Biochimica et Biophysica Acta xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Role of H1 linker histones in mammalian development and stem cell differentiation

Chenyi Pan, Yuhong Fan *

[a] School of Biology, Georgia Institute of Technology, Atlanta, GA 30332, USA
[b] The Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, GA 30332, USA

ARTICLE INFO

Article history:
Received 9 November 2015
Received in revised form 9 December 2015
Accepted 9 December 2015
Available online xxxx

Keywords:
Linker histones
H1 variants
Mammalian development
Stem cell differentiation
Chromatin
Epigenetic gene regulation

ABSTRACT

H1 linker histones are key chromatin architectural proteins facilitating the formation of higher order chromatin structures. The H1 family constitutes the most heterogeneous group of histone proteins, with eleven non-allelic H1 variants in mammals. H1 variants differ in their biochemical properties and exhibit significant sequence divergence from one another, yet most of them are highly conserved during evolution from mouse to human. H1 variants are differentially regulated during development and their cellular compositions undergo dramatic changes in embryogenesis, gametogenesis, tissue maturation and cellular differentiation. As a group, H1 histones are essential for mouse development and proper stem cell differentiation. Here we summarize our current knowledge on the expression and functions of H1 variants in mammalian development and stem cell differentiation. Their diversity, sequence conservation, complex expression and distinct functions suggest that H1s mediate chromatin reprogramming and contribute to the large variations and complexity of chromatin structure and gene expression in the mammalian genome.

1. Overview: histone H1 and its variants in mammals

The DNA of all eukaryotic nuclei is packaged into chromatin by association with histone proteins. The basic repeating unit of chromatin is the nucleosome core particle, which consists of an octamer of four core histones (H2A, H2B, H3 and H4) wrapped by 147 bp of DNA [1,2]. Linker histone H1 binds to nucleosome core particles and the linker DNA between nucleosomes to facilitate the folding of the "beads-on-a-string" extended chromatin fiber into higher order chromatin structures, the 30-nm fiber [3–6]. For gene transcription to occur, the chromatin template plays a dynamic role. Nucleosome core particles and posttranslational modifications of core histones, such as acetylation, methylation, ubiquitination, phosphorylation and sumoylation, have been shown to play critical roles in gene activation and repression. Much less is known about the role of linker histone H1 and its variants. Here we focus our discussion on recent studies about the function of mammalian H1s in development and stem cell differentiation.

The H1 histone family is the most divergent and heterogeneous group of histones among the evolutionarily conserved histone protein families. There are multiple nonallelic linker histone variants present in higher organisms which provide additional levels of regulation on chromatin structure and function [7,8]. In mammals, 11 H1 variants have been identified, including seven somatic H1s ($H1^0$, H1a, H1b, H1c, H1d, H1e and H1x) and four germ cell-specific H1s (H1t, H1T2, H1LS1 and H1oo) (Table 1). Among the multiple nomenclature systems for mammalian somatic H1 variants, the alphabetic nomenclature (H1a–e, $H1^0$ and H1x) and the alternative numeric system (H1.1–1.5, H1.0 and H1.x) are most commonly used [8–12]. Table 1 summarizes their nomenclature as well as H1 variants' distinct expression and genomic localization patterns.

The expression of H1 variants is differentially regulated during mammalian development and cellular differentiation [13–19]. H1a through H1e are the main types of H1 ubiquitously expressed in somatic cells, yet their expression is tightly regulated with distinct levels in different tissues and cell types [14,15,20,21]. $H1^0$ is expressed mainly in differentiated and nondividing cells [22], whereas the least characterized H1x appears to exhibit a G1 phase-dependent nucleolar accumulation in cultured human cell lines [23]. H1oo and H1t are germ cell-specific H1s with expression in oocytes and testis, respectively [24,25]. H1T2 and H1LS1 are two H1t-related H1s specifically expressed in spermatids [26–30].

All of the mammalian H1 genes are transcribed by RNA Pol II. The genes of major somatic H1 variants (H1a, H1b, H1c, H1d and H1e) and H1t are intronless and transcribed in S phase of cell cycle in a DNA replication-dependent manner (Table 1). Their mRNAs have characteristics of DNA replication-coupled histone mRNAs, lacking a poly(A) tail but terminated by a conserved 3′ stem-loop structure. This stem-loop hairpin is bound by the histone RNA hairpin-binding protein or stem-loop binding protein (SLBP) which is required for histone pre-mRNA

* Corresponding author at: School of Biology and the Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, IBB 2313, 315 Ferst Drive, Atlanta, GA 30332-0363, USA.
E-mail address: Yuhong.fan@biology.gatech.edu (Y. Fan).

http://dx.doi.org/10.1016/j.bbagrm.2015.12.002
1874-9399/

Table 1
Mammalian histone H1 variants.[a]

Histone H1 variants (Alt. symbols)	Mouse histone H1 genes			Human histone H1 genes			Gene intronic[b]	Expression cell types[e]	Expression dependence on DNA replication	mRNA 3′ UTR
	Gene name	Accession no.	Chr location	Gene name	Accession no.	Chr location				
H1a (H1.1)	*Hist1h1a*	NM_030609	13	*HIST1H1A*	NM_005325	6	No	Somatic	Yes	Stem-loop
H1b (H1.5)	*Hist1h1b*	NM_020034	13	*HIST1H1B*	NM_005322	6	No	Somatic	Yes	Stem-loop
H1c (H1.2)	*Hist1h1c*	NM_015786	13	*HIST1H1C*	NM_005319	6	No	Somatic	Both[g]	Both[g]
H1d (H1.3)	*Hist1h1d*	NM_145713	13	*HIST1H1D*	NM_005320	6	No	Somatic	Yes	Stem-loop
H1e (H1.4)	*Hist1h1e*	NM_015787	13	*HIST1H1E*	NM_005321	6	No	Somatic	Yes	Stem-loop
$H1^0$ (H1.0, H1f0)	*H1f0*	NM_008197	15	*H1F0*	NM_005318	22	No	Somatic[f], oocyte	No	Poly-A
H1x (H1.x, H1fx)	*H1fx*	NM_198622	6	*H1FX*	NM_006026	3	Y/N[c]	Somatic	No	Poly-A
H1oo (H1foo)	*H1foo*	NM_138311	6	*H1FOO*	NM_153833	3	Yes	Oocytes, zygote and 2-cell embryo	No	Poly-A
H1t	*Hist1h1t*	NM_010377	13	*HIST1H1T*	NM_005323	6	No	Spermatocytes spermatids	Yes	Stem-loop
H1T2	*H1fnt*	NM_027304	15	*H1FNT*	NM_181788	12	No	Spermatids	No	Poly-A
HILS1	*Hils1*	NM_018792	11	*HILS1*	AY286318	17	Y/N[d]	Spermatids	No	Poly-A

[a] Gene names and accession numbers are adapted from [192].
[b] Data derived from GenBank accessions, AceView [193] and refs. [26–29,32,55,194].
[c] Human H1FX gene encodes multiple transcript variants with or without introns (see refs. [55,193]).
[d] Human *HILS1* gene consists of two exons and a small 106-bp intron and mouse *Hils1* gene is intronless (ref. [27]).
[e] Expression data derived from refs. [14,22,24–29,39,45,46,93,194–197].
[f] $H1^0$ is enriched in differentiated cells (see ref. [22]).
[g] The *H1c* gene encodes two versions of mRNAs, polyadenylated and stem-loop forms, allowing for both replication-dependent and replication-independent expression (see ref. [36]).

processing and enhances translation [31–35]. The transcription of these six H1 genes involves typical cell cycle regulation of histone genes: transcription initiation, 3′ end processing and mRNA stability. Interestingly, *H1c* gene also produces a polyadenylated form of mRNA besides the 3′ stem-loop form mRNA, allowing for independent regulation of expression in dividing and nondividing cells [36]. These six H1 genes (*H1a–e* and *H1t*) are linked and clustered together with core histone genes on mouse chromosome 13 and the orthologous human chromosome 6 [14,21,37–40]. This genomic organization is conserved from mouse to human. The other five H1 genes, $H1^0$, *H1x*, *H1oo*, *H1t2* and *Hils1*, however, are expressed in a DNA replication-independent manner, with mRNAs polyadenylated and lacking the 3′ stem-loop structure. In addition, they are not clustered with core histones but scattered in the mammalian genomes (Table 1).

All metazoan H1s share the same tripartite domain structure that includes a short, flexible N-terminal tail domain (NTD), a central globular domain (GD) with a winged-helix motif and a long, lysine-rich C-terminal tail domain (CTD) [41–43]. The H1 globular domain is highly conserved, while the N- and C-terminal regions of H1 variants contain more sequence divergence. Comparison of the sequences of mouse and human H1 genes reveals that each mouse H1 variant gene is more similar to its human ortholog than to other mouse H1 variant genes (Fig. 1, Table 2) [44]. Table 2 summarizes the sequence comparisons among all H1 variants in mice and humans. Mouse and human $H1^0$ (H1.0) are particularly conserved with 94.8% of homology, whereas they have diverged significantly from other H1 variants with homology less than 45% (Table 2) [44]. On the other hand, H1b (H1.5), H1c (H1.2), H1d (H1.3) and H1e (H1.4) are not only highly conserved from mouse to human with over 85% sequence identity, but also share most similarities with each other, with 76–86% sequence identity. Their sequence similarity suggests that H1b, H1c, H1d and H1e may be more functionally related to each other than to other H1s. Among the 7 somatic H1s, H1x is the least conserved and the most divergent, with 70.9% identity between mouse and human H1x and less than 37% similarities with other somatic H1s. H1x also contains a lower content of basic amino acid residues [45,46]. Germ cell-specific H1 variants are generally less conserved and more divergent than somatic H1s (Table 2). Mouse H1t shares a moderate identity of 61% with human H1t and 26–55% with somatic H1s, whereas other germ cell-specific H1s (H1oo, H1T2 and HILS1) have only 10–27% sequence identity with other H1 variants. The sequence conservation of the individual H1 variants suggests that H1 variants may have distinct roles in chromatin function and gene regulation in various cellular and developmental processes.

Consistent with the sequence heterogeneity of H1 variants, cumulative evidence suggests that H1 variants differ in their biochemical properties, affinities for chromatin, capabilities in chromatin compaction and binding partners [42,47–62]. The globular and C-terminal domains are required for high-affinity binding of H1 to chromatin [42,47–49]. The globular domain is suggested to bind to nucleosomes at the dyad and the linker DNA with symmetric or asymmetric models [42,63–67], whereas the CTD is likely to bind to the linker DNA non-specifically. The globular domain of H1 variants may bind to nucleosomes in distinctive structural modes, leading to varied higher order structures [67]. Atomic force microscopy, *in vitro* biochemical assays and fluorescence recovery after photobleaching (FRAP) studies have suggested that H1b, H1d and H1e, the somatic H1s with longer C-terminal tails, display higher affinity for chromatin than H1a, H1c and $H1^0$, the somatic H1s with shorter C-terminal tails [51–53]. H1b and H1d are sometimes categorized into the intermediate group of chromatin binding affinity, so is $H1^0$, the most lysine-rich H1 with a short C-terminal tail [51–53]. The CTD appears to be the key determinant for the chromatin binding affinity of somatic H1s [48,53]. The function of CTD in condensing chromatin is related to its length, the density of basic residues, the number of S/TPXK sites and its specific amino acid composition, as well as the intrinsic protein disorder in the CTD [53,68]. The N-terminal tail appears to be dispensable for chromatin binding, nevertheless, its deletion or swapping between different H1 variants alters the binding affinity of the respective H1 variant for chromatin [47,48,50]. Not surprisingly, different H1 variants also display differential *in vivo* binding dynamics in oocytes and during ES cell nuclear transfer [69]. The binding of H1 to chromatin is also regulated by post-translational modifications [54,55,70] and histone chaperones [71,72]. In addition to binding to DNA and nucleosomes, H1 variants interact with a variety of cellular proteins, which contributes to their diverse functions in various cellular processes [58–62].

Germ cell-specific H1 variants differ dramatically from somatic H1s in amino acid sequences and biochemical properties. H1t exhibits lower binding affinity for DNA and condenses chromatin to a lesser extent than somatic H1s [73–75], which may be attributed to the absence of the S/TPXK motifs, the sites for DNA binding and phosphorylation in CTD [76–78], and the single amino acid substitution of lysine observed in somatic H1s by glutamine in the H1t globular domain [79]. H1T2 is distinctive from H1t in that it is highly enriched with arginine residues and S/TPXK sites [28]. The oocyte-specific H1, H1oo, is the longest variant, with an NTD containing multiple potential phosphorylation sites and an extraordinarily long C-terminal tail rich in acidic amino acid residues [24]. Both the N-terminal and globular domains of H1oo

Please cite this article as: C. Pan, Y. Fan, Role of H1 linker histones in mammalian development and stem cell differentiation, Biochim. Biophys. Acta (2015), http://dx.doi.org/10.1016/j.bbagrm.2015.12.002