HISTONE PROTEASES: THE TALE OF TAIL CLIPPERS

JOGESWAR SATCHIDANANDA PUROHIT123*, MADAN M. CHATURVEDI2, & PRAGNYA PANDA3

1. Department of Zoology, Smt. C.H.M. College, University of Mumbai, Ulhasnagar, Post box No 17, 421003, Thane, Maharashtra. Phone: (+91) 9764036065 Fax: (+91 251) 2731169
2. Department of Zoology, University of Delhi, North Campus, Delhi
3. School of Biotechnology, KIIT University, Bhubaneswar, Orissa

E. mail: sachin.jogesh@gmail.com, jspurohit@zoology.du.ac.in

ABSTRACT

In the histone code hypothesis histone proteolysis is depicted as a novel kind of irreversible post-translational modifications of histones. Site specific proteolysis of histone is hypothesized as a process for removal of epigenetic signature of histones. However, though there are many reports depicting histone proteolysis, there are very few reports on characterization of a histone protease. In the present review we tend to summarize classify the existing histone proteases and we also discuss the role of these proteases in modulation of chromatin structure and function.

KEY WORDS: Histones, Histone specific protease, chromatin, epigenetic marks

INTRODUCTION

In order to be packaged inside the tiny nucleus, the eukaryotic DNA associates with histone and non histone proteins to form a dynamic polymer called chromatin. This is achieved by hierarchical packaging of DNA at multiple levels to form either a 10 nm or a 30 nm or a higher order chromatin structure. Since all nuclear events like replication, transcription, recombination and repair use chromatin as template, there is a requirement of differential unpacking and subsequent repacking of the chromatin, with minute temporal and spatial precision. This is accomplished, in part, by generating differentially folded and unfolded chromatin domains. The regulated folding or unfolding of chromatin domains is called chromatin remodelling, which is brought about broadly by two processes: ATP dependent chromatin remodelling or covalent modifications of histones.

Histones are subject to a diverse array of posttranslational modifications, namely acetylation, methylation, phosphorylation, ADP-ribosylation, ubiquitination, biotinylation, sumoylation, propionylation and butyrylation etc. Many of the histone modifications act as docking sites and cause further recruitment of protein complexes that bring about broadly two effects: a permissive chromatin domain for binding of factors; or an occlusive chromatin domain refractory for factor binding. Figure-1 depicts the complete activation and inactivation marks of core histones. Accordingly, there are modifications that coexist and work sequentially in a cooperative manner but are incompatible with others. Say, for example, when histone H3 K4 residue is methylated, there occurs synergistic acetylation of H3 K14, phosphorylation of Serine 10 H3. All are involved in transcriptional activation, and all of them are incompatible with the inhibitory H3 K9 methylation. Further, the complexity increases as lysines can be mono, di or tri methylated. Accordingly, they encode different signals for activation. Taken together, a language of histone modification can be created with the available information, which was initially framed as histone code.
Distinct modifications, on one or more histone tails, act sequentially to form a “histone code”, that is read by chromo and bromo-domain containing proteins. This creates specific combinations of modifications on a local chromatin domain. In the histone code it is also described that histone methylation is permanent (since at that time histone demethylases were not known), and hence was considered to be permanent epigenetic marks which primarily acts as transcriptional memory and further sets the modification of the daughter cells horizontally. However, with the discovery of enzymes responsible for erasing methylation marks from histones, it is high time the histone code should be re-explored. One line of hope in this context is that still there are some histone methylation sites for which demethylases are not yet discovered (H3 K79, H4 K20), and it still remains a question mark whether demethylases exist for erasing all the methylation marks of histones or not.

Further, in the histone code hypothesis, histone specific proteases are proposed to remove methylation marks from histone tails by proteolytic clipping of the histone tails and they could act as demethylases (a model is depicted in Figure 2). However, though there are many reports depicting existence of histone proteases, there are very few reports on characterization of a histone specific protease. Here, we intend to illustrate different histone proteases reported till date along with their role in modulation of chromatin structure and function.

![Fig.1: The complete post-translational modification map of core histones.](image1)

Fig.1: The complete post-translational modification map of core histones. The complete modification map of all core histone N-terminal or C-terminal regions is represented along with the notion of activation or a repression signal.

![Fig.2: A model depicting the proteolysis site of histone H3 and H4.](image2)

Fig.2: A model depicting the proteolysis site of histone H3 and H4. When histone H3 is cleaved at 23rd position it is a repressive signal as the clipped histone H3 lacks all the sites for acetylation and the H3K4 site for methylation. In contrast it can now be methylated at K27 which is a repressive signal. Similarly, cleavage of histone H4 at 16th position also abolishes all its acetylation sites and could renders it to heterochromatinization due to presence of the repressive signal K20 which can be methylated.
The proteolytic processing of histones

Proteolysis of histones, though rare, is believed to be another mode of irreversible post-translational modification assigned to histones. The N-terminal/C-terminal tails of histones are more susceptible to proteolysis while the globular domains are relatively resistant to this modification, when histones exist in the form of chromatin. Proteolysis of histones emerged as a problem for chromatin isolators. Subsequently, a number of reports started unravelling the role of histone proteolysis (see later). In few cases the protease responsible for the proteolysis were also partially/homogeneously purified and characterized. In the histone code hypothesis, these proteases were hypothesized to be responsible for modulating epigenetic marks (see Figure-2). However, an experimental demonstration of removal epigenetic marks from histone tails by such a protease still remains elusive. Nevertheless, the search for such a protease is on.

Classification of histone proteases

A number of reports depict characterization of histone specific proteases. By investigating the available information on the reaction conditions for histone proteases and the specificity for the histone substrates we categorize histone specific proteases on the basis of (i) optimum pH requirement for function and (ii) specificity to histone subtypes.

(i) Classification of histone proteases depending on pH

Depending on the pH required for optimum activity, histone proteases could be classified into three groups. Acid proteases: work at optimum pH 4.0-5.5; Neutral proteases: work at a pH range of 7.0 to 8.0; and alkaline protease: functioning at an approximate pH of 9.0-10.0.

(a) Acidic proteases

These proteases have been investigated to be capable of clipping histones at pH range of 1.75 to 5.5. A chromatin bound protease activity was readily detectable in avian erythrocyte chromatin, when the chromatin was incubated at pH 3. However, the histone degradation was minimal at neutral pH. It was further seen that the protease efficiently degraded chromatin bound histones, but free histones were relatively resistant for proteolysis. The protease was hypothesized to have role in histone turnover.

Regenerating rat liver nuclei have been shown to contain a thiol-protease with a pH optimum of 5.5. This enzyme has been hypothesized to participate in the degradation of ribosomal proteins and histones which are synthesized in excess.

A protease from calf-thymus nucleus is reported, which degraded the nucleoprotein complexes at an optimum pH 4.4. However, this acid protease was neither purified nor further characterized and later described as a contaminant cytoplasmic enzyme.

(b) Neutral proteases

Neutral proteases cleave histones at neutral pH and are hypothesized to be either chromatin bound or microsomal in origin. A neutral protease activity was demonstrated in chicken erythrocytes, which was loosely bound to chromatin as it was readily extracted with 0.25M ammonium sulphate. It was shown to degrade histones with selective preference for H2A. It was later confirmed to be of lysosomal origin. However, the protease was never purified.

From Xenopus liver nuclei, a proteolytic activity specific for histones is reported. The protease showed optimal activity at neutral pH. This protease was associated to chromatin. However, this protease activity was absent in many other tissues of Xenopus and rat liver, hence depicting a tissue specific and organism specific regulation. Later, it was also demonstrated to be a cytoplasmic neutral protease.

From rat liver chromatin a neutral protease is partially characterized. It possessed a molecular weight of 25 kDa and was able to degrade H1 more rapidly at pH 8.0 as compared to other chromatin associated proteins.

Kurecki et al. studied neutral proteases from the chromatin of different animal tissues and described that the neutral proteases degrade histone H1 and H3 when present along with nucleohistone complex; however after dissociation from the chromatin, they were capable of degrading all the histone fractions. The protease was purified to homogeneity from calf thymus chromatin. It was a 15.4 kDa protein.

(c) Alkaline proteases
These proteases show optimum activity at a pH range of 8.0-10.0. A protease which was capable of hydrolyzing casein with optimum pH 10.0 (alkaline protease) was reported to exist, being bound to chromatin of various normal and tumorous tissues of rats. It was speculated to be functional in hydrolysis of non-histone proteins and H1 histones. It had a molecular mass of approximately 18.0 kDa, and was a chymotrypsin-like protease as it was inhibited by soybean trypsin inhibitor and Chymostatin.

A protease activity associated with the Micrococcal nuclease-solubilized chromatin from mouse seminiferous tubules is characterized. The protease degraded all subtypes of histones with an optimum pH 9.0. Soluble chromatin from neonatal and prepubertal mice lacked this proteolytic activity until three to four weeks after birth. The protease activity was localized in the dinucleosomes and higher oligomers but was absent in mononucleosome populations, suggesting its association with the linker DNA. Rat testis-soluble chromatin apparently lacked such a protease activity. The developmental expression and in situ localization of the protease suggested that the protease could have a role in histone displacement during mouse spermiogenesis.

(ii) Classification of proteases depending on specificity

Depending on the substrate specificity the histone proteases can be classified into broadly two groups:

(a) Histone subtype non-specific protease

These proteases degrade histone as well as other non-histone proteins and do not show specificity to any histone subtype. An alkaline protease isolated from the chromatin of rat nuclei, which was named as protein A, degraded non-histone proteins and histone H1. The preference for hydrolysis was more for casein to histones. The other protease, i.e., the neutral protease (named as protein B) degraded other core histones in preference to non-histone proteins. The degradation of histones by non-specific proteases from calf thymus depended on the association of histones with DNA; wherein, only histone H1 was accessible to proteolysis in the nucleo-histone complex due to its exposure to the aqueous environment. But when histones were free from chromatin complex, all histones other than H1 were susceptible to proteolysis. A neutral protease was purified from rat liver chromatin having molecular weight ~2,00 kDa. This protease was hypothesized to play a role in the turnover of histone and non-histone proteins in the nucleus. In addition, it was speculated to have role in gene de-repression due to the removal of histones from DNA.

Some proteases have been reported to have more preference for a histone subtype. A protease purified from the calf thymus chromatin, degraded histone H1 more preferentially than histone H3. However, along with histones, it was capable of degrading high mobility group (HMG) proteins also. Another protease activity from avian erythroid cells was reported, which was capable of cleaving histones at pH 3.0 with high activity, and was less active at pH 9; was almost inactive at neutral pH. In acidic condition i.e. pH 3, histone H1 and H2A were being efficiently degraded by above mentioned protease, when total histones were used as substrates.

A non-lysosomal calcium activated neutral protease (CANP) from the calf thymus was partially purified. This protease required high concentration of calcium ions for its activity. Histone H2A, H2B and H3 were susceptible towards this protease, however it did not degrade histone H1 and H4. The unique property of this protease was that a histone, cleaved by generating peptide fragment by the specific protease, would become refractory for further cleavage by the same. The reason for this is however unknown. It is hypothesized that CANP was probably capable of recognizing some higher order structure of its substrate and once cleaved, as the higher order structure was lost, hence, it was not able to cleave further.

From the egg extract of sea urchin, a germinal histone specific protease was partially purified and characterized. The protease was specific to a repeat sequence of germinal cell histone sequence Ser-Pro-Lys-Lys (the SPKK motif). Since the sperm histone H1 and H2B are rich in SPKK motifs, the SPKK protease was hypothesized to be responsible for the unpacking of sperm chromatin due to histone degradation and in transcriptional activation during fertilization.

(b) Histone sub-type specific proteases

Histone H2A-specific protease activity is reported in chromatin of calf thymus. During the analysis of isolated total histones, one extra
A neutral protease B specific to histone H1 has been partially purified from rat liver chromatin by sepharose 6B and DEAE-sephadex column chromatography. The neutral protease B degrades all the core-histones. However, the H1 was digested only when it was complexed with equal amount of DNA. A nucleotide and pyrophosphate dependent histone H1-specific proteolytic activity has been demonstrated in permeabilized human lymphocytes.

In the total histones isolated from BHK cells infected with Foot and mouth disease virus (FMDV), a new polypeptide (Pi) was seen in SDS PAGE with concomitant disappearance of histone H3 (FMDV). After 2 hours of infection with FMDV, increase in the viral RNA synthesis takes place followed by the proteolytic cleavage of H3 to Pi. The protein Pi co-migrates along with histone H2A/H2B in the acetic acid/urea polyacrylamide gel. In the FMDV infected BHK cells, the histone H3 degradation could be blocked by treatment with cycloheximide immediately after infection, but after hours of infection the addition of cycloheximide could not inhibit histone H3 cleavage. It was concluded that after infection, the translation of histone H3 specific protease takes place immediately. After 3 hours of infection, histone H3 is totally replaced by the protein Pi in the nucleosomes. This concludes that HMDV infection creates a specific modification in the nucleus of infected cells. Similar degradation is observed in IB-RS2 (a swine-derived cell line) cells infected with FMDV virus, suggesting that the protease converting H3 to Pi is coded by the virus and not by the host cells. When BHK-21 cells are infected with FMDV it encodes he protease 3C. Subsequently, due to processing of histone H3 by protease 3C, the N-terminal tails of histone H3 are now no longer available for acetylation. Thus transcription of host cells are inhibited. This cleavage is mapped between Leu-20 and Ala-21 of the N-terminal of histone H3.

Two distinct forms of histone H3 were demonstrated in the micro-nuclei of Tetrahymena thermophila, those are H3^F (slower migrating) and H3^S (faster migrating). H3^F is unique only to micronuclei and is obtained from H3^S by proteolytic processing. The removal of 6 amino acids from the end terminal tail of H3^S for obtaining H3^F is determined by partial proteolytic peptide mapping. It was studied during conjugation. After 10 hr conjugation H3^S could be detected in micronuclei. So, the proteolytic processing is a physically regulated phenomenon. By immuno-fluorescence analysis it was described that the old micronuclei are prototypically removed from Tetrahymena during the period of conjugation. It is confirmed from the biochemical analysis of old micronuclei that polypeptides are N-termally cleaved due to proteolysis of core histones and are having
molecular weight 1-2 KD less than that of core histone in the senescent micronuclei. Microsequence analysis confirms that the polypeptides are removed due to the N-terminal tail processing by protease activity. This phenomenon strongly suggests about the transcriptional inactivation and highly condensed chromatin formation during the micro nuclear development of Tetrahymena.

Proteolytic degradation of histone H3 and H4 in the nucleoprotein of Cycad pollen is reported. When degraded nucleoprotein was extracted with ethanolic HCl, two extra bands appeared in the SDS PAGE in addition to H3 and H4. They were named as P1 and P2. By N-terminal amino acid sequencing of P1, the cleavage site was mapped to be between Lys-23 and Ala-24 in the histone H3 and the cleavage site of P2 was mapped to be between Lys-16 and Arg-17 in histone H4. It was proposed that the proteolytic enzyme responsible for the cleavage was having similar specificity like trypsin. It has been speculated that the rate of proteolytic degradation of histone H3 is much greater than that of histone H4.

Many reports suggest the synthesis and turnover of histone H1 in non-proliferating cells. A histone H1-specific protease was purified from the calf thymus. This protease preferentially cleaves histone H1 when total isolated histones are used as substrate. The nuclear extract also showed a chromatin bound protease that specifically cleaves histone H3 as its substrate. The roles of these proteases were speculated in protein turnover as well as in controlling gene expression. Two histone H1 like protein; Hc1 and Hc2 are identified from an obligate intracellular pathogen Chlamydia trachomatis. The Hc1 and Hc2 are two lysine-rich proteins of molecular weight 18 and 32kDa respectively, which are identical with histone H1. These proteins are expressed in the late phase of life cycle, at the time of DNA compaction for forming a dense structure followed by down-regulation of transcription as well as metabolic processes. The gene encoding the protease is identified to be EUO. It was further demonstrated that the product of the EUO gene was responsible for cleaving Hc1 leading to nuclear de-condensation. The EUO gene product was purified along with GST (Glutathione-s-transferase) and this purified fusion protein was able to digest Hc1 in vitro by incubation at 37°C for 1 hour.

From rat liver chromatin histone a histone H1 specific protease was partially purified by chromatography through sepharose-6B and DEAE-sephadex columns. It is a neutral protease B of molecular weight 25,000Da capable of degrading histone H1 at pH 8 in presence of DNA and its activity was further increased in presence of urea or by pretreatment of DNA with heat. By fluorescence studies it was detected that a concentration of 2M urea is enough for changing the globular structure of histone H1 in the chromatin by making it more susceptible for the action of the protease.

A chromatin-bound proteolytic activity with unique specificity for histone H2A is demonstrated in calf thymus nuclei. The cleaved product of histone H2A migrates as a single band in SDS-PAGE and is known as ch2A. The cleavage removes 15 amino acids from the C-terminal end of intact histone H2A, between Val-114 and Leu-115, yielding H2A1–114 in addition to a free pentadecapeptide in the high ionic strength condition. The enzyme catalyzing this cleavage is partially purified and named as ‘H2A-specific protease’. The H2A-specific protease remains selectively associated with H1-containing nucleosomes, and produces H2A1–114 from acetylated histone octamer-DNA reconstituted in the physiological condition. However, the fundamental properties of the H2A-specific protease, including its primary structure and biological role, remain obscure.

Two truncated forms of monoubiquitinated H2A were purified from the nuclei of acute myeloid leukemia OCI/AML1a cells. However, the enzyme(s) catalyzing these cleavages are not known. Histone H3 specific cleavage is observed during mouse embryonic stem cell differentiation. The protease responsible for the cleavage was identified to be Cathepsin L. It cleaves histone H3 after 21\textsuperscript{st} residue from the N-terminus and progressively removes several residues between amino acids 21 and 27 from the N-terminus. The H3 cleavage may be regulated by covalent modification status of histone tail itself. The H3 clipping activity is stimulated during a specific period during differentiation of the embryonic stem cells (ESC). The identification and characterization of developmentally-regulated H3 cleavage by Cathepsin L during ESC differentiation is an important step in understanding limited nuclear histone proteolysis as a potential mode of transcriptional regulation. It suggests that mouse ESCs employ a regulated histone H3 proteolysis mechanism that may serve to alter
epigenetic signatures upon differentiation. Cathepsin L was originally described as a lysosomal protease. However, Cathepsin L has been shown to localize in nuclei where it plays a role in the proteolytic processing of transcription factor CDP/Cux. Biochemical studies have also shown that procathepsin L is localized in the nucleus in ras-transformed mouse fibroblasts.

Recently, Kouzarides and co-workers identified a protease activity in Saccharomyces cerevisiae that cleaves histone H3 after Ala21 and has a recognition site Gln19-Leu20-Ala21. This activity is induced under conditions of nutrient deprivation (stationary phase) and sporulation. Genes activated under these conditions are sensitive to H3 tail clipping and, as a result, the level of expression from these genes is compromised. However, the endopeptidase activity is yet to be identified and purified. The endopeptidase activity was found to be sensitive to the modification status of the H3 tail. The presence of trimethylated H3K4, an active mark, is inhibitory for clipping, whereas the repressive dimethylated H3R2 mark is not. The selectivity for the clipping of tails with repressive marks is consistent with the idea that clipping may be necessary for the removal of transcription inhibitors at the onset of gene expression. Repressive protein complexes bound to the H3 tail at promoter regions may thus be removed to allow activator protein complexes to take over during the induction process.

Role of histone proteases

Proteases play important role in the synthesis and degradation of cellular proteins. In cell, the concentration of synthesis and degradation of histone proteins are hypothesized to be balanced by histone proteases. On the basis of different experimental analysis, several researchers have speculated the role of histone proteases; however their exact role still remains elusive. Many investigators have discussed the possible role of these proteases in histone turn over, de-repression of gene expression, removal of somatic histone during spermatogenesis, apoptosis and removal of epigenetic marks from histones.

Histone proteases histone turnover

The rate of histone turnover is remarkably more in proliferating tissues in comparison to non-proliferating tissues. Tissues like thymus and intestinal mucosa exhibit higher rate of histone turn over in comparison to other tissues, like brain and kidney. The turnover of histones which are present as nucleosomes in reticulocytes are reported to be higher than that of erythrocytes with concomitant existence of higher proteolytic activity in the chromatin bound histones of reticulocytes than that of erythrocytes. This implies the possible role of histone proteases in histone turn over. Several other investigators had similar conclusions.

Histone proteases transcriptional regulation

Histones proteases are hypothesized also to have role in gene regulation. The dormant nuclei from chicken erythrocytes having condensed chromatin were when fused with HeLa cells, the mass of the erythrocyte nuclei increased with concomitant chromatin activation and it followed subsequent translation cycle. This suggested the fact that, some factor(s) present in the cytoplasm of HeLa cells, activated the dormant chicken erythrocytic nuclei, by making it transcriptionally and translationally active. When the fusion was allowed in the presence of proteolytic inhibitors, this reactivation of erythrocyte nuclei was lost, indicating the role of histone proteases in the activation of erythrocytic chromatin. In vitro experiments suggest that histone H2A-specific protease plays important role in the decondensation of chromatin as it is demonstrated that the proteolytically processed cH2A:H2B dimer show lower affinity towards H3:H4 tetramer (described earlier). These mechanisms imply the activation of transcription due to unfolding of chromatin. In Tetrahymena, the histone proteolysis inactivates the senescent micro nuclei. During Tetrahymena micro-nuclear development, histone H1 is rapidly dephosphorylated and proteolysis of the core histone along with dephosphorylated histone H1 leads to the formation of highly condensed chromatin followed by transcriptional silencing. The activation of transcription occurs in the estrogen-primed rat uterus due to stimulation of histone proteases.

Histone proteases in spermatogenesis

During fertilization, proteases play an important role in the degradation of somatic histone. The role of histone specific proteases in the transformation of nucleoprotein to
nucleoprotamin complex in the isolated rat spermatid nuclei is reported. These histone proteases are responsible for the dissociation of protein-protein interaction due to removal of somatic histones from the mouse testicular chromatin. During the initial stages of spermatogenesis, histones are replaced transiently by transition proteins (TP, TP2, TP3 and TP4) and are later replaced by protamines. The mechanism of histones displacement requires their proteolysis for removal. A protease activity associated with mouse testicular chromatin has also been observed, which might be involved in proteolytic removal of somatic histones during spermatogenesis.

**Histone proteases in DNA repair**

During gamma-irradiation or DNase I treatment, selective degradation of histone H1 occurs due to the scaffold-associated proteinase. It is shown that the histone H1-specific proteinase (HSP) are activated in gamma irradiated nuclei and target the repair enzymes to the damaged portion of DNA for its subsequent repair, suggesting some role of histone proteases in DNA repair.

**Role of histone protease during aging**

Substantial qualitative and quantitative changes in histones have not been reported during aging. However, some recent observations have demonstrated the proteolytic processing of histone H3 and histone H1 in Japanese quail. It has been further demonstrated that progesterone induces the cleavage of histone H3 in Japanese quail, suggesting that the protease specific to histone H3 may be regulated by progesterone. We have observed an additional band in the histones prepared either from nuclei or purified nucleosome-core from liver of old rats. This band migrated between histone H2A and H4 (Chaturvedi M.M. Group, unpublished observation). The appearance of this band had a strong correlation with concomitant decrease in the histone H3. The clipping of the histone H3 was synergistic to declining transcriptional activity in aged rats and chicken liver tissues.

**Histone proteases in chromatin modulation**

Histone proteases were proposed with a role of erasing epigenetic marks, by proteolytically clipping the N-terminus or the C-terminus, when histone lysine demethylases were not known. However, until recently there was no report on characterization of such a histone protease having clear demonstration of modulation of epigenetic mark. Cathepsin L was the first protease characterized till date which could be correlated with removal of epigenetic marks from histone. It was demonstrated that Cathepsin L proteolytically clipped the tails those histone H3 which had signal of both activation and repression marks. The clipped histone H3 could then be docked with repressive signals essential for differentiation of the stem cells. We personally feel that there could be many histone proteases still to explore responsible for modulation of epigenetic marks. The discoveries of moonlighting proteins also have added to the list to search for.

**REFERENCES**


