Trends in Genetics

CellPress REVIEWS

Review

# Birth and Death of Histone mRNAs

William F. Marzluff[1,2,3,*] and Kaitlin P. Koreski[1,3]

**In metazoans, histone mRNAs are not polyadenylated but end in a conserved stem-loop. Stem-loop binding protein (SLBP) binds to the stem-loop and is required for all steps in histone mRNA metabolism. The genes for the five histone proteins are linked. A histone locus body (HLB) forms at each histone gene locus. It contains factors essential for transcription and processing of histone mRNAs, and couples transcription and processing. The active form of U7 snRNP contains the HLB component FLASH (FLICE-associated huge protein), the histone cleavage complex (HCC), and a subset of polyadenylation factors including the endonuclease CPSF73. Histone mRNAs are rapidly degraded when DNA replication is inhibited by a 3′ to 5′ pathway that requires extensive uridylation of mRNA decay intermediates.**

## Replication-Dependent Histone mRNAs: A Novel Set of Cell Cycle-Regulated mRNAs

Histone mRNAs are tightly regulated and are present in high levels only in S-phase to provide the histone proteins necessary for packaging the newly replicated DNA. Histones are among the most evolutionarily conserved proteins in eukaryotes. They form the fundamental unit of chromatin, the nucleosome, which packages the newly replicated chromosomal DNA. There are two major classes of histone proteins, the canonical replication-dependent histones and the histone variants. The replication-dependent histones are synthesized during S-phase, and comprise the bulk of the histones in the chromatin in multicellular organisms. The metazoan replication-dependent histone mRNAs are not polyadenylated but end instead in a conserved stem-loop, while in plants, and most single-cell eukaryotes, the replication-dependent histone mRNAs are polyadenylated [1]. After extensive deep sequencing analysis, the metazoan replication-dependent histone mRNAs remain the only known eukaryotic cellular mRNAs that are not polyadenylated.

There are also histone variants, for example H3.3, H2a.Z, macroH2a, and $H1^0$, which have specific functions and are typically constitutively expressed from polyadenylated mRNAs. An exception is histone H2a.X, which in vertebrates is synthesized in large amounts during S-phase from an mRNA ending in a stem-loop, but the same gene expresses a longer polyadenylated mRNA outside S-phase [2,3].

### SLBP Binds to the 3′ Ends of Histone mRNAs

**Stem-loop binding protein** (SLBP; see Glossary) provides the general functions of the poly(A) tail, including participating in translation and mRNA degradation (Figure 1, Key Figure). SLBP is a novel RNA-binding protein, and replication-dependent histone mRNAs are its only known target [4]. The structure of the SLBP–stem-loop RNA complex has been determined [5] (Box 1) and this complex at the 3′ end of the mRNA is involved in all steps of histone mRNA metabolism (i.e., processing, nuclear export, translation, and degradation of histone mRNA). There are no homologs of SLBP in species whose histone mRNAs are polyadenylated [6].

### Trends

The genes for all five replication-dependent histone mRNAs are linked in metazoans. They encode the only eukaryotic cellular mRNAs that are not polyadenylated.

SLBP, which binds to the stem-loop at the 3′ end of histone mRNA, is required for all steps of histone mRNA metabolism.

Factors for coordinating expression of the genes for the five histone proteins and the processing of histone mRNAs are concentrated in the HLB. The HLB is present constitutively and histone gene expression is activated by phosphorylation of NPAT, a crucial factor for HLB formation, by cyclin E/Cdk2.

The active form of U7 snRNP contains a novel set of factors, including FLASH and a complex of polyadenylation factors, the HCC.

Uridylation of histone mRNA maintains the proper length of the 3′ end. When DNA replication is completed or inhibited, histone mRNA is rapidly degraded by a 3′ to 5′ pathway initiated by 3′hExo, a component of the histone mRNP. Degradation is dependent on translation, and requires Upf1 as well as uridylation of the degradation intermediates by TUT7.

[1]Integrative Program for Biological and Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
[2]Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
[3]Curriculum in Genetics and Molecular Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

*Correspondence:
marzluff@med.unc.edu (W.F. Marzluff).

The SLBP RNA-binding domain can readily be identified by informatic approaches. Likely SLBP homologs have been identified in several single-cell eukaryotes including *Volvox*, *Dictyostelium*, and *Chlamydomonas*. These species also have stem-loops close to the 3′ end of their histone mRNAs, suggesting that histone mRNAs ending in stem-loops may have been present in some single-cell eukaryotes, and thus this property has been lost in evolution [6].

## Organization and Coordinate Expression of the Histone Genes

The high demand for histone protein in S-phase is met by the coordinated expression of multiple histone genes. In metazoans, the genes for the five replication-dependent histone proteins are tightly linked. These clusters may occur as tandem repeats containing a copy of each canonical histone gene (e.g., *Drosophila* and embryonic sea urchin histone genes) or in jumbled clusters where the genes have no repeating organization (e.g., the HIST1 cluster, ~60 genes) and the HIST2 cluster (~10 genes) in mammals (Figure 2A). *D. melanogaster*, which has a genome 1/30 the size of that of mammals, contains ~100 copies of a tandemly arrayed 5 kb repeat with each repeat containing one copy of the genes encoding the canonical histones (H2a, H2b, H3, and H4) and the linker histone H1. By contrast, none of the genes for histone variants H3.3, H2a.Z, and $H1^0$ are physically linked but are present as isolated single-copy genes.

The replication-dependent histone genes are present in a specialized nuclear domain, the HLB (Box 2). The HLB creates an optimal environment for efficient transcription and histone pre-mRNA processing (3′-end formation) by concentrating factors necessary for histone mRNA biogenesis [7]. Formation of the HLB requires **nuclear protein at the ataxia-telangiectasia locus** (NPAT; Mxc in *Drosophila*) which is essential for the expression of all five classes of histone genes [8,9]. **FLICE-associated huge protein** (FLASH) and **U7 small nuclear ribonucleoprotein** (U7 snRNP), both of which are required for processing, are also found in the HLB.

NPAT was discovered as a cyclin E substrate which is localized at the histone loci in mammals and is essential for histone gene expression [10,11]. Activation of histone gene expression requires phosphorylation of NPAT by cyclin E/CDK2. The *Drosophila* ortholog of NPAT was identified as the previously described homeotic gene, **multi sex combs** (*mxc*) [12]. Like NPAT, Mxc is also phosphorylated by cyclin E in S-phase [12] and is essential for histone gene expression [8]. A fourth component of the HLB, **muscle wasted** (Mute), appears to act as a negative regulator of histone gene expression in *Drosophila* [13,14]. The HLB, containing these four core factors, is present throughout the cell cycle and is dissembled at mitosis.

## Biosynthesis of Histone mRNAs

Replication-dependent histone mRNAs do not contain introns and require only one RNA processing step, endonucleolytic cleavage after the stem-loop, to form the 3′ end of histone mRNA. The 5′ caps on histone mRNAs contain 6-methyl (me) adenosine in a novel cap structure, $^{7Me}Gppp^{me}A_{O\text{-}me}N_{O\text{-}me}$ [15]. The significance of the 6-meA on the histone mRNA cap is not clear but, with the current excitement about multiple roles for 6-meA in RNA metabolism, this unusual cap may play a role in histone mRNA metabolism. Cleavage to form the 3′ end is defined by two sites in the RNA; the stem-loop bound by SLBP 5′ of the cleavage site, and the **histone downstream element** (HDE) which basepairs with U7 snRNP 3′ of the cleavage site (Box 3). Cleavage occurs very rapidly after transcription, and in both mammals and *Drosophila* transcription terminates shortly after the processing signal [16,17].

U7 snRNA is a small (<70 nt) RNA containing a 2,2,7 trimethyl G cap. Like the spliceosomal snRNPs, U7 snRNP can exist as a core snRNP and a larger holo-snRNP which contains factors essential for processing. The core U7 snRNP consists of U7 snRNA bound to five Sm proteins,

## Glossary

**FLICE-associated huge protein (FLASH):** an HLB protein necessary for histone pre-mRNA processing.
**Histone cleavage complex (HCC):** a complex of polyadenylation factors required for histone pre-mRNA processing. Components include symplekin, CstF64, CPSF100, CPSF160, and the nuclease CPSF73.
**Histone downstream element (HDE):** a purine-rich sequence in the pre-mRNA that basepairs with the U7 snRNA.
**Histone 3′ exonuclease (3′hExo/ERI1):** a 3′ exonuclease that forms a ternary complex on the stem-loop with SLBP.
**Multi sex combs (Mxc):** the *Drosophila* ortholog of NPAT.
**Muscle wasted (Mute):** a protein required for muscle development that is a component of the *Drosophila* HLB where it serves as a putative repressor of histone gene expression.
**Nuclear protein at the ataxia-telangiectasia locus (NPAT):** in mammals, NPAT is required for HLB formation and histone gene expression.
**Stem-loop binding protein (SLBP):** binds to the 3′ end of histone mRNA and is required for histone mRNA metabolism (processing, nuclear export, translation, degradation).
**Terminal uridylyl transferases:** members of the family of noncanonical poly(A) polymerases that can add oligo(U) tails to RNAs.
**U7 small nuclear ribonuclear protein (U7 snRNP):** composed of a U7 snRNA and a novel heptameric Sm ring made up of Sm proteins B', D3, E, F, G, and Sm-like proteins (LSm) 10 and 11.
**Yin Yang 1-associated protein-related protein (YARP):** a transcriptional repressor and a homolog of *Drosophila* Mute.