occurs in its three SPKK motifs (S/T-P-X-K/R), which are located in the CTD. The addition of a phosphate group decreases the positive net charge, which can induce a conformational change in the protein. Previous FTIR studies with the isolated CTD[266] showed that phosphorylation of this domain did not significantly affect the secondary structure when free in solution and only a slight increase in β-structure was observed due to phosphorylation. The FTIR analyses of the entire and fully phosphorylated histone $H1^0$ showed that phosphorylation induced an increase of β-structure in aqueous solution from 21% in the unphosphorylated protein up to a considerable 35% in the phosphorylated protein. This structural change was also observed in previous studies with H1e subtype[267] and indicates that phosphorylation causes a conformational change in the CTD stabilised by the globular domain, which induces the acquisition of β-structure in histone $H1^0$.

Interaction with DNA induced the folding of the isolated CTD, which increased its proportion of α-helix. These results have been observed for the C-terminal domains of H1e, $H1^0$ and H1t[235, 267], and also for the entire histone H1e. FTIR analysis of the entire histone $H1^0$ bound to DNA also showed that the binding of the unphosphorylated protein to DNA induced an increase in the proportion of α-helix from a 23% in the protein free in solution to a 34%. These results confirmed that DNA binding to the unphosphorylated

$H1^0$ induced its folding in α-helix conformation, as it happened with the isolated CTDs and the entire H1e. This would be a consequence of the folding of the terminal domains upon DNA interaction and, thus, $H1^0$ can also be classified as an intrinsically disordered protein.

As mentioned before, different linker histone subtypes fold in the same way upon interaction with DNA and, thus, this structural change could be a common feature for different linker histone subtypes, which suggests that intrinsic disorder could be determinant in the regulation of linker histone function and chromatin structure. This idea is supported by intramolecular fluorescent resonance energy transfer (FRET) studies, which indicate that the CTD is unstructured in aqueous solution and undergoes nucleosome-directed folding, indicating that the H1 CTD is an intrinsically disordered domain[306].

Previous studies showed that the H1 CTD does not fold cooperatively and has the characteristics of the molten globule state[450] in the presence of crowding agents, although it becomes cooperatively folded upon interaction with DNA[235]. The similarity of the secondary structure motifs of the CTD in crowded conditions and bound to DNA suggests that the intermediate molten globule state could help to increase the rate of the transition toward the native DNA-bound, kinetically favouring H1 binding. This transition may be driven by increased hydrophobicity after charge compensation of the positive charge of the abundant Lys residues of the C-terminus by the DNA phosphates[451]. The H1 globular domain is thought to be responsible for the location of H1 by interacting with a specific site in the nucleosome. The dependence of the tightening of the CTD intermediate state in a molten globule structure on charge compensation may afford a mechanism to avoid the interference of the strong electrostatic binding of the C-terminus with binding site recognition by the globular domain because binding of the C-terminus to unspecific sites, such as nucleosomal DNA, already partially charge-compensated by core histones, may lead to partial folding and lower affinity[451].

DNA binding of the phosphorylated $H1^0$ considerably increased the amount of β-structure in contrast with the results obtained for unphosphorylated $H1^0$ bound to DNA. In this case, there was an effect of the protein/DNA ratio and the proportion of β-structure increased up to a 54% when r(w/w)=0.7. This increase was accompanied by a significant decrease in the amount of α-helix, from 34% to 15% and the amount of random coil and open loops decreased a 10% with respect to the unphosphorylated $H1^0$ bound to DNA.

As mentioned above, the CTD corresponds to a 52% of the entire protein. Nonetheless, the β-structure increased up to a 54% in the phosphorylated protein bound to DNA, which is higher than the contribution of the CTD. This fact indicates that phosphorylation, coupled with DNA binding, induces the folding of the CTD of histone $H1^0$ in an all-β conformation within the entire protein. Moreover, the concomitant decrease in α-helix could be explained by the perturbation of the globular domain, which would increase its proportion of β-structure due to the propagation of this secondary structure element from the CTD.

These results clearly resembled those obtained for the isolated CTDs[235,267] and indicate that the observed structural change is due to phosphorylation. Similar results were obtained in previous studies with H1e. When unphosphorylated, binding to DNA induced an increase of α-helix in H1e, whereas phosphorylated H1e bound to DNA increased twice its amount of β-structure, at the expense of the amount of α-helix. As it happened with $H1^0$, phosphorylation of the protein in solution only increased the percentage of β-structure, without affecting that of α-helix.

The folding of the isolated CTD and the entire protein due to DNA binding and/or phosphorylation may involve some degree of charge neutralization, leading to a less polar environment in the protein that would favour its folding. We also performed FTIR spectroscopic analyses of unphosphorylated and

phosphorylated histone $H1^0$ in the presence of the anionic detergent SDS, considering that the interaction with hydrophobic ligands could also trigger the folding of the protein as efficiently as DNA does. Due to its physico-chemical properties, SDS is capable of establishing electrostatic and hydrophobic interactions.

FTIR analysis of the unphosphrorylated $H1^0$ in the presence of SDS at a molar ratio 14:1 (SDS/$H1^0$) showed a 30% of α-helix and a 21% of β-structure. These proportions on secondary structure were similar to those obtained for the unphosphorylated protein bound to DNA, indicating that anionic SDS is capable of mimicking the environment found in the complexes with DNA.

At a molar ratio of 14:1 SDS/ $H1^0$3P, interaction with anionic SDS led to an increase of β-structure up to 55% in the triphosphorylated $H1^0$, accompanied by a decrease in α-helix up to 13%. Thus, phosphorylated $H1^0$ showed similar structural changes to those already observed in the binding with DNA at r(w/w)=0.7. The CTD corresponds to the 52% of the protein. Thus, the increase of β-structure to a 55% in the presence of anionic SDS suggest that the CTD of phosphorylated histone $H1^0$ was in an all-β conformation and that the β-structure had propagated into the globular domain, as already happened in the interaction with DNA.

Studies carried out on $H1^0$ in the presence of anionic SDS were based on previous FTIR analyses of the isolated CTD in the presence of neutral and anionic detergents at physiological salt concentration. The unphosphorylated CTD in the presence of neutral detergents Brij 35 and Triton X-100 and anionic SDS at a molar ratio 1:1, became extensively folded upon interaction, with proportions of secondary structure elements similar to those found in the complexes of the isolated CTD with DNA. Increasing the molar ratio (detergent/CTD) had no structural effect in the interaction with neutral detergents but it increased the folding of the CTD in the presence of SDS.

Partial and full phosphorylation effects in the structure of the CTD in the presence of these detergents were also studied. In the presence of neutral detergents, all the phosphorylated species had similar secondary structures. As already happened in the complexes with DNA, phosphorylation increased the amount of β-structure and decreased that of α-helix. In the presence of SDS, both the level of phosphorylation and the detergent/CTD molar ratio affected the proportions of secondary structure motifs. The greatest effects were observed at a molar ratio of 7:1 SDS/CTD: the unphosphorylated CTD showed proportions of secondary structure motifs similar to those found in the DNA complexes; and the phosphorylated CTD became an all-β protein with 83% β-structure[423].

Lysine residues are clearly amphipathic, with positive net charge and four methylene groups. Charge compensation of Lys residues by interaction with DNA phosphates may play a role in the formation of the hydrophobic core, either interacting with hydrophobic ligands or among themselves. The anionic nature of SDS favours the compensation of the positive charge of histones, allowing the formation of a more hydrophobic environment that leads to the folding of the protein. Neutral detergents do not compensate the net charge. Thus, in this case, the folding of the phosphorylated protein upon interaction with Brij 35 or Triton X-100 is only modulated by the effect of the presence of one or more phosphate groups, which does not change the net charge enough to boost the folding of the protein as efficiently as anionic SDS does. However, interaction of the CTD and the entire $H1^0$ with SDS required a minimum concentration of detergent plus maximal phosphorylation of the protein to mimic the effects on secondary structure obtained when bound to DNA, suggesting that the folding pathway of histone H1 may be mediated by the formation of an hydrophobic core and by involving charge compensation upon DNA interaction[423].

The all-β conformation of the CTD within the phosphorylated $H1^0$ in the interaction with SDS allowed the formation of ribbon-like fibres of the amyloid type, as judged by strong birefringence in the

presence of Congo red and thioflavine T (ThT) fluorescence enhancement. In the unphosphorylated $H1^0$, only amorphous aggregates of varied morphology were observed. Thus, formation of amyloid fibre may require the joint effects of hydrophobic interactions and some degree of charge neutralization, together with the all-β potential associated with full phosphorylation.

The concomitant increase of β-structure and decrease of α-helix of H1 due to full phosphorylation that occurs in the presence of SDS or in the complexes with DNA has been described as a common structural change in fibril formation of several amyloidogenic proteins[452,453,454]. Moreover, histone H1 forms amyloid fibres in the presence of phosphatidylserine[455] and has been found associated with amyloid-like fibres[280,456]. Some amyloid proteins, such as the Aβ peptide and Tau, which are involved in protein-misfolding diseases, have been described as intrinsically disordered proteins. As mentioned, the fact that DNA binding induces the folding of histone H1 allows to classify it also as an intrinsically disordered protein.

Histone H1 has many features in common with amyloidogenic proteins. In most of the cases, the formation of amyloid fibres requires a structural reorganization of the native structure or the precursor protein, involving breakage and/or partial unfolding of the native structure[457]. In the case of histone $H1^0$, the formation of amyloid fibres requires the acquisition of an all-β conformation following phosphorylation, together with a hydrophobic environment provided by SDS. The increase of protein-protein interactions would lead to the oligomerization of all-β subunits and, afterwards, to the formation of the fibres in a similar way as the Aβ peptide[458].

Proteomic characterization and identification of novel post-translational modifications in chicken erythrocyte linker histones

Chicken erythrocyte native chromatin was fractionated into soluble and insoluble fractions and linker histones were analysed by different proteomic methods in order to observe their relative proportions in each fraction and to determine novel post-translational modifications.

In the HPCE analyses, the acetylated species of H5 and secondary peaks following those of H5 and H1 species/subtypes due to basal phosphorylation were also identified. Regarding the basal level of phosphorylation, there were no significant differences between the soluble and the insoluble fractions of native chromatin.

The most remarkable difference between chromatin fractions was the relative amount of H5 and H1. In the HPCE profile of the insoluble fraction, the relative absorbance of histones H1 with respect to that of histone H5 was higher than in the soluble fraction. This fact was confirmed by the H1/H5 ratio in each fraction calculated in the SDS gel: H1/H5 ratio in the insoluble chromatin fraction was around 0.8 whereas it was 0.31 in the soluble fraction.

Bates and Thomas first determined that H1 and H5 together constituted 1.3 moles per mole of the octamer: 0.9 moles were contributed by H5 (70%) and 0.4 moles were contributed by H1 subtypes (30%)[427]. Since H5 replaces H1 in mature chicken erythrocytes, the 30% of nucleosomes are supposed to contain two linker histones. Our results show that the ratio H1/H5 was higher in the insoluble chromatin fraction. This fact suggests that the stoichiometry of histones H1/H5 could vary along native chromatin, with major presence of two linker histone molecules in some chromatin regions than others. Thus, fractionation of chromatin by nuclease digestion and posterior centrifugation, will determine whether a chromatin fragment is part of the soluble fraction (chromatin fragments with lower proportion of nucleosomes with two linker histones) or the insoluble fraction (chromatin fragments with higher proportion of nucleosomes with two linker histones).

In the present study, histone H5 and histones H1 were further separated by RP-HPLC in order to carry out tandem mass spectrometry analysis (LC-ESI-MS/MS) to map the post-translational modifications of all linker histones subtypes in each native chromatin fraction.

Unlike the study of PTMs on core histones, the PTM pattern of histone H1 has been analysed only in few species and subtypes. Recent studies have permitted to identify multiple PTMs in human, mouse, rat, chicken and fruit-fly H1 variants[242,388,432,459] and most other published work focuses mainly on mapping phosphorylation sites[459].

Previous studies on post-translational modifications on chicken linker histones were carried out on bulk chromatin, considering that linker histones PTMs might be uniform along the chromatin molecule. In contrast to that, we fractionated native chromatin from chicken erythrocytes into soluble and insoluble fractions. We could map several novel PTMs and, moreover, our analysis also allowed to find differences in specific PTMs between the linker histones extracted from insoluble and soluble fractions of native chromatin, indicating the existence of a different pattern of post-translational modifications between those two fractions.

The finding of novel post-translational modifications in chicken linker histones leads to the question of whether which of those sites have been found post-translationally modified in other species linker histones. Figures D.1. and D.2 show a multiple sequence alignment in Clustal O (1.1.1.) of the linker histones sequences in humans (*Homo sapiens*), mouse (*Mus musculus*), rat (*Rat norvergicus*) and chicken (*Gallus gallus*). Only the part of the sequences with more conserved PTMs is shown. The PTMs found for each subtype are highlighted with colours. Figure D.1. shows the alignment of the closely related subtypes $H1^0$ (humans, mouse, rat) and H5 (chicken), whereas Figure D.2. shows the alignment for somatic histone H1 counterparts.

N^{α}-acetylation, which is the most common post-translational modification in eukaryotic proteins, was identified in H5T1 from soluble and insoluble chromatin fractions. Both the acetylated and unacetylated forms of this peptide were found in high amounts, suggesting that the unacetylated peptide is also present in mature erythrocytes. Previous studies have shown that the extent of N^{α}-acetylation on $H1^0$ increased with age in rat tissues, whereas the ratio between peptides did not change with age in chicken erythrocytes[376]. In histones H1, all subtypes were found to be N^{α}-acetylated in both fractions. Furthermore, phosphorylation in S3 and S7 (H5, in both fractions) and equivalent position T3 (H1.02 and H1.1L, soluble fractions) were also identified. (S/T)3 phosphorylation, together with of N^{α}-acetylation resembles the SET domain found in several histones subtypes[242,253,460]. These modifications were identified in combination with N^{α}-acetylation and as a single modification of the peptide (Figures D.1 and D.2). In histones H1, S3 is the residue that is phosphorylated in the canonical motif of GSK3 kinase (**S**-X-X-X-S), which is involved in the regulation of metabolism, the cytoskeleton, and gene expression (GSK3B in *G. gallus*). S7 is the residue that is phosphorylated in the canonical motif (S/T-X-X-X-**S**) of CK1 kinase.

In histone H5, two novel post-translational modifications were identified: K14 acetylation, which was found in both fractions; and T84 phosphorylation; which was only found in the soluble fraction. M31 oxidation was also identified. Oxidation of methionine residues is a common modification in proteins and, thus, we do not consider this as a truly novel PTM. Nevertheless, it has never been reported for avian H5. T84 phosphorylation is adjacent to K85, which is actually a part of the primary DNA binding site in chicken H5. Hence, the negative charge due to the addition of a phosphate group could neutralize the positive charge of the lysine residue and modulate the binding to DNA in this site. Phosphorylation is known to reduce the binding affinity of linker histones for DNA, relaxing the chromatin. T84 is part of the short peptide LKQT in H5, which is coincident with the canonical phosphorylation site of NK6 kinase (NIMA-related NK6). The NIMA-related kinases are a family of serine/threonine kinases implicated in cell

cycle control. NIMA kinase was first found in *Aspergillus nidulans* and is involved in multiple aspects of mitotic progression including the timing of mitotic entry, chromatin condensation, spindle organization and cytokinesis. Evidence suggests that NEKs perform functions similar to those of NIMA[461]. NEK6 plays an important role in mitotic cell cycle progression and it is required for chromosome segregation at metaphase-anaphase transition, robust mitotic spindle formation and cytokinesis[462]. Human Nek6 transcripts are ubiquitously expressed, with lower cell cycle-dependent expression in the early G1 phase. In HeLa cells, hNek6 showed both nuclear and cytoplasmic localizations and, in insect cells, recombinant hNek6 protein effectively phosphorylated histones H1 and H3. This fact suggests that Nek6 is a mitotic histone kinase which regulates chromatin condensation[463].

Despite NEK6 function in *Gallus gallus* has not been described, chicken erythrocytes are terminally differentiated cells in which chromatin is condensed into the 30nm-fibre. Chicken erythrocytes do not enter in mitosis; neither they need a control of the cell-cycle progression. Thus, phosphorylation in H5T84 by NK6 could just act regulating chromatin condensation. T84 was found in H5 from the solubilised chromatin fraction, not in the insoluble one. Since phosphorylation reduces the affinity for DNA, this PTM could act allowing a major solubilisation of chromatin by micrococcal nuclease digestion. Furthermore, $H1^0$ in human, mouse and rat also contain the canonical motif of NEK6 phosphorylation in T84 equivalent residues (Figure D.1), which suggests that phosphorylation in this position may play a role in the regulation of DNA binding, despite no phosphorylation has been yet reported for this position in these species.

Most of the PTMs were found in the N-terminal and the globular domains of histone H5. Some of them, like K12-acetylation, have already been identified in previous studies. In our case, this PTM was only found in H5 extracted from the insoluble fraction of native chromatin, suggesting, as mentioned before, that linker histone PTMs are not uniformly distributed in erythrocyte chromatin.

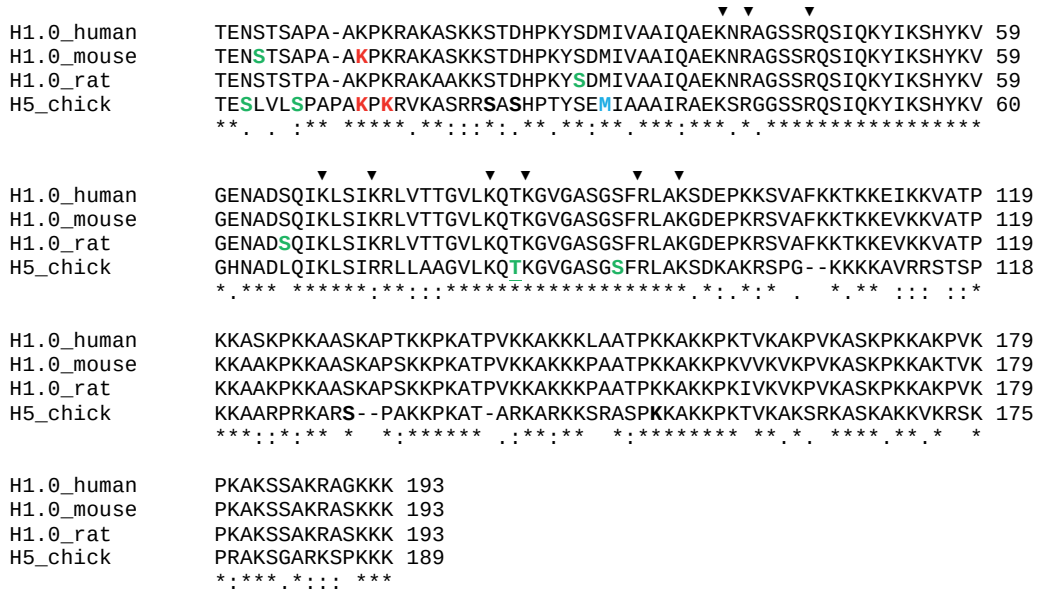


Figure D.1. Multiple sequence alignment of linker histones $H1^0$ and H5 counterparts and their PTMs. Sequence alignment by input in Clustal O (1.1.1.) Post-translational modifications for human and mouse linker histones were taken from Wisniewski *et.al.* (2007)[242]. Post-translational modifications for rat linker histones were taken from Sarg *et.al.* (2013)[388]. The PTMs for chicken are the ones reported in this study, all the PTMs mapped are highlighted in the sequence and they appear underlined when were only found in one of the fractions. The amino acid modifications are highlighted as follows: grey outlined=N^{α}-acetylation; red=lysine acetylation; green=phosphorylation; light blue=oxidation. Ubiquitination and deamidation have not been considered. DNA binding sites of histone H5 are shown with a black triangle (▾).

In chicken histone H1 subtypes, six novel post-translational modifications were identified. Five novel PTMs were identified in the N-terminal domain and at the beginning of the globular domain of histone H1 subtypes: T3-phosphorylation (H1.02 subtype), K17-acetylation, K34-acetylation, S39-phosphorylation and K46-acetylation (positions in H1.01). One more novel PTM was identified in the CTD: phosphorylation in S104 (shared peptide by all H1 subtypes). T3-phosphorylation had been previously described in H1.03, H1.1L and H1.1R, but not in H1.02. K17-acetylation was only identified in the insoluble fraction of H1.02 subtype. This PTM was found in chicken erythrocytes[432] in H1.1R and equivalent positions have been found to be acetylated in human, mouse and rat (Figure D.2.). K34-acetylation and S39-phosphorylation were found in a peptide shared by all histone H1 subtypes extracted from both fractions of chromatin. In H1.02, S39 phosphorylation was only found in the insoluble fraction. S39 in histones H1 is part of the canonical phosphorylation site of CK2 kinase (**S**/**T**-X-X-E). CK2 is found in all eukaryotes examined, usually in one or two copies. CK2 is constitutively active, phosphorylating on largely acidic sites, and is known for having a very large number of substrates and a variety of functions at different stages in the cell cycle[464].

K46-acetylation was found in a peptide that matched all histone subtypes but only in the insoluble fraction. These results also show that the patterns of PTMs between the soluble and insoluble chromatin fragments are also different for histone H1 subtypes, as already happened for histone H5. Equivalent positions of K46 had been reported to be also acetylated in several mouse and human histones H1[242].

Novel S104 phosphorylation was identified a peptide shared by all histone H1 subtypes purified from the insoluble fraction of chromatin, except for histone H1.02. This site is part of the canonical phosphorylation site of GSK3 kinase (**S**-X-X-X-S). GSK3 (Glycogen synthase kinase 3) is a serine/threonine protein kinase that has been identified to phosphorylate more than forty different proteins in a variety of different pathways[465]. GSK3 is active in central intracellular signaling pathways, including cellular proliferation, migration, inflammation and immune responses, glucose regulation, and apoptosis. All histone H1 subtypes in human, mouse and rat (except H1b) contain a GSK3 motif in S50 (hH1c numbering). This position has not been found phosphorylated in none of these subtypes and, interestingly, it is located next to K51 (hH1c numbering). K51 is acetylated in human and mouse H1c-e subtypes and this position is part of the primary DNA binding site. Thus, potential phosphorylation in this position could also modulate the interaction with DNA.

Interestingly, GSK3 activity has been associated with both pathological features of Alzheimer's disease, the abnormal accumulation of Amyloid-β (Aβ) and the hyperphosphorylation of tau proteins, which leads to the formation of the intracellular tangles[466,467]. The mechanism whereby GSK3 promotes Aβ production remains to be clarified; perhaps it could be associated with the phosphorylation of the amyloid precursor protein by GSK3 or the binding of GSK3 to presenilin-1[467].

The Aβ peptide is an intrinsically disordered protein (IDP), as well as histone H1, as already discussed above. In Chapter One of the present study, fully phosphorylated histone $H1^0$ formed amyloid fibres through the acquisition of an all-β conformation following phosphorylation, together with a hydrophobic environment. In Chapter Two, histones H1 from chicken native chromatin have been found phosphorylated in a consensus GSK3 motif in serine residues located in the terminal domains of the protein (S3 and S104), which are the parts of H1 that fold upon interaction with DNA, allowing to describe histone H1 as an IDP. Hence, phosphorylation of chicken H1 by GSK3 could also play a role in the folding of the terminal domains following phosphorylation and have an effect on histone H1 binding to DNA.

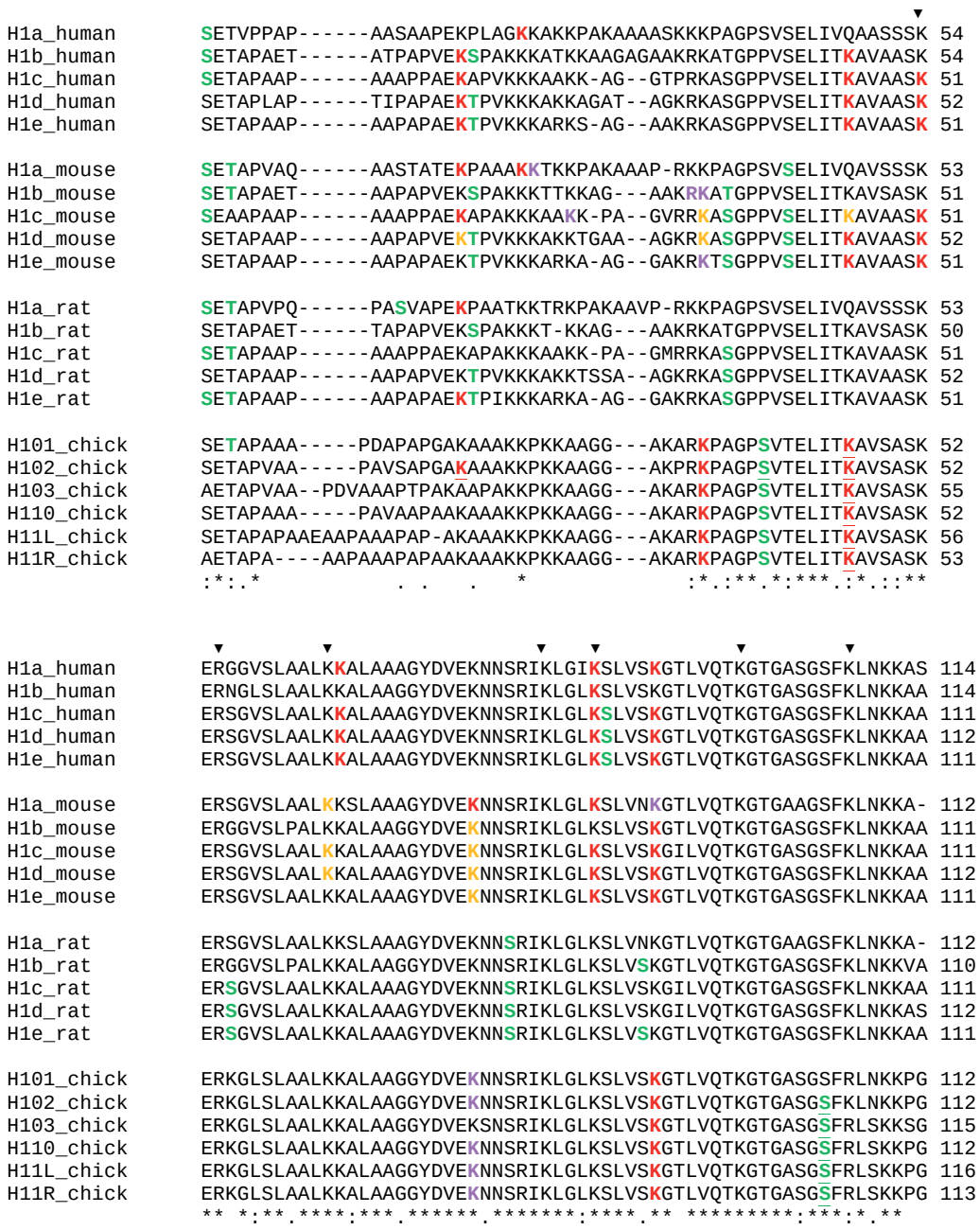


Figure D.2. Multiple sequence alignment of linker histones H1 counterparts and their PTMs. Sequence alignment by input in Clustal O (1.1.1.). Only the part of the alignment with more equivalent PTMs between subtypes is shown, which corresponds to the N- and globular domains. Post-translational modifications for human and mouse linker histones were taken from Wisniesky *et.al.* (2007)[242]. Post-translational modifications for rat linker histones were taken from Sarg *et.al.* (2013)[388]. The PTMs for chicken are the ones reported in this study, all the PTMs mapped are highlighted in the sequence and they appear underlined when were only found in one of the fractions. The amino acid modifications are highlighted as follows: grey outlined=N^{α}-acetylation; red=lysine acetylation; green=phosphorylation; orange=acetylation and/or monomethylation; magenta=dimethylation. Ubiquitination and deamidation have not been considered. DNA binding sites are shown with a black triangle (▾).

Di-methylation in K75 was also identified in a peptide shared by all histone H1 subtypes and it was present in both chromatin fractions. This position was reported as dimethylated/formylated in H1.03 subtype by Snijders *et.al.*[432]. Equivalent position of K75 has been reported to be acetylated and/or mono-methylated in mouse histones H1[242].

Equilavent acetylations were found in chicken histones H11L-K191 and H1.1R-K185, in both the soluble and the insoluble fraction of chromatin. K191 is adjacent to an SPKK motif (SPAK) in both subtypes, which is a consensus site for CDK2 phosphorylation. Sjniders *et.al.* [432] reported this residue to be also methylated. Thus, in their case, methylation next to this phosphorylation site suggests that the binding of DNA in this position could be regulated by a methyl-phos switch. Actually, previous studies already mention the possibility to extend to linker histones this binary switch[75], that has been commonly described and further studied in core histones (especially in H3). Experimental evidence has been reported for human histone H1e[253], where adjacent lysine K25 residue to the phosphoserine S26 was found to be methylated. When methylated, histone H1 binds the heterochromatin protein (HP1), whereas phosphorylation in S26 blocks this binding[281]. In mammals, histone H1 and the heterochromatin protein 1α (HP1α) interact both in vivo and in vitro. HP1α is mainly associated with centromeric chromatin and it is thought to control the formation and propagation of the heterochromatin through the interaction with methylated H3K9[468]. The binding between H1 and HP1α is regulated by a "methyl-phos switch" in which phosphorylation of histone H1 acts blocking HP1α binding to chromatin. This process has been identified in H1e from HeLa cells[281] and H1b from mouse fibroblasts[469].

As it can be observed in Figures D.1. and D.2., most of the modifications found to be conserved between linker histone subtypes occur in the N-terminal and the globular domain. Acetylation seems to be the most conserved PTM among these species. In all histone subtypes, the observed acetylation sites are located most frequently in the globular domain and, in human and mouse H1, acetylation mainly occurs on residues that have been considered to be directly involved in DNA binding[242]. In some cases, acetylated residues in some counterparts were found to be monomethylated and/or acetylated. As acetylation and methylation can interfere with DNA binding, it is possible that these modifications can significantly change the binding properties of H1 conferring specific properties in chromatin to some H1 subtypes. In core histones, these PTMs mainly act by recruiting DNA remodelling complexes. In histones H1, the fact that these PTMs are conserved in different species and that they are present in the globular domain of H1 and, thus, close to the nucleosome core, suggests that they could act synergistically with the PTMs present in core histones or by directly affecting the binding to DNA.

In the present study, two acetylation sites mapped in chicken histones H1 were found to be conserved in mouse and human counterparts: K46 and K90 (Figure D.2.). Acetylation on K46 is one of the novel PTMs identified in the present study. Other PTMs mapped in this study have equivalent modified positions in mouse and human. In some cases, the modification in the equivalent position is conserved while in others is not. Novel acetylation (K34) in chicken histones H1 was identified in this study. In mouse histones, the equivalent residue is reported to be acetylated and/or monomethylated in H1c and H1d whereas it was dimethylated in H1e and Hb. Something similar happens to the novel dimethylation found in chicken H1 K75, in mouse linker histones this residue is also modified; however, the PTMs found were acetylation and/or methylation. No PTM has been mapped for this position in human histones. All this PTMs are located in the globular domain and, hence, they potentially may regulate the interaction of linker histones with DNA and with other proteins.

Another interesting observation is about lysine K62 (H1c numbering) present in mouse histone subtypes H1a, H1c, H1d and H1e, where it is reported to be acetylated/monomethylated. In the same human counterparts, this lysine was not modified but the adjacent K63 (H1c numbering) was reported to be acetylated. This PTM was found dimethylated in chicken erythrocytes by Snijders *et.al.*[432]. The fact that

equivalent residues are modified among such divergent species suggests that the modification of this site may play an important and conserved role in linker histones. In fact, K62 and K63 are part of DNA binding sites in mouse/human/chicken H1s linker histones.

Several non-CDK sites were found to be phosphorylated in the N-terminal and globular domains of mouse and humans histones. Terminally differentiated cells, such as chicken erythrocytes, contain low basal level of phosphorylation. In the present study, we described several phosphorylation sites within the N-terminal and globular domains. The novel phosphorylation identified in S39 at the beginning of the globular domain is also present in nearby and equivalent positions in mouse and human. Histones H1 of these two species are phosphorylated in S35 and S40 (mouse H1.2 numbering). These results show that many of these non-CDK phosphorylations, located in the GH1, are conserved through evolution.

Proteomic characterization of linker histones after *ex vivo* chromatin phosphorylation and effects of linker histone phosphorylation on chromatin aggregation

In the present study, native chromatin from chicken erythrocyte nuclei was purified and phosphorylated *ex vivo* with CDK2 kinase, which specifically phosphorylates the SPKK motifs present in linker histones. The aim was to study the effects of linker histone phosphorylation on chromatin aggregation. For this purpose, we first extracted the linker histones from the *ex vivo* phosphorylated chromatin, in order to identify by proteomic techniques the different histone subtypes and to determine the level of phosphorylation and the position of phosphate groups before and after *ex vivo* phosphorylation with CDK2 kinase at different times of reaction.

HPCE analyses were performed for initial and qualitative identification of the phosphorylated species that appeared after *ex vivo* phosphorylation of chromatin. In the HPCE profiles, a progressive increase in the number of secondary peaks was observed with the increase of phosphorylation time. Interestingly, the changes in the HPCE separation profile occurred differently for histones H5 and H1: at shorter phosphorylation times, most of the secondary peaks due to phosphorylation were detected in the part of the HPCE spectra that corresponded to histone H5 variants, whereas secondary peaks for the six H1 subtypes did not appear before 1 hour of phosphorylation. Moreover, the evolution of the profiles seemed to indicate that phosphorylation lead to an increase, first, of the mono-phosphorylated species and, later on, to the appearance of di- and tri-phosphorylated species.

The total linker histone fraction (H5+H1) was separated by Reverse Phase High Performance Liquid Chromatography (RP-HPLC) and, for each phosphorylation time, two fractions of protein were obtained (H5 and H1)**.** In order to determine the level of phosphorylation, the samples were analysed by MALDITOF-MS. With the purpose of determining the positions of phosphate groups' incorporation in the SPKK motifs of histones after *ex vivo* phosphorylation, H5 and H1 were also analysed by Tandem MS (LC-ESI-MS/MS). The performance of these two techniques requires the enzymatic cleavage of the proteins into shorter peptides than can be properly detected.

As already discussed, identification of histone H1 subtypes in the MALDITOF-MS analyses was difficult due to a high background and the overlapping of the peaks. In Tandem MS, the sequences coverages obtained allowed to map several SPKK motifs in both linker histones, despite the CTDs of the proteins, where most of the SPKK sites are present, were the less covered regions in the analyses. These drawbacks were attributed to limitations of the method, especially in the case of histones H1; since

isolation of every single H1 subtype by RP-HPLC from the H1s fraction was not possible. Hence, the whole six chicken histone H1 fraction was analysed together by MALDITOF-MS and by Tandem MS.

Our analyses allowed to observe a progressive increase of the phosphorylation levels with the increase of phosphorylation time, in both histones H5 and H1s by MALDITOF-MS. Mono-, di- and even tri-phosphorylated species could be identified. Data obtained in the MALDITOF-MS analyses was used for calculation of the percentages (%) of the phosphorylated species of the C-terminal domain of H5 present in each sample. The results confirmed that, after 3 hours of phosphorylation, the total percentage of phosphorylated species was around 37% and that it increased up around 54% of phosphorylated species after overnight phosphorylation. The results also suggested that this increase first occurred in the amount of monophosphorylated species and, then, leading to the appearance of diphosphorylated species. These percentages were calculated with regard to the total amount of the different species identified in the samples. In most cases, only up to two phosphate groups could be identified by MALDITOF-MS.

Histone H5 contains five CDK2 sites per H5 molecule that are susceptible to be phosphorylated. Hence, despite the amount of phosphorylated species remarkably increased in comparison to the unphosphorylated ones after overnight phosphorylation, the percentage of histone H5 phosphorylation when considering all the potential CDK sites to be phosphorylated only corresponded to a 14.25% in this sample. This data indicated that linker histones extracted from *ex vivo* phosphorylated chromatin mainly consisted of many partially phosphorylated molecules.

In the Tandem MS analyses, several peptides containing phosphorylated SPKK motifs were identified in both linker histones. In linker histones from native chromatin fraction, none of the CDK sites appeared to be phosphorylated, whereas some of them appeared to be modified after *ex vivo* phosphorylation. In the Tandem MS analysis of histone H5, SPKK motif in position S148 was identified in all the samples and it appeared phosphorylated after 1 hour. In the Tandem MS analysis of histone H1 subtypes, SPKK motifs were identified for H1.03, H1.1L and H1.1R. All the SPKK motifs were identified, except S171 (H1.1R numbering). H1.03T16 was found phosphorylated after 15 minutes; H1.1LS192 and H1.1RS186 were phosphorylated after 1 hour; H1.03S155, H1.1LS155 and H1.1RS153 were found phosphorylated after 3 hours.

These results indicated that the increase of phosphorylation levels was due to CDK2 *ex vivo* phosphorylation of linker histones within chromatin and corroborated that an increase in the phosphorylation time, lead to the progressive phosphorylation of the SPKK sites.

Considering the physicochemical properties of histones H1 and their high sequence similarity, the results obtained in the present study could have been improved by separating every single H1 subtype by Hydrophilic-interaction chromatography (HILIC)[364]. Isolation of every histone H1 subtype from the mixture would have eased the identification of peaks in the posterior MALDITOF-MS analysis. HILIC has been implemented and optimized for the separation of both histones and a wide range of solutes[374,375,470,471]. However, its implementation for the specific isolation of pooled chicken erythrocyte H1 histone subtypes still has to be developed.

In the present study, the analysis of histone peptides caused the formation of many short and polar peptides that went undetected by LC-ESI-MS/MS, especially in the C-terminal domain of the protein, which are very lysine-rich regions and contain most of the SPKK motifs that we tried to *ex vivo* phosphorylate. A new proteomic approach, capillary-electrophoresis electrospray-ionization mass spectrometric (CESI-MS)[388,389] has been shown to overcome this disadvantage, and hence, this new technique represents a promising alternative for histone PTMs characterization.

Phosphorylation of SPKK motifs also regulates the interaction with HP1α. Specific phosphorylation H1 by CDK2 during late G1/S phase from the cell-cycle allows to stop the interaction, relaxing the chromatin to start the replication process[261]. HP1α interaction with H1 CTD depends on the net charge and the disruption of it is not related to phosphorylation in a specific site: phosphorylation of any CDK2 site disrupts this interaction.

Linker histone phosphorylation has also been shown to be one of the early alterations in the progression of head and neck squamous cell carcinomas and a potential cellular proliferation biomarker in head, neck and cervical cancers[472]. Recent studies by LC-MS profiling revealed a statistically significant increase in phosphorylation of the CDK2 in p-T146 of H1 linker histone subtypes associated with progressive bladder carcinogenesis and its invasiveness[472]. Cell cycle analysis of histone H1 phosphorylation also showed an increase of phosphorylation from G0/G1 phase to M phase, supporting this increase in H1 phosphorylation as a proliferative biomarker during bladder carcinogenesis.

Another example of the importance of phosphorylation of histone H1 within chromatin is associated, as mentioned before, to the interaction of phosphorylated H1 with αβ-amiloid and amiloidogenic α-synuclein, which are the main effectors of Alzheimer and Parkinson diseases, respectively. Phosphorylated H1 by CDK5 can be translocated to the cytoplasm during S and M phases[278], where it preferentially interacts with those two proteins[280].

Once *ex vivo* phosphorylation of linker histones within chromatin was confirmed, we studied the effect of this phosphorylation on chromatin aggregation. The induction of aggregation was carried out by an increase in $MgCl_2$ concentration up to 1.6mM $MgCl_2$ of *ex vivo* phosphorylated chromatin and it was analysed by Dynamic Light Scattering (DLS).

Fragmentation of chromatin by nuclease digestion results in individual fibres available for characterization. In general, irregularity dominates the architecture of isolated native chromatin fibres, reflecting the heterogeneity in linker DNA length, the composition of histone variants, histone PTMs, and interaction with non-histone proteins[446]. Hence, chromatin results in a heteromorphic and dynamic organization that can be affected by changes in the molecule itself or in its environment. Besides that, sections of the transcriptionally inert chicken erythrocyte nuclei show a uniform compaction of chromatin, which resolves into irregularly structured 30-nm fibres when allowed to swell by, for instance, decreasing the ionic strength to physiological salt conditions. It is not yet clear why these properties lead to the appearance of distinct 30-nm fibres, but it could be that these conditions could favour intrafibre over interfibre interactions[446].

DLS measurements showed that phosphorylation had a clear effect on chromatin aggregation of the samples. The most remarkable differences were observed in the hydrodynamic diameters (d.nm) between the control and the corresponding phosphorylated sample: the diameter sizes of the aggregated molecules peak were always higher in the unphosphorylated chromatin than in the phosphorylated one, suggesting that the aggregates were larger. Furthermore, these differences became more remarkable with the increase of the incubation/phosphorylation time.

Our results also showed that the size of the chromatin fragments also played a role in the impairment of chromatin aggregation at 1.6mM $MgCl_2$ following *ex vivo* phosphorylation of chromatin. Under the same experimental conditions, the difference between the unphosphorylated and the phosphorylated sample (Δpeaks) after overnight incubation/phosphorylation was of ~69nm in the aggregation experiments of short chromatin fragments; whereas in larger chromatin fragments this difference increased up to ~654nm. This can be attributed to the fact that compacted polynucleosome fragments exhibit more internal flexibility than compacted high molecular weight chromatin fibres do[473]. The intensity distributions in short chromatin fragments supported these observations, since the intensities of the

second peak (aggregates) in the DLS measurements were always higher in the incubation controls, indicating that those were more aggregated than the phosphorylated samples. This variation in the percentages of intensity distributions were mainly observed in the data obtained for short chromatin fragments, not for larger chromatin fragments, which did not show significant differences between peak intensities (%).

As already discussed in Chapter One, the neutralization of linker DNA charges can be attributed to electrostatic interaction with linker histones, which can affect nucleosomal spacing and folding of the chromatin fiber[474]. Linker histones also influence chromatin structure especially through their C-terminal and globular domains. SPKK motifs in the highly basic C-terminal region are critical for the macromolecular events involved in chromatin condensation[229]. Then, histone H1 phosphorylation on its SPKK motifs has also a major influence on chromatin aggregation and compaction.

Phosphorylation has shown to have an effect on the affinity of histones H1 and their CTD for DNA[266,267]. Competition assays of the two species (unphosphorylated/phosphorylated) of the CTD and the entire H1e for DNA fragments showed a decrease in the relative affinity for DNA due to phosphorylation.

In these studies, phosphorylation also showed to have an effect on histones H1 and their CTDs capacity to aggregate DNA fragments[266,267]. Histone H1e has five SPKK motifs. The more extensively phosphorylated species (tri-, tetra- and pentaphosphorylated) showed a higher aggregation capacity than the less phosphorylated species (mono-, diphosphorylated). The most extensively phosphorylated species were also the ones that showed higher proportions of β-structure, what could favour the formation of DNA-protein aggregates induced by protein-protein interactions.

In Chapter Three, the proteomic analyses of linker histones extracted from this *ex vivo* phosphorylated chromatin showed that the number of phosphorylated species of linker histones increased with the time of phosphorylation, achieving more than 50% of partially phosphorylated molecules, with one or two phosphate groups. In Chapter Four, our results showed that phosphorylation of linker histones within chromatin impaired the aggregation of the fragments, in comparison to those that had not been phosphorylated.

Hence, the level of phosphorylation of our chromatin fragments could be compared with that of the less phosphorylated H1e molecules in the studies of aggregation of DNA fragments. In our case, partial phosphorylated chicken linker histones within chromatin aggregated less and formed smaller aggregates that native chromatin, being this effect increased with incubation/phosphorylation time and size of the fragments. Whether extensively *ex vivo* phosphorylation of linker histones within chromatin would lead to the opposite, as it happens with DNA, should be further studied.

Post-translational modifications of core and linker histones are directly related with the regulation of H1 affinity for chromatin and phosphorylation of linker histones plays an important role in many processes related with chromatin regulation and dynamics. Terminal differentiation of most cell types is related to structural changes in chromatin conformation. H1 phosphorylation reduces its affinity for chromatin without implying its release. Hence, the phosphorylation effect of H1 within chromatin could be related to global structural changes in chromatin whilst H1 is bound[273].

Conclusions

1. In the entire $H1^0$ in aqueous solution, the presence of the globular domain seems to stabilize the β-structure of the CTD, which suggests the presence of some degree of structural coupling between the CTD and the globular domain.

2. Phosphorylation of the three SPKK motifs in the CTD induced an increase of the percentage of β-structure in the entire histone $H1^0$ in aqueous solution, mainly determined by a conformational change of the C-terminal domain.

3. The binding of unphosphorylated histone $H1^0$ to DNA induced a significant increase in the proportion of α-helix, as happened with the isolated CTD. Thus, histone $H1^0$ can be also classified as an intrinsically disordered protein, with coupled binding and folding.

4. The binding of $H1^0$3P to DNA induced a remarkable increase in the amount of β-structure, which was dependent on the protein/DNA ratio

5. The amount of β-structure (54%) was higher than the contribution of the CTD to the entire protein (52%), suggesting that phosphorylation, coupled with DNA binding, induced the folding of the CTD in an all-β conformation in the entire protein.

6. In the complexes $H1^0$3P-DNA there was a remarkable decrease of α-helix in comparison with $H1^0$0P bound to DNA, which could be due to the loss of α-helix in the globular domain associated to the propagation of the β-structure from the CTD towards the rest of the protein.

7. In the presence of SDS, $H1^0$0P an $H1^0$3P folded with percentages of secondary structure motifs similar to those found in when bound to DNA.

8. The presence of SDS at a molar ratio 14:1 (SDS/protein) led to an increase of β-structure of the triphosphorylated protein up to 55% indicating that the CTD within histone $H1^0$ was also in an all-β conformation and that the β-structure had propagated, in part, to the globular domain.

9. The entire $H1^0$ formed ribbon-like amyloid fibres in the presence of SDS, when the CTD was fully phosphorylated ($H1^0$3P) and adopted an all-β conformation, suggesting that amyloid fibre formation requires the joint effects of hydrophobic interactions and some degree of charge neutralization, together with the all-β potential associated with full phosphorylation.

10. Secondary peaks, due to basal phosphorylation of H5 and H1 subtypes were identified in the HPCE profiles of soluble and insoluble chicken erythrocyte chromatin. No significant differences were observed between the two HPCE profiles.

11. The relative amount of histone H1 subtypes with respect to histone H5 was higher in the insoluble chromatin fraction. H1s/H5 ratio in the insoluble chromatin fraction was 0.8 whereas it was 0.3 in the soluble fraction.

12. Eight novel PTMs, including lysine acetylation (H5K14ac, H1K34ac, H1K46ac, H1.02K17) and serine/threonine phosphorylation (H5T84phos, H1.02T3phos, H1S39phos, H1S104phos) were identified in chicken erythrocyte linker histones H5 (two novel PTMs) and H1 (six novel PTMs) by nanoLC-ESI-MS/MS.

13. Fractionation of chicken native chromatin showed that there were differences in specific PTMs between the histones extracted from insoluble and soluble fractions of native chromatin.

14. The number of phosphate groups increased with time due to *ex vivo* chromatin phosphorylation with CDK2. This increase was observed by the radioactive phosphorylation assay, HPCE analysis, MALDITOF-MS and Tandem MS.

15. The HPCE profiles suggested that phosphorylation lead to an increase, first, of the mono-phosphorylated species and, then, to the appearance of di- and tri-phosphorylated species.

16. The percentages (%) of the phosphorylated species of the C-terminal domain of H5 at different times were calculated from the MALDITOF-MS results. After overnight phosphorylation, the percentage of phosphorylated species achieved 54%.

17. All the SPKK motifs identified in the sample of native chromatin were always not phosphorylated, whereas some of them were modified after *ex vivo* phosphorylation. This result indicates that the identified phosphorylations were incorporated due to CDK2 phosphorylation.

18. In the Tandem MS analysis of histone H5, SPKK motif in position S148 was identified in all the samples and appeared phosphorylated after 1 hour.

19. In the Tandem MS analysis of histone H1 subtypes, all the SPKK motifs were identified for H1.03, H1.1L and H1.1R, except S171 (H1.1R numbering). H1.03T16 was phosphorylated after 15 minutes; H1.1LS192 and H1.1RS186 were phosphorylated after 1 hour; H1.03S155, H1.1LS155 and H1.1RS153 were phosphorylated after 3 hours.

20. The increase of $MgCl_2$ concentration up to 1.6mM led to the appearance of a second peak of aggregated molecules in the DLS measurements.

21. DLS measurements at 1.6mM $MgCl_2$ of native chromatin showed that chromatin aggregation depends on the presence of divalent cations (Mg^{2+}) in the medium and the temperature of incubation.

22. The most remarkable differences associated to *ex vivo* phosphorylation of linker histones within chromatin was a decrease in the hydrodynamic diameter of the aggregation peak measured by DLS at 1.6mM $MgCl_2$. The differences became greater with the increase of phosphorylation time.

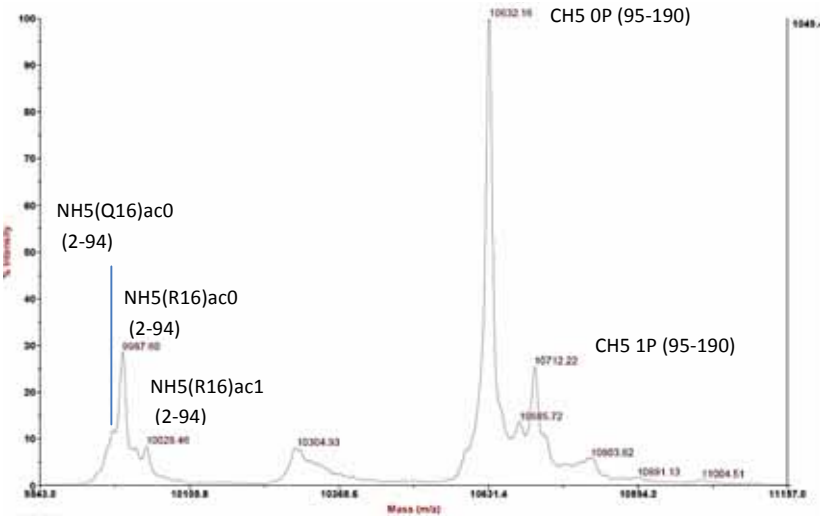
23. For short chromatin fragments (~1500bp), the DLS measurements showed that, after overnight phosphorylation, the particles of the aggregated population were ~69nm shorter in the phosphorylated chromatin than in the unphosphorylated controls. In these samples, the intensity distributions (%) of the second peak (aggregates) were always higher in the unphosphorylated chromatin, also indicating that those were more aggregated than the phosphorylated samples.

24. For large chromatin fragments (>8576bp), the DLS measurements showed the appearance of a third peak of even larger aggregates in the samples that had been incubated/phosphorylated overnight. The difference in size of the particles of the most aggregated population (peak 3) of this pair of samples (Δpeaks) was ~654nm.

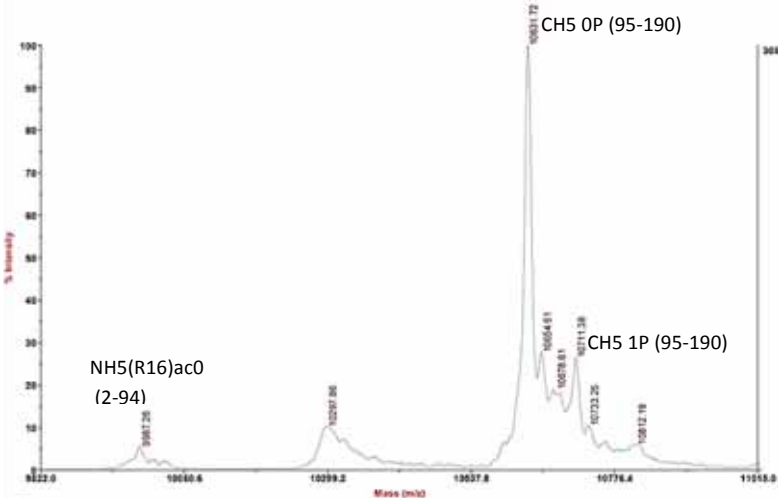
25. All these results led to the conclusion that, at 1.6mM $MgCl_2$, linker histones *ex vivo* phosphorylation impaired chromatin aggregation.

Appendix

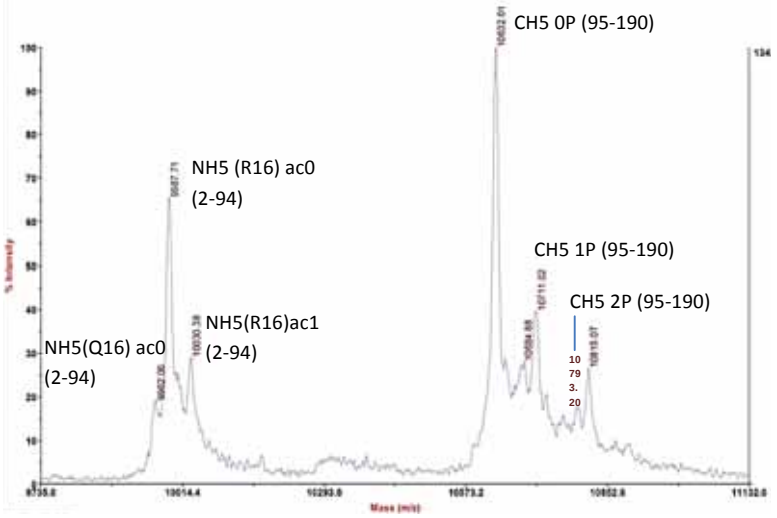
MALDI spectrum of *alfa*-chymotrypsin cleaved H5 histone fraction extracted from chicken erythrocyte's soluble chromatin phosphorylated *ex vivo* for 30 minutes.



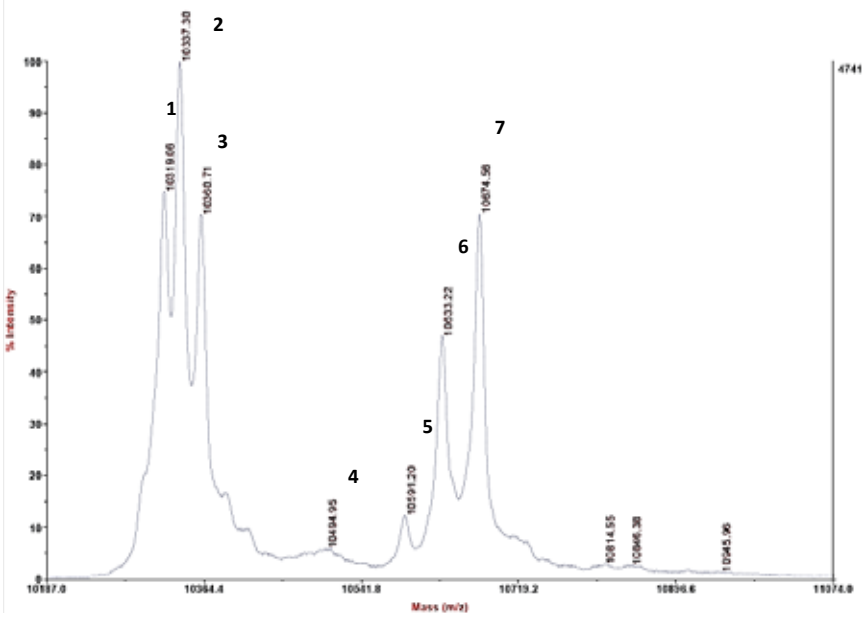
MALDI spectrum of *alfa*-chymotrypsin cleaved H5 histone fraction extracted from chicken erythrocyte's soluble chromatin phosphorylated *ex vivo* for 1 hour.



MALDI spectrum of *alfa*-chymotrypsin cleaved H5 histone fraction extracted from chicken erythrocyte's soluble chromatin phosphorylated *ex vivo* for 5 hours.

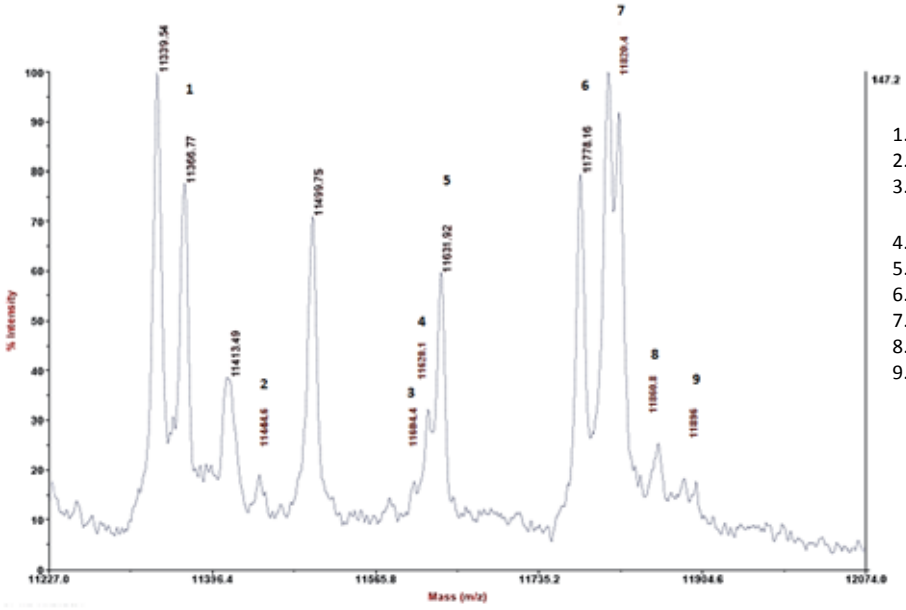


MALDI spectrum of *alfa*-chymotrypsin cleaved H1 histone fraction extracted from chicken erythrocyte's soluble chromatin phosphorylated *ex vivo* for 15 minutes. The spectrum corresponds to the region in which the peptide of the N-terminal domain plus the globular domain of histones was detected.



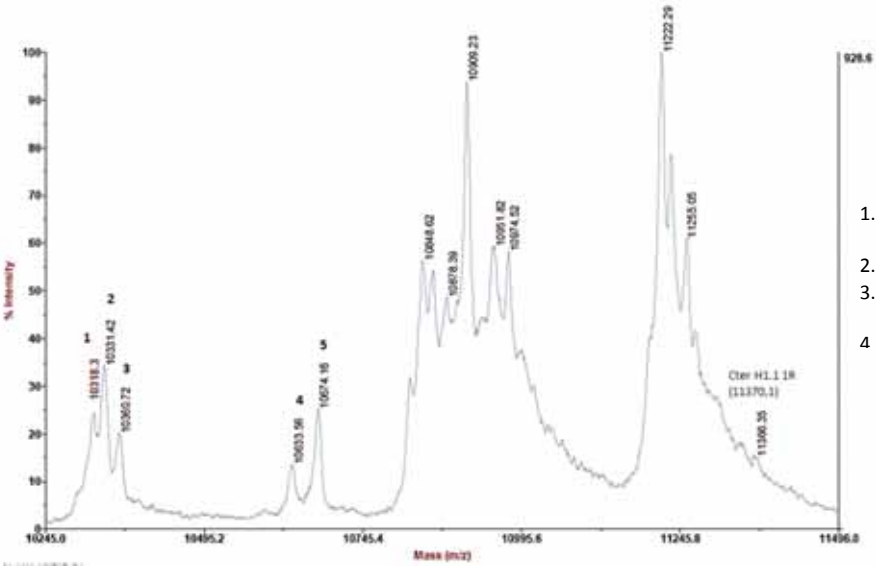
1. NH1.1Rac0/NH1.02ac0 (10322,05)
2. NH1.01ac1 (10337,96)
3. NH1. 1Rac1/NH1.02ac1 (10363,05)
4. NH1.01ac1-2P (10497,96)
5. NH1.03ac0 (10593,37)
6. NH1.1Lac0/NH1.03ac1 (10635,36)
7. NH1.1L ac1 (10676,36)

MALDI spectrum of *alfa*-chymotrypsin cleaved H1 histone fraction extracted from chicken erythrocyte's soluble chromatin phosphorylated *ex vivo* for 15 minutes. The spectrum corresponds to the region in which the peptide of the C-terminal domain of histones was detected.



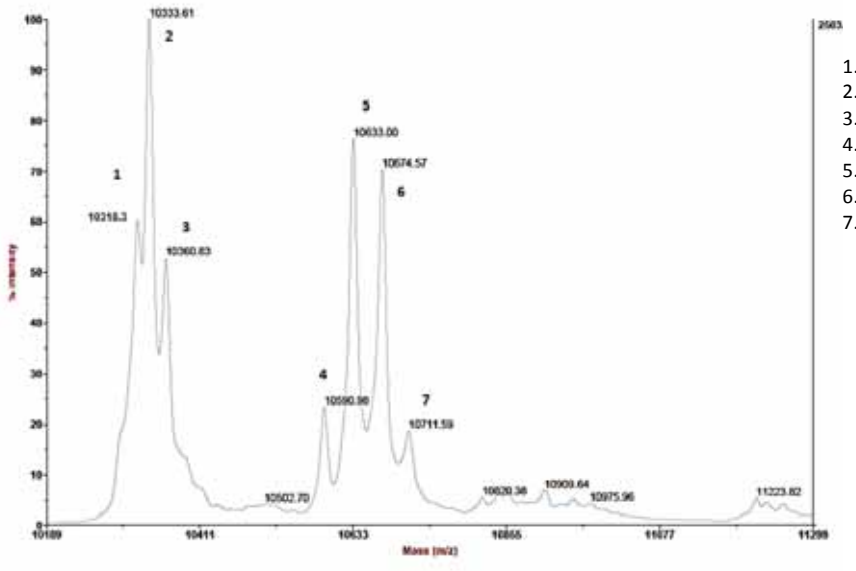
1. CH1.1R (11370,1)
2. CH1.1R -1P (11444,6)
3. CH1.02 -2P/CH1.1R -3P (11604.4/11609.1)
4. CH1.10 (11622,3)
5. CH1.01 (11636,4)
6. CH1.1L (11781,6)
7. CH1.03 (11820,7)
8. CH1.1 1L- 1P (11860,65)
9. CH1.03 - 1P (11900,7)

MALDI spectrum of *alfa*-chymotrypsin cleaved H1 histone fraction extracted from chicken erythrocyte's soluble chromatin phosphorylated *ex vivo* for 30 minutes. The spectrum corresponds to the region in which the peptide of the N-terminal domain plus the globular domain of histones was detected. The C-terminal domains of H1 histones were not detected in the MALDI assay, presumably due to degradation (see the peaks in the m/z region 10745.4-11496.0)



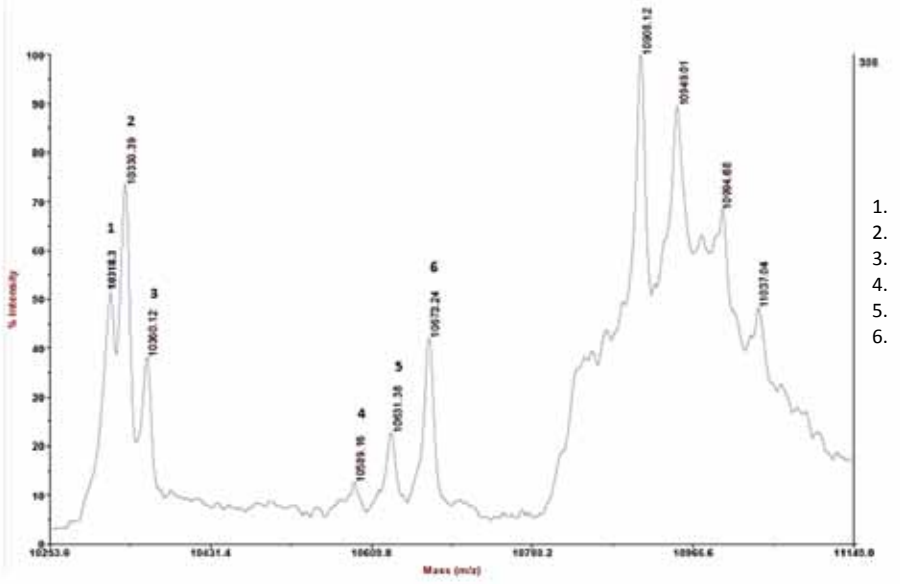
1. NH1.1Rac0/NH1.02ac0 (10322,05)
2. NH1.01 ac1 (10337,96)
3. NH1.1Rac1/NH1.02ac1 (10363,05)
4. NH1.1Lac0/NH1.03ac1

MALDI spectrum of *alfa*-chymotrypsin cleaved H1 histone fraction extracted from chicken erythrocyte's soluble chromatin phosphorylated *ex vivo* for 1 hour. The spectrum corresponds to the region in which the peptide of the N-terminal domain plus the globular domain of histones was detected. The C-terminal domains of H1 histones were not detected in the MALDI assay.



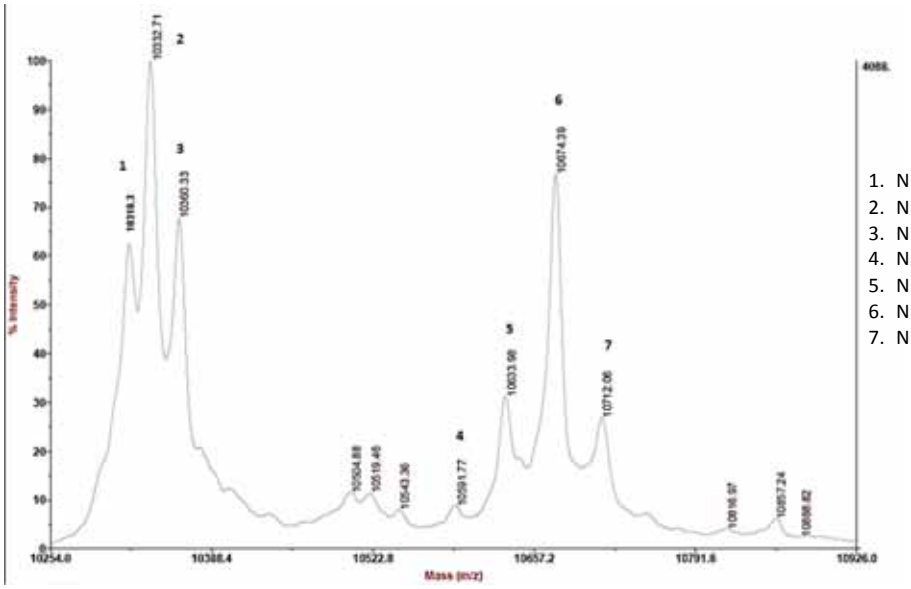
1. NH1.1Rac0/NH1.02ac0 (10322,05)
2. NH1.01 ac1 (10337,96)
3. NH1.1Rac1/NH1.02ac1 (10363,05)
4. NH1.03ac0 (10593,37)
5. NH1.1Lac0/NH1.03ac1 (10635,36)
6. H1.1L ac1 (10676,36)
7. NH1.1Lac0-1P/NH1.03ac1-1P

MALDI spectrum of *alfa*-chymotrypsin cleaved H1 histone fraction extracted from chicken erythrocyte's soluble chromatin phosphorylated *ex vivo* for 3 hours. The spectrum corresponds to the region in which the peptide of the N-terminal domain plus the globular domain of histones was detected. The C-terminal domains of H1 histones were not detected in the MALDI assay.



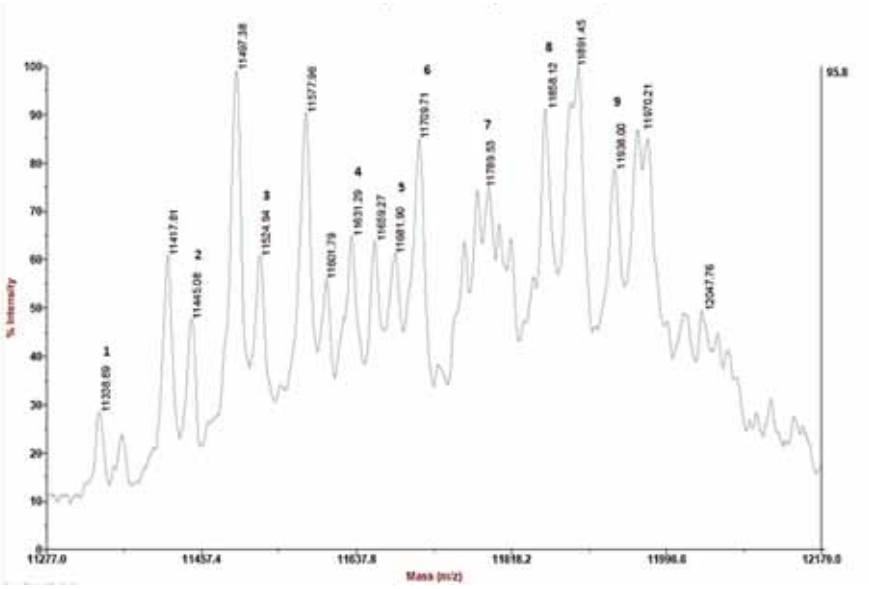
1. NH1.1Rac0/NH1.02ac0 (10322,05)
2. NH1.01 ac1 (10337,96)
3. NH1.1Rac1/NH1.02ac1 (10363,05)
4. NH1.03ac0 (10593,37)
5. NH1.1Lac0/NH1.03ac1 (10635,36)
6. H1.1L ac1 (10676,36)

MALDI spectrum of *alfa*-chymotrypsin cleavaged H1 histone fraction extracted from chicken erythrocyte's soluble chromatin phosphorylated *ex vivo* overnight. The spectrum corresponds to the region in which the peptide of the N-terminal domain plus the globular domain of histones was detected.



1. NH1.1Rac0/NH1.02ac0 (10322,05)
2. NH1.01 ac1 (10337,96)
3. NH1.1Rac1/NH1.02ac1 (10363,05)
4. NH1.03ac0 (10593,37)
5. NH1.1Lac0/NH1.03ac1 (10635,36)
6. NH1.1L ac1 (10676,36)
7. NH1.1Lac0 -1P/ NH1.03ac0-1P (10714,36)

MALDI spectrum of *alfa*-chymotrypsin cleaved H1 histone fraction extracted from chicken erythrocyte's soluble chromatin phosphorylated *ex vivo* for 5 hours. The spectrum corresponds to the region in which the peptide of the C-terminal domain of histones was detected.



1. CH1.1R (11370,1)
2. CH1.1R -1P/CH1.02 (11444,6)
3. CH1.02 -1P (11524.18)
4. CH1.01 (11636,4)
5. CH1.02 – 3P (11684,18)
6. CH1.01 -1P (11715,37)
7. CH1 .1L (11781,6)
8. CH1.1L- 1P (11860,65)
9. CH1.1L- 2P (11940,65)

References

[1] Tsuzuki J, Loeb J. In vitro phosphorylation of chicken erythrocyte histone by various protein kinases: absence of phosphorylation from F1 histone. Exp Cell Res. 1974. 88(2):303-10.

[2] Carruthers LM, Hansen JC. The core histone N termini function independently of linker histones during chromatin condensation. J Biol Chem. 2000. 275(47):37285-90.

[3] Woodcock CL, Ghosh RP. Chromatin Higher-order Structure and Dynamics. Cold Spring Harb Perspect Biol. 2010. 2(5): a000596.

[4] Wolffe, AP. Chromatin Structure and Function. San Diego, CA: Academic Press, 1992. ISBN 012 761 9119

[5] Paux E, Sourdille P, Salse J, Saintenac C, Choulet F, Leroy P, Korol A, Michalak M, Kianian S, Spielmeyer W, Lagudah E, Somers D, Kilian A, Alaux M, Vautrin S, Bergès H, Eversole K, Appels R, Safar J, Simkova H, Dolezel J, Bernard M, Feuillet C. A Physical Map of the 1-Gigabase Bread Wheat Chromosome 3B. Science. 2008 Oct 3;322(5898):101-4.

[6] van Holde K, Zlatanova J, Arents G, Moudrianakis E. Elements of chromatin structure: histones, nucleosomes, and fibres. Chromatin Structure and Gene Expression. Editor Sarah C. R. Elgin. IRL Press at Oxford University Press, 1995. ISBN 019 963 5757

[7] Anderson JD, Widom J. Sequence and Position-Dependence of the Equilibrium Accessibility of Nucleosomal DNA Target Sites. J Mol Biol. 2000 Mar 3;296(4):979-87.

[8] van Holde K, Zlatanova J. Scanning Chromatin: A New Paradigm? J Biol Chem. 2006. 5;281(18):12197-200. Review.

[9] Hewish DR, Burgoyne LA. Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. Biochem Biophys Res Commun. 1973 May 15;52(2):504-10

[10] White CL, Suto RK, Luger K. Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. EMBO J. 2001 Sep 17;20(18):5207-18

[11] Thoma F, Koller T, Klug A. Involvement of Histone H1 in the Organization of the Nucleosome and of the Salt-Dependent Superstructures of Chromatin. J Cell Biol. 1979 Nov;83(2 Pt 1):403-27.

[12] Woodcock CL, Grigoryev SA, Horowitz RA, Whitaker N. A Chromatin Folding Model that Incorporates Linker Variability Generates Fibres Resembling the Native Structures. Proc Natl Acad Sci U S A. 1993 Oct 1;90(19):9021-5.

[13] Wu C, Bassett A, Travers A. A variable topology for the 30-nm chromatin fibre. EMBO Rep. 2007 Dec;8(12):1129-34.

[14] Robinson PJJ, Fairall L, Huynh Van AT, Rhodes D. EM measurements define the dimensions of the "30-nm" chromatin fibre: Evidence for a compact, interdigitated structure. PNAS. 2006. 103 (17) 6506-6511

[15] Finch JT, Klug A. Solenoidal Model for Superstructure in Chromatin. Proc Natl Acad Sci U S A. 1976 Jun;73(6):1897-901.

[16] Felsenfeld G, McGhee JD. Structure of the 30 nm Chromatin Fibre. Cell. 1986 Feb 14;44(3):375-7

[17] Worcel A, Strogatz S, Riley D. Structure of chromatin and the linking number of DNA. Proc Natl Acad Sci U S A. 1981 Mar;78(3):1461-5

[18] Woodcock CL, Frado LL, Rattner JB. The Higher-Order Structure of Chromatin: Evidence for a Helical Ribbon Arrangement. J Cell Biol. 1984 Jul;99(1 Pt 1):42-52

[19] Williams SP, Athey BD, Muglia LJ, Schappe RS, Gough AH, Langmore JP. Chromatin fibres are left-handed double helices with diameter and mass per unit length that depend on linker length. Biophys J. 1986 Jan; 49(1):233-48

[20] Smith MF, Athey BD, Williams SP, Langmore JP. Radial density distribution of chromatin: evidence that chromatin fibres have solid centers. J Cell Biol. 1990 Feb;110(2):245-54

[21] Daban JR, Bermúdez A. Interdigitated solenoid model for compact chromatin fibres. Biochemistry. 1998 Mar 31;37(13):4299-304

[22] Laemmli UK, Kas E, Poljak L, Adachi Y. Scaffold-Associated Regions: Cis-acting Determinants of Chromatin Structural Loops and Functional Domains. Curr Opin Genet Dev. 1992 Apr;2(2):275-85. Review

[23] Fussner E, Ching RW, Bazett-Jones DP. Living without 30nm chromatin fibers. Trends Biochem Sci. 2011 Jan;36(1):1-6

[24] Bian Q, Belmont AS. Revisiting higher-order and large-scale chromatin organization. Curr Opin Cell Biol. 2012 Jun;24(3):359-66

[25] Richmond TJ, Finch JT, Rushton B, Rhodes D, Klug A. Structure of the nucleosome core particle at 7 Å resolution. Nature. 1984 Oct 11-17;311(5986):532-7.

[26] Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997 Sep 18;389(6648):251-60.

[27] Pruss D, Reeves R, Bushman FD, Wolffe AP. The influence of DNA and nucleosome structure on integration events directed by HIV integrase. J Biol Chem. 1994 Oct 7;269(40):25031-41

[28] Hagerman PJ. Sequence dependence of the curvature of DNA: a test of the phasing hypothesis. Biochemistry. 1985 Dec 3;24(25):7033-7.

[29] Klug A, Lutter LC. The helical periodicity of DNA on the nucleosome. Nucleic Acids Res. 1981 Sep 11;9(17):4267-83.

[30] Hayes JJ, Clark DJ, Wolffe AP. Histone contributions to the structure of DNA in the nucleosome. Proc Natl Acad Sci U S A. 1991 Aug 1;88(15):6829-33.

[31] Arents G, Moudrianakis EN. Topography of the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA. Proc Natl Acad Sci U S A. 1993 Nov 15;90(22):10489-93

[32] Arents G, Burlingame RW, Wang BC, Love WE, Moudrianakis EN. The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. Proc Natl Acad Sci U S A. 1991 Nov 15;88(22):10148-52.

[33] Wang BC, Rose J, Arents G, Moudrianakis EN. The octameric histone core of the nucleosome. Structural issues resolved. J Mol Biol. 1994 Feb 11;236(1):179-88

[34] Sivolob AV, Khrapunov SN. Translational positioning of nucleosomes on DNA: the role of sequence-dependent isotropic DNA bending stiffness. J Mol Biol. 1995 Apr 14;247(5):918-31

[35] Meersseman G, Pennings S, Bradbury EM. Chromatosome positioning on assembled long chromatin. Linker histones affect nucleosome placement on 5S rDNA. J Mol Biol. 1991 Jul 5;220(1):89-100

[36] Chipev CC, Wolffe AP. Chromosomal organization of Xenopus laevis oocyte and somatic 5S rRNA genes in vivo. Mol Cell Biol. 1992 Jan;12(1):45-55

[37] Ura K, Wolffe AP, Hayes JJ. Core histone acetylation does not block linker histone binding to a nucleosome including a Xenopus borealis 5S rRNA gene. J Biol Chem. 1994 Nov 4;269(44):27171-4

[38] Wang Q, Albert FG, Fitzgerald DJ, Calvo JM, Anderson JN. Sequence determinants of DNA bending in the ilvlH promoter and regulatory region of Escherichia coli. Nucleic Acids Res. 1994 Dec 25;22(25):5753-60

[39] Godde JS, Kass SU, Hirst MC, Wolffe AP. Nucleosome assembly on methylated CGG triplet repeats in the fragile X mental retardation gene 1 promoter. J Biol Chem. 1996 Oct 4;271(40):24325-8

[40] Thåström A, Lowary PT, Widlund HR, Cao H, Kubista M, Widom J. Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. J Mol Biol. 1999 Apr 30;288(2):213-29.

[41] Lowary PT, Widom J. New DNA sequence rules for high-affinity binding to histone octamer and sequence-directed nucleosome positioning. J Mol Biol. 1998;276:19–42.

[42] Routh A, Sandin S, Rhodes D. Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. Proc Natl Acad Sci U S A. 2008 Jul 1;105(26):8872-7.

[43] F. Dong, J.C. Hansen, K.E. van Holde. DNA and protein determinants of nucleosome positioning on sea urchin 5 S rRNA gene sequences in vitro. Proc. Natl Acad. Sci. USA, 87 (1990), pp. 5724–5728

[44] Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 A resolution. J Mol Biol. 2002 Jun 21;319(5):1097-113.

[45] Segal E, Fondufe-Mittendorf Y, Chen L, Thåström A, Field Y, Moore IK, Wang JP, Widom J. A genomic code for nucleosome positioning. Nature. 2006 Aug 17;442(7104):772-8

[46] Ward R, Bowman A, El-Mkami H, Owen-Hughes T, Norman DG. Long distance PELDOR measurements on the histone core particle. J Am Chem Soc. 2009 Feb 4;131(4):1348-9

[47] Cox M, Nelson DR, Lehninger AL. Lehninger Principles of Biochemistry. W.H. Freeman, San Francisco (2005). ISBN 0-7167-4339-6.

[48] Bhasin M, Reinherz EL, Reche PA. Recognition and classification of histones using support vector machine. J Comput Biol. 2006 Jan-Feb; 13(1):102-12.

[49] van Holde K. Chromatin. Springer-Verlag, New York (1988)

[50] Wells D, McBride C. A comprehensive compilation and alignment of histones and histone genes. Nucleic Acids Res. 1989; 17 Suppl:r311-46

[51] Wells D, Brown D. Histone and histone gene compilation and alignment update. Nucleic Acids Res. 1991 Apr 25;19 Suppl:2173-88.

[52] Finch JT, Lutter LC, Rhodes D, Brown RS, Rushton B, Levitt M, Klug A. Structure of nucleosome core particles of chromatin. Nature. 1977 Sep 1;269(5623):29-36.

[53] Klug A, Rhodes D, Smith J, Finch JT, Thomas JO. A low resolution structure for the histone core of the nucleosome. Nature. 1980 Oct 9;287(5782):509-16.

[54]Burlingame RW, Love WE, Wang BC, Hamlin R, Nguyen HX, Moudrianakis EN. Crystallographic structure of the octameric histone core of the nucleosome at a resolution of 3.3 A. Science. 1985 May 3;228(4699):546-53.

[55] Isenberg I. Histones. Annu Rev Biochem. 1979;48:159-91. Review.

[56]Godfrey JE, Eickbush TH, Moudrianakis EN. Reversible association of calf thymus histones to form the symmetrical octamer (H2AH2BH3H4)2: a case of a mixed-associating system. Biochemistry. 1980 Apr 1;19(7):1339-46

[57] Luger, K. Nucleosomes: Structure and Function. eLS (2001)

[58] Kuo MH, Brownell JE, Sobel RE, Ranalli TA, Cook RG, Edmondson DG, Roth SY, Allis CD. Transcription-Linked Acetylation by Gcn5p of Histones H3 and H4 at Specific Lysines. Nature. 1996. 383, 269-272.

[59] Grunstein M, Hecht A, Fisher-Adams G, Wan J, Mann RK, Strahl-Bolsinger S, Laroche T, Gasser S. The Regulation of Euchromatin and Heterochromatin by Histones in Yeast. J. Cell Sci. Suppl. 1995. 19, 29-36.

[60] Jenuwein T, Allis CD. Translating the Histone Code. Science. 2001. 293, 1074-1080.

[61] Turner BM. Cellular Memory and the Histone Code. Cell. 2002. 111, 285-291.

[62] Sawan C, Herceg Z. Histone modifications and cancer. Adv Genet. 2010;70:57-85. Review

[63] Fisher-Adams G, Grunstein M. Yeast Histone H4 and H3 N-Termini have Different Effects on the Chromatin Structure of the GAL1 Promoter. EMBO J. 1995. 14, 1468-1477.

[64] Kouzarides T. Chromatin Modifications and their Function. Cell. 2007. 128, 693-705.

[65] Zhang L, Eugeni EE, Parthun MR, Freitas MA. Identification of Novel Histone Post-Translational Modifications by Peptide Mass Fingerprinting. Chromosoma. 2003. 112, 77-86.

[66] Zhang K, Tang H, Huang L, Blankenship JW, Jones PR, Xiang F, Yau PM, Burlingame AL. Identification of Acetylation and Methylation Sites of Histone H3 from Chicken Erythrocytes by High-Accuracy Matrix-Assisted Laser Desorption Ionization-Time-of-Flight, Matrix-Assisted Laser Desorption Ionization-Postsource Decay, and Nanoelectrospray Ionization Tandem Mass Spectrometry. Anal. Biochem. 2006. 306, 259-269.

[67] Zhang K, Williams KE, Huang L, Yau P, Siino JS, Bradbury EM, Jones PR, Minch MJ, Burlingame AL. Histone Acetylation and Deacetylation: Identification of Acetylation and Methylation Sites of HeLa Histone H4 by Mass Spectrometry. Mol. Cell. Proteomics. 2002. 1, 500-508.

[68] Franklin TB, Mansuy IM. The prevalence of epigenetic mechanisms in the regulation of cognitive functions and behaviour. Curr Opin Neurobiol. 2010. 20(4):441-9

[69] Esteller M. Cancer Epigenomics: DNA Methylomes and Histone-Modification Maps. Nat. Rev. Genet. 2007. 8, 286-298.

[70] Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. Nat. Rev. Cancer 2007. 7(11), 823–33.

[71] Li X, Zhao X. Epigenetic regulation of mammalian stem cells. Stem Cells Dev. 2008. 17(6), 1043–52.

[72] Swigut T, Wysocka J. H3K27 demethylases, at long last. Cell. 2007. 131(1), 29–32

[73] Shilatifard A. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. Curr. Opin. Cell Biol. 2008. 20(3), 341–8.

[74] Varier RA, Timmers HT. Histone Lysine Methylation and Demethylation Pathways in Cancer. Biochim. Biophys. Acta. 2011. 1815, 75-89.

[75] Fischle W, Wang Y, Allis CD. Binary Switches and Modification Cassettes in Histone Biology and Beyond. Nature. 2003. 425, 475-479.

[76] Santos-Rosa H, Schneider R, Bernstein BE, Karabetsou N, Morillon A, Weise C, Schreiber SL, Mellor J, Kouzarides T. Methylation of Histone H3 K4 Mediates Association of the Isw1p ATPase with Chromatin. Mol. Cell. 2003. 12, 1325-1332.

[77] 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. J. Biol. Chem. 2005. 280, 17732-17736.

[78] Nelson CJ, Santos-Rosa H, Kouzarides T. Proline isomerization of histone H3 regulates lysine methylation and gene expression. Cell. 2006. 126(5):905-16.

[79] Sanders SL, Portoso M, Mata J, Bahler J, Allshire RC, and Kouzarides T. Methylation of Histone H4 Lysine 20 Controls Recruitment of Crb2 to Sites of DNA Damage. Cell. 2004. 119, 603-614.

[80] Csordas A. On the Biological Role of Histone Acetylation. Biochem. J. 1990.265, 23-38.

[81] Turner BM. Histone Acetylation and Control of Gene Expression. J. Cell. Sci. 1991. 99 (Pt 1), 13-20.

[82] Brownell JE, Allis CD. Special HATs for Special Occasions: Linking Histone Acetylation to Chromatin Assembly and Gene Activation. Curr. Opin. Genet. Dev. 1996. 6, 176-184.

[83] Taunton J, Hassig CA, Schreiber SL. A Mammalian Histone Deacetylase Related to the Yeast Transcriptional Regulator Rpd3p. Science. 1996. 272, 408-411.

[84] Hayashi R, Wada H, Ito K, Adcock IM. Effects of glucocorticoids on gene transcription. Eur J Pharmacol. 2004 Oct 1;500(1-3):51-62. Review.

[85] Prevelige PE Jr, Fasman GD. Structural Studies of Acetylated and Control Inner Core Histones. Biochemistry. 1987. 26, 2944-2955.

[86] Wang X, Moore SC, Laszckzak M, Ausio J. Acetylation Increases the Alpha-Helical Content of the Histone Tails of the Nucleosome. J. Biol. Chem.2000. 275, 35013-35020.

[87] Strahl BD, Allis CD. The Language of Covalent Histone Modifications. Nature. 2000. 403, 41-45.

[88] Sung MT, Dixon GH. Modification of Histones during Spermiogenesis in Trout: A Molecular Mechanism for Altering Histone Binding to DNA. Proc. Natl. Acad. Sci. U. S. A. 1970. 67, 1616-1623.

[89] Oliva R, Bazett-Jones DP, Locklear L and Dixon G. Histone hyperacetylation can induce unfolding of the nucleosome core particle. Nucleic Acids Res. 1990. 18:2739-2747.

[90] Wolffe AP, Hayes JJ. Chromatin Disruption and Modification. Nucleic Acids Res. 1999. 27, 711-720.

[91] Ura K, Kurumizaka H, Dimitrov S, Almouzni G, Wolffe AP. Histone Acetylation: Influence on Transcription, Nucleosome Mobility and Positioning, and Linker Histone-Dependent Transcriptional Repression. EMBO J. 1997. 16, 2096-2107.

[92] Khochbin S, Wolffe AP. Developmentally Regulated Expression of Linker-Histone Variants in Vertebrates. Eur. J. Biochem. 1994. 225, 501-510.

[93] Maas NL, Miller KM, DeFazio LG, Toczyski DP. Cell Cycle and Checkpoint Regulation of Histone H3 K56 Acetylation by Hst3 and Hst4. Mol. Cell. 2006. 23, 109-119.

[94] Qin S, Parthun MR. Recruitment of the Type B Histone Acetyltransferase Hat1p to Chromatin is Linked to DNA Double-Strand Breaks. Mol. Cell. Biol. 2006. 26, 3649-3658.

[95] Ozdemir A, Spicuglia S, Lasonder E, Vermeulen M, Campsteijn C, Stunnenberg HG, Logie C. Characterization of lysine 56 of histone H3 as an acetylation site in Saccharomyces cerevisiae. J Biol Chem. 2005. 280 (28): 25949–52.

[96] Masumoto H, Hawke D, Kobayashi R, Verreault A. A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. Nature. 2005. 436 (7048): 294–8.

[97]Driscoll R, Hudson A, Jackson SP. Yeast Rtt109 promotes genome stability by acetylating histone H3 on lysine 56. Science. 2007. 315 (5812): 649–52.

[98] Han J, Zhou H, Horazdovsky B, Zhang K, Xu RM, Zhang Z. Rtt109 acetylates histone H3 lysine 56 and functions in DNA replication. Science. 2007. 315 (5812): 653–5.

[99] Das C, Lucia MS, Hansen KC, Tyler JK. CBP/p300-mediated acetylation of histone H3 on lysine 56. Nature. 2009. 459 (7243): 113–7.

[100] Han J, Zhou H, Li Z, Xu RM, Zhang Z. Acetylation of lysine 56 of histone H3 catalyzed by RTT109 and regulated by ASF1 is required for replisome integrity. J Biol Chem. 2007. 282 (39): 28587–96.

[101] Wurtele H, Kaiser GS, Bacal J, St-Hilaire E, Lee EH, Tsao S, Dorn J, Maddox P, Lisby M, Pasero P, Verreault A. Histone H3 lysine 56 acetylation and the response to DNA replication fork damage. Mol Cell Biol. 2012. 32 (1): 154–72.

[102] Wurtele H, Tsao S, Lépine G, Mullick A, Tremblay J, Drogaris P, Lee EH, Thibault P, Verreault A, Raymond M. Modulation of histone H3 lysine 56 acetylation as an antifungal therapeutic strategy. Nat Med. 2010. 16 (7): 774–80.

[103] Wolffe AP. Chromatin, Third Edition: Structure and Function. Academic Press. 1998. ISBN-0127619151

[104] Mahadevan LC, Willis AC, Barratt MJ. Rapid Histone H3 Phosphorylation in Response to Growth Factors, Phorbol Esters, Okadaic Acid, and Protein Synthesis Inhibitors. Cell. 1991. 65, 775-783.

[105] Wei Y, Mizzen CA, Cook RG, Gorovsky MA, Allis CD. Phosphorylation of Histone H3 at Serine 10 is Correlated with Chromosome Condensation during Mitosis and Meiosis in Tetrahymena. Proc. Natl. Acad. Sci. U. S. A.1998. 95, 7480-7484.

[106] Murnion ME, Adams RR, Callister DM, Allis CD, Earnshaw WC, Swedlow JR. Chromatin-Associated Protein Phosphatase 1 Regulates Aurora-B and Histone H3 Phosphorylation. J. Biol. Chem. 2001. 276, 26656-26665.

[107] Chadee DN, Peltier CP, Davie JR. Histone H1(S)-3 Phosphorylation in Ha-Ras Oncogene-Transformed Mouse Fibroblasts. Oncogene. 2002. 21, 8397-8403.

[108] Hsu JY, Sun ZW, Li X, Reuben M, Tatchell K, Bishop DK, Grushcow JM, Brame CJ, Caldwell JA, Hunt DF, Lin R, Smith MM, Allis CD. Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes.Cell. 2000;102(3):279-91.

[109] Giet R, Glover DM. Drosophila Aurora B kinase is required for histone H3 phosphorylation and condensin recruitment during chromosome condensation and to organize the central spindle during cytokinesis. J. Cell Biol. 2001. 152, 669–681

[110] de la Barre AE, Gerson V, Gout S, Creaven M, Allis CD, Dimitrov S. Core histone N-termini play an essential role in mitotic chromosome condensation. EMBO J. 2000. 19, 379-391.

[111] De Souza CP, Osmani AH, Wu LP, Spotts JL, Osmani SA. Mitotic histone H3 phosphorylation by the NIMA kinase in Aspergillus nidulans. Cell. 2000. 102, 293-302.

[112] Van Hooser A, Goodrich DW, Allis CD, Brinkley BR, Mancini MA. Histone H3 phosphorylation is required for the initiation, but not maintenance, of mammalian chromosome condensation. J. Cell Sci. 1998. 111, 3497-3506.

[113] Crosio C, Fimia GM, Loury R, Kimura M, Okano Y, Zhou H, Sen S, Allis CD, Sassone-Corsi P. Mitotic phosphorylation of histone H3: spatio-temporal regulation by mammalian Aurora kinases. Mol. Cell. Biol. 2002. 22, 874–885.

[114] Kallio MJ, McCleland ML, Stukenberg PT, Gorbsky GJ. Inhibition of aurora B kinase blocks chromosome segregation, overrides the spindle checkpoint, and perturbs microtubule dynamics in mitosis. Curr. Biol. 2002. 12, 900–905.

[115] Ota T, Suto S, Katayama H, Han ZB, Suzuki F, Maeda M, Tanino M, Terada Y, Tatsuka M. Increased Mitotic Phosphorylation of Histone H3 Attributable to AIM-1/Aurora-B Overexpression Contributes to Chromosome Number Instability. Cancer Res. 2002 62, 5168-5177); (Katayama H, Brinkley WR, Sen S. The Aurora Kinases: Role in Cell Transformation and Tumorigenesis. Cancer Metastasis Rev. 2003. 22, 451-464.

[116] Nowak SJ, Corces VG. Phosphorylation of Histone H3 Correlates with Transcriptionally Active Loci. Genes Dev.2000. 14, 3003-3013.

[117]Sung MT, Dixon GH. Modification of Histones during Spermiogenesis in Trout: A Molecular Mechanism for Altering Histone Binding to DNA. Proc. Natl. Acad. Sci. U. S. A. 1970. 67, 1616-1623.

[118] Jackson V, Shires A, Tanphaichitr N, Chalkley R. Modifications to Histones Immediately After Synthesis. J. Mol. Biol.1976. 104, 471-483.

[119] Dimitrov S, Dasso MC, Wolffe, AP. Remodeling Sperm Chromatin in Xenopus Laevis Egg Extracts: The Role of Core Histone Phosphorylation and Linker Histone B4 in Chromatin Assembly. J. Cell Biol. 1994. 126, 591-601.

[120] Kaufman PD, Botchan MR. Assembly of Nucleosomes: Do Multiple Assembly Factors Mean Multiple Mechanisms? Curr. Opin. Genet. Dev.1994. 4, 229-235.

[121] Wade PA, Pruss D, Wolffe AP. Histone Acetylation: Chromatin in Action. Trends Biochem. Sci. 1997. 22, 128-132.

[122] Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. J Biol Chem.1998. 273 (10): 5858–68.

[123] Paull TT, Rogakou EP, Yamazaki V, Kirchgessner CU, Gellert M, Bonner WM. A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. Curr. Biol. 2000. 10 (15): 886–95.

[124] Celeste A, Petersen S, Romanienko PJ, Fernandez-Capetillo O, Chen HT, Sedelnikova OA, Reina-San-Martin B, Coppola V, Meffre E, Difilippantonio MJ, Redon C, Pilch DR, Olaru A, Eckhaus M, Camerini-Otero RD, Tessarollo L, Livak F, Manova K, Bonner WM, Nussenzweig MC, Nussenzweig A. Genomic instability in mice lacking histone H2AX. Science 2002. 296 (5569): 922–7.

[125] Shroff R, Arbel-Eden A, Pilch D, Ira G, Bonner WM, Petrini JH, Haber JE, Lichten M. Distribution and dynamics of chromatin modification induced by a defined DNA double-strand break. Curr Biol. 2004 Oct 5;14(19):1703-11.

[126] Rogakou EP, Boon C, Redon C, Bonner WM. Megabase chromatin domains involved in DNA double-strand breaks in vivo. J Cell Biol. 1999. 146 (5): 905–16.

[127] Stewart GS, Wang B, Bignell CR, Taylor AM, Elledge SJ. MDC1 is a mediator of the mammalian DNA damage checkpoint. Nature. 2003. 421 (6926): 961–6.

[128] Bekker-Jensen S, Mailand N. Assembly and function of DNA double-strand break repair foci in mammalian cells. DNA Repair (Amst). 2010 Dec 10;9(12):1219-28. Review.

[129] Kimura Y, Tanaka K. Regulatory mechanisms involved in the control of ubiquitin homeostasis. J Biochem 2010. 147 (6): 793–8.

[130] Hochstrasser M. Origin and Function of Ubiquitin-like Protein Conjugation. Nature. 2009. 458 (7237): 422–429.

[131] West MH, Bonner WM. Histone 2B can be Modified by the Attachment of Ubiquitin. Nucleic Acids Res.1980. 8, 4671-4680.

[132] Kleinschmidt AM, Martinson HG. Structure of Nucleosome Core Particles Containing uH2A (A24). Nucleic Acids Res.1981. 9, 2423-2431.

[133] Gushchin DI, Ebralidze, KK, Mirzabekov AD. Structure of Nucleosomes. Localization of the H2A and H2B Histone Segments Interacting with DNA using DNA-Protein Crosslinking. Mol. Biol. (Mosk).1991. 25, 1400-1411.

[134] Parseghian MH, Henschen AH, Krieglstein KG, Hamkalo BA. A Proposal for a Coherent Mammalian Histone H1 Nomenclature Correlated with Amino Acid Sequences. Protein Sci.1994. 3, 575-587.

[135] Moore SC, Jason L and Ausió J. The elusive structural role of ubiquitinated histones. Biochem Cell Biol. 2002. 80(3):311-9.

[136] Hay RT. SUMO: a history of modification Mol. Cell. 2005. 18 (1): 1–12. Review.

[137] Nathan D, Ingvarsdottir K, Sterner DE, Bylebyl GR, Dokmanovic M, Dorsey JA, Whelan KA, Krsmanovic M, Lane WS, Meluh PB, Johnson ES, Berger SL. Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications. Genes Dev.2006. 20 (8): 966–76.

[138] Cuthbert GL, Daujat S, Snowden AW, Erdjument-Bromage H, Hagiwara T, Yamada M, Schneider R, Gregory PD, Tempst P, Bannister AJ, Kouzarides T.Histone Deimination Antagonizes Arginine Methylation. Cell. 2004. 118, 545-553.

[139] Belenky P, Bogan KL, Brenner C. NAD+ metabolism in health and disease. Trends Biochem. Sci. 2007. 32 (1): 12–9.

[140] Corda D, Di Girolamo M. NEW EMBO MEMBER'S REVIEW: Functional aspects of protein mono-ADP-ribosylation. EMBO J. 2003. 22 (9): 1953–8.

[141] Mathis G, Althaus FR. Uncoupling of DNA Excision Repair and Nucleosomal Unfolding in Poly(ADP-Ribose)-Depleted Mammalian Cells. Carcinogenesis. 1990. 11, 1237-1239.

[142] Allan J, Hartman PG, Crane-Robinson C, Aviles FX. The structure of histone H1 and its location in chromatin. Nature 1980; 288:675–679.

[143] Furrer P, Bednar J, Dubochet J, Hamiche A, Prunell A. DNA at the entry-exit of the nucleosome observed by cryoelectron microscopy. J. Struct. Biol. 1995; 114:177–183.

[144] Sivolob A, Prunell A, Linker histone-dependent organization and dynamics of nucleosome entry/exit DNAs. A. J. Mol. Biol. 2003; 331:1025–1040.

[145] Langowski J. Polymer Chain Models of DNA and Chromatin. Eur. Phys. J. E. Soft Matter. 2006.19, 241-249.

[146] Thomas JO, Rees C, Finch JT. Cooperative Binding of the Globular Domains of Histones H1 and H5 to DNA. Nucleic Acids Res. 1992. 20, 187-194.

[147] Dou Y, Mizzen CA, Abrams M, Allis CD, Gorovsky MA. Phosphorylation of Linker Histone H1 Regulates Gene Expression in Vivo by Mimicking H1 Removal. Mol. Cell.1999. 4, 641-647.

[148] Zlatanova J, Caiafa P, van Holde K. Linker Histone Binding and Displacement: Versatile Mechanism for Transcriptional Regulation. FASEB J.2000. 14, 1697-1704.

[149] Koop R, Di Croce L, Beato M. Histone H1 Enhances Synergistic Activation of the MMTV Promoter in Chromatin. EMBO J. 2003. 22, 588-599.

[150] Pennings, S., Meersseman, G., Bradbury EM. Linker Histones H1 and H5 Prevent the Mobility of Positioned Nucleosomes. Proc. Natl. Acad. Sci. U. S. A. 1994. 91, 10275-10279.

[151] Izaurralde E, Kas E, Laemmli UK. Highly Preferential Nucleation of Histone H1 Assembly on Scaffold-Associated Regions. J. Mol. Biol. 1989. 210, 573-585.

[152] Dimitrov S, Almouzni G, Dasso M, Wolffe AP. Chromatin transitions during early Xenopus embryogenesis: Changes in histone H4 acetylation and in linker histone type. Dev. Biol. 1993; 160:214–227.

[153] Lennox RW, Oshima RG, Cohen LH. The H1 histones and their interphase phosphorylated states in differentiated and undifferentiated cell lines derived from murine teratocarcinomas. J Biol Chem 1982; 257:5183-5189.

[154] Hall JM, Cole RD. Modulation in proportions of H1 histone subfractions by differential changes in synthesis and turnover during butyrate treatment of neuroblastoma cells. Biochemistry 1985; 24:7765-7771.

[155] Domínguez V, Piña B, Suau P. Histone H1 subtype synthesis in neurons and neuroblasts. Development 1992; 115:181-185.

[156] Liao LW, Cole RD. Differences among fractions of H1 histones in their interactions with linear and superhelical DNA: circular dichroism. J Biol Chem 1981; 256:10124-10128.

[157] Khadake JR, Rao MR. DNA- and chromatincondensing properties of rat testes H1a and H1t compared to those of rat liver H1bdec; H1t is a poor condenser of chromatin. Biochemistry 1995; 34:15792-15801.

[158] Talasz H, Sapojnikova N, Helliger W, Lindner H, Puschendorf B. In vitro binding of H1 histone subtypes to nucleosomal organized mouse mammary tumor virus long terminal repeat promotor. J Biol Chem 1998; 273:32236-32243.

[159] Gunjan A, Alexander BT, Sittman DB, Brown DT. Effects of histone H1 variant overexpression on chromatin structure. J Biol Chem 1999; 274:37950-37956.

[160] Ponte I, Vidal-Taboada JM, Suau P. Evolution of the H1 histone class: evidence for the functional differentiation of the subtypes. Mol Biol Evol 1998; 15:702-708.

[161] Brown DT, Alexander BT, Sittman DB. Differential effect of H1 variant overexpression on cell cycle progression and gene expression. Nucleic Acids Res 1996; 24:486-493.

[162] Steinbach OC, Wolffe AP, Rupp RA. Somatic linker histones cause loss of mesodermal competence in Xenopus. Nature1997; 389:395-399.

[163] Albig W, Meergans T, Doenecke D. Characterization of the H1.5 gene completes the set of human H1 subtype genes. Gene. 1977; 184: 141–148.

[164] Eick S, Nicolai M, Mumberg D, Doenecke D. Human H1 Histones: Conserved and Varied Sequence Elements in Two H1 Subtype Genes. Eur. J. Cell Biol.1989. 49, 110-115

[165] Lennox RW. Differences in Evolutionary Stability among Mammalian H1 Subtypes. Implications for the Roles of H1 Subtypes in Chromatin. J. Biol. Chem. 1984. 259, 669-672.

[166] Lennox RW, Cohen LH. The histone H1 complements of dividing and nondividing cells of the mouse. J. Biol. Chem. 1983; 258: 62–268.

[167] Meergans T, Albig W, Doenecke D. Varied expression patterns of human H1 histone genes in different cell lines. DNA Cell Biol.1997; 16:1041–1049.

[168] Dong Y, Sirotkin AM, Yang YS, Brown DT, Sittman DB, Skoultchi AI. Isolation and characterization of two replication-dependent mouse H1 histone. Nucleic Acids Res.1994; 22:1421–1428.

[169] Franke K, Drabent B, Doenecke D. Expression of murine H1 histone genes during postnatal development. Biochim. Biophys. Acta 1998; 1398: 232–242.

[170] Orrego M, Ponte I, Roque A, Buschati N, Mora X, Suau P. Differential affinity of mammalian histone H1 somatic subtypes for DNA and chromatin. BMC Biol. 2007. 5(1):22.

[171] Jiang T, Zhou X, Taghizadeh K, Dong M, Dedon PC. N-formylation of lysine in histone proteins as a secondary modification arising from oxidative DNA damage. Proc. Natl. Acad. Sci. U. S. A. 2007; 104:60–65.

[172] Alami R, Fan Y, Pack S, Sonbuchner TM, Besse A, Lin Q, Greally JM, Skoultchi AI, Bouhassira EE. Mammalian linker-histone subtypes differentially affect gene expression in vivo. Proc Natl Acad Sci USA 2006; 100:5920-5925.

[173] Parseghian MH, Newcomb RL, Winokur ST, Hamkalo BA. The distribution of somatic H1 subtypes is non-random in active vs. inactive chromatin. Chromosome Res 2000; 8:405-424.

[174] Kamakaka RT, Thomas JO. Chromatin structure of transcriptionally competent and repressed genes. EMBO J 1990; 9:3997-4006.

[175] Wolffe AP. Histone H1. Int. J. Biochem. Cell Biol. 1997. 29, 1463-1466.

[176] Tanaka M, Hennebold JD, Macfarlane J, Adashi EY. A Mammalian Oocyte-Specific Linker Histone Gene H1oo: Homology with the Genes for the Oocyte-Specific Cleavage Stage Histone (Cs-H1) of Sea Urchin and the B4/H1M Histone of the Frog. Development. 2001. 128, 655-664.

[177] Tanaka Y, Kato S, Tanaka M, Kuji N, Yoshimura Y. Structure and Expression of the Human Oocyte-Specific Histone H1 Gene Elucidated by Direct RT-Nested PCR of a Single Oocyte. Biochem. Biophys. Res. Commun. 2003. 304, 351-357.

[178] Martianov I, Brancorsini S, Catena R, Gansmuller A, Kotaja N, Parvinen M, Sassone-Corsi P, Davidson I. Polar Nuclear Localization of H1T2, a Histone H1 Variant, Required for Spermatid Elongation and DNA Condensation during Spermiogenesis. Proc. Natl. Acad. Sci. U. S. A. 2005. 102, 2808-2813.

[179] Tanaka H, Iguchi N, Isotani A., Kitamura K, Toyama Y, Matsuoka Y, Onishi M, Masai K, Maekawa M, Toshimori K, Okabe M, Nishimune Y. HANP1/H1T2, a Novel Histone H1-Like Protein Involved in Nuclear Formation and Sperm Fertility. Mol. Cell. Biol. 2005. 25, 7107-7119.

[180] Yan W, Ma L, Burns KH, Matzuk MM. HILS1 is a Spermatid-Specific Linker Histone H1-Like Protein Implicated in Chromatin Remodeling during Mammalian Spermiogenesis. Proc. Natl. Acad. Sci. U. S. A. 2003. 100, 10546-10551.

[181] Wang ZF, Sirotkin AM, Buchold GM, Skoultchi AI, Marzluff WF. The Mouse Histone H1 Genes: Gene Organization and Differential Regulation. J. Mol. Biol. 1997. 271, 124-138.

[182] Albig W, Drabent B, Kunz J, Kalff-Suske M, Grzeschik KH, Doenecke, D. All Known Human H1 Histone Genes Except the H1(0) Gene are Clustered on Chromosome 6. Genomics. 1993. 16, 649-654.

[183] 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 1Mb. Genome Res.1996. 6, 688-701.

[184] Happel N, Schulze E, Doenecke D. Characterisation of Human Histone H1x. Biol. Chem. 2005. 386, 541-551.

[185] Brannan CI, Gilbert DJ, Ceci JD, Matsuda Y, Chapman VM, Mercer JA, Eisen H, Johnston LA, Copeland NG, Jenkins NA. An Interspecific Linkage Map of Mouse Chromosome 15 Positioned with Respect to the Centromere. Genomics. 1992. 13, 1075-1081.

[186] Khochbin S. Histone H1 diversity: bridging regulatory signals to linker histone function. Gene. 2001; 271 (1): 1-12.

[187] Graziano V, Gerchman SE, Schneider DK, Ramakrishnan V. Histone H1 is located in the interior of the chromatin 30-nm filament. Nature 1994; 368 (6469):351-354.

[188] Routh A, Sandin S, Rhodes D. Nucleosome repeat length and linker histone stoichiometry determine chromatin fibre structure. Proc Natl Acad Sci U S A. 2008. 105(26):8872-7.

[189] Fan Y, Nikitina T, Zhao J, Fleury TJ, Bhattacharyya R, Bouhassira EE, Stein A, Woodcock CL, Skoultchi AI. Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. Cell 2005. 123:1199-1212.

[190] Öberg C, Izzo A, Schneider R, Wrange Ö, Belikov S. Linker Histone Subtypes Differ in Their Effect on Nucleosomal Spacing In Vivo. J Mol Biol. 2012. 419(3-4):183-97.

[191] Thomas JO. Histone H1: location and role. Curr Opin Cell Biol. 1999. 11(3):312-7.

[192] Noll M, Kornberg RD. Action of micrococcal nuclease on chromatin and the location of histone H1. J. Mol. Biol. 1977. 109.:393–404.

[193] Woodcock CL, Skoultchi AI, Fan Y. Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. Chromosome Res. 2006. 14.17–25.

[194] Hashimoto H, Takami Y, Sonoda E, Iwasaki T, Iwano H, Tachibana M, Takeda S, Nakayama T, Kimura H, Shinkai Y. Histone H1 null vertebrate cells exhibit altered nucleosome architecture. Nucleic Acids Res. 2010. 38.:3533–3545

[195] Perišić O, Collepardo-Guevara R, Schlick T. Modeling studies of chromatin fibre structure as a function of DNA linker length. J Mol Biol. 2010. 403(5):777-802.

[196] Dorigo B, Schalch T, Kulangara A, Duda S, Schroeder RR, Richmond TJ. Nucleosome Arrays Reveal the Two-Start Organization of the Chromatin Fibre. Science. 2004. 306, 1571-1573.

[197] Sancho M, Diani E, Beato M, Jordan A. Depletion of human histone H1 variants uncovers specific roles in gene expression and cell growth. PLoS Genet. 2008;4(10):e1000227. Epub 2008 Oct 17

[198] Schlissel MS, Brown DD, The transcriptional regulation of Xenopus 5S RNA genes in chromatin: The roles of active stable transcription complexes and histone H1. Cell 1984; 37:903–913.

[199] Laybourn PJ, Kadonaga JT, Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II. Science (Washington, D.C.) 1991; 254:238–245.

[200] Zlatanova J, Van Holde K, Histone H1 and transcription: still an enigma? J Cell Sci 1992; 103:889-895.

[201] Bouvet P, Dimitrov S, Wolffe AP. Specific regulation of Xenopus chromosomal 5S rRNA gene transcription in vivo by histone H1. Genes Dev 1994; 8:1147-1159.

[202] Shen X, Gorovsky MA, Linker histone H1 regulates specific gene expression but not global transcription in vivo. Cell 1996; 86:475-483.

[203] Vermaak D, Steinbach OC, Dimitrov S, Rupp RA, Wolffe AP. The globular domain of histone H1 is sufficient to direct specific gene repression in early Xenopus embryos. Curr Biol 1998; 8:533-536.

[204] Koop R, Di Croce L, Beato M. Histone H1 enhances synergistic activation of the MMTV promoter in chromatin. EMBO J 2002; 22:588-599.

[205] Roque A, Orrego M, Ponte I, Suau P. The preferential binding of histone H1 to DNA scaffold-associated regions is determined by its C-terminal domain. Nucleic Acids Research 2004; 32 (20):6111-6119.

[206] Mirkovitch J, Mirault ME, Laemmli UK. Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. Cell. 1984. 39 (1): 223–32.

[207] Strick R, Laemmli UK. SARs are Cis DNA Elements of Chromosome Dynamics: Synthesis of a SAR Repressor Protein. Cell. 1995. 83, 1137-1148.

[208] Adams CC, Workman JL. Nucleosome Displacement in Transcription. Cell. 1993. 72, 305-308.

[209] Yang G, Leuba SH, Bustamante C, Zlatanova J, van Holde K. Role of Linker Histones in Extended Chromatin Fibre Structure. Nat. Struct. Biol. 1994. 1, 761-763.

[210] Guo CY, Mizzen C, Wang Y, Larner JM. Histone H1 and H3 dephosphorylation are differentially regulated by radiation-induced signal transduction pathways. Cancer Res. 2000. 60:5667–5672.

[211] Talasz H, Helliger W, Sarg B, Debbage PL, Puschendorf B, Lindner H. Hyperphosphorylation of histone H2A.X and dephosphorylation of histone H1 subtypes in the course of apoptosis. Cell Death Differ. 2002; 9,:27–39.

[212] Horn PJ, Carruthers LM, Logie C, Hill DA, Solomon MJ, Wade PA, Imbalzano AN, Hansen JC, Peterson CL. Phosphorylation of linker histones regulates ATP-dependent chromatin remodeling enzymes. Nature Struct. Biol. 2002; 9:263–267.

[213] Hartman PG, Chapman GE, Moss T, Bradbury EM. Studies on the role and mode of operation of the very-lysine-rich histone H1 in eukaryotic chromatin: the three structural regions of the histone H1 molecule. Eur. J.Biochem. 1977; 77:45-51.

[214] Kasinsky HE, Lewis JD, Dacks JB, Ausió J. Origin of H1 linker histones. FASEB J. 2001; 15:34–42.

[215] Patterton HG, Landel CC, Landsman D, Peterson CL, Simpson RT. The biochemical and phenotypical characterization of Hho1p, the putative linker histone H1 of Saccharomyces cerevisiae. J. Biol. Chem. 1998; 273:7268–7276.

[216] Ramesh S, Bharath MM, Chandra NR, Rao MR A K52Q Substitution in the Globular Domain of Histone H1t Modulates its Nucleosome Binding Properties. FEBS Lett.2006. 580, 5999-6006.

[217] Allan J, Mitchell T, Harborne N, Bohm L, Crane-Robinson C. Roles of H1 domains in determining higher order chromatin structure and H1 location. J Mol Biol. 1986;187(4):591-601.

[218] Vila R, Ponte I, Collado M, Arrondo JL, Jimenez MA, Rico M, Suau P. DNA-Induced Alpha-Helical Structure in the NH2-Terminal Domain of Histone H1. J. Biol. Chem. 2001. 276, 46429-46435.

[219] Vila R, Ponte I, Jimenez MA, Rico M, Suau P. An Inducible Helix-Gly-Gly-Helix Motif in the N-Terminal Domain of Histone H1e: A CD and NMR Study. Protein Sci. 2002. 11, 214-220.

[220] Cerf C, Lippens G, Muyldermans S, Segers A, Ramakrishnan V, Wodak SJ, Hallenga K, Wyns L. Homo- and Heteronuclear Two-Dimensional NMR Studies of the Globular Domain of Histone H1: Sequential Assignment and Secondary Structure. Biochemistry.1993. 32, 11345-11351.

[221] Ramakrishnan V, Finch JT, Graziano V, Lee PL, Sweet RM. Crystal Structure of Globular Domain of Histone H5 and its Implications for Nucleosome Binding. Nature. 1993. 362, 219-223.

[222] Zhou YB, Gerchman SE, Ramakrishnan V, Travers A, Muyldermans S. Position and Orientation of the Globular Domain of Linker Histone H5 on the Nucleosome. Nature.1998. 395, 402-405.

[223] Brown DT, Izard T, Misteli T. Mapping the Interaction Surface of Linker Histone H1(0) with the Nucleosome of Native Chromatin in Vivo. Nat. Struct. Mol. Biol. 2006. 13, 250-255.

[224] Subirana JA. Analysis of the Charge Distribution in the C-Terminal Region of Histone H1 as Related to its Interaction with DNA. Biopolymers.1990. 29, 1351-1357.

[225] Felsenfeld G, McGhee JD. Structure of the 30 Nm Chromatin Fibre. Cell. 1986. 44, 375-377.

[226] Butler PJ. A Defined Structure of the 30 Nm Chromatin Fibre which Accommodates Different Nucleosomal Repeat Lengths. EMBO J.1984. 3, 2599-2604.

[227] Clark DJ, Hill CS, Martin SR, Thomas JO. Alpha-Helix in the Carboxy-Terminal Domains of Histones H1 and H5. EMBO J.1988. 7, 69-75.

[228] Moran F, Rodriguez AT, Suau P, Montero F. Kinetic Analysis of Psi-DNA Structure Formation Induced by Histone H1 and its C-Terminal Domain. Biophys. Chem. 1981. 33, 133-141.

[229] Lu X, Hansen JC. Identification of Specific Functional Subdomains within the Linker Histone H10 C-Terminal Domain. J. Biol. Chem.2004. 279, 8701-8707.

[230] Bharath MM, Ramesh S, Chandra NR, Rao MR. Identification of a 34 Amino Acid Stretch within the C-Terminus of Histone H1 as the DNA-Condensing Domain by Site-Directed Mutagenesis. Biochemistry.2002. 41, 7617-7627.

[231] Vila R, Ponte I, Jimenez MA, Rico M, Suau P. A helix-turn motif in the C-terminal domain of histone H1. Protein Sci. 2000;9(4):627-36.

[232] Rall SC, Cole RD. Amino acid sequence and sequence variability of the amino-terminal regions of lysine-rich histones. J Biol Chem. 1971. 246(23):7175-90.

[233] von Holt C, de Groot P, Schwager S, Brandt W. Histone Genes. Wiley-Interscience, 1984. New York.

[234] Suzuki M, Gerstein M, Johnson T. An NMR Study on the DNA-Binding SPKK Motif and a Model for its Interaction with DNA. Protein Eng. 1993. 6, 565-574.

[235] Roque A, Iloro I, Ponte I, Arrondo JL, Suau P. DNA-Induced Secondary Structure of the Carboxyl-Terminal Domain of Histone H1. J. Biol. Chem.2005. 280, 32141-32147.

[236] Dyson HJ, Wright PE. Intrinsically Unstructured Proteins and their Functions. Nat. Rev. Mol. Cell Biol. 2005. 6, 197-208.

[237] Uversky VN. Natively Unfolded Proteins: A Point Where Biology Waits for Physics. Protein Sci. 2002. 11, 739-756.

[238] Uversky VN. Protein Folding Revisited. A Polypeptide Chain at the Folding-Misfolding-Nonfolding Cross-Roads: Which Way to Go? Cell Mol. Life Sci. 2003. 60, 1852-1871.

[239] Hill CS, Martin SR, Thomas JO. A Stable Alpha-Helical Element in the Carboxy-Terminal Domain of Free and Chromatin-Bound Histone H1 from Sea Urchin Sperm. EMBO J.1989. 8, 2591-2599.

[240] Erard M, Lakhdar-Ghazal F, Amalric F. Repeat Peptide Motifs which Contain Beta-Turns and Modulate DNA Condensation in Chromatin. Eur. J. Biochem. 1990. 191, 19-26.

[241] Jiang T, Zhou X, Taghizadeh K, Dong M, Dedon PC. N-formylation of lysine in histone proteins as a secondary modification arising from oxidative DNA damage. Proc. Natl. Acad. Sci. U. S. A. 2007. 104:60–65

[242] Wisniewski JR, Zougman A, Krüger S, Mann M. Mass spectrometric mapping of linker histone H1 variants reveals multiple acetylations, methylations, and phosphorylation as well as differences between cell culture and tissue. Mol. Cell. Proteomics. 2007. 6:72–87.

[243] Sarg B, Helliger W, Talasz H, Forg B, Lindner HH. Histone H1 phosphorylation occurs site-specifically during interphase and mitosis: identification of a novel phosphorylation site on histone H1. J Biol Chem. 2006. 281(10):6573-80.

[244] Green GR, Lee HJ, Poccia DL. Phosphorylation weakens DNA binding by peptides containing multiple "SPKK" sequences. J Biol Chem. 1993;268(15):11247-55

[245] InCiesla J, Fraczyk T, Rode W. Phosphorylation of basic amino acid residues in proteins: important but easily missed. Acta Biochim Pol. 2011. 58 (2): 137–47.

[246] Deutscher J, Saier MH Jr. Ser/Thr/Tyr protein phosphorylation in bacteria - for long time neglected, now well established. J Mol Microbiol Biotechnol. 2005. 9 (3–4): 125–31.

[247] Zor T, Mayr BM, Dyson HJ, Montminy MR, Wright PE. Roles of Phosphorylation and Helix Propensity in the Binding of the KIX Domain of CREB-Binding Protein by Constitutive (c-Myb) and Inducible (CREB) Activators. J. Biol. Chem. 2002. 277, 42241-42248.

[248] Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK. The Importance of Intrinsic Disorder for Protein Phosphorylation. Nucleic Acids Res. 2004. 32, 1037-1049.

[249] Baatout S, Derradji H. About Histone H1 Phosphorylation during Mitosis. Cell Biochem. Funct. 2006. 24, 93-94.

[250] Morgan DO. Cyclin-Dependent Kinases: Engines, Clocks, and Microprocessors. Annu. Rev. Cell Dev. Biol.1997. 13, 261-291.

[251] Bharath MM, Chandra NR, Rao MR. Molecular modeling of the chromatosome particle. Nucleic Acids Res. 2003; 31: 4264–4274

[252] Paulson JR, Patzlaff JS, Vallis AJ. Evidence that the endogenous histone H1 phosphatase in HeLa mitotic chromosomes is protein phosphatase 1, not protein phosphatase 2A. J. Cell Sci. 1996. 109 :1437–1447.

[253] Garcia BA, Busby SA, Barber CM, Shabanowitz J, Allis CD, Hunt DF. Characterization of phosphorylation sites on histone H1 isoforms by tandem mass spectrometry. J Proteome Res. 2004. 3(6):1219-27

[254] Talasz H, Helliger W, Puschendorf B, Lindner H. In vivo phosphorylation of histone H1 variants during the cell cycle. Biochemistry 1996. 35:1761–1767.

[255] Bradbury EM, Inglis RJ, Matthews HR, Sarner N. Phosphorylation of very lysine-rich histone in Physarum polycephalum. Correlation with chromosome condensation. Eur. J. Biochem. 1973. 33:131–139.

[256] Boggs BA, Allis CD, Chinault AC. Immunofluorescent studies of human chromosomes with antibodies against phosphorylated H1 histone. Chromosoma. 2000. 108:485–490.

[257] Th'ng JP, Guo XW, Swank RA, Crissman HA, Bradbury EM. Inhibition of Histone Phosphorylation by Staurosporine Leads to Chromosome Decondensation. J. Biol. Chem. 1994. 269, 9568-9573.

[258] Gurley LR, Walters RA, Tobey RA. Sequential phosphorylation of histone subfractions in the Chinese hamster cell cycle. J. Biol. Chem. 1975. 250:3936–3944.

[259] Yasuda H, Matsumoto Y, Mita S, Marunouchi T, Yamada M. A mouse temperature-sensitive mutant defective in H1 histone phosphorylation is defective in deoxyribonucleic acid synthesis and chromatin condensation. Biochemistry. 1981. 20:4414–4419.

[260] Halmer L, Gruss C, Effects of cell cycle dependent histone H1 phosphorylation on chromatin structure and chromatin replication. Nucleic Acids Res. 1996; 24:1420–1427.

[261] Alexandrow MG, Hamlin JL, Chromatin decondensation in S-phase involvesrecruitment of CDK2 by Cdc45 and histone H1 phosphorylation. J. Cell Biol. 2005; 168:875–886.

[262] Dasso M, Dimitrov S, Wolffe AP. Nuclear Assembly is Independent of Linker Histones. Proc. Natl. Acad. Sci. U. S. A. 1994. 91, 12477-12481.

[263] Ohsumi K, Katagiri C, Kishimoto T. Chromosome Condensation in Xenopus Mitotic Extracts without Histone H1. Science. 1993. 262, 2033-2035.

[264] Hendzel MJ, Lever MA, Crawford E, Th'ng JP. The C-Terminal Domain is the Primary Determinant of Histone H1 Binding to Chromatin in Vivo. J. Biol. Chem. 2004. 279, 20028-20034.

[265] Dyson HJ, Wright PE. Coupling of Folding and Binding for Unstructured Proteins. Curr. Opin. Struct. Biol. 2002. 12, 54-60

[266] Roque A, Ponte I, Arrondo JLR, Suau P. Phosphorylation of the carboxy-terminal domain of histone H1: effects on secondary structure and DNA condensation. NAR 2008. 36(14): 4719-4726.

[267] Teruel N, Ponte I, Suau P. Efectes de la fosforilació sobre l'estructura i la interacció amb el DNA de la histona H1 i el seu domini C-terminal. PhD thesis report, 2011.

[268] Banks GC, Deterding LJ, Tomer KB, Archer TK. Hormone-mediated dephosphorylation of specific histone H1 isoforms. J Biol Chem. 2001. 276(39):36467-73.

[269] Lee HL, Archer TK. Prolonged glucocorticoid exposure dephosphorylates histone H1 and inactivates the MMTV promoter. EMBO J. 1998;17(5):1454-66.

[270] Bhattacharjee RN, Archer TK. Transcriptional silencing of the mouse mammary tumor virus promoter through chromatin remodeling is concomitant with histone H1 phosphorylation and histone H3 hyperphosphorylation at M phase. Virology. 2006. 346(1):1-6.

[271] Horn PJ, Carruthers LM, Logie C, Hill DA, Solomon MJ, Wade PA, Imbalzano AN, Hansen JC, Peterson CL. Phosphorylation of linker histones regulates ATP-dependent chromatin remodeling enzymes. Nat Struct Biol. 2002. 9(4):263-7.

[272] Kysela B, Chovanec M, Jeggo PA. Phosphorylation of linker histones by DNA-dependent protein kinase is required for DNA ligase IV-dependent ligation in the presence of histone H1. Proc Natl Acad Sci U S A. 2005. 102(6):1877-82.

[273] Yellajoshyula D, Brown DT. Global modulation of chromatin dynamics mediated by dephosphorylation of linker histone H1 is necessary for erythroid differentiation. Proc Natl Acad Sci U S A. 2006. 103(49):18568-73.

[274] Mishra S, Saleh A, Espino PS, Davie JR, Murphy LJ. Phosphorylation of histones by tissue transglutaminase. J Biol Chem. 2006. 281(9):5532-8.

[275] Tsiftsoglou AS, Pappas IS, Vizirianakis IS. Mechanisms involved in the induced differentiation of leukemia cells. Pharmacol Ther. 2003. 100(3):257-90.

[276] Matushansky I, Radparvar F, Skoultchi AI. Reprogramming leukemic cells to terminal differentiation by inhibiting specific cyclin-dependent kinases in G1. Proc Natl Acad Sci U S A. 2000;97(26):14317-22.

[277] Matushansky I, Radparvar F, Skoultchi AI. CDK6 blocks differentiation: coupling cell proliferation to the block to differentiation in leukemic cells. Oncogene. 2003. 22(27):4143-9.

[278] Bleher R, Martin R. Nucleo-cytoplasmic translocation of histone H1 during the HeLa cell cycle. Chromosoma. 1999. 108(5):308-16.

[279] Tsai LH, Delalle I, Caviness VS Jr, Chae T, Harlow E. p35 is a neural-specific regulatory subunit of cyclin-dependent kinase 5. Nature. 1994. 371(6496):419-23.

[280] Duce JA, Smith DP, Blake RE, Crouch PJ, Li QX, Masters CL, Trounce IA. Linker histone H1 binds to disease associated amyloid-like fibrils. J Mol Biol. 2006. 361(3):493-505.

[281] Daujat S, Zeissler U, Waldmann T, Happel N, Schneider R. HP1 Binds Specifically to Lys26-Methylated Histone H1.4, Whereas Simultaneous Ser27 Phosphorylation Blocks HP1 Binding. J. Biol. Chem. 2005. 280, 38090-38095.

[282] Sarg B, Helliger W, Hoertnagl B, Puschendorf B, Lindner H. The N-terminally acetylated form of mammalian histone H1(0), but not that of avian histone H5, increases with age. Arch Biochem Biophys. 1999;372(2):333-9.

[283] Lindner H, Sarg B, Hoertnagl B, Helliger W. The microheterogeneity of the mammalian H1(0) histone. Evidence for an age-dependent deamidation. J Biol Chem. 1998;273(21):13324-30.

[284] Goytisolo FA, Gerchman SE, Yu X, Rees C, Graziano V, Ramakrishnan V, Thomas JO. Identification of Two DNA-Binding Sites on the Globular Domain of Histone H5. EMBO J.1996. 15, 3421-3429.

[285] Lake RS, Salzman NP. Occurrence and Properties of a Chromatin-Associated F1-Histone Phosphokinase in Mitotic Chinese Hamster Cells. Biochemistry. 1972. 11, 4817-4826.

[286] Panzeter PL., Realini CA, Althaus FR. Noncovalent Interactions of Poly(Adenosine Diphosphate Ribose) with Histones. Biochemistry. 1992. 31, 1379-1385.

[287] Poirier GG, de Murcia G, Jongstra-Bilen J, Niedergang C, Mandel P. Poly(ADP-Ribosyl)Ation of Polynucleosomes Causes Relaxation of Chromatin Structure. Proc. Natl. Acad. Sci. U. S. A. 1982. 79, 3423-3427.

[288] D'erme M, Zardo G, Reale A, Caiafa P. Co-Operative Interactions of Oligonucleosomal DNA with the H1e Histone Variant and its Poly(ADP-Ribosyl)Ated Isoform. Biochem. J. 1996. 316 (Pt 2), 475-480.

[289] Zardo G, Caiafa P. The Unmethylated State of CpG Islands in Mouse Fibroblasts Depends on the Poly(ADP-Ribosyl)Ation Process. J. Biol. Chem. 1998. 273, 16517-16520.

[290] Zardo G, Marenzi S, Caiafa P. H1 Histone as a Trans-Acting Factor Involved in Protecting Genomic DNA from Full Methylation. Biol. Chem. 1998. 379, 647-654.

[291] Clarke S. Propensity for spontaneous succinimide formation from aspartyl and asparaginyl residues in cellular proteins. Int. J., Peptide Protein Res. 1987. 30, 808-821

[292] Travers A. The location of the linker histone on the nucleosome. Trends Biochem Sci. 1999. 24(1):4-7. Review.

[293] Wright HT. Sequence and Structure Determinants of the Nonenzymatic Deamidation of Asparagine and Glutamine Residues in Proteins. Protein Eng.1991. 4, 283-294.

[294] Pham AD, Sauer F. Ubiquitin-activating/conjugating Activity of TAFII250, a Mediator of Activation of Gene Expression in Drosophila. Science. 2000. 289, 2357-2360.

[295] Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y. Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1. Cell. 2004. 119, 941-953.

[296] Syed SH, Goutte-Gattat D, Becker N, Meyer S, Shukla MS, Hayes JJ, Everaers R, Angelov D, Bednar J, Dimitrov S. Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome Proc. Natl. Acad. Sci. USA. 2010. 107(21):9620-5.

[297] Dimitrov SI, Russanova VR, Pashev IG. The globular domain of histone H5 is internally located in the 30 nm chromatin fibre: An immunochemical study. EMBO J. 1987. 6(8):2387–2392.

[298] Russanova VR, Dimitrov SI, Makarov VL, Pashev IG. Accessibility of the globular domain of histones H1 and H5 to antibodies upon folding of chromatin. Eur J Biochem . 1987. 167(2):321–326.

[299] Zlatanova J, Seebart C, Tomschik M. The linker-protein network: Control of nucleosomal DNA accessibility. Trends Biochem Sci. 2008. 33(6):247–253.)

[300] Simpson RT. Structure of the Chromatosome, a Chromatin Particle Containing 160 Base Pairs of DNA and all the Histones. Biochemistry.1978. 17, 5524-5531.

[301] Hayes JJ, Wolffe AP. Preferential and Asymmetric Interaction of Linker Histones with 5S DNA in the Nucleosome. Proc. Natl. Acad. Sci. U. S. A. 1993. 90, 6415-6419.

[302] Hayes, J. J., Pruss, D., and Wolffe, A. P. (1994) Contacts of the Globular Domain of Histone H5 and Core Histones with DNA in a "Chromatosome". Proc. Natl. Acad. Sci. U. S. A. 91, 7817-7821.

[303] Lee KM, Hayes JJ. Linker DNA and H1-Dependent Reorganization of Histone-DNA Interactions within the Nucleosome. Biochemistry. 1998. 37, 8622-8628.

[304] Guschin D, Chandler S, Wolffe AP. Asymmetric Linker Histone Association Directs the Asymmetric Rearrangement of Core Histone Interactions in a Positioned Nucleosome Containing a Thyroid Hormone Response Element. Biochemistry. 1998. 37, 8629-8636.

[305] Usachenko SI., Gavin IM, Bavykin SG. Alterations in Nucleosome Core Structure in Linker Histone-Depleted Chromatin. J. Biol. Chem.1996. 271, 3831-3836.

[306] Caterino TL, Fang H, Hayes JJ. Nucleosome Linker DNA Contacts and Induces Specific Folding of the Intrinsically Disordered H1 Carboxyl-Terminal Domain. Mol Cell Biol. 2011. 31(11): 2341–2348.

[307] Hieb AR, D'Arcy S, Kramer MA, White AE, Luger K. Fluorescence strategies for high-throughput quantification of protein interactions. Nucleic Acids Res. 2012. 40(5): e33.

[308] Hansen JC, Lu X, Ross ED, Woody RW. Intrinsic protein disorder, amino acid composition, and histone terminal domains. J. Biol. Chem. 2006. 281:1853–1856.

[309] Uversky VN. Seven lessons from one IDP structural analysis. Structure. 2010. 18:1069–1071.

[310] Roque A, Ponte I, Suau P. Role of charge neutralization in the folding of the carboxy-terminal domain of histone H1. J. Phys. Chem. B. 2009. 113:12061–12066.

[311] Fang H, Clark DJ, Hayes JJ. DNA and nucleosomes direct distinct folding of a linker histone H1 C-terminal domain. Nucleic Acids Res. 2012. 40(4):1475-84).

[312] Sheng S, Czajkowsky DM, Shao Z. Localization of linker histone in chromatosomes by cryo-atomic force microscopy. Biophys. J. 2006. 91:L35–L37.

[313] Meyer S, Becker N, Syed SH, Goutte-Gattat D, Shukla MS, Hayes J, Angelov D, Bednar J, Dimitrov S, Everaers R. From crystal and NMR structures, footprints and cryo-electron-micrographs to large and soft structures: nanoscale modeling of the nucleosomal stem. Nucleic Acids Res. 2011;39:9139–9154.

[314] Clark DJ, Thomas JO. Salt-dependent co-operative interaction of histone H1 with linear DNA. J Mol Biol. 1986. 187(4):569–580.);

[315] Shintomi K, Iwabuchi M, Saeki H, Ura K, Kishimoto T, Ohsumi K. Nucleosome assembly protein-1 is a linker histone chaperone in Xenopus eggs. Proc Natl Acad Sci USA. 2005. 102(23):8210–8215.

[316] Saeki H, Ohsumi K, Aihara H, Ito T, Hirose S, Ura K, Kaneda Y. Linker histone variants control chromatin dynamics during early embryogenesis. Proc Natl Acad Sci USA. 2005. 102(16):5697–5702.)

[317] Quina AS, Buschbeck M, Di Croce L. Chromatin Structure and Epigenetics. Biochem. Pharmacol. 2006. 72, 1563-1569.

[318] Fierz B, Muir TW. From Chromatin as an expansive canvas for chemical biology. Nat Chem Biol. 2012. 17;8(5):417-27

[319] Jacobson RH, Ladurner AG, King DS, Tjian R. Structure and function of a human TAFII250 double bromodomain module. Science. 2000. 288(5470):1422-5.

[320] Jacobs SA, Khorasanizadeh S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. Science. 2002. 295(5562):2080-3.

[321] Nakayama J, Rice JC, Strahl BD, Allis CD, Grewal SI. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science. 2001. 292(5514):110-3.

[322] Sun ZW, Allis CD. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature. 2002. 418(6893):104-8.

[323] Tanny JC, Erdjument-Bromage H, Tempst P, Allis CD. Ubiquitylation of histone H2B controls RNA polymerase II transcription elongation independently of histone H3 methylation. Genes Dev. 2007 Apr 1;21(7):835-47.

[324] Fischle W, Wang Y, Allis CD. Binary switches and modification cassettes in histone biology and beyond. Nature. 2003. 425(6957):475-9

[325] 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. 2005. 438(7071):1116-22.

[326] Cheung WL, Ajiro K, Samejima K, Kloc M, Cheung P, Mizzen CA, Beeser A, Etkin LD, Chernoff J, Earnshaw WC, Allis CD. Apoptotic phosphorylation of histone H2B is mediated by mammalian sterile twenty kinase. Cell. 2003. 113(4):507-17.

[327] Ahn SH, Henderson KA, Keeney S, Allis CD. H2B (Ser10) phosphorylation is induced during apoptosis and meiosis in S. cerevisiae. Cell Cycle. 2005 Jun;4(6):780-3. Epub 2005 Jun 14.

[328] Ahn SH, Diaz RL, Grunstein M, Allis CD. Histone H2B deacetylation at lysine 11 is required for yeast apoptosis induced by phosphorylation of H2B at serine 10. Mol Cell. 2006. 24(2):211-20.

[329] Wang GG, Allis CD, Chi P. Chromatin remodeling and cancer, Part I: Covalent histone modifications. Trends Mol Med. 2007. 13(9):363-72. Review.

[330] Ruthenburg AJ, Li H, Patel DJ, Allis CD. Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol. 2007. 8(12):983-94. Review.

[331] Allis CD , Jenuwein T, Reinberg D, Caparros. 2006. Edited by John Matlock ISBN-10: 0879697245 | ISBN-13: 978-0879697242

[332] Grewal SI, Elgin SC. Heterochromatin: New Possibilities for the Inheritance of Structure. Curr. Opin. Genet. Dev.2002.12, 178-187.

[333] Perry CA, Allis CD, Annunziato, AT. Parental Nucleosomes Segregated to Newly Replicated Chromatin are Underacetylated Relative to those Assembled De Novo. Biochemistry. 1993. 32, 13615-13623.

[334] Jackson V. Deposition of Newly Synthesized Histones: New Histones H2A and H2B do Not Deposit in the Same Nucleosome with New Histones H3 and H4. Biochemistry. 1987. 26, 2315-2325.

[335] Jackson V. Deposition of Newly Synthesized Histones: Hybrid Nucleosomes are Not Tandemly Arranged on Daughter DNA Strands. Biochemistry. 1988. 27, 2109-2120

[336] Ng HH, Xu RM, Zhang Y, Struhl K. Ubiquitination of Histone H2B by Rad6 is Required for Efficient Dot1-Mediated Methylation of Histone H3 Lysine 79. J. Biol. Chem. 2002. 277, 34655-34657.

[337] Briggs SD, Xiao T, Sun ZW, Caldwell JA, Shabanowitz J, Hunt DF, Allis CD, Strahl BD. Gene Silencing: Trans-Histone Regulatory Pathway in Chromatin. Nature. 2002. 418, 498.

[338] McNairn AJ, Gilbert DM. Epigenomic Replication: Linking Epigenetics to DNA Replication. Bioessays. 2003. 25, 647-656.

[339] Barth A. Fine-Structure Enhancement--Assessment of a Simple Method to Resolve Overlapping Bands in Spectra. Spectrochim. Acta A Mol. Biomol. Spectrosc. 2000. 56, 1223-1232.

[340] Kumosinski TF, Unruh JJ. Quantitation of the Global Secondary Structure of Globular Proteins by FTIR Spectroscopy: Comparison with X-Ray Crystallographic Structure. Talanta. 1996. 43, 199-219.

[341] Mayer E. FTIR Spectroscopic Study of the Dynamics of Conformational Substates in Hydrated Carbonyl-Myoglobin Films Via Temperature Dependence of the CO Stretching Band Parameters. Biophys. J. 1994. 67, 862-873.

[342] Byler DM, Susi H. Examination of the Secondary Structure of Proteins by Deconvolved FTIR Spectra. Biopolymers. 1986. 25, 469-487.

[343] Susi H, Timasheff SN, Stevens L. December 1967. Infrared Spectre and Protein Conformations in Aqueous Solutions: The amide I band in H_2O and D_2O solutions. JBC,1967.242; 23;5460-5466.

[344] Su X, Ren C, Freitas MA. Mass spectrometry-based strategies for characterization of histones and their post-translational modifications. ExpertRevProteomics. 2007. 4(2):211-25. Review.

[345] Lindner HH. Analysis of histones, histone variants, and their post-translationally modified forms. Electrophoresis. 2008. 29(12):2516-32

[346] Gurley LR, London JE, Valdez JG. High-performance capillary electrophoresis of histones. J Chromatogr. 1991. 559(1–2):431–443

[347] Lindner H, Helliger W, Dirschlmayer A, Jaquemar M, Puschendorf B. High-performance capillary electrophoresis of core histones and their acetylated modified derivatives. Biochem J. 1992. 283(Pt 2):467–471

[348] Lindner H, Helliger W, Dirschlmayer A, Talasz H, Wurm M, Sarg B, Jaquemar M, Puschendorf B. Separation of phosphorylated histone H1 variants by high-performance capillary electrophoresis. J Chromatogr. 1992. 608(1–2):211–216

[349] Lindner H, Helliger W, Sarg B, Meraner C. Effect of buffer composition on the migration order and separation of histone H1 subtypes. Electrophoresis. 1995. 16(4):604–610

[350] Lindner H, Wurm M, Dirschlmayer A, Sarg B, Helliger W. Application of high-performance capillary electrophoresis to the analysis of H1 histones. Electrophoresis. 1993. 14(5–6):480–485

[351] Aguilar C, Hofte AJ, Tjaden UR, van der Greef J. Analysis of histones by on-line capillary zone electrophoresis–electrospray ionisation mass spectrometry. J Chromatogr A. 2001. 926(1):57–67)

[352]Wiktorowicz JE, Colburn JC. Separation of cationic proteins via charge reversal in capillary electrophoresis. Electrophoresis. 1990. 11(9):769–773

[353] Mizzen CA, McLachlan DR. Capillary electrophoresis of histone H1 variants at neutral pH in dynamically modified fused-silica tubing. Electrophoresis. 2000. 21(12):2359–2367

[354] Koutzamani E, Loborg H, Sarg B, Lindner HH, Rundquist I. Linker histone subtype composition and affinity for chromatin in situ in nucleated mature erythrocytes. J Biol Chem. 2002. 22;277(47):44688-94.

[355] Sidoli S, Cheng L, Jensen ON. Proteomics in chromatin biology and epigenetics: Elucidation of post-translational modifications of histone proteins by mass spectrometry. J Proteomics. 2012. 27;75(12):3419-33. Review.

[356] Gurley LR, D'Anna JA, Blumenfeld M, Valdez JG, Sebring RJ, Donahue PR, Prentice DA, Spall WD. Preparation of histone variants and high-mobility group proteins by reversed phase high-performance liquid chromatography. J Chromatogr. 1984. 297:147–165

[357] Gurley LR, Prentice DA, Valdez JG, Spall WD. Histone fractionation by high-performance liquid chromatography on cyanoalkysilane (CN) reverse-phase columns. Anal Biochem. 1983. 131(2):465–477);

[358] Gurley LR, Prentice DA, Valdez JG, Spall WD. High performance liquid chromatography of chromatin histones. J Chromatogr. 1983. 266:609–627)

[359] Lindner H, Helliger W, Puschendorf B. Histone separation by high-performance liquid chromatography on C4 reverse-phase columns. Anal Biochem. 1986. 158(2):424–430

[360] Lindner H, Helliger W, Puschendorf B. Separation of Friend erythroleukaemic cell histones and high-mobility-group proteins by reversed-phase high-performance liquid chromatography. J Chromatogr. 1988. 450(3):309–316

[361] Helliger W, Lindner H, Hauptlorenz S, Puschendorf B. A new HPLC isolation procedure for chicken and goose erythrocyte histones.Biochem J. 1988. 255(1):23–27

[362] Lindner H, Helliger W, Puschendorf B. Separation of rat tissue histone H1 subtypes by reverse-phase HPLC Identification and assignment to a standard H1 nomenclature. Biochem J. 1990. 269(2):359–363

[363] Mizzen CA, Alpert AJ, Lévesque L, Kruck TP, McLachlan DR. Resolution of allelic and non-allelic variants of histone H1 by cation-exchange-hydrophilic-interaction chromatography. J Chromatogr B Biomed SciAppl. 2000. 744(1):33-46

[364] Alpert AJ. Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. J Chromatogr. 1990. 19;499:177-96

[365] Deterding LJ, Banks GC, Tomer KB, Archer TK. Understanding global changes in histone H1 phosphorylation using mass spectrometry.Methods. 2004. 33, (1):53–58

[366] Mann M, Hendrickson RC, Pandey A. Analysis of proteins and proteomes by mass spectrometry. Annu Rev Biochem. 2001. 70:437-73. Review.

[367] Hillenkamp F, Karas M. Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. Methods Enzymol. 1990. 193:280-95

[368] Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. Science. 1989. 246(4926):64-71

[369] Jackson PS, Gurley LR. Analysis of nucleoproteins by direct injection of dissolved nuclei or chromosomes into a high-performance liquid chromatographic system. J. Chromatogr. 1985. 326:199–216

[370] Beck HC, Nielsen EC, Matthiesen R, Jensen LH, Sehested M, Finn P, Grauslund M, Hansen AM, Jensen ON.Quantitative proteomic analysis of post-translational modifications of human histones. Mol. Cell Proteomics. 2006. 5(7):1314–1325

[371] Du YC, Gu S, Zhou J, Wang T, Cai H, Macinnes MA, Bradbury EM, Chen X. The dynamic alterations of H2AX complex during DNA repair detected by a proteomic approach reveal the critical roles of Ca^{2+}/calmodulin in the ionizing radiation-induced cell cycle arrest. Mol. Cell Proteomics. 2006. 5(6):1033–1044

[372] Bonenfant D, Coulot M, Towbin H, Schindler P, van Oostrum J. Characterization of histones H2A and H2B variants and their post-translational modifications by mass spectrometry. Mol. Cell. Proteomics. 2005. 5(3):541–552

[373] Naldi M, Andrisano V, Fiori J, Calonghi N, Pagnotta E, Parolin C, Pieraccini G, Masotti L. Histone proteins determined in a human colon cancer by highperformance liquid chromatography and mass spectrometry. J. Chromatogr. A. 2006. 1129(1):73–81.

[374] Lindner H, Sarg B, Meraner C, Helliger W. Separation of acetylated core histones by hydrophilic-interaction liquid chromatography. J. Chromatogr. A. 1996. 743(1):137–144

[375] Lindner H, Sarg B, Helliger W. Application of hydrophilic-interaction liquid chromatography to the separation of phosphorylated H1 histones. J. Chromatogr. A. 1997;782(1):55–62)

[376] Sarg B, Helliger W, Hoertnagl B, Puschendorf B, Lindner H. The N-terminally acetylated form of mammalian histone H1(o), but not that of avian histone H5, increases with age. Arch Biochem.Biophys. 1999. 372(2):333–339

[377] Sarg B, Green A, Soderkvist P, Helliger W, Rundquist I, Lindner HH. Characterization of sequence variations in human histone H1.2 and H1.4 subtypes. FEBS. J. 2005. 272(14):3673–3683

[378] Lindner H, Sarg B, Helliger W. Capillary electrophoresis analysis of histones, histone variants, and their post-translationally modified forms: a review. J. Capill. Electrophor.Microchip. Technol. 2003. 8(34):59–67

[379] Lindner H, Wesierska-Gadek J, Helliger W, Puschendorf B, Sauermann G. Identification of ADP-ribosylated histones by the combined use of high-performance liquid chromatography and electrophoresis. J. Chromatogr. 1989. 472(1):243–249

[380] Roepstorff P, Fohlman J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. Biomed.Mass Spectrom. 1984. (11): 601

[381] Zhang K, Yau PM, Chandrasekhar B, New R, Kondrat R, Imai BS, Bradbury ME. Differentiation between peptides containing acetylated or tri-methylated lysines by mass spectrometry: an application for determining lysine 9 acetylation and methylation of histone H3. Proteomics. 2004. 4(1):1–10

[382] Eyers, C; Eyers, PA; Beynon, RJ; Gaskell, S; Johnson, H. Rigorous determination of the stoichiometry of protein phosphorylation using mass spectrometry. J Am Soc Mass Spectrom. 2009. (20) 2211-20

[383] Haydon CE, Lewis TS, Resing KA, Ahn NG, Schweppe RE. The characterization of protein post-translational modifications by mass spectrometry. Acc. Chem. Res. 2003. 36(6)453-61

[384] Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature. 2003. 13;422(6928):198-207

[385] Wolters DA, Washburn MP, Yates JR 3rd. An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem. 2001. 1;73(23):5683-90

[386] McLafferty FW, Breuker K, Jin M, Han X, Infusini G, Jiang H, Kong X, Begley TP. Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. FEBS J. 2007. 274(24):6256-68.)

[387] Bonaldi T, Imhof A, Regula JT. A combination of different mass spectroscopic techniques for the analysis of dynamic changes of histone modifications. Proteomics. 2004. 4(5):1382-96

[388] Sarg B, Faserl K, Kremser L, Halfinger B, Sebastiano R, Lindner HH. Comparing and Combining CE-ESI-MS and nano-LC-ESI-MS for the Characterization of Post-translationally Modified Histones. Mol Cell Proteomics. 2013 May 29. [Epub ahead of print]

[389] Faserl K, Sarg B, Kremser L, Lindner H. Optimization and evaluation of a sheathless capillary electrophoresis-electrospray ionization mass spectrometry platform for peptide analysis: comparison to liquid chromatography-electrospray ionization mass spectrometry. Anal Chem. 2011;83(19):7297-305.

[390] Chu B. Laser Light scattering: Basic Principles and Practice. Academic Press, 1992. ISBN 0-12-174551-1

[391] www.malvern.com/labeng/technology/dynamic_light_scattering/dynamic_light_scattering.htm

[392] Berne BJ, Pecora R. Dynamic Light Scattering. Courier Dover Publications; 2000. ISBN 0-486-41155-9

[393] Pecora R. Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy, Plenum Press, 1985.

[394] Johnson CS Jr, Gabriel DA. Laser Light Scattering, Dover Publications, Inc., New York, 1981

[395] Malvern Instruments. Dynamic light scattering: Common Terms Defined. Inform White Paper, 2011.

[396] Stojanović Z, Marković S. 2012, Determination of Particle Size Distributions by Laser Diffraction. Technics – New Materials. vol. 67 (11-20.) Special Edition. 2012.

[397]Mudroch A, Azcue JM, Paul Mudroch P. Manual of physico-chemical analysis of aquatic sediments Physico-chemical analysis of aquatic sediments. Boca Raton [etc.] Lewis Publishers, 1997. ISBN 1566701554

[398] Syvitski JPM. Principles, Methods and Application of Particle Size Analysis. Cambridge University Press, 2007. ISBN 978-0-521-04461-5

[399] www.malvern.com/labeng/technology/laser_diffraction/laser_diffraction.htm

[400] Bandekar J, Krimm S. Vibrational analysis of peptides, polypeptides, and proteins: Characteristic amide bands of β-turns. Biophysics. Proc. Natl. Acad. Sci. USA. 1979. 76(2):774-777

[401] Arrondo JLR, Muga A, Castresana J, Goñi FM. Quantitative studies of the structure of proteins in solution by Fourier-Transform Infrared Spectroscopy. Prog. Biophys. Molec. Biol. 1993. 59: 23-56

[402] Susi H, Timasheff SN, Stevens, L. Infrared Spectra and Protein Conformations in Aqueous Solutions: The amide I band in H2O and D2O solutions. JBC. 1967. 242 (23): 5460-5466.)

[403] Arrondo JL, Echabe I, Iloro I, Hernando MA, de la Cruz F, Goñi FM. A bacterial TrwC relaxase domain contains a thermally stable alpha-helical core. J Bacteriol. 2003. 185(14):4226-32.

[404] Arrondo JL, Goñi FM. Structure and dynamics of membrane proteins as studied by infrared spectroscopy. Prog Biophys Mol Biol. 1999. 72(4):367-405.

[405] Arrondo JL, Castresana J, Valpuesta JM, Goñi FM. Structure and thermal denaturation of crystalline and noncrystalline cytochrome oxidase as studied by infrared spectroscopy. Biochemistry. 1994 Sep 27;33(38):11650-5.

[406] Vila R, Ponte I, Collado M, Arrondo JL, Suau P. Induction of secondary structure in a COOH-terminal peptide of histone H1 by interaction with the DNA: an infrared spectroscopy study. J Biol Chem. 2001. 17;276(33):30898-903

[407] Goormaghtigh E, Ruysschaert JM, Raussens V. Evaluation of the Information Content in Infrared Spectra for Protein Secondary Structure Determination. Biophys. Jour.2006. 90: 2946–57);

[408] Krimm S, Abe Y. Intermolecular Interaction Effects in the Amide I Vibrations of β-Polypeptides. Proc. Nat. Acad. Sci. USA. 1972. 69 (10): 2788-2792.);

[409] Byler DM, Susi H. Examination of the secondary structure of proteins by deconvolved FTIR spectra. Biopolymers. 1986. 25: 469–487

[410] Nilsson MR. Techniques to study amyloid fibril formation in vitro. Methods. 2004. 34, 151–160

[411] Bracken C, Iakoucheva LM, Romero PR, Dunker AK. Combining prediction, computation and experiment for the characterization of protein disorder. Curr. Opin. Struct. Biol. 2004. 14, 570–576

[412] Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. Nat. Rev. Mol. Cell Biol. 2005. 6, 197–208.

[413] Gunasekaran K, Tsai CJ, Kumar S, Zanuy D, Nussinov R. Extended disordered proteins: targeting function with less scaffold. Trends Biochem. Sci. 2003. 28, 81–85

[414] Cerf C, Lippens G, Ramakrishnan V, Muyldermans S, Segers A, Wyns L, Wodak SJ, Hallenga K. Homo- and heteronuclear two-dimensional NMR studies of the globular domain of histone H1: full assignment, tertiary structure, and comparison with the globular domain of histone H5. Biochemistry. 1994. 20;33(37):11079-86

[415] Richards RC, O'Neil DB, Thibault P, Ewart V. Histone H1, an antimicrobial protein of atlantic salmon (Salmo salar). Biochem. Biophys. Res. Commun. 2001. 284, 549–555

[416] Rose FR, Bailey K, Keyte JW, Chan WC, Greenwood D, Mahida YR. Potential role of epithelial cell-derived histone H1 proteins in innate antimicrobial defense in the human gastrointestinal tract. Infect. Immun. 1998. 66, 3255–3263.),

[417] Brix K, Summa W, Lottspeich F, Herzog V. Extracellularly occurring histone H1 mediates the binding of thyroglobulin to the cell surface of mouse macrophages. J Clin Invest. 1998. 15;102(2):283-93.

[418] Holers VM, Kotzin BL. Human peripheral blood monocytes display surface antigens recognized by monoclonal antinuclear antibodies. J Clin Invest. 1985. 76(3):991-8.

[419] Bolton SJ, Perry VH. Histone H1; a neuronal protein that binds bacterial lipopolysaccharide. J. Neurocytol. 1997. 26, 823–831

[420] Class R, Lindman S, Fassbender C, Leinenbach HP, Rawer S, Emrich JG, Brady LW, Zeppezauer M. Histone H1 suppresses tumor growth of leukemia cells in vitro, ex vivo and in an animal model suggesting extracellular functions of histones. Am J Clin Oncol. 1996. 19(5):522-31.

[421] Zhao H, Shambhunath B, Tuominen EKJ, Kinnunen KJ. Interactions of histone H1 with phospholipids and comparison of its binding to giant liposomes and human leukemic T cells. Biochemistry . 2004a. 43, 10192–10202

[422] Kleine TJ, Lewis PN, Lewis SA. Histone-induced damage of a mammalian epithelium: the role of protein and membrane structure. Am. J. Physiol. 1997. 273, C1925–C1936

[423] Roque A, Teruel N, López R, Ponte I, Suau P. Contribution of hydrophobic interactions to the folding and fibrillation of histone H1 and its carboxy-terminal domain. J Struct Biol. 2012. 180(1):101-9.

[424] Engel MF, Khemtémourian L, Kleijer CC, Meeldijk HJ, Jacobs J, Verkleij AJ, de Kruijff B, Killian JA, Höppener JW. Membrane damage by human islet amyloid polypeptide through fibril growth at the membrane. Proc Natl Acad Sci USA. 2008. 22;105(16):6033-8

[425] Griffin MD, Mok ML, Wilson LM, Pham CL, Waddington LJ, Perugini MA, Howlett GJ. Phospholipid interaction induces molecular-level polymorphism in apolipoprotein C-II amyloid fibrils via alternative assembly pathways. J Mol Biol. 2008. 4;375(1):240-56

[426] Martins IC, Kuperstein I, Wilkinson H, Maes E, Vanbrabant M, Jonckheere W, Van Gelder P, Hartmann D, D'Hooge R, De Strooper B, Schymkowitz J, Rousseau F. Lipids revert inert Abeta amyloid fibrils to neurotoxic protofibrils that affect learning in mice. EMBO J. 2008. 9;27(1):224-33

[427] Bates DL, Thomas JO. Histones H1 and H5: one or two molecules per nucleosome? Nucleic Acids Res. 1981. 9(22): 5883–5894.

[428] Kowalski A, Pałyga J. Chromatin compaction in terminally differentiated avian blood cells: the role of linker histone H5 and non-histone protein MENT. Chromosome Res. 2011. 19(5): 579–590.

[429] Sun JM, Ali Z, Lurz R, Ruiz-Carrillo A. Replacement of histone H1 by H5 in vivo does not change the nucleosome repeat length of chromatin but increases its stability.EMBO J. 1990. 9(5):1651-8.

[430] Tsuzuki J, Loeb J. In vitro phosphorylation of chicken erythrocyte histone by various protein kinases: absence of phosphorylation from F1 histone. Exp Cell Res. 1974. 88(2):303-10.

[431] Faserl K, Sarg B, Kremser L, Lindner H. Optimization and evaluation of a sheathless capillary electrophoresis-electrospray ionization mass spectrometry platform for peptide analysis: comparison to liquid chromatography-electrospray ionization mass spectrometry. Anal Chem. 2011. 1;83(19):7297-305

[432] Snijders AP, Pongdam S, Lambert SJ, Wood CM, Baldwin JP, Dickman MJ. Characterization of post-translational modifications of the linker histones H1 and H5 from chicken erythrocytes using mass spectrometry. J Proteome Res. 2008. 7(10):4326-35

[433] Stadtman ER, Van Remmen H, Richardson A, Wehr NB, Levine RL. Methionine oxidation and aging. Biochim Biophys Acta. 2005. 17;1703(2):135-40.

[434] Polevoda B, Sherman F. N-terminal acetyltransferases and sequence requirements for N-terminal acetylation of eukaryotic proteins. J. Mol. Biol. 2003. 325, 595– 622

[435] Cohen P, Frame S. The renaissance of GSK3. Nat. Rev. Mol. Cell Biol. 2001; 2:769-76

[436] Gurley L R, Valdez JG, Buchanan JS. Characterization of the mitotic specific phosphorylation site of histone H1. Absence of a consensus sequence for the p34cdc2/cyclin B kinase. J. Biol. Chem. 1995. 246, 27653– 27660

[437] Shannon MF, Wells JR. Characterization of the six chicken histone H1 proteins and alignment with their respective genes. J Biol Chem. 1987. 15;262(20):9664-8

[438]Coles LS, Robins AJ, Madley LK, Wells JR. Characterization of the chicken histone H1 gene complement. Generation of a complete set of vertebrate H1 protein sequences. J Biol Chem. 1987. 15;262(20):9656-63

[439] Delcuve GP, Davie JR. Chromatin structure of erythroid-specific genes of immature and mature chicken erythrocytes. Biochem J. 1989. 263(1): 179–186

[440] Mann M, Jensen ON. Proteomic analysis of post-translational modifications. Nat Biotechnol. 2003.21(3):255-61.

[441] Shannon MF, Wells JR. Characterization of the six chicken histone H1 proteins and alignment with their respective genes. J Biol Chem. 1987. 15;262(20):9664-8

[442] Kinkade JM Jr. Qualitative species differences and quantitative tissue differences in the distribution of lysine-rich histones. J Biol Chem. 1969. 25;244(12):3375-86.

[443] Panyim S, Bilek D, Chalkley R. An electrophoretic comparison of vertebrate histones. J Biol Chem. 1971. 10; 246(13):4206-15.

[444] Roche J, Gorka C, Goeltz P, Lawrence JJ. Association of histone H1(0) with a gene repressed during liver development. Nature. 1985. 14-20;314(6007):197-8

[445] Carruthers LM, Hansen JC. The core histone N termini function independently of linker histones during chromatin condensation. J Biol Chem. 2000. 275(47):37285-90.

[446] Woodcock CL, Ghosh RP. Chromatin Higher-order Structure and Dynamics. Cold Spring Harb Perspect Biol. 2010. 2(5): a000596.

[447]Widom J. Physicochemical studies of the folding of the 100 A nucleosome filament into the 300 A filament. Cation dependence. Mol Biol. 1986. 190(3):411-24.

[448]Sen D, Crothers DM. Condensation of chromatin: role of multivalent cations. Biochemistry. 1986. 25(7):1495-503

[449]Hansen JC. Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. Annu Rev Biophys Biomol Struct. 2002. 31:361-92.

[450] Privalov PL. Intermediate states in protein folding. J Mol Biol. 1996. 24; 258(5):707-25. Review.

[451] Roque A, Ponte I, Suau P. Macromolecular crowding induces a molten globule state in the C-terminal domain of histone H1. Biophys J. 2007. 15;93(6):2170-7

[452] Aguzzi A, Polymenidou M. Mammalian prion biology: one century of evolving concepts. Cell. 2004 Jan 23;116(2):313-27. Review.

[453]Lührs T, Ritter C, Adrian M, Riek-Loher D, Bohrmann B, Döbeli H, Schubert D, Riek R. 3D structure of Alzheimer's amyloid-beta(1-42) fibrils. Proc Natl Acad Sci U S A. 2005. 102(48):17342-7.

[454] Sunde M, Serpell LC, Bartlam M, Fraser PE, Pepys MB, Blake CC. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. J Mol Biol. 1997;273(3):729-39.

[455]Zhao H, Tuominen EK, Kinnunen PK.Formation of amyloid fibers triggered by phosphatidylserine-containing membranes. Biochemistry. 2004 Aug 17;43(32):10302-7.

[456] Gorbenko GP, Kinnunen PK. The role of lipid-protein interactions in amyloid-type protein fibril formation. Chem Phys Lipids. 2006. 141(1-2):72-82. Review.

[457] Fandrich M, Forge V, Buder K, Kittler M, Dobson CM, Diekmann S. Myoglobin Forms Amyloid Fibrils by Association of Unfolded Polypeptide Segments. Proc. Natl. Acad. Sci. U. S. A. 2003. 100, 15463-15468.

[458] Tseng BP, Esler WP, Clish CB, Stimson ER, Ghilardi JR, Vinters HV, Mantyh PW, Lee JP, Maggio JE. Deposition of monomeric, not oligomeric, Abeta mediates growth of Alzheimer's disease amyloid plaques in human brain preparations. Biochemistry. 1999 Aug 10;38(32):10424-31.

[459] Bonet-Costa C, Vilaseca M, Diema C, Vujatovic O, Vaquero A, Omeñaca N, Castejón L, Bernués J, Giralt E, Azorín F. Combined bottom-up and top-down mass spectrometry analyses of the pattern of post-translational modifications of Drosophila melanogaster linker histone H1. J Proteomics. 2012. 75(13):4124-38.

[460] Tweedie-Cullen RY, Brunner AM, Grossmann J, Mohanna S, Sichau D, Nanni P, Panse C, Mansuy IM.Identification of combinatorial patterns of post-translational modifications on individual histones in the mouse brain. PLoS One. 2012;7(5):e36980. Epub 2012.

[461] Lizcano JM, Deak M, Morrice N, Kieloc, A, Hastie CJ., Dong L, Schutkowski M, Reimer U, Aless, DR. Molecular basis for the substrate specificity of NIMA-related kinase-6 (NEK6). Evidence that NEK6 does not phosphorylate the hydrophobic motif of ribosomal S6 protein kinase and serum- and glucocorticoid-induced protein kinase in vivo. J. Biol. Chem. 2002. 277(31):27839-49.

[462] Yin MJ, Shao L, Voehringer D, Smeal T, Jallal B.The serine/threonine kinase Nek6 is required for cell cycle progression through mitosis. J. Biol. Chem. 2004. 278 (52): 52454–60.

[463] Hashimoto Y, Akita H, Hibino M, Kohri K, Nakanishi M.Identification and characterization of Nek6 protein kinase, a potential human homolog of NIMA histone H3 kinase. Biochem. Biophys. Res. Commun. 2002. 293 (2): 753–8.

[464] Bosc DG, Lüscher B, Litchfield DW. Expression and regulation of protein kinase CK2 during the cell cycle. Mol Cell Biochem. 1999. 191(1-2):213-22.

[465] Jope RS, Johnson GV. The glamour and gloom of glycogen synthase kinase-3. Trends Biochem Sci 2004. 29 (2): 95–102

[466] Aplin AE, Jacobsen JS, Anderton BH, Gallo JM. Effect of increased glycogen synthase kinase-3 activity upon the maturation of the amyloid precursor protein in transfected cells. Neuroreport. 1997. 8(3):639-43.

[467] Jope RS, Yuskaitis CJ, Beurel E. Glycogen synthase kinase-3 (GSK3): inflammation, diseases, and therapeutics. Neurochem Res. 2007. 32(4-5):577-95.

[468] Bannister AJ., Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T. Selective Recognition of Methylated Lysine 9 on Histone H3 by the HP1 Chromo Domain. Nature. 2001. 410, 120-124.

[469] Lake RS, Salzman, NP. Occurrence and Properties of a Chromatin-Associated F1-Histone Phosphokinase in Mitotic Chinese Hamster Cells. Biochemistry. 1972. 11, 4817-4826.

[470] Preinerstorfer B, Schiesel S, Lämmerhofer M, Lindner W. Metabolic profiling of intracellular metabolites in fermentation broths from beta-lactam antibiotics production by liquid chromatography-tandem mass spectrometry methods. J Chromatogr A. 2010. 15;1217(3):312-28

[471] Melmer M, Stangler T, Schiefermeier M, Brunner W, Toll H, Rupprechter A, Lindner W, Premstaller A. HILIC analysis of fluorescence-labeled N-glycans from recombinant biopharmaceuticals. Anal Bioanal Chem. 2010. 398(2):905-14

[472] Telu KH, Abbaoui B, Thomas-Ahner JM, Zynger DL, Clinton SK, Freitas MA, Mortazavi A. Alterations of histone h1 phosphorylation during bladder carcinogenesis. J Proteome Res. 2013. 5;12(7):3317-26

[473] Härd T, Nielsen PE, Norden B. Molecular flexibility of extended and compacted polynucleosomes. A steady-state fluorescence polarization study. Eur Biophys J. 1988;16(4):231-41.

[474] Caterino TL, Hayes JJ. Structure of the H1 C-terminal domain and function in chromatin condensation. Biochem Cell Biol. 2011; 89(1):35-44. Review