CONTRIBUTIONS to SCIENCE, **2** (2): 225-235 (2002)
Institut d'Estudis Catalans, Barcelona

Expression, structure and evolution of H1 linker histones

Imma Ponte, Roger Vila and Pere Suau*
Departament de Bioquímica i Biología Molecular. Universitat Autònoma de Barcelona

Abstract

H1 linker histones bind to linker DNA and contribute to the stabilization of both the nucleosome and chromatin higher-order structure. We have studied aspects of the expression, structure and evolution of H1 linker histones. H1 presents multiple isoforms. In rat cerebral cortex neurons, the subtypes H1a,b,c,d are replaced, to varying degrees, by H1e during postnatal development. In vivo labelling experiments have shown the importance of differential turnover and synthesis rates in defining the H1a-e subtype composition of chromatin. H1° is regulated in a different way, its expression being linked to neuronal terminal differentiation; it is also the only histone subtype which is regulated by external signals in specific neuronal populations. We estimated the rates of nucleotide substitution for several H1 subtypes. The rates of nonsynonymous substitution differed among subtypes by almost one order of magnitude. Such a wide variation in the tolerance of amino acid substitutions is consistent with the functional differentiation of the subtypes. We showed, using CD, ^{1}H-NMR and IR spectroscopy, that the H1 terminal domains acquire substantial amounts of secondary structure upon interaction with DNA. In H1°, we characterized an *a*-helical element in the N-terminal domain and a helix-turn motif in the C-terminal domain, both adjacent to the globular domain. The N-terminal domain of H1e contains two α-helices separated by a Gly-Gly motif, which behaves as a flexible linker.

Resum

Hem estudiat aspectes de l'expressió, l'estructura i l'evolució de les histones de la classe H1, que s'uneixen al DNA internucleosomal i participen en l'estabilització del nucleosoma i de la superestructura de la cromatina. Hem estudiat l'expressió dels subtipus de l'H1 a les neurones de l'escorça cerebral de rata. Durant el desenvolupament postnatal, els subtipus H1a, b, c i d són substituïts en diferents proporcions per l'H1e, que es converteix en el subtipus majoritari de les neurones madures. Diversos estudis de marcatge en viu han mostrat la importància del recanvi diferencial, juntament amb les tasses de síntesi, per a determinar les proporcions dels subtipus a la cromatina. L'expressió de l'H1° està associada a la diferenciació terminal neuronal. També és l'únic subtipus que respon a senyals externs en determinades poblacions neuronals. Hem estimat les velocitats de substitució nucleotídica no sinònima de diversos subtipus de l'H1. Les velocitats de substitució no sinònima varien al voltant d'un ordre de magnitud. Aquestes diferències de tolerància de les substitucions d'aminoàcids estan a favor de la diferenciació funcional dels subtipus de l'H1. Hem demostrat, mitjançant DC, ^{1}H-NMR i espectroscòpia d'IR, que els dominis terminals de les H1 adquireixen estructura secundària quan interaccionen amb el DNA. Hem caracteritzat un element α-helicoïdal en el domini N-terminal de l'H1° i un motiu hèlice-colze en el domini C-terminal, ambdós contigus al domini globular. El domini N-terminal de l'H1e conté dues hèlices α separades per un motiu Gly-Gly, que fa de frontissa.

Keywords: Chromatin, histone H1, expression, structure, evolution

Histone H1 binds to the linker DNA in chromatin fibres. It is currently accepted that H1 may have a regulatory role in transcription through the modulation of chromatin higher-order structure. In vitro experiments with reconstituted chromatin have shown that H1 can repress promoters containing the RNA start site in the linker DNA, and that some sequence-specific transcription factors can counteract the H1-mediated repression. Preferential binding to scaffold-associated regions and participation in nucleosome positioning have also been proposed as possible mechanisms by which H1 might contribute to transcriptional regulation [1-3]. Histone H1 has

*Author for correspondence: Pere Suau, Departament de Bioquímica i Biología Molecular, Facultat de Ciències, Universitat Autónoma de Barcelona. 08193 Bellaterra, Catalonia (Spain). Tel. 34 935811391. Fax: 34 93 5811264. Email: Pere.Suau@uab.es

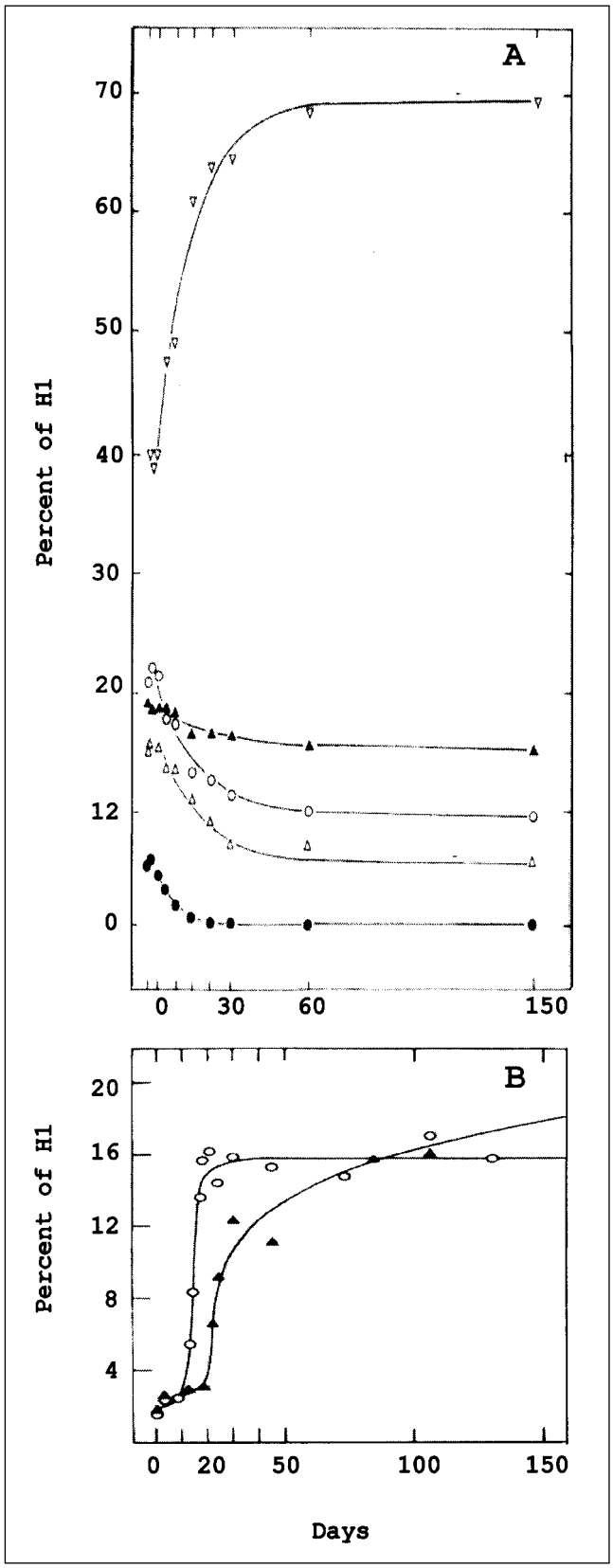

Figure 1. Change in the proportions of H1 subtypes during postnatal development in cerebral cortex neurons. (A) The time courses of the changes of H1a-e subtype proportions were approximated by single exponential functions. (B) Accumulation of H1°. (o) neuronal nuclei, (▲) glial nuclei.

been described as a general transcriptional repressor because it contributes to chromatin condensation, which limits the access of the transcriptional machinery to the DNA. However, experiments in vivo indicate that linker histones participate in complexes that can either activate or repress specific genes. Regulated expression of H1 during *Xenopus* development has a specific role in the differential expression of oocyte and somatic 5S rRNA genes [4]. In *Tethrahymena*, H1 does not have a major effect on global transcription, but can act as either a positive or negative gene-specific regulator of transcription in vivo [5]. Previous work has clearly established that the globular domain of H1 is sufficient to direct specific gene repression in early *Xenopus* embryos [6]. However, other gene-specific effects, such as activation of the mouse mammary tumor virus [7] or the activation or repression of specific genes in *Tethrahymena* are regulated by phosphorylation localized to the tail-like domains [8].

H1 has multiple isoforms [9]. The H1 complement has been best characterized in mammals, for which six somatic subtypes, designated H1a-e and H1°, and a male germ-line specific subtype, H1t, have been identified [10-12]. Although the precise function of H1 isoforms has yet to be determined, several observations suggest their functional differentiation. Some subtypes, particularly H1°, are non-randomly distributed in cell nuclei [13-16]. The subtypes differ in the extent of phosphorylation and in turnover rate [17]. In vitro evidence supports the idea that the subtypes differ in their ability to condense chromatin [18-19]. An earlier electrophoresis study suggested that subtypes H1a-e also differed in evolutionary stability [6].

H1 histones have a characteristic three-domain structure. The central domain is globular, while the N- and C-terminal domains are highly basic and have little or no structure in solution [3]. The C-terminus may, however, acquire a substantial proportion of α-helical structure upon interaction with DNA [20]. The distinct composition and conformation of the H1 domains suggests that they could play multifunctional roles in chromatin structure. The study of the detailed structure of the terminal domains, free in solution and bound to DNA, may provide insight into the binding of H1 in chromatin and thus contribute to the understanding of H1 function. Here, we summarize our studies on the expression, structure and evolution of H1 subtypes.

Expression of histone H1 subtypes in the central nervous system

Rat cerebral cortex neurons contain the five histone H1 subtypes, H1a-e, and the subtype H1° present in other mammalian somatic tissues. The four subtypes H1a-d decay exponentially during postnatal development and are partially or totally replaced by H1e, that becomes the major H1 subtype in adults [21]. The changes affecting the H1a-e complement are already apparent at birth, and the subtype composition is close to steady state by postnatal day 30 (Figure 1A). H1° is regulated in a different way. It accumulates in a

restricted period between postnatal days 8 and 18, coinciding with neuronal terminal differentiation, and its concentration remains stable thereafter [22,23] (Figure 1B).

The presence of changes in H1 subtype proportions during postnatal development indicates that the rates of synthesis and/or degradation must be different in neurons and in their germinal cells (neuroblasts).

Mammalian cortical neurons do not divide after birth [24], and the incorporation of newly synthesized chromatin can thus be followed by postnatal labelling. The study of histone synthesis in neuroblasts is not so straightforward because these germinal cells cannot be isolated for biochemical analysis. We devised a labelling format that allows access to histone synthesis in neuroblasts. It consists in injecting gravid rats with [^{14}C]Lys during the proliferation period of the fetuses' cortical neurons. The histone incorporated in dividing neuroblasts is then recovered from the neurons of the newborn [25].

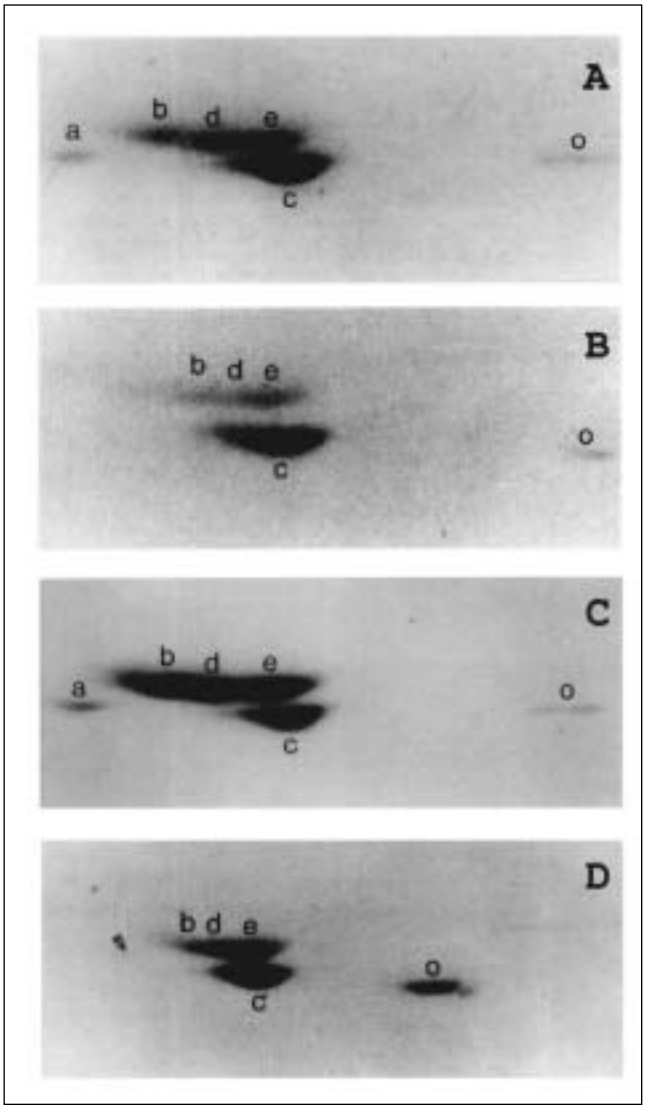

Figure 2. Mass and synthesis patterns of the histone H1 subtypes of rat brain cortical neurons and their germinal cells. (A) Fluorography showing the synthesis pattern of germinal cells obtained by prenatal labelling during the period of proliferation of the cortical neurons. (B) Synthesis pattern of neurons from newborn animals. (C) H1 fraction from newborn animals stained with Coomassie blue. (D) Synthesis pattern of cortical neurons from 30-day-old rats.

All subtypes are incorporated in dividing neuroblasts. The pattern is dominated by H1c with about 50% of total H1 (Figure 2). The subtype composition of newly replicated chromatin is basically determined by the relative synthesis rates of the subtypes. However, as soon as the chromatin assembly process is completed, the representation of the different subtypes will either increase or decrease according to the balance of synthesis and turnover rates. The pattern obtained by prenatal labelling can thus be considered a good approximation of the composition of newly replicated chromatin.

The neuron synthesis pattern obtained by postnatal labelling is clearly distinct from that of neuroblasts (Figure 2). H1a is not synthesized in neurons, H1b is almost undetectable, and H1d decreases to very low levels. In contrast, H1c and H1e are incorporated at levels equivalent to those of neuroblasts. Therefore, while the loss of H1a, H1b and H1d during postnatal development is a consequence of decreased synthesis, the accumulation of H1e and the decrease of H1c are due to the relatively higher turnover of H1c. It has been suggested that H1c would be much less efficient than the other subtypes in promoting chromatin higher-order structure [12]. If this was so, a clearer expression of the properties of H1c should be found in newly replicated chromatin before its concentration has been levelled down by its high turnover. In contrast, the subtype composition of mature chromatin is dominated by H1e, the subtype with the slowest turnover.

The regulation of H1^{o} differs from that of other H1 subtypes. The basal levels of H1^{o} present in germinal cells are not affected by neuronal determination and remain stable until the onset of neuronal terminal differentiation around postnatal day 9 (P9). Over the next ten days, the amount of H1^{o} increases seven-fold in the rat and fourteen-fold in the mouse [22] (Figure 1). Comparison of the in vivo synthesis patterns of immature and mature neurons shows that the accumulation of H1^{o} is due to an increased level of synthesis [25] (Figure 2).

To gain insight into the possible role of H1^{o} in nerve tissue, we analysed the central nervous system (CNS) of transgenic mice carrying a 3.2 kb fragment of H1^{o} upstream promoter sequences linked to the *E. coli* reporter gene *lac Z*. This fragment is sufficient to direct expression of the heterologous gene in the CNS, with spatial and temporal patterns that agree well with the developmental regulation of the endogenous protein [26]. The transgene is induced during postnatal development, coinciding with neuronal terminal differentiation. At P9, the earliest time at which the transgene product can be detected, positive neurons are observed in the granular layer of the cerebellar cortex and the granular and pyramidal fields of the hippocampus. The transgene is then induced in other CNS areas, such as the neocortex, thalamus, hypothalamus, olfactory bulb, globus pallidus, superior and inferior colliculus, substantia nigra, pontine nuclei and brain-stem.

The first positive glial cells in areas of white matter were observed at P21, in agreement with previous biochemical

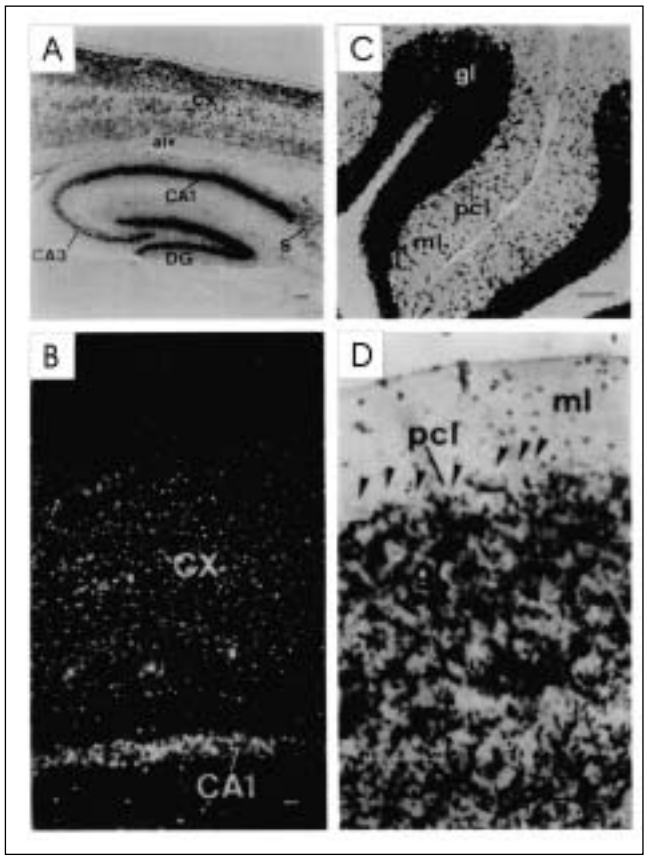

Figure 3. Expression of histone H1^o in adult mice. (A) H1^o-lac Z expression in the cerebellar cortex. (B) Distribution of H1^o-like immunoreactivity in the cerebellar cortex. (C) H1^o-lac Z expression in the cerebral cortex and hippocampus. (D) H1^o-like immunoreactivity in the cerebral cortex and hippocampus. alv, alveus; cx, cerebral cortex; CA1, CA3, pyramidal fields of the hippocampus; DG, dentate gyrus; gl, granular layer; pcl, Purkinje cell layer; ml, molecular layer. Purkinje cells (arrow heads) are negative.

observations showing that the accumulation of H1^o in glial cells of rat cerebral cortex is delayed by approximately ten days with respect to neurons [22].

A period of latency between the exit of the cell cycle and overt differentiation can be defined in the majority of CNS neuronal populations. For instance, mouse cerebral cortex neurons proliferate from gestational day 16 (E16) to shortly before birth, and terminally differentiate from about P9 to P30. Such a clear distinction between the periods of neuronal determination and differentiation has made it possible to show that activation of the H1^o promoter during postnatal development is unrelated to determination and quiescence as these take place during prenatal neurogenesis.

The coincidence of the temporal and regional distributions of endogenous protein and transgene activity shows that the onset of H1^o accumulation in differentiating neurons is, at least in part, under transcriptional control (Figure 3). However, other factors, such as post-transcriptional mechanisms, the balance of synthesis and turnover rates of the whole set of subtypes, and binding site availability [25], could also be involved in determining the final level of H1^o in the different cell types.

Although H1^o is widely expressed throughout the CNS, some neuronal classes do not express H1^o at detectable levels. Examples of these are Purkinje cells and the pyramidal neurons of the mitral layer of the olfactory bulb. Therefore, differentiation appears to be a necessary, but not sufficient, condition for H1^o transcriptional activation. H1^o does not contribute to the specificity of the phases of the cell cycle as it is synthesized at low basal levels throughout the cycle, nor does it contribute to the specificity of the quiescent state per se; the latter is shown by the absence of H1^o transcriptional activation in immature neurons, in which quiescence is not yet accompanied by differentiation. However, the H1^o promoter does have a property that is not shared by any other mammalian histone promoter, namely, it responds to terminal differentiation in specific cell types. This property suggests that the effects of H1^o on chromatin structure might be relevant to the expression of the neuronal differentiated phenotype.

H1^o is under hormonal regulation in several mouse tissues, such as prostate and *vas deferens* [27]. Testosterone appears to be necessary to maintain normal levels of H1^o in these tissues. We studied the gonadal hormone regulation of H1^o in the arcuate area of the hypothalamus. In females, the number of immunoreactive neurons in the arcuate increases during postnatal development. In contrast, the percentage of immunoreactive neurons in males is low at all ages studied (Figure 4). The expression of H1^o in the ventromedial part of the arcuate is reversibly and negatively regulated during the estrous cycle by the level of plasma estradiol. Early postnatal androgenization of females suppresses the increment in the number of immunoreactive neurons in the arcuate dur-

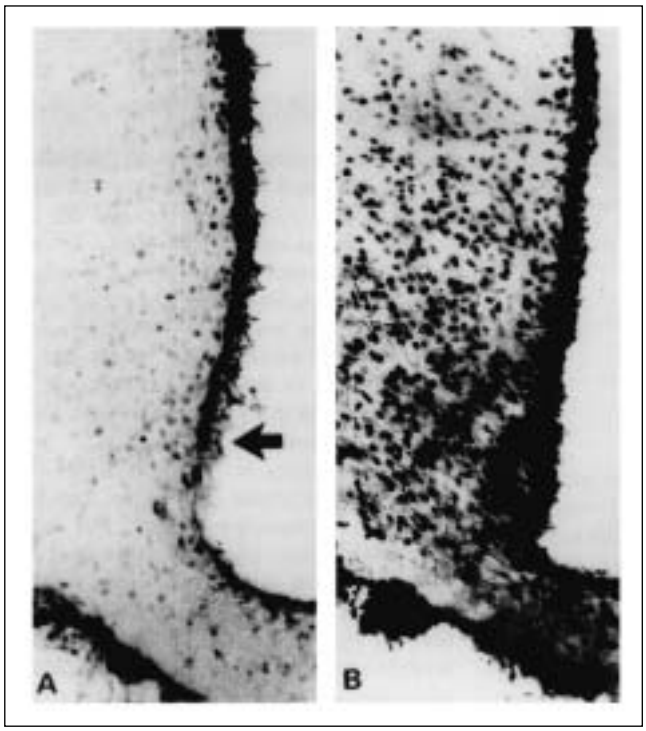

Figure 4. Distribution of H1^o immunoreactive cell nuclei in the mediobasal hypothalamus. The arrow in A is located in the third ventricle and points to the arcuate area. (A) 90-day-old male. (B) 90-day-old female in diestrus.

ing postnatal development, and leads to permanently decreased levels of H1° immunoreactivity in post-puberal females. Gonadal hormone regulation thus leads to sexual dimorphism in H1° expression [14].

In light of these results, H1° appears to be the only histone subtype capable of responding to external signals under normal physiological conditions.

To gain insight into the possible participation of H1° in regulating transcriptional activity in mature neurons, we studied H1° immunoreactivity in neurons of the supraoptic nucleus (SON) of the hypothalamus [28]. The neurosecretory cells of the SON produce the neurohormones vasopressin and oxytocin and are easily activated by osmotic stimulation. The increase in the overall transcriptional activity of SON neurons induced by osmotic stimulation is accompanied by a reduction in overall H1° immunoreactivity. Suppresion of neuronal stimulation causes an overexpression of H1°.

Immunogold labelling showed that H1° immunoreactivity was associated with chromatin fibres, the higher accumulation being on condensed chromatin. Stimulation of SON neurons induces the decondensation of chromatin throughout the cell nucleus. This suggests that differential expression of H1° could be involved in gene regulation through the modulation of chromatin higher order-structure.

Evolution

The expression of H1 subtypes is under distinct regulatory patterns. A major question about H1 subtypes is whether their expression patterns have co-evolved with functional differentiation. An earlier electrophoresis study suggested that subtypes H1a-e differed in evolutionary stability [12]. We estimated the rates of nucleotide substitution of several vertebrate H1 subtypes [29]. The rates of sequence divergence reflect the degree of functional constraint on the nucleotide and amino acid sequences. Thus, comparison of the rates for different subtypes might give some insight into their functional significance in the various cellular and developmental contexts in which they are expressed. We considered those subtypes which could be clearly identified as homologous in different lineages on the basis of the comparison of the amino acid sequences. These include mammalian and amphibian H1°, avian H5 and the mammalian subtypes H1a-e and H1t.

The H1 subtypes show a wide variety of substitution rates, from 0.13×10^{-9} nsyn/site/year in H1e to 1.16×10^{-9} nsyn/site/year in H1t. The individual domains cover an even wider range of substitution rates, from effectively 0.00 in the globular domains of mammalian H1d and H1° to 1.88×10^{-9} nsyn/site/year in the N-terminal domain of H1t. H1 is the fastest-evolving histone class. However, the H1 subtypes, except for H1t and H1a, can still be considered to be highly-conserved proteins. H1e, the most conserved subtype, evolves at a rate which is only about twice as fast as those of vertebrate H2A and H2B (Table 1).

The terminal domains are generally more variable than the central globular domain. The absence of tertiary structure in the terminal domains may contribute to the generally higher tolerance of amino acid substitutions in the terminal regions of the molecule. However, this and other general effects must be modulated by specific constraints operating on the individual domains in order to explain the non-systematic maintenance of the rate ratios among the domains of the molecule in the different subtypes. The most extreme

Table 1. Rates of synonymous and nonsynonymous substitution in histone H1 subtypes

	Entire Protein		N-Terminal Domain		Globular Domain		C-Terminal Domain	
	Nonsynonymous Rate	Synonymous Rate	Nonsynonymous Rate	Synonymous Rate	Nonsynonymous Rate	Synonymous Rate	Nonsynonymous Rate	Synonymous Rate
Mammalian H1° and *Xenopus* H1°								
Rat H1° vs. human H1°	0.14 (0.04)	1.81 (0.32)	0.29 (0.18)	5.40 (1.62)	0.00 (0.00)	1.60 (0.44)	0.21 (0.07)	1.34 (0.44)
Mouse H1° vs. human H1°	0.13 (0.04)	1.88 (0.32)	0.00 (0.00)	1.88 (1.01)	0.00 (0.00)	1.88 (0.49)	0.27 (0.08)	1.85 (0.46)
Mouse H1° vs. *Xenopus* H1°a	0.20 (0.03)	—	0.21 (0.09)	—	0.07 (0.02)	—	0.30 (0.05)	—
Mouse H1° vs. *Xenopus* H1°b	0.22 (0.03)	—	0.15 (0.07)	—	0.11 (0.03)	—	0.32 (0.05)	—
Rat H1° vs. *Xenopus* H1°a	0.19 (0.03)	—	0.20 (0.09)	—	0.07 (0.02)	—	0.28 (0.05)	—
Rat H1° vs. *Xenopus* H1°b	0.20 (0.03)	—	0.15 (0.07)	—	0.11 (0.03)	—	0.30 (0.05)	—
Human H1° vs. *Xenopus* H1°a	0.17 (0.02)	—	0.21 (0.09)	—	0.07 (0.02)	—	0.26 (0.04)	—
Human H1° vs. *Xenopus* H1°b	0.19 (0.03)	—	0.15 (0.07)	—	0.11 (0.03)	—	0.27 (0.04)	—
Average	0.18 (0.01)	1.85 (0.22)	0.17 (0.03)	2.64 (0.95)	0.06 (0.01)	1.74 (0.33)	0.28 (0.02)	1.79 (0.32)
Avian H5								
Chicken H5 vs. duck H5	0.34 (0.07)	2.07 (0.34)	1.31 (0.44)	2.46 (1.13)	0.28 (0.09)	1.85 (0.48)	0.21 (0.07)	2.16 (0.52)
Mammalian H1t								
Rat H1t vs. human H1t	1.20 (0.13)	—	1.88 (0.41)	—	0.47 (0.12)	—	1.62 (0.23)	—
Rat H1t vs. *rhesus* H1t	1.12 (0.12)	—	1.67 (0.38)	—	0.40 (0.11)	—	1.56 (0.22)	—
Mouse H1t vs. human H1t	1.22 (0.13)	—	1.85 (0.42)	—	0.52 (0.12)	—	1.63 (0.23)	—
Mouse H1t vs. *rhesus* H1t	1.13 (0.12)	—	1.59 (0.37)	—	0.44 (0.12)	—	1.56 (0.22)	—
Average	1.16 (0.06)		1.75 (0.20)		0.46 (0.06)		1.59 (0.11)	
Mammalian H1a								
Mouse H1a vs. human H1a	0.61 (0.09)	—	1.03 (0.29)	12.64 (11.25)	0.23 (0.08)	13.47 (9.10)	0.79 (0.15)	—
Mammalian H1b								
Mouse H1b vs. human H1b	0.20 (0.05)	10.06 (2.41)	0.15 (0.10)	—	0.16 (0.07)	6.44 (1.70)	0.25 (0.07)	10.75 (4.07)
Mammalian H1c								
Mouse H1c vs. human H1c	0.30 (0.06)	5.71 (0.87)	0.54 (0.20)	4.75 (1.69)	0.03 (0.03)	4.36 (1.03)	0.43 (0.10)	7.95 (2.20)
Mammalian H1d								
Mouse H1d vs. human H1d	0.37 (0.06)	6.88 (1.13)	0.60 (0.22)	5.39 (1.98)	0.00 (0.00)	5.63 (1.40)	0.56 (0.11)	9.13 (2.79)
Mammalian H1e								
Mouse H1e vs. human H1e	0.14 (0.04)	4.02 (0.57)	0.15 (0.10)	2.74 (0.99)	0.00 (0.00)	3.62 (0.35)	0.21 (0.07)	4.97 (1.05)
Rat H1e vs. human H1e	0.11 (0.03)	4.16 (0.59)	0.26 (0.14)	2.97 (1.06)	0.03 (0.03)	2.65 (0.64)	0.15 (0.06)	6.81 (1.64)
Average	0.13 (0.03)	4.09 (0.41)	0.21 (0.09)	2.81 (0.72)	0.02 (0.02)	3.13 (0.36)	0.18 (0.04)	5.89 (0.97)

Note.—Rates are in units of substitutions per site per 10^9 years. The average is the arithmetic mean, and values in parentheses are the standard deviations. In the comparisons between mammalian H1° and *Xenopus* H1° and between rodent and primate H1t, the sequence difference was too great to allow the use of the Jukes-Cantor correction.

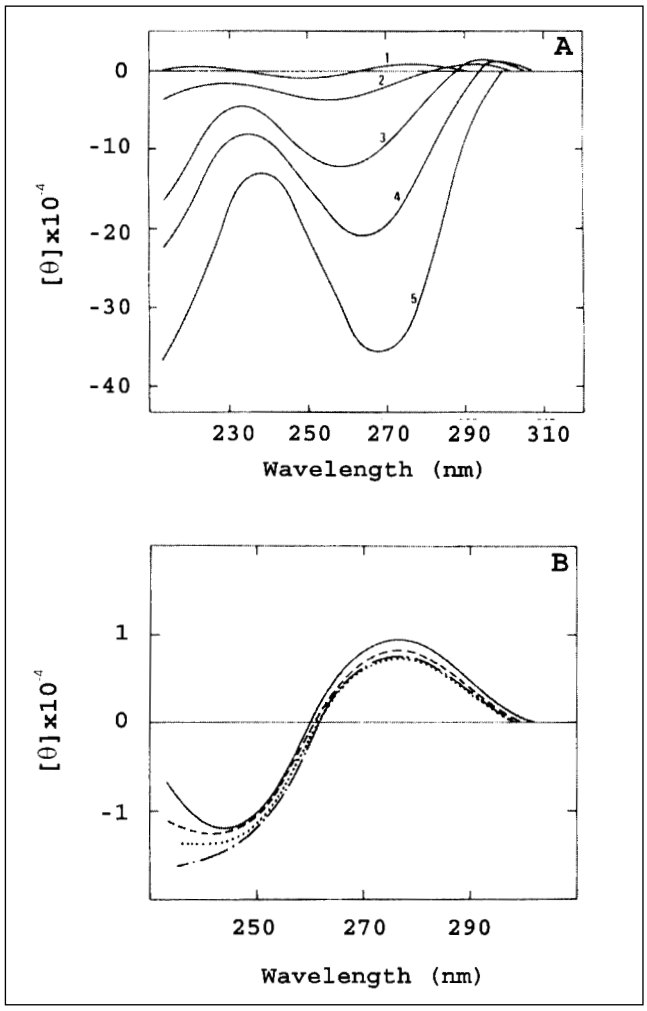

Figure 5. Circular dichroism spectra of complexes of the structural domains of histone H1 with DNA. (A) Spectra of complexes of the C-terminal domain of H1 with DNA as a function of ***r***, the weight ratio peptide DNA in 0.14 M NaF, 2 mM phosphate buffer, pH 7.0. 1, native DNA; 2, ***r*** = 0.1; 3, ***r*** = 0.3; 4, ***r*** = 0.4; 5, ***r*** = 0.5. (B) Spectra of complexes of DNA with the globular domain (GH1) and a peptide containing the N-terminal domain and the globular domain of H1 (NTB). (——) native DNA; (.–.–.) GH1-DNA complex, ***r*** = 0.4; (.....) NTB-DNA complex, ***r*** = 2.0; (-----) NTB-DNA complex, ***r*** = 4.0.

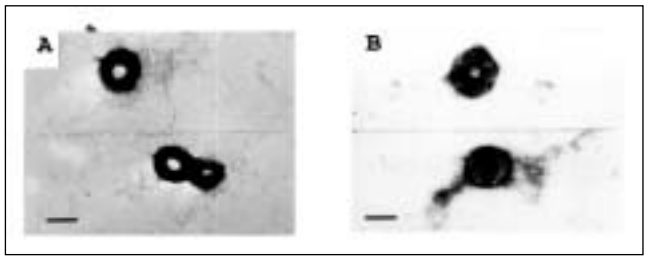

Figure 6. Electron micrographs of complexes of the isolated C-terminal domain of histone H1 with DNA in 140 mM NaCl, 1 mM phosphate buffer, 0.8 mM Na_2EDTA, pH 7.4. The samples were contrasted with uranyl acetate. (A) ***r*** = 0.4, positive staining; (B) ***r*** = 0.4, negative staining. The bar represents 0.1 μm.

case is the C-terminal domain of H5, which is even more conserved than the globular domain. Considering that the role of H5 is apparently restricted to tight chromatin condensation and inactivation in some nucleated erythrocytes, the higher evolutionary stability of the C-terminus compared with the other domains might be accounted for by the observations indicating that the C-terminus is the main protein region responsible for chromatin condensation [30-32].

The wide range of substitution rates in H1 genes favours the functional differentiation of the subtypes. Furthermore, the non-systematic maintenance of the rate ratios of the molecule's domains is consistent with the assumption that the domains have specific functions. Several features, such as the terminal domain lengths, the amount of secondary structure in the C-terminal domain, and the specific phosphorylation patterns of the subtypes, could be relevant to domain functional modulation. It should be noted that in addition to nucleotide substitution, the mechanisms leading to short insertions/deletions have been important in the evolution of H1 coding and flanking regions [33, Ponte et al., in preparation]

Interaction of the C-terminal domain of histone H1 with DNA

We studied the interaction of the C-terminal domain of H1 using circular dichroism (CD), precipitation curves and electron microscopy. As the intact H1 molecule, its C-terminal domain induces the so-called psi state of DNA, which is characterized by a non-conservative CD spectrum. This kind of spectrum is attributed to ordered aggregation of the DNA molecules. On a molar basis, intact H1 and its C-terminal domain gave spectra of similar intensity. Neither the globular domain nor an N-terminal fragment, that includes both the globular and the N-terminal domains, had any effect on the conservative CD of DNA. From these results it was concluded that the condensation of DNA mediated by histone H1 was mainly due to its C-terminal domain (Figure 5). This suggests that the C-terminal domain is also involved in the condensation of the chromatin fibre [30].

The C-terminal domain may contribute to the bending of the spacer DNA in the 30 nm filament. Several authors have described a salt-dependent transition from non-cooperative to cooperative binding of H1 to either linear or supercoiled DNA in the range 20-50 mM NaCl [34-37]. We used precipitation curves to demonstrate salt-dependent binding cooperativity of the isolated C-terminal domain to linear DNA [32]. The transition towards cooperative binding occurs in the range 30-60 mM NaCl. The protein/DNA ratio (by mass), ***r***, of fully saturated complexes was 0.7, which is equivalent to a +/- ratio of 1.0. The DNA binding sites for the C-terminal domain and whole H1 were calculated from the stoichiometry of insoluble complexes. Values of 22 and 33 base pairs were found for the C-terminal domain and whole H1, respectively.

The isolated C-terminal domain of H1 condenses the DNA in toroidal particles (Figure 6). They are best observed at physiological salt concentrations (140 mM NaCl), but also at very low salt (10 mM NaCl). In the latter case, a small excess of C-terminal peptide is necessary for toroids to become the predominant structure, whereas at physiological salt toroids are observed at ***r*** values well below charge saturation [32].

This is consistent with the results of precipitation curves showing salt-dependent binding cooperativity. The average inner radius of the particles is of the order of 19.5 nm, at both low and physiological salt conditions.

Toroidal DNA is wound circumferentially in a single direction with a minimum radius of curvature bounding the hole in the middle. Manning calculated that if the phosphate charge is neutralized, the DNA molecules spontaneously adopt a bent conformation with a radius of curvature of about 17.0 nm [38]. Wilson and Bloomfield [39], using Manning's counterion condensation theory, calculated a critical value of 90% charge compensation for toroid formation, and observed that saturation with Mg^{2+} (88% charge neutralization) does not induce toroids. Therefore, the observation of toroids with inner radii of about 19.5 nm suggests that the H1 C-terminal domain is capable of neutralizing most of the charge of its DNA binding site.

The C-terminal domain binds to linker DNA in chromatin. The ability of the C-terminal domain to achieve almost full electrostatic neutralization of the DNA charge might have a profound effect on the properties of the linker.

Structure of the terminal domains of histone H1 in solution and bound to DNA

H1 terminal domains have, in general, little structure in solution. However, the C-terminal domain acquires a substantial amount of α-helical structure in the presence of secondary-structure stabilizers such as trifluoroethanol (TFE) or $NaClO_4$ [20-40]. It is important to establish the structural properties of the N- and C-terminal domains of linker histones in solution and when bound to DNA in order to understand the role they play in chromatin higher-order structure and gene regulation.

We have studied the conformational properties of several peptides belonging to the N-terminal and C-terminal domains by CD, ^{1}H-NMR and IR spectroscopy:

Ac-EPKRSVAFKKTKKEVKKVATPKK (CH-1). This belongs to the C-terminal domain of histone H1^o (residues 99-121) and is adjacent to the central globular domain.

Ac-TENSTSAPAAKPKRAKASKK-NH_2 (NH-1) , which corresponds to the complete N-terminal domain of H1^o (residues 1-20), and Ac-PAAKPKRAKASKKSTDHPKYSDM-NH_2 (NH-2), which partially overlaps with NH-1, but lacks the first 7 residues and extends 10 residues into the adjacent globular domain (residues 8-29).

Ac-EKTPVKKKARKAAGGAKRKTSG-NH_2 (NE-1), which corresponds to the positively-charged region of the N-terminal domain of H1e (residues 15-36).

The C-terminal domain

In aqueous solution, CH-1 behaved as a mainly unstructured peptide, although turn-like conformations in rapid equilibrium with the unfolded state could be present. Addition of TFE resulted in a substantial increase of the helical content (Figure 7). The helical limits, as indicated by (*i*, *i*+3) nuclear Overhauser effect (NOE) cross correlations and significant up-field conformational shifts of the C^α protons, spanned from Pro100 to Val116, with Glu99 and Ala117 as N- and C-caps, respectively. A structure calculation performed on the basis of distance constraints derived from NOE cross peaks in 90% TFE confirmed the helical structure of this region (Figure 7). The helical region had a marked amphipathic character, due to the location of all positively-charged residues on one face of the helix and all the hydrophobic residues on the opposite face. The peptide has a TPKK motif at the C-terminus, following the α-helical region. The ob-

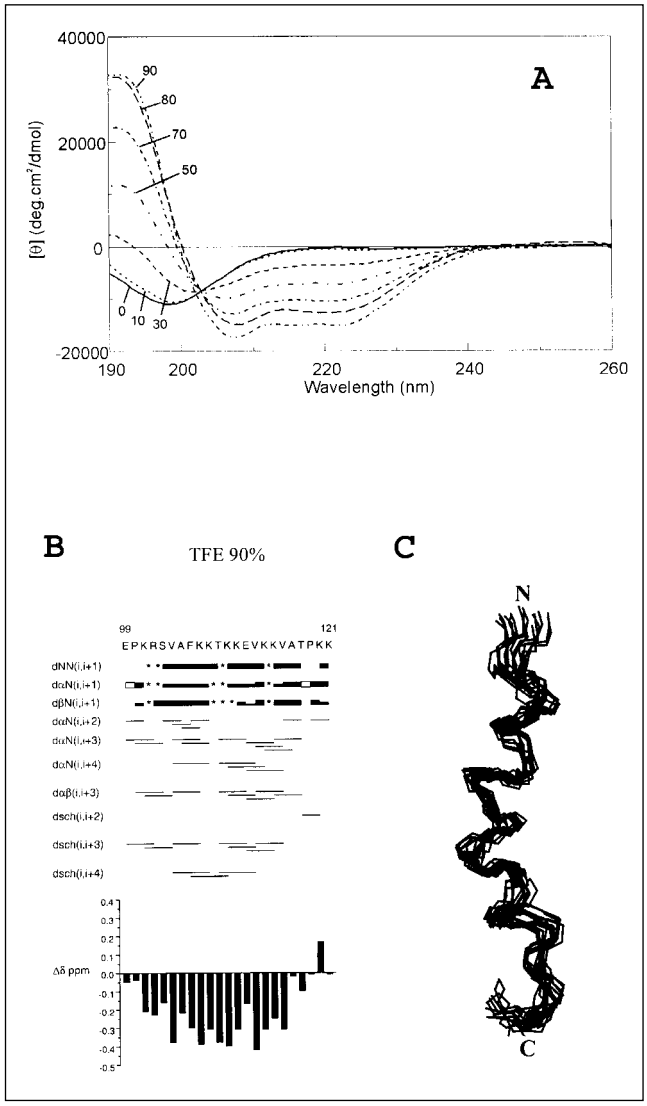

Figure 7. Structure of the CH-1 peptide. (A) TFE-dependent conformational transition of the CH-1 peptide measured by CD. The numbers refer to the TFE concentration in percentage by volume. (B) Summary of the NOE connectivities of CH-1 in 90% TFE solution. The thickness of the lines reflects the intensity of the sequential NOE connectivities. An asterisk (*) indicates an unobserved NOE connectivity due to signal overlapping, closeness to the diagonal or overlapping with the solvent signal. An open box indicates a $d\alpha\delta(i,i+1)$ NOE connectivity, where $i+1$ is proline, dsch indicates NOE connectivities involving side chains. The conformational shifts with respect to random coil values, $C^\alpha H\Delta\delta$, are shown as a function of the sequence number (bottom). (C) Superposition of the best 15 structures of the CH-1 peptide calculated by distance geometry methods on the basis of the distance constraints derived from observed NOE cross correlations in 90% TFE solution.

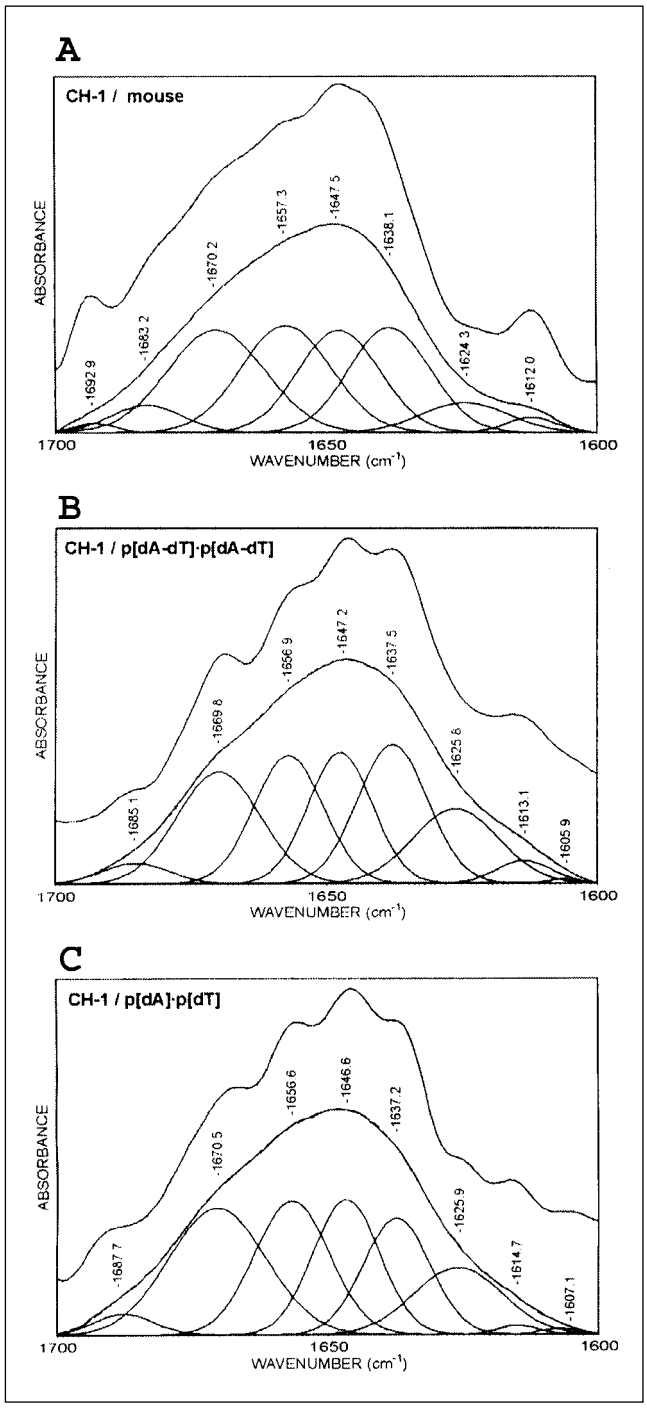

Figure 8. Amide I' band decomposition of the spectra of the CH-1 peptide bound to DNA. (A) Mouse DNA. (B) Poly(dA-dT)·poly(dA-dT). (C) Poly(dA)·poly(dT). The reconstitution of the absorption spectra from the component bands is indicated by dashed lines. The Fourier decomposition of the absorption spectra, that was used to fix the position of the component bands, is included. The DNA contribution to the spectra of the complexes was subtracted as described in the text. Peptide/DNA ratio = 0.7 (w/w).

served NOE connectivities suggest that the TPKK sequence adopts a type (I) β-turn conformation, a σ-turn conformation or a combination of both, in fast equilibrium with unfolded states [41] (Figure 7). Sequences of the (S/T)P(K/R)(K/R) kind have been proposed as DNA binding motifs [42].

We used IR spectroscopy to identify the features of the low-resolution structure of the bound peptide [43]. IR spectroscopy is particularly well suited to the study of complexes with basic peptides, since it is not affected by turbidity. The DNA contribution to the spectra of peptide-DNA complexes was subtracted using a DNA sample of the same concentration. The asymmetric component of the phosphate vibration at 1087 cm^{-1} was used to check the accuracy of the subtraction. The difference spectra of complexes of different peptide/DNA ratio were basically identical, indicating that the amide I' region was not significantly affected by DNA spectral changes (Figure 8).

In aqueous (D_2O) solution the amide I' is dominated by component bands at 1643 cm^{-1} and 1662 cm^{-1}, which have been assigned to random coil conformations and turns, respectively. In accordance with previous NMR results, the latter component was interpreted as arising in turn-like conformations in rapid equilibrium with unfolded states. The peptide becomes fully structured either in 90% TFE or upon interaction with DNA. In these conditions, the contributions of turn (1662 cm^{-1}) and random coil components virtually disappear. In TFE, the spectrum is dominated by the α-helical component (1654 cm^{-1}). The band at 1662 cm^{-1} shifts to 1670 cm^{-1}, and has been assigned to 3_{10} helical structure. The amide I' band of the complexes with the DNA retains the components that were attributed to the 3_{10} helix and the TPKK turn (Figure 8). In the complexes with DNA, the α-helical component observed in TFE splits into two components at 1657 cm^{-1} and 1647 cm^{-1}. Both components are inside the spectral region of α-helical structures. Our results support the presence of inducible helical and turn elements, both sharing the character of DNA-binding motifs.

The C-terminal domain is presumably involved in the organization of the linker DNA in chromatin. The CH-1 peptide is adjacent to the globular domain and may affect the conformation of the initial part of the linker DNA at the entry/exit of the nucleosome. This could have important consequences, particularly, in those cases where gene regulation is mediated by the tail-like domains [7,8].

The N-terminal domain

The first half of the N-terminal domain of H1° is devoid of basic residues, and it is not expected to interact strongly with DNA. In contrast, the second half of the domain is highly basic, with 1 Arg and 5 Lys residues. The proximity of this region to the globular domain suggests that it may contribute to the binding stability of the globular domain in chromatin [31]. The clustering of the basic residues of the N-terminal domain in the vicinity of the globular domain, leaving the rest of the domain free of basic residues, is a common feature of H1 subtypes [44].

We have studied two peptides: one comprising the 20 residues of the N-terminal domain (NH-1), as defined by trypsin cutting at Lys20, and the other of 23 residues, spanning 10 residues into the globular domain, and lacking the first 7 residues of the protein (NH-2) [45]. NH-1 and NH-2 have little structure in aqueous solution. The IR spectrum of both peptides is dominated by the random coil and, to a lesser extend, by turn conformations, probably in rapid equilibrium with the unfolded state. Small amounts of α-helix are

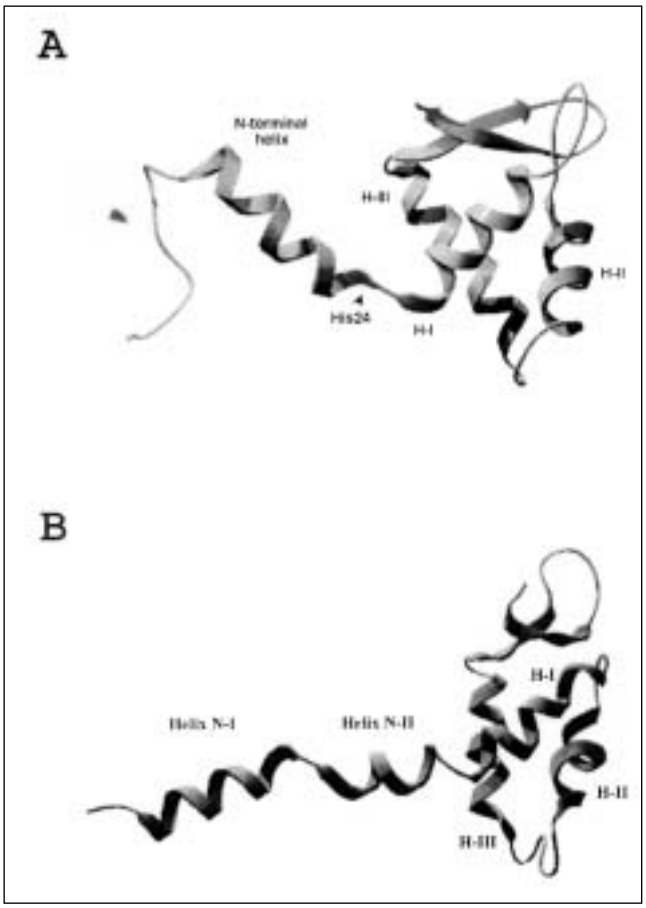

Figure 9. Structure of the N-terminal domain. (A) NMR structure of the folded N-terminal domain of H1° connected to the crystal structure of the globular domain of H5. (B) NMR structure of the folded N-terminal domain of H1e connected to the NMR structure of the globular domain of chicken histone H1 globular domain.

also present. TFE induces a significant increase in the amount of α-helical structure in both peptides. NH-2 has a much higher helical propensity than does NH-1, with 30% helical structure compared with the 18% of NH-1. When the amount of α-helix for the entire NH-2 peptide is corrected for the length of the helical region (13 residues), a fractional helicity of 57% is obtained. This relatively high helical propensity explains the abundance of medium range NOE cross-correlations in this region.

The NMR analysis shows that the α-helix is induced in the positively-charged second half of the N-terminal domain. In NH-1, the helical region spans from Lys11 to Arg19. In NH-2, the helical element also begins at Lys11, but it spans until Asp23, with His24 as the C-cap, thus surpassing the limit of the globular domain as defined by trypsin cutting at Lys20. The N-terminal helix element is thus close to helix I of the globular domain, which starts at Lys26. The extension of the structural limit of the N-terminal domain beyond the proteolytic limit, together with the proximity of the N-terminal helical element to helix I, supports the view that the N-terminal basic cluster and the globular domain of histone H1 may act in concert in chromatin structure [31]. The proximity of the N-terminal basic cluster to the DNA at the entry/exit point of the nucleosome is supported by the cross-linking of His25 of histone H5, an avian subtype closely related to H1°, to the terminal regions of chromatosomal DNA [46]. Figure 9A shows a model structure of the N-terminal domain of H1° connected to the globular domain of H5 [47].

The distribution of the basic residues in the N-terminal helix is not amphipathic. However, the basic residues are distributed so that the first three basic residues are on one face of the helix and the next three on the other face. The basic residues are thus distributed in two clusters with opposite orientation. This arrangement, which is generated by helical compaction, seems suited to the cross-linking of non-consecutive segments of DNA and could have a role in the organization of the nucleosomal or the linker DNA in chromatin.

The question is whether the interaction with DNA induces a significant increase in the helicity of the N-terminal peptides. IR spectroscopy shows that this is indeed the case. The percentages of α-helix in NH-1 and NH-2 in the complexes are 18-20% and 27-29%, respectively. These percentages are close to those observed in TFE solution and agree well with the different helical propensities of the two peptides. They reflect an important stabilization of the helical structure when compared with the low percentages present in water (6-7%) and demonstrate that the DNA displaces the coil (disordered) $\rightleftarrows$ helix equilibrium toward helix formation. As a low resolution technique, IR spectroscopy cannot, in general, specify which particular residues are in helical conformation, but, because coulombic interactions between basic residues and DNA phosphates promote helical compaction in model and natural peptides, it is likely that the minimal region with helical structure in NH-2 is that containing the basic cluster, between Lys11 and Lys20; however, as in TFE, the helical structure could extend to Asp23.

The peptide NE-1 comprises the positively-charged region of the N-terminal domain of H1e, which is the most abundant subtype in many mammalian somatic cells. In aqueous solution, NE-1 behaved as a mainly unstructured peptide. Addition of TFE resulted in a substantial increase in the helical content. In 90% TFE, the peptide was structured in two α-helices spanning from Thr17 to Ala27 (helix N-I) and from Gly29 to Thr34 (helix N-II), with fractional helicities of 54% and 44%, respectively (Figure 9B). Both helical elements have a marked amphipathic character, with all the basic residues on one face of the helix and the apolar residues on the other. The two helical elements are separated by a Gly-Gly motif, which behaves as a flexible linker between the helical regions. Gly-Gly motifs at equivalent positions are found in many vertebrate H1 subtypes. The wide range of relative orientations of the helical axes allowed by the Gly-Gly motif may facilitate the tracking of the phosphate backbone by the helical elements or the simultaneous binding of two non-consecutive DNA fragments in chromatin [47].

Summary

Although the structure of the nucleosome is well known, the further folding of nucleosome arrays of is poorly understood.

Despite recent progress, the structural role of linker histones also remains very unclear. There are several models for describing how histone H1 association affects DNA at the entry/exit of the nucleosome. H1 also makes contact with core histones that could modify the nucleosome core. Histone H1 presents a great variety of subtypes, differing largely in sequence and probably also in structure, which together with post-transcriptional modifications, offer great possibilities for functional modulation.

Acknowledgements

This work has been financed in part by the *Ministerio de Educación y Ciencia* (DGICYT, PB98-0896) and the *Generalitat de Catalunya* (SGR2001/00199).

References

[1] Zlatanova, J., and Van Holde, K., 1992. Histone H1 and transcription: still an enigma? *J. Cell Sci.* 103:889-895.

[2] Khochbin, S., and Wolffe, A.P., 1994. Developmentally regulated expression of linker-histone variants in vertebrates. *Eur. J. Biochem* 225:501-510.

[3] Wolffe, A.P., Khochbin A.P.S. and Dimitrov S. 1997. What do linker histones do in chromatin? *BioEssays* 19:249-255.

[4] Bouvet, P., Dimitrov, S., and Wolffe, A.P. 1994. Specific regulation of *Xenopus* chromosomal 5S rRNA gene transcription in vivo by histone H1. *Genes Dev.* 8:1147-1159.

[5] Shen, X. and Gorovsky, M.A. 1996. Linker histone H1 regulates specific gene expression but not global transcription in vivo. *Cell* 86:475-483.

[6] Vermaak, D., Steinbach, O.C., Dimitrov, S., Rupp, R.A.W. and Wolffe A.P. 1988. The globular domain of histone H1 is sufficient to direct specific gene repression in early *Xenopus* embryos. *Curr. Biol.* 8:533-536.

[7] Lee, H.L. and Archer, T.K. 1998. Prolonged glucocorticoid exposure dephosphorylates histone H1 and inactivates MMTV promoter. *EMBO J*: 17:1454-1466.

[8] Dou, Y., Mizzen, C.A., Abrams, M., Allis, C.D. and Gorovsky, M.A. 1999. Phosphorylation of linker histone H1 regulates gene expression in vivo by mimicking H1 removal. *Mol. Cell* 4:641-647.

[9] Parseghian, M.H. and Hamkalo, B.A. 2001. A compendium of the histone H1 family of somatic subtypes: An elusive cast of characters and their characteristics. *Biochem Cell Biol.* 79:289-304.

[10] Panyim, S. and Chalkley, R. 1969. A new histone found only in tissues with little cell division. *Biochem. Biophys. Res. Commun.* 37:1042-1049.

[11] Bucci, L.R., Brock, W.A. and Meistrich, M.L. 1982. Distribution and synthesis of histone I subfractions during spermatogenesis in the rat. *Exp. Cell Res.* 140:111-118.

[12] Lennox, R.W. 1984. Differences in evolutionary stability among mammalian H1 subtypes. *J. Biol. Chem.* 259:669-672.

[13] Banchev, T., Sebreva, J., Zlatanova, J. and Tsanev, R., 1988. Immunofluorescent localization of histone H1^{o} in the nuclei of proliferating and differentiating Friend cells. *Exp. Cell Res.* 177:1-8.

[14] García-Segura, L.M., Luquín, S., Martínez, P., Casas, M.T. and Suau, P. Differential expression and gonadal hormone regulation of histone H1^{o} in the developing and adult rat brain. *Dev. Brain Res.* 73:63-70.

[15] Parseghian, M.H.D., Harris, D.A., Rishwain, D.R. and Hamkalo, B.A. 1994. Characterization of a set of antibodies specific for three human histone H1 subtypes. *Chromosoma* 103:192-208.

[16] Ponte, I., Martínez, P., Ramírez, A., Jorcano, J.L., Monzo, M. and Suau, P. 1994. Transcriptional activation of histone H1^{o} during neuronal terminal differentiation. *Dev. Brain Res.* 80:35-44.

[17] Lennox, R.W., Oshimam, R.G. and Cohen, L.H. 1982. The H1 histones and their interphase phosphorylated states in differentiated and undifferentiated cell lines derived from murine teratocarcinomas. *J. Biol. Chem.* 257:5183-5189.

[18] Liao, L.W. and Cole, R.D. 1981. Differences among fractions of H1 histones in their interactions with linear and superhelical DNA: circular dichroism. *J. Biol. Chem.* 256:10124-10128.

[19] Kadake, J.R. and Rao, M.R.S. 1995. DNA- and chromatin-condensing properties of rat testes H1a and H1t compared to those of rat liver H1bdec: H1t is a poor condenser of chromatin. *Biochemistry* 34:15792-15801.

[20] Clark, D.J., Hill, C.S., Martin, S.R. and Thomas, J.O. 1988. α-helix in the carboxy-terminal domain of histones H1 and H5. *EMBO J.* 7:69-75.

[21] Piña, B., Martínez, P. and Suau, P. 1987. Changes in H1 complement in differentiating rat-brain cortical neurons. *Eur. J. Biochem.* 164: 71-76.

[22] Piña, B., Martínez, P., Simón, L. and Suau, P. 1984. Differential kinetics of histone H1ºaccumulation in neuronal and glial cells from rat cerebral cortex during postnatal development. *Biochem. Biophys. Res. Commun.* 123:697-702.

[23] Piña, B. and Suau, P. 1987. Changes in the proportions of histone H1^{o} subtypes in brain cortical neurons. *FEBS Lett.* 210: 161-164.

[24] Jacobson, M., Developmental Neurobiology, Plenum Press, New York, 1991.

[25] Domínguez, V., Piña, B. and Suau, P. 1992. H1 subtype synthesis in neurons and neuroblasts. *Development* 115:181-185.

[26] Ponte, I., Martínez, P., Ramírez, A., Jorcano, J.L., Monzó, M. and Suau, P. 1994. Transcriptional activation of histone H1^{o} during neuronal terminal differentiation. *Dev. Brain Res.* 80:35-44.

[27] Gjerset, R., Gorka, C., Hasthorpe, S., Lawrence, J.J. and Eisen, H. 1982. Developmental and hormonal reg-

ulation of protein H1° in rodents. *Proc. Natl. Acad. Sci. USA.* 79:2333-2337.

[28] Lafarga, M., García-Segura, L.M., Rodríguez, J.R. and Suau, P. 1995. Expression of histone H1° in transcriptionally-activated supraoptic neurons. *Mol. Brain Res.* 29:317-324.

[29] Ponte, I., Vidal-Taboada, J.M. and Suau, P. 1998. Evolution of the vertebrate H1 histone class: evidence for the functional differentiation of the subtypes. *Mol. Biol. Evol.* 15:702-708.

[30] Morán, F., Montero, F., Azorín, F. and Suau, P. 1985. Condensation of DNA by the C-terminal domain of H1. A circular dichroism study. *Biophys. Chem.* 22:125-129.

[31] Allan, J., Mitchell, T., Harborne, N., Bohm, L. and Crane-Robinson, C. 1986. Roles of H1 domains in determining higher-order chromatin structure and H1 location. *J. Mol. Biol.* 187:591-601.

[32] Rodríguez, A.T., Pérez, L., Morán, F., Montero, F. and Suau, P. 1991. Cooperative interaction of the C-terminal domain of histone H1 with DNA. *Biophys. Chem.* 39:145-152.

[33] Ponte, I., Monsalves, C., Cabañas, M., Martínez, P. and Suau, P. 1996. Sequence simplicity and evolution of the 3' untranslated region of the histone H1° gene. *J. Mol Evol.* 43:125-134.

[34] Renz, M. and Day, L.A. 1976. Transition from non-cooperative to cooperative and selective binding of histone H1 to DNA. *Biochemistry*. 15:3220-3228.

[35] Watanabe, F. 1986.Cooperative interaction of histone H1 with DNA *Nucleic Acids Res.* 14:3573-3585

[36] Clark, D.J. and Thomas, J.O. 1986. Salt-dependent cooperative interaction of histone H1 with linear DNA. *J. Mol. Biol.* 187:569-580

[37] Clark, D.J. and Thomas, J.O. 1988. Differences in the binding of H1 variants to DNA. Cooperativity and linker-length related distribution. *Eur J Biochem.* 178:225-233

[38] Manning, G.S. 1979. Theory of H1-mediated control of higher orders of structure in chromatin. *Biopolymers* 18:2929-2942

[39] Wilson, R.W. and Bloomfield, V.A. 1979. Counterion-induced condensation of deoxyribonucleic acid. a light-scattering study. *Biochemistry* 18:2192-2196

[40] Hill, C.S., Martin, S.R. and Thomas, J.O. 1989. A stable alpha-helical element in the carboxy-terminal domain of free and chromatin-bound histone H1 from sea urchin sperm. *EMBO J.* 8:2591-2599.

[41] Vila, R., Ponte, I., Jiménez, M.A., Rico, M. and Suau, P. 2000. A helix-turn motif in the C-terminal domain of histone H1. *Protein Sci.* 9:627-636.

[42] Suzuki, M. 1989. SPKK, a new nucleic acid-binding unit of protein found in histone. *EMBO J.* 8:797-804.

[43] Vila, R., Ponte, I., Collado, M., Arrondo, J.L.R. and Suau, P. 2001. Induction of secondary structure in a COOH-terminal peptide of histone H1 by interaction with the DNA. *J. Biol Chem.* 276:30898-30903.

[44] Bohm, L. and Mitchell, T.C. 1985. Sequence conservation in the N-terminal domain of histone H1. *FEBS Lett.* 193:1-4.

[45] Vila, R., Ponte, I., Collado, M., Arrondo, J.L.R., Jiménez, M.A., Rico, M. and Suau, P. 2001. DNA-induced α-helical structure in the NH_2-terminal domain of histone H1. *J. Biol. Chem.* 276:46429-46435.

[46] Mirzabekov, A.D., Pruss, D.V. and Ebralidse, K.K. 1990. Chromatin superstructure-dependent crosslinking with DNA of the histone H5 residues Thr1, His25 and His62. *J. Mol Biol.* 211:479-491.

[47] Ramakrishnan, V., Finch, J.T., Graciano, V., Lee, P.L. and Sweet, R.M. 1993. Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature* 362:219-223.

[48] Vila, R., Ponte, I., Jiménez, M.A., Rico, M. and Suau, P. 2002. An inducible helix-Gly-Gly-helix motif in the N-terminal domain of histone H1e: A CD and NMR study. *Protein Sci.* 11:214-220.

About the authors

The authors belong to the Department of Biochemistry and Molecular Biology of Autonomous University of Barcelona.

Prof. P. Suau graduated in Sciences (Biology) at the University of Barcelona. He obtained his PhD from the same University in 1975, working at Departamento de Química Macromolecular de la ESIIB. After four years spent in different laboratories in France and the UK, working on chromatin structure and protein-nucleic acid interactions, he joined the UAB in 1979.

Dr. I. Ponte graduated in biology at the University of Barcelona. She obtained her PhD from the same university, working at the Consejo Superior de Investigaciones Científicas, and joined the UAB in 1991.

Dr. R. Vila obtained his PhD at the UAB under the supervision of I. Ponte and P. Suau.

The group's activity is focussed on the study of chromatin structure and gene expression. A second area of interest is the cloning and regulation of glutamate receptor subunits.